Evaluation and Comparison of Habitat and Resource Equivalency Analysis as Used to Conduct OPA NRDA

JUNE 2022

American
Petroleum
Institute

Special Notes

API publications necessarily address problems of a general nature. With respect to particular circumstances, local, state, and federal laws and regulations should be reviewed. The use of API publications is voluntary. In some cases, third parties or authorities having jurisdiction may choose to incorporate API standards by reference and may mandate compliance.

Neither API nor any of API's employees, subcontractors, consultants, committees, or other assignees make any warranty or representation, either express or implied, with respect to the accuracy, completeness, or usefulness of the information contained herein, or assume any liability or responsibility for any use, or the results of such use, of any information or process disclosed in this publication. Neither API nor any of API's employees, subcontractors, consultants, or other assignees represent that use of this publication would not infringe upon privately owned rights.

Users of this standard should not rely exclusively on the information contained in this document. Sound business, scientific, engineering, and safety judgment should be used in employing the information contained herein.

API is not undertaking to meet the duties of employers, manufacturers, or suppliers to warn and properly train and equip their employees, and others exposed, concerning health and safety risks and precautions, nor undertaking their obligations to comply with authorities having jurisdiction.

API publications may be used by anyone desiring to do so. Every effort has been made by the Institute to ensure the accuracy and reliability of the data contained in them; however, the Institute makes no representation, warranty, or guarantee in connection with this publication and hereby expressly disclaims any liability or responsibility for loss or damage resulting from its use or for the violation of any authorities having jurisdiction with which this publication may conflict.

API publications are published to facilitate the broad availability of proven, sound engineering and operating practices. These publications are not intended to obviate the need for applying sound engineering judgment regarding when and where these publications should be utilized. The formulation and publication of API publications is not intended in any way to inhibit anyone from using any other practices.

Any manufacturer marking equipment or materials in conformance with the marking requirements of an API standard is solely responsible for complying with all the applicable requirements of that standard. API does not represent, warrant, or guarantee that such products do in fact conform to the applicable API standard.

Contents

Page

Executive Summary ... 1
Acronyms ... 3
1 Introduction ... 4
1.1 Oil Spill-related NRDA Defintions and Concepts ... 5
1.2 A Perspective on Restoration Scaling Methods ... 6
1.3 Reading and Using this Report ... 6
2 An Overview of REA and HEA ... 7
2.1 Hypothetical Spill Scenario ... 7
2.2 REA for Hypothetical Spill Scenario ... 9
2.3 Habitat Equivalency Analysis ... 17
3 The Relationship of REA and HEA to Economic Principles ... 25
3.1 Why Practitioners Should Understand the Economic Foundation of REA and HEA ... 25
3.2 The Basic Economics of Compensation ... 26
3.3 REA, HEA, and Economics ... 27
3.4 Practical Implications of the Link between Economics and REA and HEA ... 47
4 Resource Equivalency Analysis in Depth ... 48
4.1 Implementing REA ... 49
4.2 Evaluating Alternative Approaches for Implementing REA ... 54
4.3 Common REA Issues and Sensitivities ... 59
4.4 REA in Cooperative versus Litigation Settings ... 60
4.5 REA Summary ... 60
5 Habitat Equivalency Analysis in Depth ... 61
5.1 Typically Available Information ... 66
5.2 Estimating HEA Injury ... 68
5.3 Estimating HEA Credit ... 77
5.4 Common HEA Pitfalls ... 79
5.5 Sensitivities ... 82
5.6 Cooperative versus Litigious HEA ... 86
5.7 HEA Summary ... 86
6 REA and HEA: A Comparison of Strengths, Weaknesses, and Appropriate Uses ... 87
7 Emerging Issues ... 89
7.1 The Habitat-Based Resource Equivalency Method ... 89
7.2 Landscape-Level Effects ... 91
7.3 Discounting in HEA and REA ... 92
7.4 The Treatment of Uncertainty in REA and HEA ... 96
7.5 Managing NRDA Costs: A Decision Approach ... 105
7.6 Climate Change ... 110
7.7 A Better Framework for Cooperative Assessment? ... 111
8 References ... 112
Annex A—Comprehensive Bibliography ... 117
Annex B—Quick Reference Summary of HEA and REA Models ... 126
Annex C—ENTRIX Unpublished Report to API: Assessment of Beached Bird Modeling Methods ... 117

Contents

Page

Figures

Ex. 1 Graphic Representation of REA 1
Ex. 2 HEA Illustrated as Services Flowing from Two Discrete Sites 2
2-1 Graphic Representation of REA 10
2-2 Illustration of Iterative Process Used to Scale Restoration 14
2-3 HEA Iillustrated as Services Flowing from Two Discrete Sites 18
3-1 Two Indifference Curves Reflecting Baseline and Injured Bundles of Goods and Services 30
3-2 The Value of Ecological Services and Compensation for an Oil Spill 31
3-3 Economic Value-to-value Scaling of Compensatory Restoration 33
3-4 Valuing the Effects of a Spill: REA and HEA Versus Value-to-value Scaling 34
3-5 Value-to-value Scaling of Compensatory Restoration with Multiple Time Periods 42
3-6 Linear Approximations to the Value Curve 43
4-1 Graphic Representation of REA 49
4-2 Predictive Power of REA Approaches as a Function of Extinction Risk 59
5-1 HEA Illustrated as the Regional Provision of a Composite Service 61
5-2 HEA Illustrated as Composite Service Flowing from Two Discrete Sites 62
5-3 Relationships Between Service Indicators, Services, and the Composite Service 64
5-4 HEA Debit Relative to a Decreasing Baseline Service Level 65
5-5 Example Grouping of Shoreline Habitat Types for NRDA Use 68
5-6 Schematic of a Typical OPA NRDA Conducted via HEA 69
5-7 Similar Projects Generate Different Service Increments When Base Acres are Mismatched 81
5-8 Increasing Recovery Time Results in a Modest Increase in Debit When the Initial Service Loss is Small 83
5-9 Increasing Recovery Time Results in a Large Increase in Debit When the Initial Service Loss is Large 84
5-10 The Shape of the Recovery Curve Influences the Importance of Loss and Recovery Assumptions 85
7-1 Monte Carlo Distribution of NRD Probabilities 98
7-2 The Relationship Between Utility of Services and the Amount of Services in a Representative Time Period 102
7-3 The Cost of Information 108
7-4 Optimizing a Study 109
B-1 Graphic Representation of REA 126
B-2 Graphic Representation of HEA 127

Tables

2-1 Data Required to Project the With-Spill-and-Restoration Population Level through Time 12
2-2 Summary of Wetland Functions and Services 19
2-3 HEA Debit Calculations 22
2-4 Acre Years of Credit 23
4-1 Direct Injury as Estimated via Static Arithmetic 51
4-2 Indirect Injury Associated with the F1 Generation 52
4-3 Input Parameters for Density-independent Scenarios 57
4-4 Input Parameters for Density-dependent Scenarios 57
4-5 Predictive Power of REA Approaches 58
5-1 Data Commonly Available for Use in HEA 67
5-2 Summary of Services and Impacts Described by Trustees 80

Evaluation and Comparison of Habitat and Resource Equivalency Analysis as Used to Conduct OPA NRDA

Executive Summary

Resource Equivalency Analysis (REA) and Habitat Equivalency Analysis (HEA) are models designed to identify the amount of habitat restoration that compensates the public for ecological injuries caused by, in the scope of this report, an oil spill. They are a component of almost all Natural Resource Damage Assessments (NRDAs) conducted under the Oil Pollution Act of 1990 (OPA). Sections 1 and 2 of this report provide an introduction to OPA NRDA and introduce, by way of example, the basic use of REA and HEA. The example, which is intended to mimic an actual OPA NRDA, also introduces several of the key caveats associated with the application of REAs and HEAs in a real-world setting.

REA is most commonly used when individual members of a population of animals experience spill-related mortality and/or a spill-related reduction in reproduction. The basic premise underlying REA is that if a spill results in the loss of individuals through mortality and/or a reduction in reproduction, the public can be compensated via a restoration project that creates individuals that otherwise would not have existed (Figure Ex. 1). When the “with-spill-and-restoration” population projection is below the baseline, a debit accumulates. When the “with-spill-and-restoration” population projection exceeds the baseline, a credit accumulates. Injuries are identified as fully compensated (i.e., the restoration project is “scaled”) when the debit (the discounted value of the orange area in Figure Ex. 1) is equal to the credit (the discounted value of the green area in Figure Ex. 1).

Multiple ways to implement REA have been developed. These differ in their technical rigor and can result in damage estimates that differ by millions of dollars. Sections 3 and 4 of this report use economic and ecological theory to identify and discuss theoretically appropriate approaches to REA and compare and contrast alternative approaches to that standard that have been applied in actual cases. Even within the confines of a theoretically-appropriate REA, there exist several issues related to the calculation of a net change in population level, impacts to nuisance species, and impacts to species with non-zero extinction risks that may require event-specific adjustments. These issues are discussed in Section 7.

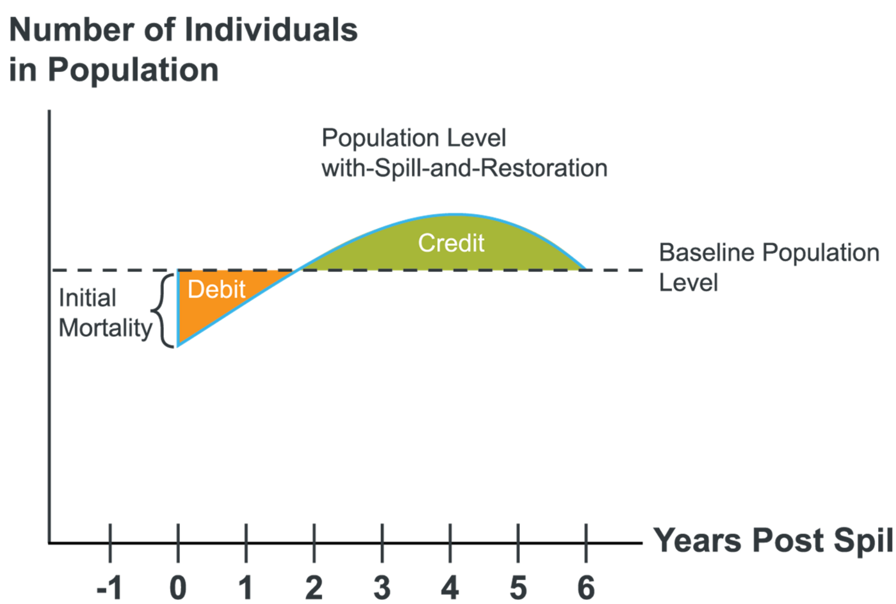


Figure Ex. 1—Graphic Representation of REA

HEA is most commonly used when habitats are impacted and a similar habitat will be restored. The basic premise underlying HEA is that all of the various services flowing from a habitat can combined into a single composite service. If a spill results in a reduction in the quantity of the composite service produced by the impacted habitat, compensation can be achieved via a restoration project that creates an offsetting quantity of composite service that otherwise would not exist. HEA is often conceptualized as debits and credits flowing from two discrete sites. Debits accrue when the level of composite service provided by the injured site is below its baseline level; credits accrue when the level of composite service provided by the restored site is above its baseline level (Figure Ex. 2). Compensation is achieved when the debit (discounted value of the area represented in orange in Figure Ex. 2) is equal to the credit (discounted value of the green area in Figure Ex. 2).

In practice, there is no definitive way to define or measure the composite level of service flowing from a habitat. As a result, the specific approach taken to implementing HEA varies significantly across applications. That variation can generate damage estimates that differ by millions of dollars even when all assumptions related to ecology, biology, and toxicology are identical. As such, Sections 3 and 5 of this report use economic and ecological theory to identify and discuss the theoretical considerations associated with HEA. In addition to these theoretical considerations, several practical concepts related to double counting, out-of-kind restoration, the definition of baseline service levels, and the use of field data are discussed in Section 5 and Section 7.

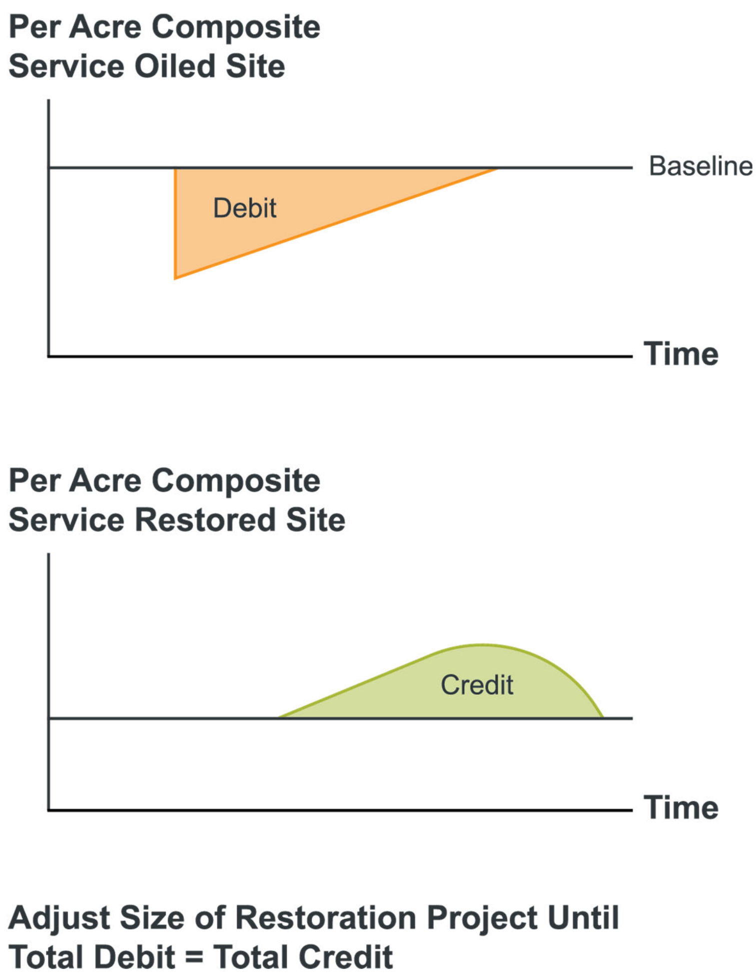


Figure Ex. 2—HEA Illustrated as Services Flowing from Two Discrete Sites

Most oil spill NRDAs use both REA and HEA to assess the range of injuries associated with an oil spill. There are no hard rules that allow a practitioner to identify the "best" way to integrate the two models. However, by combining an understanding of REA and HEA with the spill facts, informed decisions are possible. An approach to making such decisions is outlined in Section 6.

Acronyms

BCA	benefit-cost analysis
BCR	benefit-cost ratio
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act of 1980
CFR	Code of Federal Regulations
DA	decision analysis
DARP	damage assessment and restoration plan
DBY	discounted bird-year
DLY	discounted loon-year
DOI	U.S. Department of the Interior
DSAY	discounted service acre year
ENVOI	expected net value of information
EPF	ecological production function
EVOI	expected value of information
FEIS	final environmental impact statement
GDP	gross domestic product
GPS	global positioning system
HaBREM	Habitat-based Resource Equivalency Method
HEA	habitat equivalency analysis
LRM	logistic regression model
NOAA	National Oceanic and Atmospheric Administration
NRD	natural resource damage(s)
NRDA	Natural Resource Damage Assessment
OPA	Oil Pollution Act
PAH	polyaromatic hydrocarbon

PRP	potentially responsible parties
REA	resource equivalency analysis
RHV	relative habitat value
RP	responsible party
SAY	service acre year
SCAT	Shoreline Cleanup Assessment Technique
SRTP	social rate of time preference
SWF	social welfare function
SY	species years
trustees	designated natural resource trustee agencies
USFWS	U.S. Fish and Wildlife Service
WTA	willingness-to-accept
WTP	willingness-to-pay

1 Introduction

The Oil Pollution Act of 1990 (OPA), the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA), and a host of state regulations allow designated natural resource trustee agencies ("trustees"), acting on behalf of the public, to recover natural resource damages (NRDs) to compensate for injuries to natural resources that occur as the result of unpermitted releases of oil or hazardous substances. Under these statutes, funds collected from parties potentially responsible for the spill or release (potentially responsible parties or "PRPs") may only be used for natural resource restoration actions and to reimburse trustees for "reasonable" assessment costs. The process of determining the appropriate amount of restoration is known as Natural Resource Damage Assessment (NRDA). Natural resource services can be broadly categorized as ecological services and human use services.

Two models, resource equivalency analysis (REA) and habitat equivalency analysis (HEA), are the primary tools used to guide practitioners through the NRDA process for ecological services. REA is most commonly used when individual members of a population of organisms (usually a species of bird, turtle, marine mammal, or fish, but sometimes also an amount of ground water) experience spill-related mortality and/or a spill-related reduction in reproduction. For example, REA might be used to determine how many osprey nesting platforms are required to compensate the public for the spill-related death of 10 ospreys. In contrast, HEA is used when habitats are impacted as the result of an oil spill. For example, HEA might be used to determine how many acres of mangroves must be planted to compensate the public if 25 acres of mangrove habitat are lightly oiled.

This report focuses on the use of REA and HEA when assessments are conducted under OPA. In this setting, the models are typically implemented with input from both the responsible party (RP) and the trustees. It has been suggested that, when used in this manner, REA and HEA need no strong technical underpinnings; if they are useful in reaching settlements, they have served a useful function.

However, to the extent that REA and HEA form the basis of any NRD settlement, whether that settlement is reached cooperatively or not, they also introduce technical considerations into the process. This technical rigor is the only means of assuring the public that the NRDA process has identified an appropriate level of restoration, and it may be required when the trustees seek judicial approval of the settlement or if the settlement is challenged in court. Thus, the technical aspects of the models, their proper uses, and potential abuses are important for NRDA practitioners to understand. Naturally, if presented in litigation, the technical basis for the analysis will receive considerable scrutiny.

Our purpose in writing this report is to provide a general understanding of REA and HEA, including their origins, relationship to economic methods, and application in real-world spill settings. The remainder of this introduction discusses a series of spill-related definitions and concepts, provides a broad perspective on the use of REA and HEA in OPA NRDA, and acts as a guide to the remainder of the report.

1.1 Oil Spill-related NRDA Definitions and Concepts

When conducting a NRDA pursuant to OPA, the basic approach is to identify the type and scale of natural resource restoration that, when implemented, results in no net loss of ecological services as a result of the spill. NRD liability, measured in dollars, is the cost of implementing the scaled restoration projects plus trustees' assessment costs.

The process begins with **injury determination**, during which practitioners determine whether an observable or measurable adverse change in a natural resource or an impairment of a natural resource service has resulted from the incident.[1]

Injury quantification involves measuring the amount of any injuries relative to baseline. **Baseline** is the condition of natural resources and the level of ecological services that would have been provided had the incident not occurred.[2] The term **ecological services** refers to functions performed by a natural resource for the benefit of another natural resource and/or the public (15 Code of Federal Regulations [CFR] § 990.30); these are sometimes called natural resource services and/or ecosystem services.

There are two types of restoration actions that form the basis for NRDs. **Primary restoration** serves to speed recovery of natural resources and their services to baseline. Natural recovery is also a form of primary restoration. **Compensatory restoration** compensates for the **interim loss** of natural resources and services that occurs between the date of the incident and the time of recovery to baseline. Primary restoration actions undertaken by trustees usually are conducted during the spill response, as determined in coordination with other response and/or regulatory agencies. When conducting an assessment of compensatory restoration requirements following an oil spill, response actions and primary restoration actions are taken into consideration. REA and HEA apply to compensatory restoration for injuries to ecological services.

Determining the "right" amount of compensatory restoration is referred to as **restoration scaling**, and there are three broad approaches to the task.

1) **Value-to-cost Scaling**—In this approach, the analyst uses economic methods to estimate the dollar value of injuries as the public's willingness-to-pay (WTP) to avoid the injury; then, trustees spend that amount of money on restoration. Value-to-cost scaling can result in overcompensation or undercompensation because some restoration projects provide highly valued ecological services to the public in a cost-effective manner whereas others do not.

---

1 The definition of injury is at 15 Code of Federal Regulations § 990.30; injury determination is specified at §990.51(a) and (b). Determining that an injury resulted from an incident requires showing exposure of the resource to discharged oil and a pathway linking the incident to the injuries.

2 Baseline can be estimated using historical data, reference data, control data, or data on incremental changes relative to an unspecified baseline (e.g., number of dead animals), alone or in combination, as appropriate. For an extensive discussion of methods used to estimate and nuances related to baseline see Desvousges et al (2018).

2) **Value-to-value Scaling**—In this approach, the analyst uses economic methods to estimate the type and amount of compensatory restoration that creates a public value (WTP) equal to the public value (WTP) of the lost ecological services. Value-to-value scaling is the scaling approach least subject to challenge *on conceptual grounds*, as it is based on accepted principles of economics. However, the details of specific value-to-value scaling applications, particularly as they relate to potential non-use values, have been the subject of significant challenges, both in academic literature and in court.

3) **Service-to-service Scaling**—In this approach, the economic valuation component of value-to-value scaling is dropped from the analysis, and the calculations rely on ecological metrics alone. Thus, rather than equating the *monetary value* of injuries to the *monetary value* of restoration, this method equates the *amount* of injury to the amount of restoration, where the "amount" is some ecological measure of services. REA and HEA (or generally "service equivalency" methods) are service-to-service scaling models.

The OPA NRDA regulations establish [§ 990.53(d)] that, if the injured and restored resources are of the same type and quality and of comparable value, trustees must consider the service-to-service methods. The regulations then state that if the conditions under which these methods are appropriate do not hold, trustees may use value-to-value scaling. If conditions are such that value-to-value scaling is indicated, and quantification of the economic value of lost services is practical but the quantification of the economic value of restored services is not, value-to-cost scaling may be used. Thus, the OPA regulations establish a hierarchy of desired restoration projects (as they provide services related to the injury) and associated scaling methods, with service equivalency methods ranked first, value equivalency second, and value-to-cost third.

1.2 A Perspective on Restoration Scaling Methods

Under OPA, scaling models are typically implemented with input from both the RP and the trustees. In this setting, the primary role of REA and HEA may be to provide a structure for organizing negotiations and identifying the types of agreements that need to be reached before restoration can be identified and right-sized. It has been suggested that, when used in this manner, REA and HEA need no strong technical underpinnings; if they are useful in reaching settlements, they have served a useful function.

Recognizing the pragmatic value of the preceding sentence, we also note that an absence of technical rigor can impede rather than expedite negotiations. Further, technical rigor is the key means of ensuring that the NRDA process has identified an appropriate level of compensatory restoration and the goals of OPA have been met. Finally, there is always the issue of defending an assessment should it be placed into evidence in mediation or before a court.

It is our perspective that giving thought to the immediate, practical, and theoretical defensibility of methodological choices embedded in REAs and HEAs is important in NRDA practice. A simple and *ad hoc* assumption *can* be a useful one, but if it is made without considering if and how it diverges from disciplined practice, it can undermine both its own expediting intent and public confidence in the NRDA process. Thus, the technical aspects of the models, their proper uses, and potential abuses are important for all NRDA practitioners to understand.

1.3 Reading and Using this Report

The remainder of this report is divided into six sections:

— Section 2 provides a primer on HEA and REA.

— Section 3 describes the origins of the REA and HEA models and their basis in economics.

— Section 4 describes REA in detail, outlining assumptions and data used to scale restoration.

— Section 5 describes HEA in detail, outlining assumptions and data used to scale restoration.

— Section 6 distinguishes REA and HEA by highlighting their strengths and weaknesses.

— Section 7 identifies and discusses emerging issues related to the use of REA and HEA.

The REA and HEA primer in Section 2 is intended largely for those who are not familiar with the basic intent and mechanics of these models. In the primer section, we introduce a hypothetical, but realistic, oil spill scenario that is carried through the document to provide context to the discussion. A condensed summary of the HEA and REA models is provided in Annex A.

Section 3 provides the NRDA practitioner insight into the theory that underlies REA and HEA. We believe this insight can help cut through many of the uncertainties that surround REA and HEA by focusing attention on the key uncertainties not directly addressed and/or bounded by theoretical considerations.

The remainder of the text addresses REA and HEA in depth by comparing the strengths and weaknesses of the models and identifying key uncertainties and emerging issues.

While the report is written so that it would be possible to read any one section in isolation, to some extent sections build upon one another and are written with the expectation that readers work their way through the text in order. Each section is discussed in the context of a cooperative assessment or settlement negotiations,[3] which almost all OPA NRDAs to date have been. However, we also identify issues relevant to litigation. Common pitfalls, sensitivity of calculated NRDs to these pitfalls, and strategies for guarding against errors and abuses are provided for REA and HEA in Section 4 and Section 5, respectively. Emerging Issues at the frontier of REA and HEA applications are reviewed in Section 7.

A bibliography of references is provided in Annex A for further reading.

2 An Overview of REA and HEA

"Most of us forget the basics and then wonder why the specifics don't work."

— Garrison Wynn

This section provides a basic understanding of REA and HEA as applied in a typical oil spill NRDA. The intent is to familiarize the reader with the basic versions of the models and their mechanics. The complications that arise in most real-world settings are alluded to, but the reader is directed to subsequent sections for more in-depth discussions.

The hypothetical spill described herein is among the simplest that might reasonably be encountered. The categories of injury are unambiguous, the level of interaction between those categories is minimal, and the assumptions that underlie both REA and HEA, as outlined in Section 3, are reasonably consistent with the facts of the spill.

2.1 Hypothetical Spill Scenario

On the evening of January 10, 2018, the dredge *Clay Thomas,* operating in rough seas, released 1,500 barrels of Number 6 fuel oil into the Atlantic Ocean approximately 0.5 mile off the North Carolina coast.

Strong winds blew the oil quickly to shore, stranding it above the normal high-tide line. The oil came ashore unevenly on a 5-mile stretch of sandy beach along a shoreline that forms the seaward border of a

---

3 We do not primarily discuss litigation, and in making comparisons across methods, we assume the same data are used. In general, our comments about defensibility of methods apply equally to analyses undertaken by trustees or RPs. Our discussion of common practices is what we have observed in cooperative assessments, with varying degrees of collaboration in data interpretation.

waterfowl refuge area. Oil also entered an adjacent wetland in which the average width of the oiling band was 45 feet and the degree of oiling was uniform throughout.

During the response, nearshore water samples were taken in the intertidal zone. When tested, they contained no dissolved oil constituents. However, sunken oil surveys did identify oil on shallow subtidal sediments.

The cleanup process began January 16 and continued until January 26. Pooled oil was removed using oil sorbent material, and some of the more heavily oiled wetland vegetation was cut and removed. Oiled sand was removed by hand, and the wrack line was removed with it. To ensure the safety of workers, the refuge immediately inland of the oiled beaches was closed to waterfowl hunting from January 17 to January 27.

In the days following the spill, wildlife responders noted oiled and unoiled common loons, mallard ducks, and double-crested cormorants. A total of 25 dead common loons, 35 dead double-crested cormorants, and 50 dead mallards were collected during daily searches of the spill area. The area is wintering ground for the endangered piping plover; four dead adults were found during response. Wildlife responders did not note any other concentrations of birds or wildlife (typical of the area in January); nor did they report observing other live oiled birds or wildlife.

The oiled beach is used as a plover nesting habitat from April through August. In the spring following the spill, piping plover survival and productivity data were collected in a manner consistent with historical monitoring protocols. When reproductive success among plovers using beaches in the affected area was compared to success among those using surrounding beaches, productivity (fledglings produced per nesting pair) appeared normal, as did the number and proportion of adults nesting.

The only recreational service impacted by the oil spill was due to the closure of the refuge to waterfowl hunters. The refuge manager reported that hunters are required to use one of 30 blinds, and that all blinds are full during the season. When the area reopened, hunting use immediately returned to pre-spill levels; 600 waterfowl hunting trips had been canceled due to the spill-related closure.

Based on this description, the resources of concern for the hypothetical *Clay Thomas* spill include the following:

a) shoreline oiling on sand beach habitats, including the impact of response activities;

b) shoreline oiling of wetland habitat, including the impact of response activities, such as cutting and removal of oiled vegetation and trampling;

c) oiling of subtidal sediments and the potential oiling of organisms living therein;

d) common loons, mallards, cormorants, and piping plovers experiencing spill-related acute mortality;

e) recreational impacts limited to the loss of waterfowl hunting opportunities due to the spill-related closure of the refuge;

f) largely due to their documented absence and/or low level of winter activity, other types of wildlife not subject to adverse impacts;

g) water samples indicating no risk to aquatic organisms from dissolved oil constituents in the water column in offshore or intertidal habitats.

The scale of compensatory restoration for spill-related impacts to bird populations will be estimated using REA. Compensatory restoration for impacts to the aquatic habitat, sand beach, and wetland habitat will be estimated using HEA. Monetary valuation methods will be used to assess the compensatory requirements associated with loss of recreational hunting opportunity.

2.2 REA for Hypothetical Spill Scenario

The *Clay Thomas* assessment would include separate REAs for the four bird species impacted: loons, mallards, plovers, and cormorants.

The basic premise underlying each REA is that the public can be compensated for a spill-related decrease in the population level with a restoration-related increase in the population level. This is typically achieved by increasing survival and/or productivity rates.

For example, one could compensate the public for spill-related mortality among common loons by deploying loon nesting rafts—which the literature suggests increase the reproductive success of the loon population—and in doing so, increase the number of loons in the population. REA answers the question, "How many nesting rafts need to be deployed?"[4]

The relevant metric in a loon REA would be a discounted loon-year (DLY), where one loon living for 1 year is said to provide one loon-year of service. Loon-years occurring in the future are discounted[5] to reflect the observation that society places different values on a good depending on when it is provided. The resulting unit is a DLY. When assessing mallards, the unit in the mallard REA would be a discounted mallard-year, and the same rationale would apply to plovers and cormorants.

Generally, the REA process for each species can be thought of as occurring in three steps:

1) The baseline population level (i.e., the number of individuals that would have been in the population but for the spill) is projected through time using a mathematical population projection model. This is represented by the black dotted line in Figure 2-1.
2) That same mathematical model is used to project the population level through time given the effects of the spill and the effects of a restoration project. This is represented by the solid blue line in Figure 2-1.
3) An iterative process is used to identify the size of the restoration project that, when implemented, ensures that society experiences no net loss of discounted species years.

The iterative process is often discussed in terms of debits and credits, as illustrated in Figure 2-1. When the "with-spill-and-restoration" population projection is below baseline, a debit accumulates. When the "with-spill-and-restoration" projection exceeds the baseline, a credit accumulates. In theory, full compensation occurs (i.e., the restoration project is "scaled") when the debit (the discounted value of the orange area in Figure 2-1) is equal to the credit (the discounted value of the green area in Figure 2-1).[6]

---

[4] This may mean implementing restoration away from the spill area; for example, loons killed in North Carolina may nest, and so be restored, in New England. The public was compensated for ruddy duck mortalities related to the Chalk Point oil spill in Maryland by preservation of prairie pothole breeding areas in North Dakota.

[5] Discounting is a process whereby the value of services that will be received in the future is reduced. For example, if a 3 percent discount rate is assumed, a loon-year occurring next year is equivalent to 0.97 loon-years occurring this year. A loon-year occurring 2 years into the future would be worth only 0.942 loon-years occurring this year.

[6] In HEA, it is common to separate the debit and credit calculations. Because habitat restoration is often off-site and does not have a primary restoration element, this separation can be sensible. The injured resource recovers to baseline independently of the restoration resource, and so a discrete debit can be calculated. This clear separation usually does not hold for REA. The restoration project is often implemented before the population recovers to baseline and, since it is directed to the injured population, it therefore has a primary restoration element. The debit cannot be calculated separately from restoration, and the iterative scaling process is necessary.

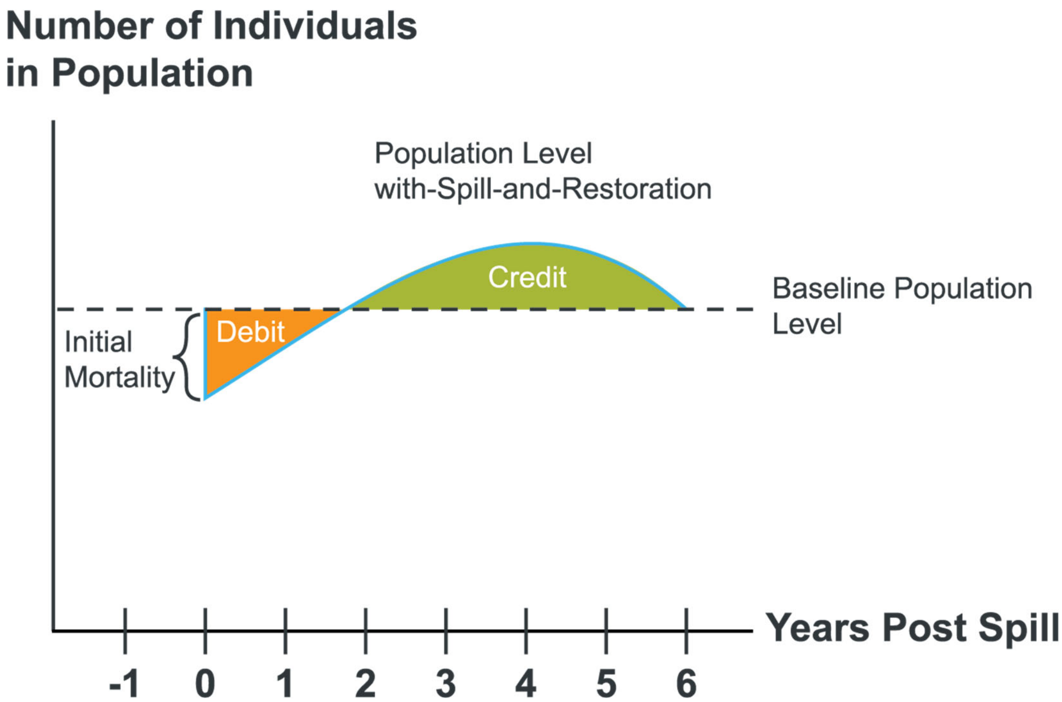


Figure 2-1—Graphic Representation of REA

However, in practice, the REA analyst must answer several questions: How does one project the baseline population level through time? What data and methods are used to estimate the spill-related level of initial mortality? Is it possible to predict how the population level will respond to the conditions that prevail after an oil spill, both before and after restoration is implemented?

General insight into these questions can be gained by reviewing the species-specific REAs that would likely be conducted for the hypothetical spill. In this section, we focus on the loon REA. At the end of the section, we point out complications that can arise when one considers REAs for the other species and direct the reader to other sections of the report where the complications are addressed in depth.

2.2.1 Common Loon REA

The first task in the loon REA is to define the population of loons at issue. Since the injury is to loons off the coast of North Carolina in January, we specify the relevant population to be the common loons that winter in the Northeast and Middle Atlantic.[7]

The second task is to project the baseline number of loons through time. Baseline is a critical issue in NRDA, as the amount of injury is calculated relative to baseline; changing the estimate of baseline has a direct impact on estimated NRDs. For many oil spill NRDAs, future population levels can be estimated by reviewing recent population trends and projecting the recently observed rate of increase or decrease into the future. This is a practical and acceptable path forward if there is no reason to believe that the current population growth rate will change over the relevant time frame[8]. Alternatively, population modeling can be used to predict the future population levels that would have prevailed under baseline conditions. The

7 There are a number of ways the population could be defined in a cooperative assessment; this is one example.

8 In many cases, practitioners make the simplifying assumption that baseline populations were in an approximate equilibrium and so calibrate underlying demographic parameters to imply zero population growth at baseline. As such, the baseline level itself need never be specified, only those positive or negative deviations from baseline resulting from restoration and injury, respectively. Unlike HEA applications, where differences related to baseline often result in divergent damage estimates, in our experience, material disagreements surrounding baseline population trajectories rarely emerge.

projected baseline depends on both the form of the model and the assumed path over time of population drivers.

In our hypothetical example, the common loon population is thought to be in a nest-site-limited equilibrium (i.e., the population has attained the highest level possible given the number and quality of nest sites available). Further, the population is expected to remain at that level indefinitely absent anthropogenic events such as oil spills and/or nest-site restoration.[9]

Third, the REA practitioner needs to estimate the magnitude of spill-related mortality (this is often referred to as acute mortality). In our hypothetical spill, 25 dead common loons were collected during the response. However, that number may not represent the number of common loons that actually died due to the spill. Some of the common loon carcasses collected may have been present on the shoreline whether or not there had been a spill; additionally, some spill-related common loon carcasses may have been consumed by scavengers before searchers could find them, and some may have been deposited in places where they could not be found. As such, it is common practice to estimate a carcass multiplier where:

$$\textit{Total Common Loon Collections} \times \textit{Carcass Multiplier} = \textit{Total Spill-related Common Loon Mortality}$$

While several methods[10, 11] have been used to estimate such multipliers, in this case, the historical average multiplier (4) is applied, and common loon mortality is estimated to be 25 × 4 = 100 common loons.

The fourth task is to construct a mathematical model to project the with-spill-and-restoration population through time. The model focuses on how survival and reproductive rates among the individuals that survived the spill will respond to post-spill conditions and restoration. That is, the model incorporates any delayed mortality due to oiling as well as any increase in survival rates due to density-dependent mechanisms and/or restoration. The model also incorporates any spill-related loss of reproduction and any increase in reproduction that may occur as the result of density-dependent mechanisms and/or restoration. Critically, the model used to project the with-spill-and-restoration population levels must focus on the individuals that survived the spill because it is the response of the “survivors” that dictates how the with-spill-and-restoration population level will evolve relative to baseline after the spill.

The model should satisfy two additional requirements. First, it should allow the practitioner to evaluate the impact of a restoration project (or projects) in various scales and/or project combinations. Second, the model should be consistent with the methods or assumptions made when the baseline population level was projected. Thus, either the REA should use the same population model to project both baseline and with-spill-and-restoration populations, or the practitioner must ensure the assumptions embedded in the baseline projection are not incompatible with the assumptions of the model used to project the with-spill-and-restoration population levels.

The technical details of constructing such a model are discussed in Section 4. The necessary inputs are identified in Table 2-1, along with input values for a hypothetical loon REA. Consistent with the assumption embedded in the baseline population projection, the model used to project the loon population levels with-spill-and-restoration assumes that the number of suitable nest sites limits population growth rates to zero at the population level that prevailed prior to the spill.

---

9 If the number of nest sites or fledges produced per nest exhibits a trend due to factors unrelated to the spill (for example, climate change), this may create an upward or downward trend in the baseline loon population. Issues such as this are discussed in Chapters 4 and 7.

10 These include beached-carcass modeling, live oiled bird (i.e., swept through) modeling, and benefits transfer. See Annex C for a detailed discussion.

11 Baseline will be discussed in more detail in Section 4 and in Section 7, where climate change is addressed.

Table 2-1—Data Required to Project the With-Spill-and-Restoration Population Level through Time

Parameters	Common Loon Estimate
Annual survival rates	76 percent from fledge to age 1; 80 percent from age 1 to 2 and from age 2 to 3; and 92 percent for age 3 plus
Maximum age of individuals in the population	30 years
Proportion of the population that is female	50 percent
Age at first breeding among females	6
Proportion of breeding age females reproducing annually	Limited by nest-site availability at the population level that prevailed in the year before and the year after the spill
Offspring produced annually per breeding female	0.62 fledges per breeding female

The fifth task is to develop potential restoration actions and quantity the effect they would have on with-spill survival and reproduction rates. For loons killed by the *Clay Thomas* spill, the baseline and post-spill populations are assumed to be limited by the number and quality of available nest sites. However, it is determined that more nest sites can be supplied by deploying loon nest rafts. Further, it is understood that females that otherwise would not have nested will use the nest rafts, and each one will produce, on average, 0.62 fledglings annually.

Next, the practitioner undertakes an iterative process to determine restoration scale. A starting point might be a project size with a hypothetical deployment of five loon nest rafts beginning 5 years after the spill and continuing for 30 years, and then a comparison with the with-spill-and-restoration population projection to the baseline projection. The size of the restoration project is then increased or decreased iteratively until the specified project compensates for the assumed injury. That is, the practitioner adjusts the size of the restoration project until the debit equals the credit.

Figure 2-2 illustrates the process. In this figure, the total length of the bar in any given year represents the net effect the spill and restoration have had on the common loon population in any given year. Red bars represent years when the with-spill-and-restoration population is smaller than it would have been under baseline. The darker portion of each red bar illustrates the discounted value of the missing loon-years. Green bars represent years when the with-spill-and-restoration population is bigger than it would have been under the baseline. The darker portion of each green bar illustrates the discounted value of those extra loon-years.

— The top panel illustrates the common loon shortfall when no restoration is implemented. In the spill year, the population is short 100 individuals (there were 100 acute mortalities), so the "year 0" debit is 100 DLYs.

  In the next year, the population is missing 86 individuals. The partial recovery to baseline occurs because of nest-site limitation. Given nest-site limitation, "each year after the spill the juvenile age class will be entirely replaced. That is, despite the fact that some breeding adults have been removed from the population, the population produces the same number of juveniles post-spill as it would have under baseline conditions" [California Department of Fish and Game (CDFG) et al. 2004]. As such, the year after the spill, there are as many young-of-year produced as there would have been at baseline. Since there is a 1-year lag from the spill date, the 86 loon-years still missing are discounted[12] 1 year, and the REA debit in "year 1" is equal to 86 × (1/1.03) = 83.4 DLYs.

In the next year, the population is missing only 76 individuals, as the first two age classes have now recovered to baseline levels. Discounting back to the year of the spill, the DLYs of debit in the second post-spill year is $76 \times 1/(1.03)^2 = 71.6$ DLYs. This process would continue until the population recovers to the baseline level 30 years post spill.

— The middle panel illustrates the effect of introducing five loon nesting rafts. These rafts are deployed in year 5 and maintained through year 30 (note the small inflection points in years 5 and 30). The rafts have the effect of adding nest sites for birds that otherwise would not have had a nest; this increases the number of young-of-year produced above baseline. This increase in reproduction relative to baseline causes the common loon population to: (1) recover to its baseline level more rapidly than it would have absent restoration; and (2) eventually increases population numbers to levels greater than would have prevailed under baseline. When the rafts are removed in year 30, the population slowly returns to its baseline level, which is dictated by the number and quality of naturally occurring nest sites.

When the length of all the dark red bars (the discounted loon-years of debit) is compared to the length of all the dark green bars (the DLYs of credit), a net debit is apparent. Thus, five loon rafts deployed in year 5 and maintained for 25 years thereafter is insufficient compensatory restoration.

— The bottom panel is similar to the middle panel, except 12 loon rafts have been deployed. The length of all the dark red bars (the DLYs of debit) is equal to the length of all the dark green bars (the discounted loon-years of credit). Thus, the REA suggests that deploying 12 loon rafts beginning 5 years after the spill and continuing for 25 years thereafter compensates the public for the spill-related loon mortalities.

12 The discount rate used in OPA NRDA is generally 3 percent. The present value in year T of a DLY occurring in year t is given by $DLY\ at\ T = (DLY\ at\ t) \times 1/(1+r)^{(T-t)}$, where r is the annual rate. Therefore, a DLY 1 year hence is "worth" (1/1.03) DLYs today. See Section 7.3 for a more detailed discussion of discounting.

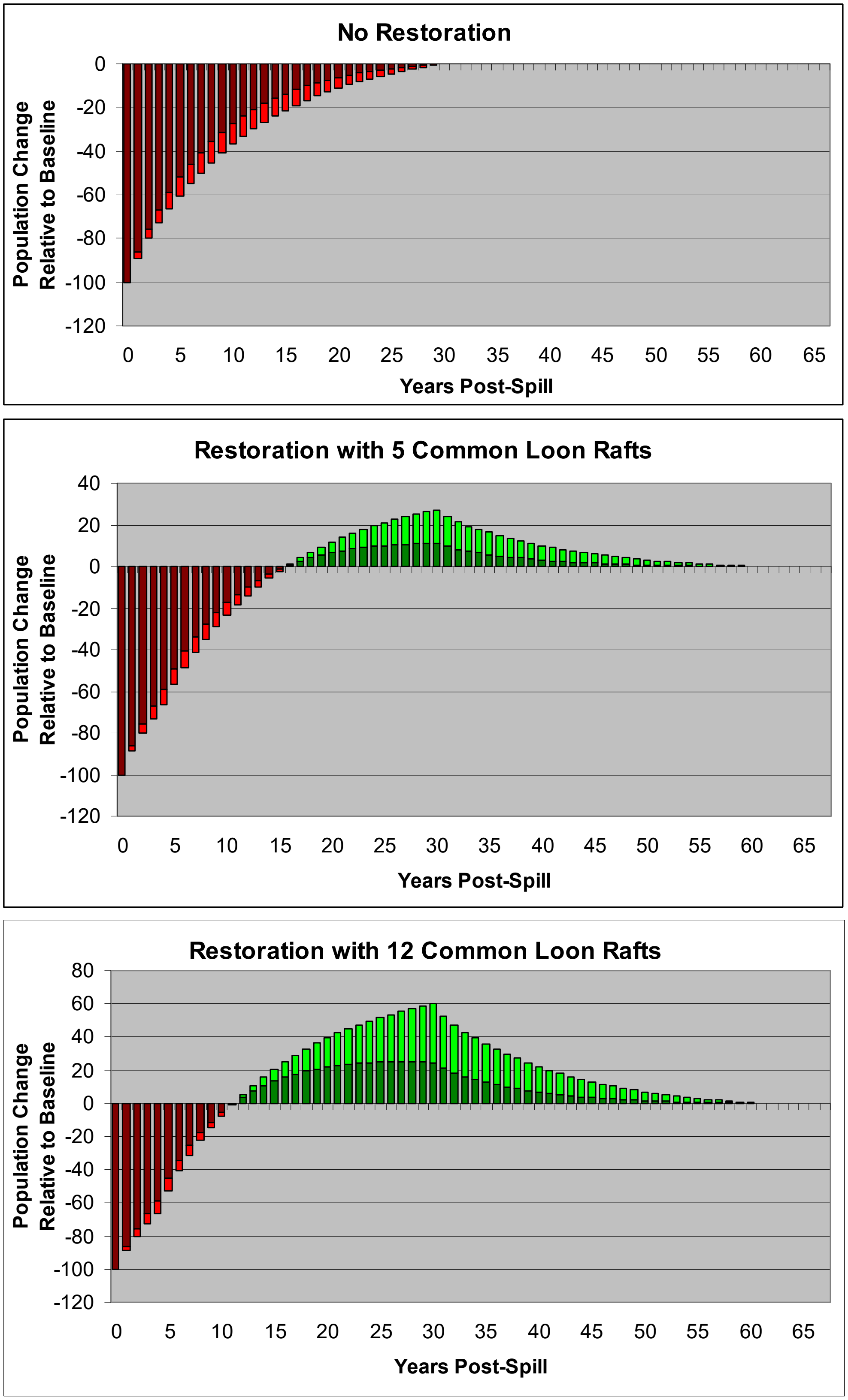


Figure 2-2—Illustration of Iterative Process Used to Scale Restoration

2.2.2 REA Implementation Issue: REA Approach

As discussed more in Section 4, the most difficult aspect of implementing REA, and the most important aspect in terms of ultimate liability, is selecting a basic computational approach. Generally, there are three options: specifying a recovery period based on professional judgment without formal modeling, what we call "static arithmetic", and population modeling. Two of these methods, professional judgment and static arithmetic, can be characterized as shortcuts that attempt to directly estimate debits and credits without projecting population levels. The third approach, population modeling (which was illustrated in Section 2.2.1), estimates debits and credits as the difference between internally consistent baseline and with-spill-and-restoration population projections.

The professional judgment approach tends to be relatively accurate if the injured populations are likely to recover within one or two breeding cycles (this may be the case for species that have relatively short lifespans, reproduce at an early age, and have many offspring per breeding attempt). Under the other scenarios we have tested, it appears that population modeling is the preferred REA approach as the static arithmetic approach, often employed by Trustees, are difficult to defend technically and tend to overestimate compensatory requirements.[13]

2.2.3 REA Implementation Issue: Net Effects

Consider the impact of the *Clay Thomas* spill on mallards. There are two routes by which the mallard population was affected by the spill. First, individual mallards died because of oiling. Second, hunting closures resulted in fewer mallards being harvested by hunters relative to the number that would have been harvested absent the spill. In this case, it is highly likely that the spill actually resulted in a net increase in the post-spill mallard population relative to baseline. As such, and as further discussed in Section 3, compensation for impacts to the mallard population would not be required.

This is not to say that oil spills are beneficial events. Rather, after the public is compensated for the loss of recreational services (in this case, waterfowl hunting opportunity), there is no residual reduction in services provided by the affected mallard population. Within the REA construct, if there is no reduction in service, the public requires no compensation (Wakefield and Davis, 2017).

Similar considerations arise for any species that is harvested commercially or recreationally. Further, the concept of "net" changes can be critical even when species are not harvested. Consider a population of colonially nesting seabirds that are nest-site limited. A spill could cause some adult mortality and cause some loss of production in the spill year. However, the spill could also cause some birds to move to an island that had not previously been colonized. This new site could become, over time, a new colony, thereby expanding the overall number of nest sites available to the population; that is, the spill has the effect of increasing carrying capacity. If the new colony would not have been established absent the spill, the net effects may be an *increase* in discounted bird-years (DBYs) provided by the with-spill population even without any compensatory restoration. Monitoring may be required in order to verify scaling assumptions.

2.2.4 REA Implementation Issue: Nuisance Species

The *Clay Thomas* spill resulted in mortality to double-crested cormorants. The 2003 Final Environmental Impact Statement (FEIS) for the management of double-crested cormorants (U.S. Fish and Wildlife Service [USFWS] and U.S. Department of Agriculture, 2003) identifies the double-crested cormorant as a species that is having a significant negative impact on both ecological processes and the human enjoyment of the environment. The FEIS further identifies as a preferred alternative that agencies within

---

[13] It has been suggested that simplified REA approaches will reduce total NRDA liability (the cost of implementing scaled restoration projects plus the trustees' assessment costs) if the reduction in transaction costs exceeds the increase in restoration costs. In our experience, and as illustrated in Section 4, this generally does not occur. The incremental cost associated with the population modeling approach to REA is nominal and the restoration overestimates associated with static arithmetic often represent significant costs.

the state of North Carolina operate under a public depredation order, which allows lethal control to reduce the adverse effect on double-crested cormorants. In 2016, the U.S. District Court for the District of Columbia vacated the depredation order on the grounds that the USFWS failed to consider a reasonable range of alternatives when extending the public depredation order through 2019. Since that time, the USFWS has issued individual take permits and continues to receive requests for increased take in response to growing conflicts (USFWS 2020).

As further discussed in Sections 3 and 4, compensation for reductions in the number of nuisance species, as determined within the technical construct of REA, should not be necessary. This is not to say that nuisance species do not produce any ecological services. However, on net, the services associated with one additional member of the population are negative. Thus, within the construct of a REA, the public does not require compensation when the population level of an invasive or nuisance species is reduced.[14]

2.2.5 REA Implementation Issue: Species of Special Concern and Uncertainty

As a result of the *Clay Thomas* spill, a number of piping plovers were killed; this species is federally listed as threatened. Applying REAs to listed species can be complicated in several ways. Some complications are procedural—for instance, there likely will be a Section 7 Consult[15] pursuant to the Endangered Species Act. Thus, some degree of analysis of injuries could be undertaken outside of the NRDA process, but that analysis may have considerable implications for the NRDA. More substantively, trustees have raised uncertainty related to extinction as a reason to implement highly conservative approaches to REA when applied to listed species. In one OPA case, they chose to truncate computation of credits at 25 years based on the logic that uncertainty prevented any projections beyond that time frame.

Uncertainty raises complex issues, which are further discussed in Sections 3, 4, and 7. However, the primary take-home message is that NRDA practitioners should recognize that most sources of uncertainty *affect the baseline and with-spill-and-restoration populations similarly*, and hence “cancel out” in REA calculations. Thus, for most forms of uncertainty, a REA based on deterministic population modeling gives answers very similar to those generated when National Oceanic and Atmospheric Administration (NOAA) guidance on formally incorporating risk and uncertainty into NRDA are followed. Hence, REA practitioners should be wary of ad hoc REA modifications designed to address “uncertainty,” particularly if those ad hoc approaches have any material impact on restoration estimates.

When species are in rapid decline for unknown reasons, it may not be possible to identify a compensatory restoration project. If this occurs, it may be appropriate to treat the species as one element of a habitat and address the overall injury using HEA.

2.2.6 REA Implementation Issue: Ecology, Biology, and Toxicology

The REA discussion in Section 2.2 has focused on issues associated with the basic mechanics of a REA. We have not addressed any of the issues associated with determination of the actual effects of spills and restoration on survival rates and reproductive success.

This approach was taken for two reasons. First, there is little uncertainty when it comes to identifying the “right” mechanical approach to REA that is consistent with economic and ecological theory. In contrast, ecological and economic theory cannot help identify the “right” approach for estimating potential changes to future survival and reproduction rates. That is, while much of the ecology, biology, and toxicology underlying OPA NRDA is necessarily idiosyncratic and uncertain, the NRDA community should be able to reach consensus as to the appropriate approach to modeling REA debits and credits. Second, while

---

14 It is not uncommon for trustees to include the injury associated with nuisance species and then restore a different, positively valued species. This has the feel of a supplemental environmental project under a penalty concept rather than compensation for service loss under NRDA.

15 Under Section 7 of the Endangered Species Act, federal agencies must consult with the USFWS when any action the agency carries out, funds, or authorizes may affect a listed endangered or threatened species or designated critical habitat.

uncertainty surrounding ecology, biology, and toxicology can translate into significant variation in NRD liability, in our experience, the variation in NRD liability associated with alternative (errant) REA mechanics often dwarfs the variation in NRD liability that may result from differing opinions related to ecology, biology, and/or toxicology.

2.3 Habitat Equivalency Analysis

HEA is most commonly used when habitats are impacted by oil and restoration of a similar habitat is anticipated. For the *Clay Thomas* spill, separate HEAs would likely be used to estimate compensatory restoration for the subtidal, sand beach, and wetland habitat impacts.

The basic premise underlying HEA is that all of the services flowing from a habitat can be aggregated into a single composite service. If a spill results in a reduction in the quantity of the composite service produced by the impacted habitat, compensation can be achieved via a restoration project that increases the amount of the composite service provided at some other site.

Focusing on any single habitat, HEA is often characterized in terms of debits and credits flowing from two discrete sites. Debits accrue when the level of composite service provided by the injured site is below baseline; credits accrue when the level of composite service provided by the restored site is above baseline (Figure 2-3). Compensation is achieved when the debit (present discounted value of the red area in Figure 2-3) is equal to the credit (present discounted value of the green area in Figure 2-3).

The unit of analysis in HEA is a service acre year (SAY) where the level of service provided by a base acre in 1 year is defined as 1 SAY. The injured and restored habitats are then judged relative to that base acre. Thus, if the base acre is a pristine site, a degraded acre of habitat may generate 50 percent services (or 0.5 SAYs); a more degraded site may provide 25 percent services (or 0.25 SAYs). SAYs occurring in the future are discounted to reflect the fact that society does not have the same value for a given service occurring in different time periods. The resulting unit is a discounted service acre year (DSAY).

But, how does a HEA practitioner project the service level provided by one site relative to another? What data and methods are used to estimate the spill-related initial loss of service? Is it possible to predict how a habitat's composite service level will respond to the conditions that prevail after an oil spill?

General insight into these questions can be gained by reviewing the HEAs that would likely be conducted for the hypothetical spill. For the *Clay Thomas* spill, we focus the primer on a HEA for wetlands. Some of the more complex issues that would be associated with a HEA applied to subtidal and/or sandy habitats are identified, and the reader is referred to subsequent sections for more in-depth discussion.

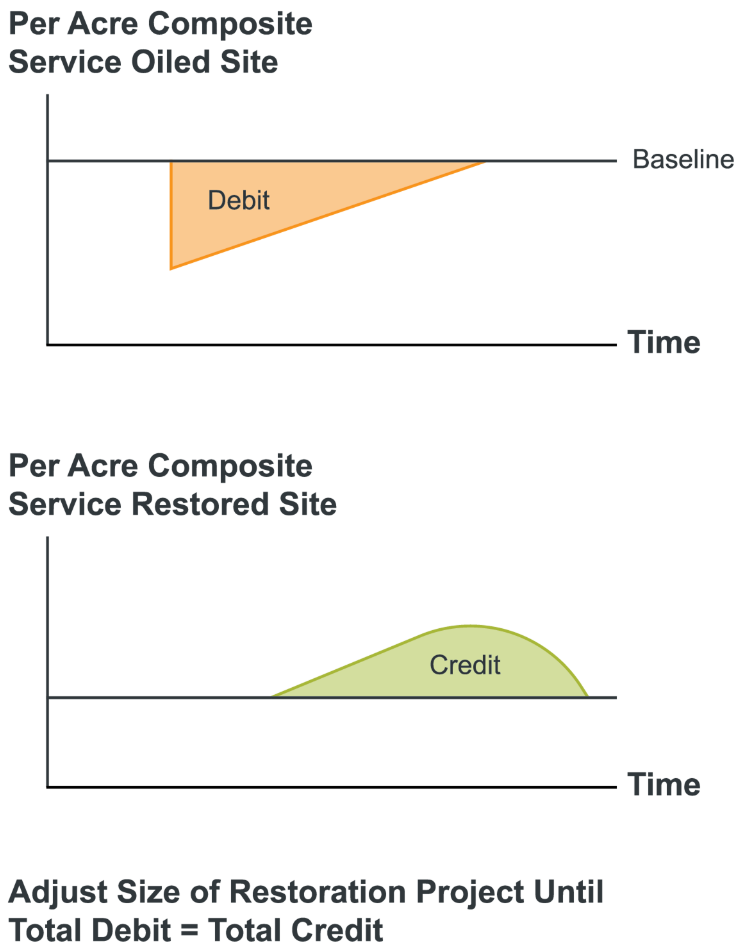


Figure 2-3—HEA Iillustrated as Services Flowing from Two Discrete Sites

2.3.1 Identifying and Aggregating Wetland Services

The first step in a HEA is often to identify the individual services that can be provided by the habitat at issue (in this case, wetland). Table 2-2 is a list of wetland services. The list, derived in part from both the Millennium Ecosystem Assessment (2005) and the National Research Council (2005), is provided to illustrate several issues and is not intended to be exhaustive.

In providing Table 2-2, a few notes are in order:

1) The services identified in Table 2-2 include more than “final ecosystem services,” which are only services directly valued by people; the table also includes intermediate services, which are those provided from one organism or habitat to another. They are included because, per the OPA NRDA regulations, intermediate services are to be included. The OPA regulations define services as “… the functions performed by a natural resource for the benefit of another natural resource and/or the public” (15 C.F.R. §990.30). The inclusion of intermediate services necessitates careful accounting of services across HEAs and REAs to ensure individual services are not double counted.

2) In the early stages of HEA, services are often conceptualized as the capacity of a habitat to contribute to the production of a service regardless of other factors. For example, a wetland may be able to sequester nutrients. Whether it actually does so will depend on whether nutrients are actually running off nearby uplands.

3) In many cases, a spill will cause some individual services to increase and others to decrease, and leave others unaffected. These differential effects necessitate reliance on the notion of a composite service.

Table 2-2—Summary of Wetland Functions and Services

Service	Function	Potential Service Indicator Metric
Habitat for biota	Marshes serve as physical habitats for organisms including birds, mammals, insects, fish, and invertebrates. The type and density of the vegetation is the primary determinant of species use.	Stem density, plant biomass, plant diversity, invertebrate abundance or diversity
Food web support	Primary production forms the base of the primary food web and the detrital food web.	Invertebrate abundance, community structure, community diversity, plant and/or invertebrate biomass
Sediment shoreline stabilization	Marsh vegetation stabilizes the soil and prevents erosion during normal tides, wave action, or storm events.	Plant biomass, stem density, below-ground biomass
Water filtration	Particles and nutrients are physically removed from water.	Plant biomass, stem density, soil character
Nutrient removal and transformation	Nutrients are converted to plant material, thereby reducing the occurrence of algal blooms and anoxic conditions.	Plant biomass, stem density, bacteria; meiofauna
Sediment/toxicant retention	Sediments and the toxicants bound to them are sequestered in wetlands. Wetlands encourage redox reactions that can detoxify many compounds.	Below-ground biomass, soil character
Carbon sequestration	Under many conditions, wetlands sequester carbon that would otherwise be released to the atmosphere.	Above-ground biomass, below-ground biomass, soil biomass, methanogenesis
Soil development and biogeochemical cycling	The soil is a living system that converts chemicals from one form to another and supports the growth of higher plants through biogeochemical cycling and the breakdown of detritus.	Plant biomass, stem density, soil character
Storm surge protection	Wetland habitat is a buffer between open waters and other habitats/infrastructure. Vegetation absorbs wave energy and wetland soils absorb water, which reduces impacts to inland habitat and property.	Stem density
Recreational and other human uses	Wetlands provide open natural space where humans recreate and may otherwise derive well-being.	Recreational trip counts

One of the key elements of any HEA, and one that can lead to significant sources of divergence in damage estimates among parties, is the method used to combine individual services into an aggregated service—that is, into a single composite service.

For example, some HEA practitioners may use above- and/or below-ground plant biomass as a proxy for the composite service level provided by a wetland. That is, they will base a wetland HEA entirely upon changes in plant biomass. Their logic may be that as the plants go, so goes the wetland and all of the services it provides. Other practitioners might focus on the abundance of the benthic community. Their logic might be that the invertebrate community integrates, and so reflects, things like soil quality, plant health, and hydrology, which are all factors that influence service levels. Still others may use functional assessment methods that formally integrate several metrics into a single measure of service provision.

While all approaches have their strengths and weaknesses, the key theoretical consideration is that the method used, whatever it may be, should represent the full suite of services provided by the wetland. That is, practitioners should not use a particularly sensitive or insensitive (and perhaps not representative of the overall effect) measure of a single wetland service as an indicator of the overall composite service level. See Section 5.2.2.2 for additional discussion on the use of service indicators.

In the case of the *Clay Thomas* spill wetland HEA, as is the case in many OPA HEAs, there will be no formal identification of a metric or equation that combines metrics used to quantify composite service provision. Instead, multiple lines of evidence will be combined to form a best professional judgment[16] as to changes in the level of composite service provision. However, to be clear, *all* such judgments are based on either an explicit (preferred) or implicit value weighting of changes to individual services based on relative service values.

2.3.2 Estimating Baseline Service Levels Relative to a Base Acre

In our hypothetical example, and in most OPA HEAs, the base acre is the injured site just prior to the spill. Because the level of service flowing from the injured site is not expected to have changed significantly over the relevant time horizon absent the spill, the baseline service level is assumed constant through time.[17] As such, baseline at the injured site is specified as a constant 100 percent.[18]

The restoration site in our hypothetical example is an abandoned industrial site covered with an invasive species of plant. The site has been judged to provide no valuable ecosystem services in its current state and is not expected to provide any valuable ecological services in the foreseeable future unless restoration is implemented. As such, the baseline level of service provided by the restored site is assumed to remain constant at zero percent.

In the hypothetical example above, services of the injured habitat are assumed to be a constant 100 percent at baseline. While this is common, it may not always be the case. Consider the example where a wetland has been degraded by historic impacts unrelated to the spill and there are pre-spill commitments that it will be enhanced in the future. If this habitat is injured by an oil spill before the proposed enhancements can be implemented, baseline is not a constant 100 percent, but instead increases through time from the pre-spill service level (degraded baseline) to the new, higher post-enhancement baseline according to some schedule. If the post-enhancement baseline is specified as eventually reaching 100 percent after some period of time during which services increase, then service losses associated with the spill are still being measured relative to the injured habitat at baseline, but the baseline increases through time.

Though less common, there may also be instances where services of both the injured and restored habitats are measured relative to some ideal or pristine condition representing 100 percent service provision. This could be the case where services are measured using a functional assessment model calibrated to a regional reference domain. In this situation, neither the injured habitat at baseline nor the restored habitat before or after restoration may achieve 100 percent services. In such cases, it may make

---

16 A former trustee referred to this as a gestalt-based HEA based on the idea that the mind “informs” what the eye sees by perceiving a series of individual elements as a whole, and this is done absent a conscious process. A recent paper by trustees and their consultants (Baker et al., 2020) characterizes the creation of a composite index in HEA as leading to “practitioner interference” in the NRDA process and proposes an alternative approach to HEA to avoid this process based on multiple REAs. The mode has the acronym HaBREM and is discussed in Section 7.

17 For historical contaminated sites, where time horizons are long, the baseline is a very complex issue. At these sites, the baseline cannot be observed for any recent time periods, which creates considerable opportunity for controversy.

18 One can imagine circumstances where baseline is changing in an OPA HEA application; this could be the case when development or restoration of adjacent lands could lead to a declining or improving baseline. Having a clear idea of the baseline is especially important when HEA analyses or assumptions from one spill are applied to another spill. A good example is the Athos I spill NRDA on the Delaware River. In this case, failure to make adjustments for different baselines led to considerable overestimation of restoration requirements.

sense to normalize service levels so that baseline services of the injured habitat are set at 100 percent (with the reference condition providing more than 100 percent services). Most practitioners find this approach to be the most intuitive.

Regardless of which methods are used, the examples above demonstrate why it is important for HEA practitioners to remain cognizant of how injured and restored services are conceptualized relative to one another, especially as it relates to baseline. More important than which approach is used (relative vs. absolute) is that the same approach be used for both injured services and restored services. Desvousges et al. (2018) provide additional insights into establishing baseline.

2.3.3 Identifying the Footprint of the Affected Area

The areal extent of shoreline oiling (i.e., the spill's physical footprint) is generally estimated using data describing visible oiling (e.g., slick maps, banding on shorelines or vegetation, subsurface/buried bands of oil, oil/sheen on water surface, and sheening of disturbed sediments). However, other data may also provide insight, such as cleanup records, buried-oil surveys, contaminant concentrations in environmental media (water, soil, sediments, biota), and chemical fingerprinting of the spilled oil.

Generally, the shoreline footprint will be broken out by habitat type and degree of impact. Shoreline type (wetland, sandy, rocky) is often taken from response records or from existing data, such as NOAA Environmental Sensitivity Index mapping. Degrees of impact are often categorized as light, moderate, and heavy, and are based largely on the degree of shoreline oiling as determined by the Shoreline Cleanup Assessment Technique (SCAT) process. Additional exposure categories may be created to address intertidal areas that were exposed to the oil before it stranded or areas that were only exposed to sheen or scattered tarballs, or may have received different treatments during the response (e.g., mechanical removal by heavy equipment versus removal by rakes and shovels).

In the *Clay Thomas* spill, the footprint is based on the 5-mile length of wetland oiling and 45-foot average width. This implies a 9-acre footprint with a uniform level of oiling throughout.[19]

2.3.4 Magnitude of Initial Service Loss

The HEA practitioner must estimate the magnitude of the initial service loss (i.e., the initial reduction in the provision of the composite service). If the practitioner has selected some field measurement or ecological index as a proxy for the composite service, the initial service loss can be calculated based on the difference between measurements in the impacted and reference areas.

When proxies are not used, HEA practitioners often have widely varying opinions as to how to estimate the magnitude of an initial service loss. Often this is not because of different data interpretation or missing data, but rather because they have very different definitions of what makes up the composite service.

For the *Clay Thomas* spill wetland HEA, a 100 percent initial service loss is assumed. This is a common assumption despite the observation that it is conservative, because many wetland services are generally not affected by oiling.

2.3.5 Recovery Curves

The HEA practitioner must project the level of composite service provided by the impacted site through time. The projection should consider the effects of response activities (if any) and natural attenuation. Several conceptual approaches have been used by NRDA practitioners; each has its own advantages and disadvantages. Importantly, the approach should embody three ideas. First, the recovery of the most severely or least impacted service is likely not a good indicator of changes in the composite service level. Second, recovery to baseline does not necessarily mean that an injured site has returned to the exact

---

[19] The treatment of the footprint of on-water oiling poses some difficulties; these are discussed in Section 5.

biophysical conditions that prevailed prior to the spill; recovery to baseline means that the site provides the same level of the composite service post-spill as it would have if the spill had not occurred. Finally, if specific indices or metrics were used to specify initial injury levels, those same indices or metrics should likely be used to project the composite service through time.

In our hypothetical spill, recovery and supporting assumptions are based on the following logic:

The initial service loss is assumed to be 100 percent; this loss decreased linearly to 60 percent over the first post-spill year. Thus, the average service loss during the first post-spill year is 80 percent of baseline. The relatively high level of loss was selected by trustees because, at one year following a spill of No. 6 fuel oil, it is not uncommon for oil to remain visible in the environment. Further, the results of a study in a Virginia marsh indicated that there would likely be a significant reduction in primary production 12 months after oiling.

The average service loss over the second post-spill year is assumed to be 50 percent of baseline. This would be consistent with the assumption that primary productivity would remain reduced. Mussels, crabs, and invertebrates would begin to return to the habitat, though their density would remain low.

The average service loss over the third post-spill year is assumed to be 30 percent of baseline as residual oil biodegrades or is transported out of the system and primary production returns to normal levels.

The average service loss over the fourth post-spill year is 15 percent of baseline as levels of primary production and densities of most invertebrate populations are approaching baseline.

From the fifth to the 10th year post-spill, the average service loss is 9, 7, 5, 3, 1, and 0 percent, respectively.

Table 2-3 reports the average SAYs of loss each year post-spill. It also reports the DSAYs of loss each year following the spill where all services are discounted back to the year of the spill using a 3-percent discount rate (for more on discounting, see Sections 3 and 7).

The loss per acre of habitat within the footprint is estimated to be 1.89 DSAYs. The total loss for the 9-acre footprint is therefore 17.01 DSAYs.

Table 2-3—HEA Debit Calculations

	Years Post-Spill										
Time period	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	More than 9	Total
SAYs lost per acre in footprint	0.8	0.5	0.3	0.15	0.09	0.07	0.05	0.03	0.01	0	2
DSAYs of loss per injured acre	0.79	0.48	0.28	0.14	0.08	0.06	0.04	0.02	0.01	0.00	1.89

2.3.6 Estimating Credit per Acre of Restoration

The credit side of the HEA model is where the increase in service provided by one unit (1 acre) of restoration is estimated. The size of the restoration project that ensures public compensation is calculated by dividing the total DSAYs of debit by the DSAYs of credit provided by 1 acre of restoration.

As was done on the debit side of the model, the increase in service is expressed relative to the base acre. Again, the focus is on the initial impact of restoration (which can be negative if restoration disturbs a site that is providing some positive level of service), and then on the change in the level of service provided through time. The issues related to the estimation of these parameters are not materially different than those related to the debit side.

The change in composite services are based on the following logic:

The abandoned industrial site covered with invasive vegetation provides no ecological services prior to restoration. As such, the construction phase, which begins at the outset of year 5, does not result in an initial reduction in services.

Over the first 10 years post-construction, services are assumed to increase linearly from 0 to 50 percent. This is consistent with the assumption that above-ground vegetation, which provides the majority of wetland services, develops relatively quickly, whereas the development of the invertebrate community and invertebrate-dependent wildlife may lag behind.

Over the next 10 years, services are assumed to increase linearly from 50 percent to 80 percent as the invertebrate community and invertebrate-dependent wildlife communities develop to their full potential.

The maximum service level is assumed to be 80 percent of the base acre because, except in rare cases, created wetlands rarely develop the highly organic soils and below-ground biomass typical of the base acre. It is assumed that service levels remain at 80 percent of the base acre for the next 10 years. At this point, the increment to marsh service will be assumed to decrease linearly to zero over the next 20 years as the marsh is lost to sea level rise and erosion.

Table 2-4 reports the average SAYs of credit each year post-spill. It also reports the DSAYs of credit each year where all services are discounted back to the year of the spill using a 3-percent discount rate. These assumptions imply 11.28 DSAYs of credit per acre of restoration.

Table 2-4—Acre Years of Credit

Relative Year Post-Spill	Uplift	Discounted Value	Relative Year Post-Spill	Uplift	Discounted Value
1	0	0	29	0.8	0.339
2	0	0	30	0.8	0.330
3	0	0	31	0.8	0.320
4	0	0	32	0.8	0.311
5	0.025	0.022	33	0.8	0.302
6	0.075	0.063	34	0.8	0.293
7	0.125	0.102	35	0.8	0.284
8	0.175	0.138	36	0.76	0.262
9	0.225	0.172	37	0.72	0.241
10	0.275	0.205	38	0.68	0.221
11	0.325	0.235	39	0.64	0.202
12	0.375	0.263	40	0.6	0.184
13	0.425	0.289	41	0.56	0.167
14	0.475	0.314	42	0.52	0.150
15	0.515	0.331	43	0.48	0.135
16	0.545	0.340	44	0.44	0.120

Relative Year Post-Spill	Uplift	Discounted Value	Relative Year Post-Spill	Uplift	Discounted Value
17	0.575	0.348	45	0.4	0.106
18	0.605	0.355	46	0.36	0.092
19	0.635	0.362	47	0.32	0.080
20	0.665	0.368	48	0.28	0.068
21	0.695	0.374	49	0.24	0.056
22	0.725	0.378	50	0.2	0.046
23	0.755	0.383	51	0.16	0.035
24	0.785	0.386	52	0.12	0.026
25	0.8	0.382	53	0.08	0.017
26	0.8	0.371	54	0.04	0.008
27	0.8	0.360	55	0	0.000
28	0.8	0.350	**Total DSAYs Credit: 11.284**		

2.3.7 Restoration Requirement Identified by HEA

The total debit is 17.01 DSAYs, which is calculated as the product of the footprint (9 acres) and the debit per acre (1.89 DSAYs per acre). The credit per acre of restoration is estimated to be 11.28 DSAYs. The total restoration requirement of approximately 1.5 acres is calculated by dividing the total debit by the DSAYs of credit per acre of restoration.

2.3.8 Critical HEA Topic: HEA Approach

As discussed in Section 5, the most difficult aspect of HEA in a cooperative setting is often deciding on the conceptual approach to defining and evaluating changes in the composite service. There is no set way to define or measure the composite level of service flowing from a habitat, and variation in the conceptual approach can generate damage estimates that differ by millions of dollars even when all assumptions related to ecology, biology, and toxicology are identical.

Once a conceptual approach has been identified, the practitioner must remain cognizant of issues related to incomplete data, baseline, and variation between the impacted and the restored sites.

2.3.9 Critical HEA Topic: Overlapping Services

One of the habitat services provided by a sand beach habitat is services to birds. However, impacts to birds are being directly addressed via multiple species-specific REAs. Thus, debits and credits generated by the REAs and the sandy shoreline HEA are not strictly additive; to treat them as such generally leads to double-recovery of restoration requirements, which is prohibited under OPA (see Section 7).

2.3.10 Critical HEA Topic: Out-of-Kind Restoration

As a result of the *Clay Thomas* spill, oil sank and bottom sediments were affected. This may appear to be an ideal situation to use HEA. However, for technical and practical reasons, it is very unusual to actually create or restore subtidal habitat. Instead, it is common to create or restore wetland to compensate for subtidal injury. This requires a weighting mechanism to translate sediment debits to wetland credits. Several approaches have been employed, and these are discussed in Section 7.

2.3.11 Critical HEA Topic: Treatment of Uncertainty

As is the case with REA, trustees have, in some cases, raised uncertainty as a reason to rely on HEA assumptions intentionally designed to overestimate restoration requirements. The NRD practitioner must understand that many forms of uncertainty affect both the injured site and the restoration site, and they may therefore "cancel out" in HEA calculations. Thus, for most forms of uncertainty, HEA gives answers very similar to those generated when NOAA guidance on formally incorporating risk and uncertainty into NRDA are followed (see Sections 5.5 and 7.4.2). Hence, HEA practitioners should be wary of ad hoc HEA modifications designed to "address uncertainty," particularly if those ad hoc approaches have any material impact on restoration estimates.

3 The Relationship of REA and HEA to Economic Principles

"He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may be cast."

— Leonardo da Vinci

3.1 Why Practitioners Should Understand the Economic Foundation of REA and HEA

Measurement in science typically involves two steps. First, one identifies the thing that is to be measured in concept. Then, one specifies the observable, empirical quantities that will represent that conceptual measure. This section describes the *conceptual* scaling model that REA and HEA seek to implement in practice. That is, we ask: What are REA and HEA trying to do?

Our conceptual ideal is an economic model of scaling compensatory restoration. This means that we evaluate REA and HEA based on how closely the scale of restoration they identify matches the amount of compensation that would be identified if a full economic assessment were conducted.

This analysis is important for three reasons. First, it specifies when REA and HEA are defensible in the sense that they comport with a coherent and well-developed theoretical approach to compensation. This is important for lawyers to know, as the analysis identifies the pitfalls in presenting a case based on REA or HEA. Second, the analysis is useful in negotiations, as it points to standards that an opposing party's model may violate and provides a rationale for making alterations. Third, all but the simplest real-world cases in which REA and HEA are applied do not conform to all the assumptions needed to make REA and HEA fully defensible. Adjustments to the basic models are often needed to address this fact. To assure NRDA practitioners and outside stakeholders that the NRDA reliably compensates the public for injuries, adjustments should conform to the underlying lack-of-fit between the economic model and REA and HEA approximations to it. Some rules of thumb or adjustments to the model will be better than others by this standard.

Why is economics the ideal approach?

— The OPA NRDA regulations explicitly emulate the economic approach. They define the value of resources exactly as most economists would, and they recognize that service-to-service scaling (REA and HEA) requires assumptions that may not hold, in which case a valuation approach is warranted. This valuation approach is the economic model for scaling restoration.

— Economics provides a well-developed and coherent set of theories and methods for determining how to compensate the public for changed circumstances.

— In principle, the economic approach is highly flexible in the case facts it can accommodate. Restored services can be totally different than injured services; services gained or lost can occur at different points in time and be made comparable with a discount rate that has an established basis; the public can be composed of very different types of people, with widely divergent preferences for ecological services; the multiple services in HEA can be aggregated into a single composite index using economic principles; and uncertainties can be dealt with in a consistent manner.

— Trustee economists and their consultants originally developed HEA and REA methods as explicit and formal (mathematical) approximations to fully specified economic methods (Mazzotta et al., 1994, Unsworth and Bishop, 1994, Jones and Pease, 1997). They have been evaluated by this standard in the literature and in practice ever since (Desvousges et al., 2018; Dunford et al., 1994). Despite not having any explicitly economic parameters other than a discount rate, HEA and REA are economic models modified by economic assumptions in order to “make the economics disappear.”

For very good reasons, REA and HEA make simplifying assumptions. If the assumptions are approximately satisfied in the case at hand, all is good, and REA and HEA results can estimate restoration scale reliably (assuming the details are implemented sensibly). The inputs to REA and HEA are largely based on chemistry, toxicology, and biology; given a spreadsheet that implements the model’s calculations, economists are not really needed. However, if the assumptions do not hold in the case at hand, and no adjustments are made to the basic model, then one is not sure the answers provide a good estimate of true NRDs.

If REA and HEA are to be applied when they do not quite fit the facts, the adjustments made must reflect the *logic* of the economic approach. Since the standard of comparison is provided by economic methods, one can judge the alternative adjustments by how well they approximate economic results. By understanding the target, one can identify some adjustments as being better than others. There are good rules of thumb and bad rules of thumb, and knowing the underlying theory allows one to tell the former from the latter.

If the simple model significantly diverges from reality, it may be worth deriving economically based adjustments, thereby creating a “hybrid” service-to-service/valuation method. How far one travels down this path is basically a question of the benefits versus the costs of a simpler approach.

This section is about principles. We assume perfect implementation of any method, unbound by budget, time, uncertainties in measurement, and all the details that good practice must confront. We will remind ourselves of that by writing “in principle” to emphasize this stance, but it should be kept in mind throughout. Whether implementing an economic approach would *actually* be better overall in terms of accuracy and cost is a matter for both analysis and debate.

The key question addressed in this section is: Under what conditions do REA and HEA give the same answers as the full economic “value-to-value” methods? That is, when are HEA and REA fully defensible? The answer to this question identifies sources of potential weakness in using REA and HEA in a litigation setting, and thereby identifies for lawyers the potential vulnerabilities of the methods. While defensibility is of course magnified in court, it plays an important role in negotiations and cooperative assessment, as well.

3.2 The Basic Economics of Compensation

The issue we address goes back to Figure 2-1 illustrating REA and Figure 2-3 illustrating HEA. At any point in time, the vertical distance between the with-spill-and-restoration services and the baseline services is the amount of species years (SYs) or SAYs lost or gained in that period. How exactly do these REA and HEA measures relate to compensating individuals for their losses?

The economic theory of compensation is approximately 100 years old. Economists developed the theory in terms of *monetary* payments either (1) to an individual (to accept a negative change or forgo a positive one) or (2) by an individual (to avoid a negative change or obtain a positive one). Every ordinary transaction is an instance of mutual compensation, with the consumer willing to pay an amount to obtain an apple and the merchant being willing to accept this payment to give it up. Extending this idea to public goods such as ecological services, and devising and refining methods for implementing it in practice, is one of the important developments in economics over the past half-century.

Some comments are in order:

— Compensation is defined at the level of the individual; compensation of the relevant public must come from some rule for adding up (aggregating) over individuals.

— The payments are for defined and specific changes in goods or services from one situation (baseline) to another, and hence both the starting point and the magnitude of the change matter. An idea such as “the value of New Jersey’s resources” is not meaningful unless one is comparing the current state versus the Atlantic Ocean being coincident with the current Pennsylvania and Delaware state boundaries.

— There are two basic ideas for compensation: WTP to either obtain a benefit or avoid a harm, and willingness-to-accept (WTA) compensation to either forgo a benefit or endure a harm.

— “Money” here is a highly useful shorthand for a bundle of real goods or services that would be bought by the individual with the money either given up in WTP or received in WTA. This point is important because it is sometimes thought that “economic value” is synonymous with “value denominated in money.” It is not. A barter economy with all exchanges being made as one real good for another is just as economic as a modern currency-based economy. Indeed, the OPA NRDA regulations (15 CFR § 990.30) define value as “…the maximum amount of <u>goods, services, or money</u> an individual is willing to give up to obtain a specific good or service, or the minimum amount of <u>goods, services, or money</u> an individual is willing to accept to forego a specific good or service” (emphasis added).

  Thus, “paying” an individual via an increase in natural resource services is just as “economic” as paying them in money.

In the “olden days” of NRDA, before the publishing of final OPA NRDA regulations and associated guidance documents, monetary compensation was the focus of NRDA economists’ attention. Despite a recognition that monies had to be spent on restoration, NRDA economists focused on a value-to-cost scaling method, which equates NRDs to the economic value of the injury measured in money. This amount is then spent on restoration. This approach changed in 1994.

Mazzotta, Opaluch, and Grigalunas (1994) and Unsworth and Bishop (1994) recognized, apparently independently, the implications for scaling methods that require damage awards be spent on resource restoration projects. They set forth the conceptual basis for compensation in-kind via restoration. In particular, these authors saw that by restoring resources that provide similar services to those injured, scaling takes place as it would in a barter-based economy, without valuation in money. It is this economic idea that defined the approaches we now call REA and HEA.

3.3 REA, HEA, and Economics

We assume the reader of this document is not trained in economics, and simplify our analysis where practical. The essence of this section is a demonstration that REA and HEA are applicable to “NRDA in the small.” By this, we mean that the assumptions under which REA and HEA are defensible serve to restrict the cases in which they can be rigorously applied. Specifically:

— The focus is on the natural resources physically injured and restored. Their broader context is usually not considered. Analytically, the affected resources are isolated from a larger world of other types of services, other habitats or resources in the landscape, and regional supplies of and demands for natural resource services.

— The resources injured and restored must provide essentially the same services, limiting restoration options.

— The changes in services brought about by injuries and restoration must be a small part of the resources in the region in which they reside.

— The affected public must be homogeneous, with the same preferences for different services. This likely limits the ability to address restoration benefits to people remote from injuries in space or in time.

REA and HEA go wrong in situations where the larger world is important and so invades the small world space where the basic HEA and REA models directly apply. The general economic model is a "large world" method. In principle, it can handle all of the more general circumstances and considerations that REA and HEA cannot. How and why this is so, we now try to make more precise.

3.3.1 Compensation and NRDA

NRD provisions are compensatory, not punitive. Compensation, conceptualized as WTP or WTA, could take the form of (1) a monetary payment, (2) a project that provides more of the ecological service that was lost, (3) a project that provides more of some other ecological service, or (4) a project that provides more of any other economic good. We are interested here in monetary payments or payments in ecological services.

Outside of the federal statutory NRD context, economists would recommend a "lesser of" rule, which asserts that the appropriate form of compensation (monetary versus restoration, and for the latter exactly what type of services are to be provided) should be based on cost effectiveness, with damages based on the least-expensive method of compensating the public.[20] Moreover, other ways of using restoration funds to enhance the well-being of the general public would be considered via a benefit-cost analysis that incorporates the opportunity costs of these funds. If the benefits of additional restoration do not outweigh its cost, there is some other (albeit unspecified) use to which the funds could be devoted to provide a greater benefit to the public than restoration. However, in the NRD context, the Ohio decision[21] gave a clear preference to restoration.

Despite this preference for restoration, to fully understand REA and HEA, we will first discuss monetary compensation, as monetary compensation is the basic ingredient of the economic valuation approach to scaling.

3.3.2 Compensating an Individual

The economic model of compensation applied in NRDA is based on individuals and their preferences for different bundles of market goods consumed and ecological services enjoyed. In this section, we are going to work at the level of a single individual. We can think of him or her as a representative or average person in the population. We also are going to focus on a single time period. When we get to compensating the public, with many individuals, and to a model with many time periods, things become more complicated.

3.3.2.1 Preference and Indifference

Let $M = (M_1, M_2, \ldots, M_N)$ be a list of regular market goods and services and $Q = (Q_1, Q_2, \ldots, Q_S)$ be a list of different ecosystem services provided by a natural resource. A possible "consumption bundle" for the individual is the pair (M, Q). Individuals are assumed to be able to rank alternative bundles according to their preferences; one bundle is at least as preferred as another—or not. Preferences vary from individual

---

[20] This is because less of society's scarce resources are devoted to compensation than otherwise, not because economists want to minimize the RP's liability.

[21] State of Ohio v. U.S. Department of the Interior, 880 F.2d 432, 455 (D.C. Cir. 1989). See Israel et al. (2020) for a legal and policy discussion and Brown (1993) for an economic critique.

to individual, and the economics approach tends to not question individuals' preferences and the choices an individual makes based on their preferences.[22]

The theory does require preferences to be rational. Rationality of preferences requires two things.

1) An individual can compare any two bundles of goods and services and decide which is preferred. This is called the "completeness" of preferences. This rules out comparison of things so different that to rank them "just doesn't compute"; an example might be including in a bundle changes in deeply felt religious or moral beliefs along with, say, a new pair of shoes. While the former may be part of a worldview on which preferences are based, they are not subject to ranking via preferences. We restrict the analysis to rankings of comparable bundles.
2) If bundle A is preferred to B, and B is preferred to C, then A is also preferred to C. This is called "transitivity" of preferences and seems like a minimal requirement of rationality.

Having rational preferences means that when faced with a set of bundles of goods and services, each individual can rank them from their most preferred to least preferred. That is, persons with rational preferences can find a best (most preferred) bundle from among the options they are provided.

To this basic definition of rational preferences, we add two more requirements. First, increasing the amount of any one good or service without diminishing others leads to a strictly more preferred bundle.[23] Second, we assume that preferences are convex; that is, if the person is indifferent between bundle x and bundle y, any composite bundle that is a mix of x and y is judged to be at least as good as either x or y. If the person strictly prefers a mix of x and y to either x or y alone, this is called "strict convexity" of preferences. It plays an important role in the economics of REA and HEA, the discount rate, and the treatment of uncertainty in NRDA, as will be discussed below.

We describe bundles over which preferences are defined as including (1) everyday economic goods and services we buy and sell and (2) ecological services. To simplify matters, we assume that all the various market goods are aggregated into one composite generalized market good, which we label M and call "money." We arbitrarily give it a price of \$1, and M becomes expenditure on current consumption.[24] For the time being, we assume there is only one ecological service, which we denote by Q.

Starting with one fixed reference bundle, among all other possible comparison bundles, an individual will rank some as better than the reference, some worse, and some equally good. An *indifference curve* plots all the bundles that are ranked "equally good" according to an individual's preferences. Indifference curves that are farther from the origin provide bundles that are preferred to bundles on indifference curves closer to the origin.

Two indifference curves are shown in Figure 3-1, which has Q on the horizontal axis and M on the vertical axis. There are two reference bundles, B and S. Relative to bundle B, bundle S has less of the ecological service and the same amount of money. We can think of bundle B as the baseline situation before an oil spill and S as the bundle after an oil spill. We label the indifference curve that runs through S with an I, for "injured"; the indifference curve that runs through B is the baseline indifference curve (B).

The indifference curves are bowed inward. This is the strict convexity of preference assumption in action.

---

22 An excellent and very readable discussion of the economic approach to linking preferences to individual values is provided by Hausman (2012).

23 We measure services as providing beneficial outcomes. If something is a negative (e.g., an invasive species), we redefine its measurement to be less of the bad (e.g., more invasive-free habitat). We will assume further that preferences are continuous (the set of bundles preferred to A and the set that A is preferred to are both closed sets—that is, containing their boundaries, without gaps). See Varian (1992).

24 We thus take out savings; donations to others might be thought of as buying gift cards at no markup.

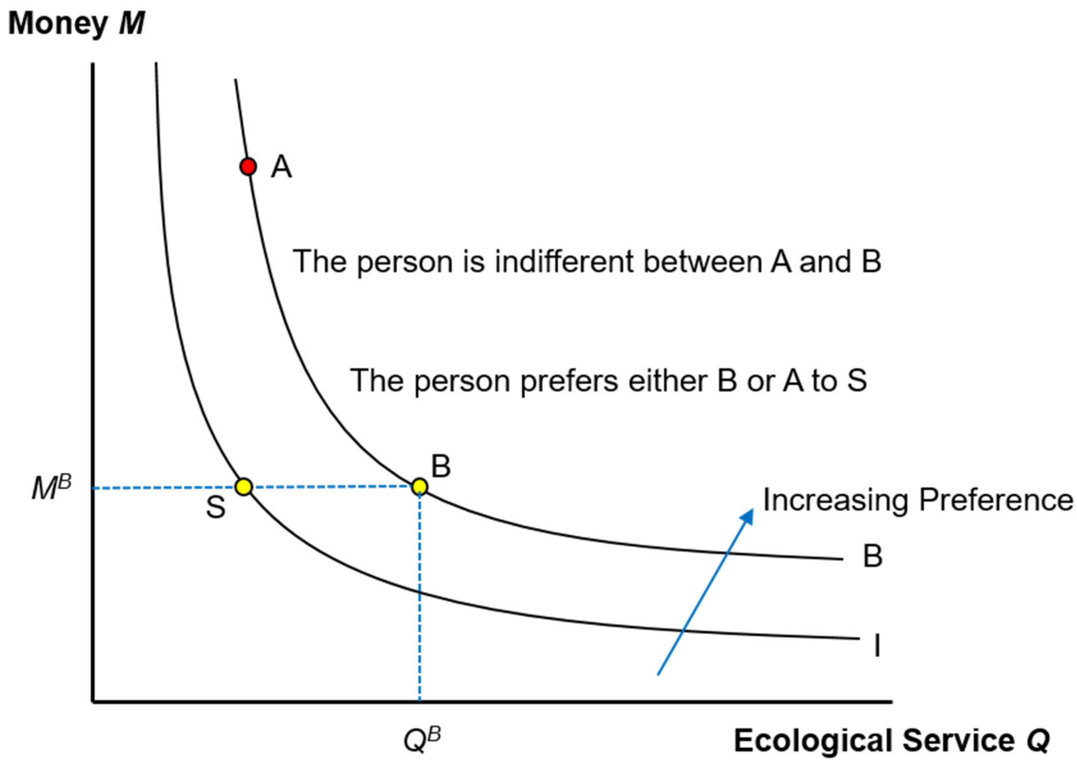


Figure 3-1—Two Indifference Curves Reflecting Baseline and Injured Bundles of Goods and Services

3.3.2.2 Service Values and Value-to-Value Scaling of Restoration

Given the two bundles and indifference curves in Figure 3-1, we illustrate in Figure 3-2 the effect of an oil spill and the subsequent actions that could be taken to compensate the individual for spill-related injury. Almost everything you need to know about the economics of HEA and REA can be gleaned from thinking carefully about this diagram.

The OPA NRDA regulations define the value of resource services. Applied to an oil spill, these are either a WTP to avoid the spill or a WTA to endure it. These payments, given or received, can be denominated in money, market goods, or ecological service. Since money and goods are the same in our model, we will focus on money and services.

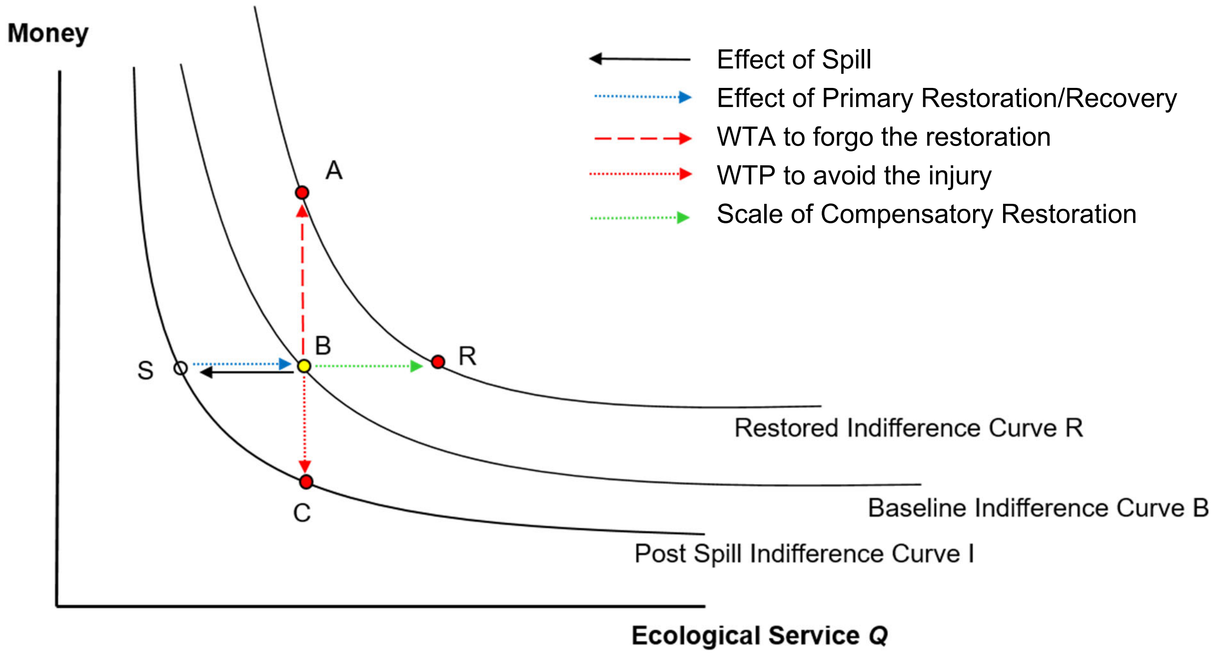


Figure 3-2—The Value of Ecological Services and Compensation for an Oil Spill

The spill scenario we will use to illustrate the principles involved in individual compensation is very simple. To illustrate compensatory restoration for the effects of a spill, we introduce the aspect of time by assuming two time periods and a zero discount rate. The person enters the first period with the baseline bundle of money income and services at point B in Figure 3-2.

The injury occurs in an instant at the beginning of the first time period, and primary restoration is accomplished instantaneously at the end of it, with no increase in services in the interim. Therefore, the length of the black arrow in Figure 3-2 is the amount of interim loss measured as a number of SAYs.

Because the discount rate happens to be zero, the length of the black arrow is the DSAYs of HEA debit. Compensatory restoration is implemented at the beginning of the second period, instantly provides a level of services that depends on scale, and lasts one period. The scaling question that arises is: How much restoration will create as much value as that lost?

The spill reduces the amount of the ecological service available (solid black arrow), resulting in bundle S, where bundle S has less ecological service but no change in money relative to bundle B. Injury quantification estimates the length of the solid black arrow.

Bundle B is preferred to S, as it is on a higher indifference curve. Primary restoration and recovery of services return the individual to the pre-spill level of utility; this is defined as any point along indifference curve B. We do not have to return the individual exactly to bundle B; anywhere along the baseline indifference curve will suffice. But for convenience, suppose recovery is back to B and the recovery is given by the solid blue arrow in the figure. WTP^I to avoid the scenario of the spill plus primary restoration must be zero. How this is accomplished with physical resource restoration is a separate question, and the costs are (assumed in NRDA) not to be borne by the people experiencing the injury.

The individual sitting at B (with the spill having not yet occurred) is willing to pay something to avoid enduring the injury. The most he or she would pay WTP^I is given by the length of the vertical (downward) red solid arrow from B to C (holding services constant and reducing money income). If the person pays less than WTP^I (and avoids the spill), he or she would be strictly better off than at baseline, while if more is paid, he or she is worse off than at baseline. WTP^I is a measure of the economic value of the injury to this person.

Turning to restoration, WTA^R is a minimum payment in money the individual would be willing to take, when they are sitting at baseline, to forgo the restoration. This is the length of the vertical red dotted arrow from B to A. The person is indifferent between getting the restoration and residing at point R, or forgoing the restoration but accepting WTA^R in return and residing at point A.

The appropriate scale of compensatory restoration is found as follows. An amount of compensatory restoration is the length of the green arrow moving horizontally to the right from the baseline at B, keeping money constant and increasing services. Value-to-value scaling equates the value of restoration (WTA^R) to the value of the injury (WTP^I). Hence, restoration is scaled when the lengths of the two red arrows are the same. If restoration was greater than the length of the green arrow, it would reach an even higher indifference curve than R, and the person would demand more than (WTA^R) to forgo it, and value equivalency would not hold. Conversely, if the length of the green arrow falls short of R, the person is on a lower indifference curve than R and the person would be willing to accept less than (WTA^R), and once again, value equivalency would not hold.[25]

There are two things to note about Figure 3-2.

1) The WTA and WTP measures of compensation involve *actual* money changing hands—either giving the person money or taking it from them. This does not happen in OPA NRDAs. Thus, the direct use of WTP^I or WTA^I as a measure of compensation is of limited relevance. As illustrated in Figure 3-2, they are relevant in defining value-to-value scaling of restoration. The requirement to spend NRD recoveries on restoration is sometimes absent (e.g., international spills, state actions), and we sometimes see monetary valuation of the injury advanced as a way to compute NRDs. It is only rationalized when the cost of the restoration benefits is too great to warrant the effort; this is "value-to-cost" scaling that can only be defended as a cheap expedient. To see how wrong value-to-cost scaling can be, let V be the WTP per acre for restoration. This is the economic benefit per unit of restoration. If C is the unit cost of restoration (e.g., cost per acre), V/C is the benefit-cost ratio (BCR) for undertaking restoration. Value-to-cost scaling only gives the correct scale of restoration if the BCR equals 1. As discussed in Byrd and Tomasi (2021), if the BCR is greater than 1, overcompensation results from value-to-cost scaling, and if it is less than 1, undercompensation results. In our experience, BCRs tend to be much larger than 1, so value-to-cost scaling results in an overestimate of NRDs.
2) Note that if the indifference curve that runs through the restored bundle becomes steep to the left of R, WTA becomes large. Indeed, as the indifference curve becomes vertical, WTA goes to infinity. This means there is no amount of restoration or money that can compensate for the injury. Steepness indicates the good on the vertical axis is not a good substitute for the one on the horizontal axis. If the things you can buy with money do not provide a good substitute for ecological services, WTA will be large. WTP, on the other hand, will be finite, as it cannot be bigger than the pre-spill amount of money.

It is an important clarification of nomenclature that "value-to-value" scaling does *not* require the use of dollars. The "value" in value-to-value scaling reflects voluntary tradeoffs made by individuals according to their preferences. What is required is the use of economic methods that can identify the trades individuals

---

[25] The approach depicted in Figure 3-2 measures the value of the spill as a WTP and the value of restoration as a WTA. These measures are called "equivalent variation" and are based on the person remaining at the baseline bundle. An alternative approach called "compensating variation" measures the value of injury as the WTA to endure it and the value of restoration as the WTP to obtain it. This places the person in the post-change situation when measuring compensatory payments. A similar value-to-value process as above can be used to scale restoration. Some economists prefer the compensating variation approach. The two are the same if, fixing the amount of services, increasing money does not alter the *slope* of the indifference curves and WTP for small changes in services is independent of the amount of income.

would make, according to their utility functions, among ecological services affected by the spill and provided by restoration.[26]

3.3.3 Compensation Using REA or HEA

How are REA and HEA linked to the economic model of compensation?

Figure 3-3 displays some of the same information that was in Figures 3-1 and 3-2, but in a different format. Figure 3-3 plots the amount of the service Q along the horizontal axis and WTP along the vertical axis. While the curve labeled WTP(Q) is always increasing (the individual is always willing to pay something to get more service), it is steep near zero and flatter to the right. This shape indicates that an increment of services is worth more when services are scarce than when they are abundant. This follows from the strict convexity of the indifference curve (and it just makes sense). We will now see why this is a fundamental issue.

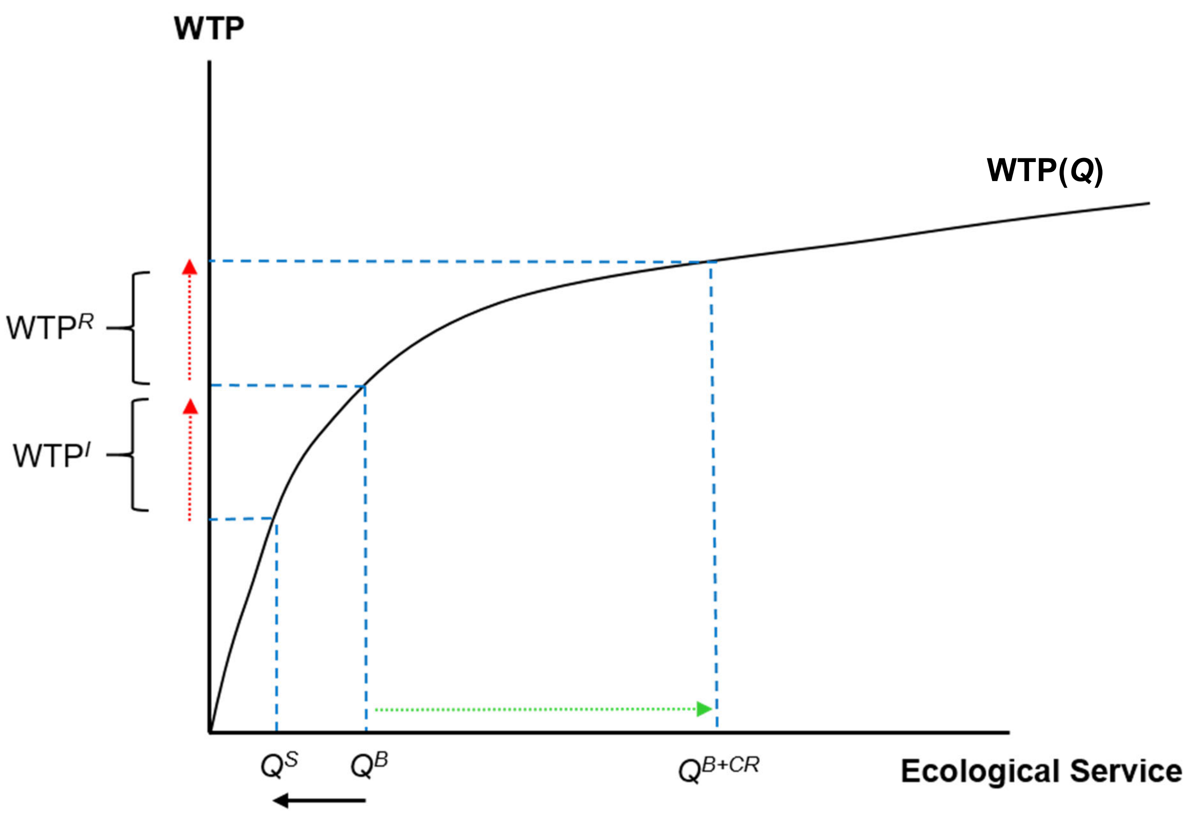


Figure 3-3—Economic Value-to-value Scaling of Compensatory Restoration

The full economic approach of value-to-value scaling is depicted in Figure 3-3. In the full economic model with many time periods, the amounts WTP^I and/or WTP^R from this figure would appear in Figures 3-1 and 3-2 for each time period, not SYs or SAYs.

To get from the economic model depicted in Figure 3-3 to REA and HEA requires (1) multiplying an amount of service change by the incremental (marginal) value of a service change measured at baseline,

[26] An example is the analysis undertaken at the Fox River, Wisconsin, site by Triangle Economic Research on behalf of the RPs; they scaled restoration in terms of "utils" (the hypothetical unit of utility) lost and gained, using essentially a value-to-value approach that did not go through dollars. For a recreation version, see Parsons and Kang (2010).

and (2) assuming that the injured and restored marginal values are the *same* at *every* level of injury and restoration. This is so values can be canceled out and scaling can be done in services alone.

This constant incremental value assumption requires the value curve to be a straight line, not curved (equivalently, indifference curves need to be straight lines and not convex). If WTP(*Q*) is curved, the HEA/REA value curve (which must be assumed linear) can only approximate the WTP curve around the baseline services. The resulting difference between REA and HEA and economic scaling is shown in Figure 3-4. While true values are read off of WTP(*Q*), REA/HEA values are read off of the approximation line. REA/HEA understates true debits and overstates true credits unless the amount of service changes is small relative to the curvature in WTP(*Q*).

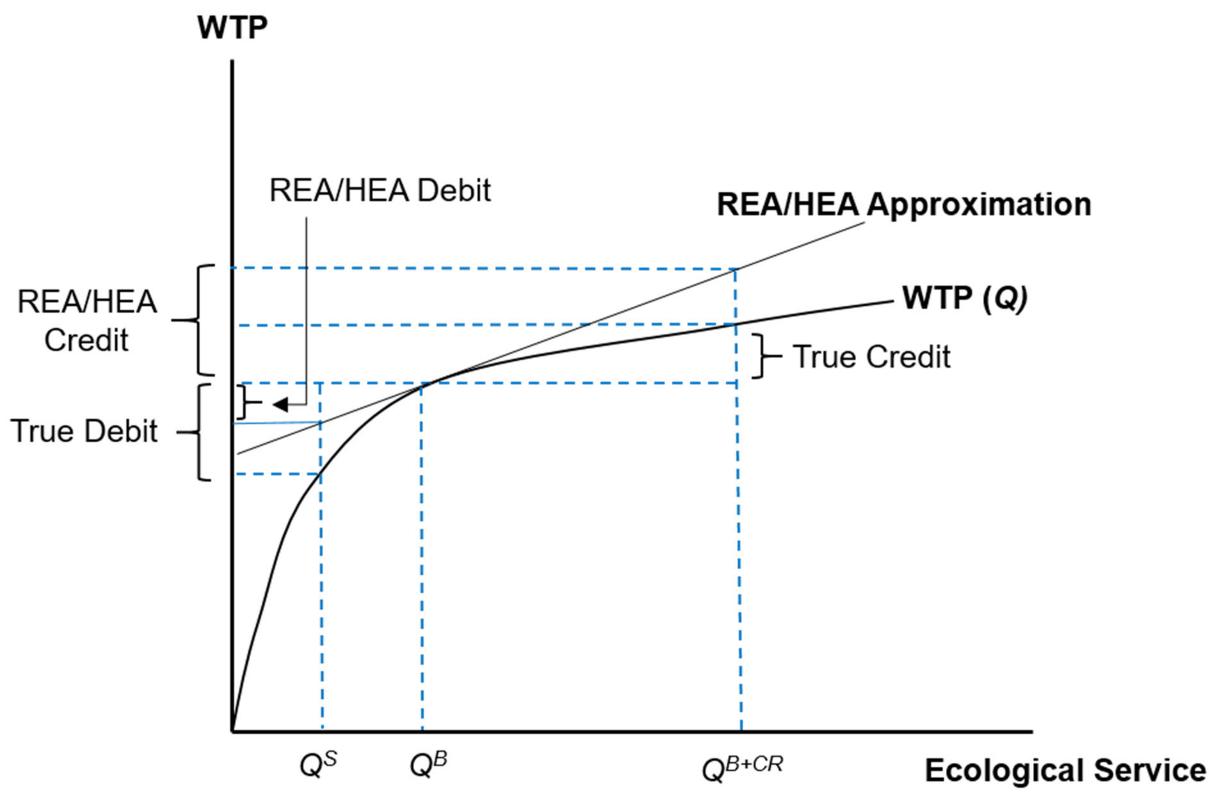


Figure 3-4—Valuing the Effects of a Spill: REA and HEA Versus Value-to-value Scaling

As mentioned above, if preferences were convex but not strictly so, the indifference curves in Figures 3-1 and 3-2 would be straight lines. In this case, the WTP curve would also be a straight line. REA and HEA would implement value-to-value scaling exactly, with no approximation errors, no matter the size of the injury and restoration.

3.3.4 Which Measure of Compensation Is Preferred?

In Figure 3-2, we showed two monetary measures of value: WTA to forego restoration and WTP to avoid the spill. We could have switched these and estimated different scales for compensatory restoration (see footnote 24). The size of the divergence depends on the shape of the indifference curves. If the indifference curves are straight lines, WTP and WTA coincide and one gets the same scale with either approach.

If the two approaches yield different answers, which is preferred? In principle, WTA to endure the spill would seem to be preferred in NRDA, which is only triggered by a spill happening and the public receiving compensation. WTP to avoid the spill (shown in Figure 3-2) seems irrelevant. However, the empirical measurement of WTA for ecological service losses presents significant practical problems. In almost all attempts to measure exact compensation with money (e.g., in the *Exxon Valdez* and Deepwater Horizon spills), one measures individual WTP to avoid a future spill that has consequences similar to the one being valued. Thus, for practical reasons, WTP is used when measuring injuries. Regarding restoration benefits, the approach in Figure 3-2 is to have a WTA to forgo such benefits. It is not clear whether this would be similarly difficult to measure, but note that having measured WTP to avoid the spill, one has in hand what the value of the restoration needs to be, and the empirical problem is to find out how much restoration corresponds to this amount.

With that caveat, we can evaluate possible rules for scaling compensatory restoration. The first is the "lesser of" rule. Suppose that one anticipates actually paying compensation in money and that WTP is less than restoration cost. Then, the monetary value approach is the cheapest way to provide exact compensation. Of course, the court in the Ohio decision gave preference to restoration-based compensation, anticipating that restoration cost would be greater than WTP in most cases.[27] The "lesser of" rule was changed to a "grossly disproportionate" rule in which restoration is preferred unless the cost of restoration is far above the WTP measure. This will often be relevant for primary restoration. But, if the restoration cost is less than WTP, as will usually be the case for compensatory restoration, providing restoration benefits is *cheaper* than providing monetary compensation, and also provides exact compensation. Both are valid economic measures, and in this case, the "lesser of" rule is consistent with a preference for restoration.

We summarize the discussion so far with the following findings.

Theoretical Finding 1: The economic model allows compensation for loss of one good (ecological services) with provision of a different good (market goods purchased with money). The amount of such compensation depends on the individual's rate of substitution between services and money. That rate of substitution is reflected in the shape of the indifference curve.

Theoretical Finding 2: Restoration of resource services can serve as an alternative, but still exact, measure of compensation and can be consistent with a fully economic model.

Theoretical Finding 3: The "lesser of" rule is economically efficient. Thus, if restoration costs are less than WTP, restoration cost is the preferred measure irrespective of any requirement under OPA that monies collected be spent on restoration. Thus, in non-OPA spills (e.g., spills in non-U.S. waters; state actions) restoration cost can be perfectly consistent with economic principles.

Theoretical Finding 4: In general, value-to-cost scaling is not consistent with the economic model of compensation. It should only be used if the BCR of restoration equals 1, or if the cost of estimating the BCR is prohibitively large. This comports with the OPA NRDA regulations.

Theoretical Finding 5: Value-to-value scaling is based in concept on the economic model of compensation and could, in principle, be used in a litigation-based case. The caveat "in principle" is vital. For services other than recreation, implementation may embody substantial errors of measurement.

Theoretical Finding 6: If spill-related service changes are "small," REA and HEA may approximate a value-to-value scaling in a one-person, one-service, timeless model. If the changes are not small, REA or HEA will underestimate NRD liability. "Small" means a regionally insignificant change in service level such that the incremental utility of services (i.e., slope of the utility) does not change over the range of injury and restoration. The service change can be large if the WTP function is approximately a straight line (WTP for a small change in services does not change as the baseline level of services varies). To be a good approximation, the amount of service change must become smaller as the WTP function becomes more curved.

---

27 See Haddad and Israel (2017).

3.3.5 Compensation When There are Many Services: The Composite Service Index

In the prior section, there is only one service on the horizontal axes in Figures 3-1 through 3-4. Clearly, natural resources provide many services. REA and HEA scale restoration only by altering the amount of the resource (REA generally works on the number of individuals in a population) or acres of habitat (HEA generally works on a quality-adjusted estimate of the acres of habitat). Thus, one has to find a way to aggregate services into a single composite index. This section discusses the economic principles behind aggregating many services into one index.

In this section, we continue to assume one person and one time period, but the symbol Q now represents a list of S different services instead of just one service. The individual's preferences are defined over consumption bundles $(M, Q_1, Q_1, \dots, Q_S)$. These objects are ranked by preferences, and indifference is defined by *sets* of indifferent bundles rather than a curve in two dimensions. It is now convenient to work not directly with preferences, but with a utility function that represents the preferences (i.e., has equivalent information). This function assigns a single number to each indifference set (or indifference curve with two goods). Higher ranked sets are assigned a higher number. The utility function is defined by $U(M, Q_1, Q_1, \dots, Q_S)$; it exists if preferences are rational. Since it assigns a single number to each bundle, utility serves as an aggregator of services and provides an index for all the services (plus money). Services can substitute for one another in the generation of individual well-being to a degree that depends on the form of the utility function and the preferences it represents. We can define indifference curves between two services, much as we did for indifference curves between market goods and the single service in the prior section. This would look like Figure 3-1, but now with two different services on the two axes with money and all the other services fixed. The same interpretations hold.

Value-to-value scaling in utility terms is exactly the same as in the prior section, with the utility function and the resulting measure of utility providing the aggregation over services. Rather than saying that WTP or WTA are distances between indifference curves, we say that they equate utility with and without the spill.

For each service, generically service Q_s, we hold the quantity of all the *other* services at their baseline levels and consider a small increment to service q_s. If this is an increment to the baseline level of service and the increment q_s is "small," we get the marginal value of service s denoted by $v_s(M, Q^B)$. The notation says that the marginal value v_s depends on (is a function of) the baseline, as well as money income. Here, *all* services are at their baseline levels.[28] We will just write v_s in place of $v_s(M, Q^B)$, and it should be remembered that this is the value at baseline levels of income and all services.

We continue to assume that the marginal utility of money is independent of service changes over the relevant range. Now, we undertake the same approximation procedure and find the scaling equation [Equation (1)]:

$$\sum_s v_s \,[Q_s^B - Q_s^I] = \sum_s v_s \,[Q_s^R - Q_s^B] \qquad (1)$$

Here, we see that the composite service is a summation across services of the marginal values for each service times the amount of change in that service. The marginal values of the services are weights in the summation. Note that, even if a service is not injured, its baseline level might still affect the value of changes among the services that are injured.

In terms of the model of this section, this finding is not a "big deal" because the baseline levels of services *not* affected likely are not changing. But it is important when time is introduced because, over time, the baseline levels of the injured and uninjured services can change and alter the values of the injured services.

---

28 That is, $v(M, Q_S^B) = lim_{q_s \to 0} \left(\frac{1}{m}\right) \{(\frac{1}{q_s})\{U(M, Q_1^B, \dots, Q_{s-1}^B, Q_s^B + q_s, Q_{s+1}^B, \dots Q_S^B) - U(M, Q_1^B, \dots, Q_s^B, \dots, Q_S^B)\}$.

Equation (2) defines a composite service index for the individual. We sum up the service changes caused by the oil spill and the restoration with a system of weights, one for each service. The weights are the marginal values of each service. The composite service index is:

$$Q^{comp} = \sum_s v_s \, Q_s \qquad (2)$$

We will frequently drop the "composite" superscript for convenience and just write Q for the composite service. Whenever there are multiple services, we mean a composite service unless we state otherwise.

In addition, the services potentially can substitute for one another in creating utility. If one service q_j goes down by some amount Δ_j and another service q_k goes up by some amount Δ_k, utility is unchanged and the amount of composite service is unchanged, only so long as $\Delta_k = \Delta_j(v_j/v_k)$. The term in parentheses defines a rate of trade between services in producing a given level of the composite index.

Finally, the composite service index can be used in an HEA as long as *each* marginal value v_s is a constant over the relevant range of injured and restored services. This means that we only need to know relative values at baseline, and these are constant, so the composite index itself does not change between baseline, injured, and restored services.

Theoretical Finding 7: With multiple services, a composite service index can be formed based on an individual's relative marginal values for the different services. With this composite, an increase in one service can, in general, substitute for a decrease in another, holding total services constant. REA and HEA require that each marginal value in the composite service is a constant within the range of variation in the level of services associated with injury and restoration. This latter condition will hold in REA if indifference curves are fixed radial expansions of one another so that the slopes of the indifference curves stay the same moving out along straight lines from the origin (such preferences are called homothetic). This isn't enough for HEA, which needs parallel straight-line indifference curves. HEA therefore has stricter assumptions needed to aggregate services than does REA.

As soon as there are multiple services being affected, being consistent with an economic model of compensation requires knowing the individual's relative valuations of the different services. But, by the assumption that marginal values do not vary, the form of the index does not change during the analysis. That is, we only need to know the relative values at baseline and do not need to re-compute the weights in the index as we move from baseline to injured to restored resource combinations.

In REA, it is assumed that all services move in fixed proportion to the amount of the resource. Therefore, in REA, the resource is its own aggregator of services (in other words, a mallard REA assumes all mallards provide the same value). In HEA, however, there is no such natural assumption (not all acres of marsh provide the same value), and we have shown that *values* need to be used to be consistent with economic principles. But the whole point of HEA is to avoid valuation.

It is common practice for NRDA practitioners to assign weights (implicitly or explicitly) to the many individual services provided by a habitat. How this assignment is made may have a material effect on NRDs and may be a source of considerable disagreement among the parties when using HEAs.

3.3.6 Multiple Services, Multiple People

The public is a large group of individuals, each with potentially different individual preferences. This presents a significant complication under some specific circumstances, but not others.

It is sometimes thought that if there are many people with different preferences, HEA and REA are inconsistent with the economic approach; this often-cited result is based on Flores and Thatcher (2002).

However, this is only true if there is more than one service, and if the mix of services changes as between injured and restored services.[29]

Theoretical Finding 8: The multiple-person situation is the same as the single-person situation, only if at least one of the following hold.

1) The persons involved have the identical preferences and can be treated using a "representative individual."
2) The persons involved have different preferences and compensation can be paid in money. Each person would be provided different amounts of money (or some other transferable good) based upon their individual WTP. Thus, if the money value approach were used, and if money were actually provided to individuals as compensation, there is no theoretical complication associated with expanding the model to many people. Resource restoration could be accompanied by side payments that exactly compensate everyone (assuming restoration is scaled correctly)
3) There is only one service.
4) If there are multiple services, they are provided as a fixed proportion of the amount of the resource (e.g., birds in REA or acres in HEA); these proportions are the same at both the injured and restored habitats, and marginal values do not change as services change (individual preferences are homothetic) over the relevant range.
5) In all circumstances, people must have the same discount rate.

Condition 1 has a long history in economic analysis of public policies. We treat a collection of people as if they were one individual. There is a restrictive assumption about individual preferences where they do not have to be identical, but if the assumption holds, assuming there is a representative is valid [Deaton and Muellbauer (1980)]. We do not delve into this further here.

Condition 2 is not applicable to NRDAs under OPA, and Condition 3 is not applicable in this section by assumption. Condition 4 achieves the same result as Condition 3 by adding the assumption that marginal values do not change as the level of services change. In general, HEA is more demanding of assumptions that need to be imposed and information needed to relax them, as it usually is the case that the mix of services changes across baseline, injured, and restored bundles.[30]

Regarding Condition 5, note that if there is a single service that varies over time, this is the same as having multiple services. When we say that people have identical preferences, this includes their preferences over *when* they consume a given level of resource. In the language of discounting (covered in Section 3.3.7), all individuals in the public must discount their utility at the same rate. Thus, in Condition 5, we add the caveat that people share the same discount rate.

Suppose none of these conditions are applicable. What goes wrong and what can be done? Restoration provides an increase in an entire list of ecological services: $\Delta Q = ([Q_1^R - Q_1^B], \ldots, [Q_s^R - Q_s^B], \ldots [Q_S^R - Q_S^B])$. Once this benefit is provided to one person, it is provided to everyone because, for the most part, ecological services are public goods.[31] Therefore, unless all members of the public have identical preferences for ecological services, any change in the mix of services between injured and restored conditions will be viewed differently, and without an ability to tailor different packages for different people, some individuals will be overcompensated and others undercompensated.

---

29 With one service (and one time period), people may have different degrees of utility change from an injury and from restoration, but still agree on the amount of restoration that is needed in compensation.

30 With Conditions 3 and 4, the result will hold, but note the closeness of the approximation in the scaling equation may vary across people; the "smallness" result of Theoretical Finding 7 will be driven by the person with the curviest preferences.

31 Public goods are defined to include goods and services where (1) no individual can be excluded from deriving utility from the good or service and (2) the fact that any one individual derives utility from the good or service in no way affects the ability of others to derive utility from the good or service. Recreation services are not pure public goods; the results here show that value-to-value scaling, readily available in recreation, solves the problem of heterogeneous preferences.

It is therefore necessary to develop a model of “public compensation” when compensation is in the form of an ecological service provided equally to all and when individual preferences differ. There are many ways to formulate such a model. Here we note two broad approaches to resolving this issue, both based on the idea of aggregating the preferences of multiple individuals in public policy decisions. The first approach is based on benefit-cost analysis (BCA) reasoning, which requires that the sum (over people) of WTP to avoid the injury equals the sum of WTA to forgo the benefits of restoration. That is, the first approach is value-to-value scaling. Note that this is a real issue for REA and HEA, whose very purpose is to avoid measuring values. The second approach is to use a social welfare function (SWF) to aggregate individuals. This has not been applied in actual NRDAs to our knowledge.[32]

The BCA approach to NRDA was introduced by Jones and Pease (1997) and discussed further by Flores and Thatcher (2002). This is based on work developed by Kaldor and Hicks in the 1930s (Hicks 1939; Kaldor 1939). The Kaldor-Hicks criterion for judging when overall public compensation has been achieved, stated in terms of NRD compensation, is as follows. The public, having experienced the decrement in services due to the spill and having been provided the increment to services from restoration, consists of “winners” (those who are overcompensated) and “losers” (those who are undercompensated). If the winners can compensate the losers *in money* such that (1) all the losers are returned to their baseline indifference curve and (2) the winners remain at or above their baseline indifference curve having paid the compensation, the public is judged to be compensated overall. If the compensation actually takes place, no one is worse off and at least some people are better off after the restoration. It is rather widely believed that if at least one person strictly prefers social situation x to situation y and everyone else is indifferent between them, society as a whole should judge x as preferred to situation y. But, under the Kaldor-Hicks criterion, the restoration action is deemed socially beneficial *even if the monetary compensation of the losers by the winners does not happen*. This is, of course, much more controversial a proposition.

Value-to-value scaling is based on the Kaldor-Hicks idea. It is an important reminder that “value-to-value” scaling does *not* require the use of dollars when there is only one person (the representative individual). What *is* required is the use of economic methods that can identify the trades individuals would make, according to their utility functions, among ecological services affected by the spill and provided by restoration.[33] With the Kaldor-Hicks approach and many persons with different preferences, the use of money is crucial. It is the medium that allows us to ensure that each person (hypothetically) is compensated in a world where (1) compensation comes in the form of public ecological services and (2) individuals have heterogeneous preferences for those ecological services.

Since individuals do, in fact, have different preferences, there exists a fundamental problem with providing public ecological services as compensation. To judge whether the public as a whole is compensated, we need to re-introduce money so as to determine if the Kaldor-Hicks criterion holds. But this is exactly what REA and HEA seek to avoid.

As noted in Section 1.2, the OPA NRDA regulations establish (§ 990.53(d)) that if the injured and restored resources are of the same type and quality, trustees must consider first the service-to-service methods. The OPA NRDA regulations then state that if the trustees have determined that service-to-service scaling is not appropriate, value-to-value scaling may be used. We have now identified another condition governing the appropriateness of HEA: People need to have the same preferences. Again, this can be relaxed for REA, where people can have different preferences as long as each person’s preferences have an appropriate form (homothetic).

---

[32] A social welfare function approach is discussed in Tomasi (2021a) and Tomasi (2021b) as a vehicle for incorporating environmental justice considerations in NRDA.

[33] An example is the analysis undertaken at the Fox River, Wisconsin, site by Triangle Economic Research on behalf of the PRPs; in this case they scaled restoration in terms of “utils” (the hypothetical unit of utility) lost and gained; essentially a value-to-value approach that did not go through dollars. For a recreation version, see Parsons and Kang (2010).

There is a class of approaches other than the Kaldor-Hicks criterion for solving this problem of heterogeneous preferences. This class involves specifying a function that aggregates individual preferences into a social preference. These specify a SWF or other type of social ordering of individual circumstances that is used to judge tradeoffs in well-being across people. Importantly, these approaches allow alternative ethical considerations into the aggregation, in which case environmental justice considerations could be formally brought into NRDA restoration evaluation. A full exploration is beyond this document and would require research into new hybrid economic and REA/HEA methods.[34]

In a cooperative setting, the heterogeneous preference complication is often ignored; REA and HEA practitioners simply assume homogenous public preferences and appeal to the fact that the people harmed by a release are often the same persons benefiting from restoration, and that the array of restored services is sufficiently close to the array of injured services to diminish the bias. However, as a restoration program departs from one achieving narrow service and geographic nexus, the issue becomes more severe. A litigation setting may demand that the issue be addressed.

3.3.7 One Person, Multiple Time Periods

As we saw in Section 2, REA and HEA focus on the difference in service level provision at each point in time between the baseline scenario and the with-spill-and-restoration scenario. These are entire time-paths of service changes; discounting is used to convert changes in any given time period to its equivalent in a base year. This section is more explicit about what the discount rate is and the assumptions needed to derive REA and HEA from a more general model with economic values.

In some sense, we have already dealt with time when we discussed multiple services. The same service at two distinct points in time (denoted by t) can just be treated as different services. Here, we are more explicit and distinguish different bio-physical services from their appearance in time and write the list of services at time t as $(Q_{1t}, \ldots, Q_{St})$, with t running from some initial past date t^0 into the indefinite future.

We will start with a single service; call this Q_t. In a model with time, one can refer back to graphics such as Figure 3-3 and Figure 3-4 for *each time period*. However, elements of those graphics now require a time subscript on each variable, so that Q_t^B is baseline services in period t, and so on.

If restoration provided compensation for injury in each time period, our analysis would be done; time would not matter to scaling. But this is not the case in reality; in some periods there is injury and no restoration, in other periods restoration and no injury, and in other periods a mix of both. What we want is that the sum over time of the benefits of restoration compensates for the sum over time of the injury. The novel thing about time is that the same amount of gain or loss in different time periods is generally not equivalent in the eyes of the individual (or the public). Because individuals value the same commodity provided at different times differently, a discount rate is needed.

Let the base year for discounting be labeled date 0. All gains and losses occurring at some date t need to be converted to a date 0 equivalent using a discount rate. The value-to-value scaling equation requires that the discounted sum of WTP in each time period to avoid the injury equals the discounted sum of WTP to obtain the restoration. This provides a basis for developing a discount rate for HEA and REA.

The appropriate discount rate is called the social rate of time preference (SRTP). This is based on the public's willingness to trade a loss of a unit of consumption (a reduction in the bundle of goods and services) in one period for a gain in another period. The basic question is: if you give up 1 unit today (which costs you a loss of current well-being of some amount C), how much do you need to get back in the future (a break-even benefit B) to make the trade worthwhile? If there is a penalty for waiting such that a given increment to consumption is worth less in the future than now, then to entice you into the trade, B

[34] See Adler (2019) for a discussion of SWFs. Chapters in the volume edited by Adler and Fleurbaey (2016) describe a variety of approaches that might be adapted to the NRDA context. As previously noted, Tomasi (2021a and 2021b) adopts this approach to including environmental justice in NRDA.

has to be bigger than C. This must be, since you need more units of consumption in the future than given up now to make up for the fact that they are each worth less. This increase in future consumption is a rate of return on the investment of C needed to justify making it. This rate of return defines the discount rate. Suppose the benefit occurs t years into the future. There will be some average rate of growth r_t between now and then such that $B = C(1 + r_t)^t$; this defines the discount rate r_t that corresponds to this span of time. Note that if a unit of consumption in the future is worth *more* than it is now, fewer units of well-being need to be conveyed to the future than given up now, and the discount rate will be *negative*. This point is extremely significant for NRDA, especially for REA and HEA, which we will address below.

There are numerous and varied reasons why the value of a unit of consumption might change over time. One reason why there is a discount rate is that the individual may be impatient about when he or she receives utility. Let δ be the rate of impatience. If U_t is utility from consuming a bundle at date t, this is related to utility received at date 0 by the equation $U_0 = U_t/(1 + \delta)^t$. This says that consuming a bundle of goods in the future is worth less than receiving that exact same bundle at date 0. Absent any other considerations as to why the value of a unit of consumption is changing over time, impatience gives a reason for a positive discount rate.

An additional reason there is a discount rate is that the value of an increment or decrement to services changes over time due to a change in the baseline amount of services. This is unrelated to impatience. If I promise to give you $1,000 worth of consumption in five years, its value to you (relative to now) is different if you just won $10 million in the lottery than if you had just lost your job and sent a child to college.

Figure 3-5 reproduces Figure 3-3, with two time periods, under the assumption that the baseline regional supply of services is falling over time (e.g., wetlands are being lost due to climate change or the number of plovers is falling due to habitat modification and increases in predators). At an early time period, the relatively abundant baseline provides a relatively high level of service provision. An injury occurs in that period, the magnitude of which is ΔQ^I; this is measured in SAYs. The corresponding loss in value is the small amount of WTP ΔV^I. Restoration occurs in the future. Suppose the restoration provides an amount of service uplift ΔQ^R (also measured in SAYs) that is exactly equal to the amount of injury ΔQ^I. We can see from Figure 3-5 that this provides a value ΔV^R that is much larger than the loss of value ΔV^I since the increment to services occurs when the resource base is degraded and therefore is more valuable. That is, if injury occurs when a service is relatively common and restoration occurs at a later date when a service is relatively scarce, compensation may require *less* of a service gain than the amount of service lost to injury.

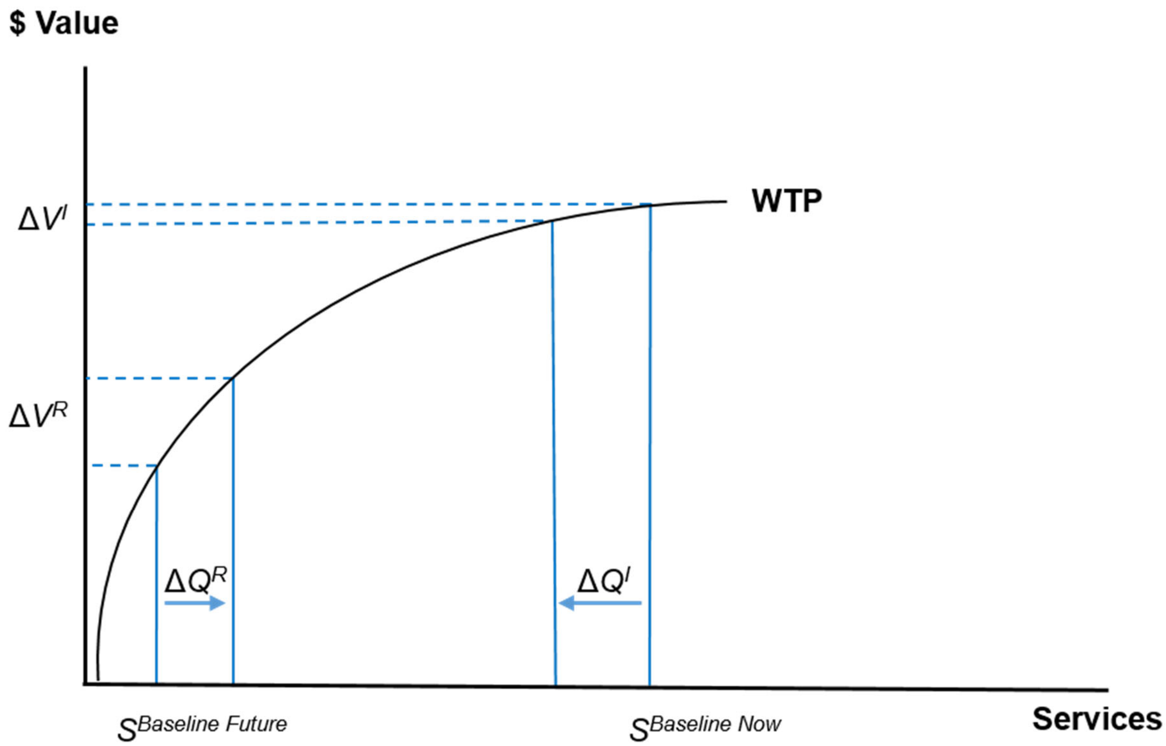


Figure 3-5—Value-to-value Scaling of Compensatory Restoration with Multiple Time Periods

If this was the only consideration in the discount rate (assuming the rate of impatience [δ] is zero), this degradation of the level of service provision at baseline implies a negative discount rate. Of course, if we switched the time periods and the injury occurs now with a low level of service provision and restoration occurs in a future, service-rich environment, the discount rate is positive.

We now turn to REA and HEA. Figure 3-4 shows that REA and HEA are a linear approximation to the value curve. Recalling the discussion of that curve, the incremental value associated with the approximation is the slope of the overall WTP curve at the baseline level of services. Let us now write this as $v(Q_t^B)$; call this v_t for short. We approximate the overall value change by $\Delta V \approx v_t \times \Delta Q_t$. In Figure 3-6, we put two different REA/HEA approximations on the overall WTP curve.

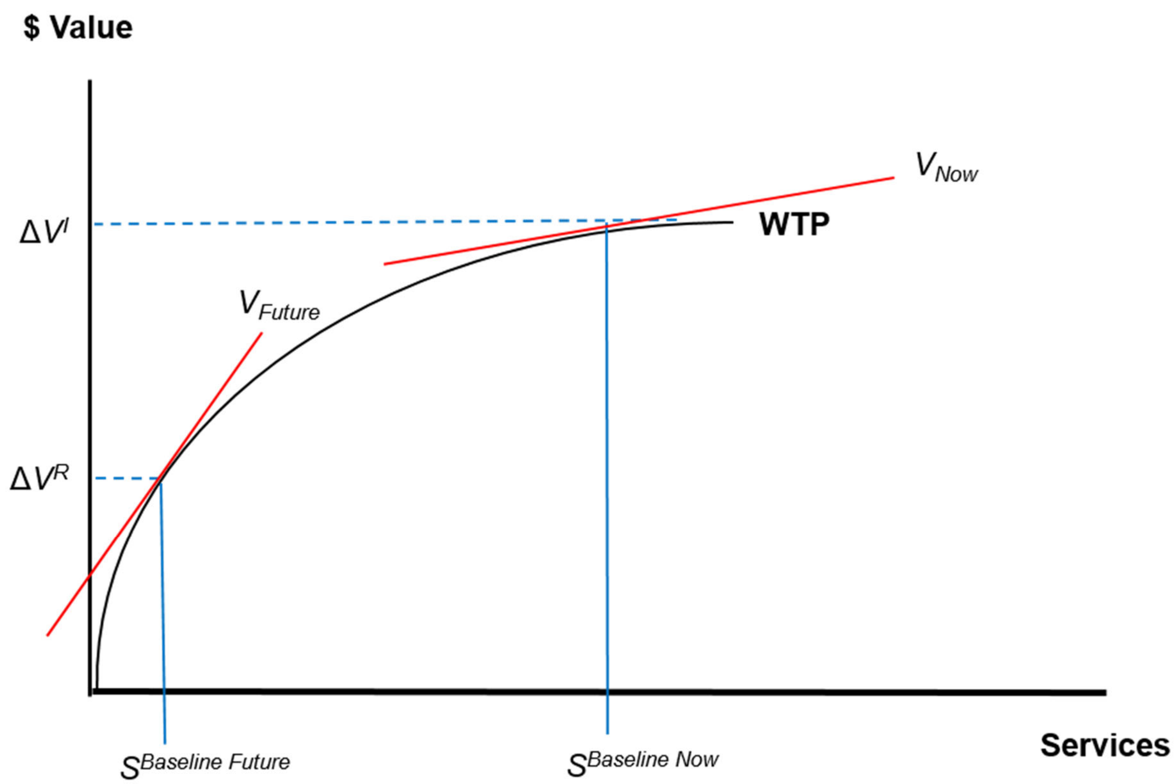


Figure 3-6—Linear Approximations to the Value Curve

The marginal value of services is now changing over time. There is not a single marginal value term v that can be canceled out of value-to-value scaling as we did for the timeless world. Instead, there is a different value v_t for each time period. There are two ways we can interpret Figure 3-6:

1) The traditional interpretation is to take a limited view of discounting and assume the discount rate reflects only impatience and is equal to δ as defined above. In that case, Figure 3-6 says that REA and HEA do not work. One must impose an assumption that the baseline is constant so there is only one marginal value term in value-to-value scaling that cancels out in every time period. This is the interpretation given by Dunford et al. (2004), English et al. (2009), and Desvousges et al. (2018). This is the source of the widely-held view that REA and HEA require a constant baseline. As indicated in the next bullet, this view is false.

2) Alternatively, the change in value over time could be embodied into the discount rate. The consumption rate of discount takes this approach. We can see from Figure 3-6 that the necessary adjustment to the discount rate involves two factors: (1) the degree of curvature of the WTP function and (2) the rate of change in the baseline over time. Let γ be a measure of the curvature of the WTP function and g_t be the average rate of growth (or degradation) in services between the base year for discounting and date t. Then the discount rate corresponding to date t is

$$r_t = \delta + \gamma g_t.$$

As discussed when talking about the REA/HEA approximation, the approximation of the overall WTP curve is itself a straight line. In this case, $\gamma = 0$; the change in baseline values does not matter regardless of the rate of change in services over time, and the discount rate equals the rate of impatience.

Theoretical Finding 9: The basic REA and HEA formulations hold only if the per-unit service value for each service is constant through time. This holds if either (1) the baseline service level and baseline income are constant over time, or (2) if either of these are changing, the utility function is linear in the element changing and its slope is independent of the other elements that are changing. However, if one knows the curvature of the WTP function, a changing baseline can be incorporated into the discount rate.

A comment on this result is in order. One often sees a HEA with a declining or rising or variable baseline. Does this invalidate the result just stated? Not necessarily. Imagine a regional amount of the resource/habitat, of which the injured and restored portions are small. The marginal values will reflect the *regional* supply of services, which may be approximately constant even if the baseline in a small area (the injured area) is falling or rising. In this scenario, marginal values may be constant over time, even if the "local" baseline, where the REA or HEA is being applied, is changing. It is when the regional baseline is changing and/or overall resource values are changing that the issues discussed above are applicable.

As a final comment, suppose that there are multiple services. If the relative amounts of these services are unchanging over time, we would have the same composite service index in each time period. The analysis is just the same as above for a single service, now applied to the composite service. But if the mix of services is changing over time, the composite needs to be recalculated each period. There is no simple fix using the discount rate, as *each service needs its own discount rate*.

3.3.8 Ecological Production of Services: The Tie to Science

The preceding results were developed in terms of natural resource services. Another relevant concept is that of the ecological production function (EPF). This function relates the services provided by a habitat/population to the attributes of the habitat/population.

Let the attributes of an injured resource be described by the list $a = (a_1, \dots, a_K)$, and A be the amount of resource, either acres in HEA or population numbers in REA. The general EPF takes the form $Q = (Q_1, \dots, Q_S) = G(A; a_1, \dots, a_K)$. This allows the services per unit of resource to be different at different scales. It is almost always assumed in a REA or HEA that $Q = AG(a_1, \dots, a_K)$, in which case services per unit of resource are constant as scale changes. Exposure to contamination will alter the attributes from baseline to impaired levels. That is, a^B will change to a^I. The change in attributes may alter services from Q^B to Q^I; this is injury in a REA or HEA.

The attributes (or a subset of them) can be thought of as service indicators or metrics. Estimating the changes $a^B - a^I = (\Delta a_1, \dots, \Delta a_K)$ caused by the spill is the province of the various biophysical sciences employed in NRDA.

Based on these estimates, one can use the EPF to compute service changes. One of the key issues in HEA is that the service indicators are (should be) closely linked to services. This creates the logical connection that:

$$(\Delta Q_1, \dots, \Delta Q_S) = Ag(\Delta a_1, \dots, \Delta a_K).$$

The ecological implications for services come to play in the per-acre function $g(\Delta a_1, \dots, \Delta a_K)$. Here, potential substitutions or synergies are identified. For example, assume one species of invertebrate declines due to contamination while another expands to fill that niche. If the service is "food for fish" and the fish involved are opportunistic feeders (not preferring one invertebrate to another once adjusted for biomass), a substitution has occurred in the way the fish food service is produced, but there has been no loss in the fish food service itself.

Theoretical Finding 10: Arithmetic manipulations of resource attributes in a HEA are meaningless in terms of the theory of compensation unless (1) the attributes are indicators of a specific service or set of services and (2) the way that attributes combine ecologically to produce final services is explicitly specified.

3.3.9 Some NRDA Implications

The preceding equations all express one idea: It is only under specific conditions that REA and/or HEA can be used to scale the amount of restoration that compensates the public for injuries to natural resources.

1) The models can "work" when restoration provides services that are of like kind and quality to those injured. This occurs when either (a) the same mix of services that was lost is provided (more likely in a REA than a HEA) or (b) the practitioner has a system of weights that convert individual services into a composite service index. To be consistent with the economic model, these weights are the relative values of the services' contributions to individual well-being.
2) The per-unit economic values of resource services must be approximately constant across the range of variation in services. This will be true if the amount of change in resource services is small compared to the regional supply of services to the public. If the change is large, practitioners must know that individual preferences for services are such that changes in the value of services depend only on the size of the change and not on the base from which it starts.
3) Incremental service values must be constant across time. This will be true if the baseline levels of natural resource services are constant through time. If not, then if you know how the service levels change through time, you may be able to adjust the discount rate to take this into account; however, this requires knowledge of or an assumption about preferences for resources and on how curved the value function is.
4) Each person must agree that the amount of restoration provided exactly compensates him or her. This will occur if you are doing an REA, and people's preferences, while they may differ, have a particular property (homothetic) that may or may not hold, but is likely to approximately hold if changes are small. In a HEA, where the mix of services provided by restoration likely differs relative to the mix of services lost due to a release, people must be assumed to have the same preferences to a close approximation. If neither of these is true, each person will disagree about the amount of restoration needed. For each to be fully compensated requires resource restoration plus an individual-specific side payment, positive or negative, that makes each individual's well-being absent the spill the same as with the spill, plus restoration, plus the side payment. If the side payments balance out such that their sum is positive, then restoration would pass a benefit-cost test. But, this requires one to know these values, and so just doing a HEA is not enough.

3.3.9.1 Substitution in NRDA

In general, relative to the baseline bundle (i.e., the mix of money and ecological services under baseline conditions), after compensation the individual may have different levels of specific ecological services in present-value terms. Thus, compensation in an economic sense does not require that each individual ecological service be restored. Rather, compensation requires that the bundle of ecological services in the with-spill-and-restoration scenario provide the same discounted level of utility as was provided by the baseline bundle of ecological services. In a HEA, this requires a set of weights be used in forming the composite service index.

The potential ability to substitute across services is part of basic economic theory. This allows one to search for cost-effective restoration actions that (1) have low per-unit costs and (2) provide services with a high rate of substitution with the injured resources. Unless HEA is augmented in some way to include an approximation for such substitution, using HEA can only increase restoration costs, all else equal. Of course, the appropriate way to augment HEA and REA with "weights" allowing substitutions is to do so using the economic concepts described above. However, HEA and REA are often applied in such a way as to ignore this potential substitution, which generally serves to increase restoration costs and NRD liability.

A corollary point arises frequently. Suppose that some but not all services are injured. It is common to just measure changes in injured services and ignore the ones that do not change. This is fine so long as the restoration provides this same set of services—no more and no less—with the same level of quality. If this is not the case, an adjustment needs to be made so that each unit of restoration has the same value as each unit of injury. Such an adjustment in a HEA is called a relative habitat value (RHV). It rarely is needed in a REA, because restoration typically affects the same population that was injured and services are reasonably assumed to be proportional to all affected resources in the same fashion.

Further, additional sources of substitution in the EPF come into play with respect to both the calculation of injury and the design of restoration. The EPF is the basis for specifying RHVs so that services can be made equivalent.

3.3.9.2 Non-Use Values in REA and HEA

Note that when the REA or HEA is applied in a manner consistent with the theory of compensation, the per-unit values of resources, $v(M_t, Q_t)$ are implicitly part of the scaling apparatus. They simply cancel out because they are constants. These unit values are defined generally by preferences (a utility function) for ecological services and include all the individual's motives for valuing resources. As such, they include what are variously called *non-use, existence, or passive-use values* to the extent they are relevant. Thus, the application of REA and/or HEA implicitly addresses such non-use values. To add these separately would constitute double-counting.

Theoretical Finding 11: REA and HEA include non-use, existence, or passive-use values.

3.3.10 The Key Theoretical Difference between REA and HEA

REA has two basic advantages over HEA when it comes to the theory outlined in the preceding text.

First, when each service is provided in fixed proportion to the amount of resource, any resource service functions as its own service aggregator. No values are needed, as in HEA, to form a composite service index that is consistent with the economic model. As such, REA appropriately assumes compensation has been achieved if the present value population with the spill and restoration is the same as it would have been at baseline.

In theory, an estimate of the total compensatory restoration requirement can be generated by implementing independent REAs for each injured population, so long as services of the populations do not interact and the selected compensatory restoration project for any one population provides no environmental service other than increasing the number of individuals in the population in question.[35]

In contrast, while HEA uses the same basic framework as REA, HEA simultaneously evaluates all of the services provided by an ecosystem. From a theoretical perspective, even when all the assumptions built into REA hold true, HEA is only consistent with the economic model of compensation under two circumstances:

1) If all services flowing from the injured ecosystem are injured in exactly the same proportion as will be provided by the restoration project (in which case it becomes REA); or
2) A system of weights allows the various injured and restored services to be converted to a single composite service. To be consistent with the full economic model, that system of weights would be the ratios of the per-unit values $v(M_t, Q_t)$ associated with each service.

In addition, REA is consistent with the economic model even when there are multiple types of people with different preferences, as long as preferences have a particular property (radially parallel indifference curves). HEA will only be consistent in this setting if there is no change in the mix of services between

[35] It is often difficult to identify projects that provide the exact environmental services needed and no others. A practical solution to this problem, which still avoids the need to estimate utility functions, would be the banking of services for subsequent sale to a RP in need of the specific banked service. For example, a wetland might be created to offset a spill-related injury to turtles. If the created wetland also contributed to groundwater recharge, the water-recharge services could theoretically be banked and sold to a RP with groundwater liability.

injured and restored habitats (it is a REA). Otherwise, it is not possible to assess exact compensation without knowing how economic values differ across people.

Finally, it is worth noting that most OPA NRDAs have used both REA and HEA to estimate compensatory requirements and then have added those requirements together to estimate the total compensatory requirement. Unless the services evaluated via REA are separate and distinct from the services provided by the habitats evaluated using HEA, simply adding the restoration requirements will result in an overestimate of compensatory requirements for any single oil spill even when all other necessary conditions hold. This issue is discussed in more detail in Section 7.

3.4 Practical Implications of the Link between Economics and REA and HEA

Here, we briefly summarize the conclusions developed in Sections 3.1 through 3.3.

- Compensation requires that the public experience no net loss in well-being (utility or degree of preference satisfaction) associated with an oil spill. Compensation is, at its core, an economic question related to individual preferences and not fundamentally a biological question. A model founded on this principle would implement value-to-value scaling in its most general form. This sets aside any issues of measurement and implementation.
- Compensation for a specific spill effect is required if the availability of some desirable ecological service is reduced as a result. Compensation is not required if the quantity of some commodity widely viewed as a nuisance/disamenity is reduced.
- If a spill results in the increased availability of some ecological services and the decreased availability of other ecological services, individuals (and therefore the public) may or may not require compensation depending on their preferences.
- When required, compensation can be accomplished by providing the public more money, more of the ecological services that were lost, more of some other ecological service, or more of any other good. In the long run, the public is better off if the most cost-effective means of compensation is implemented.
- If compensation were to be accomplished by providing money, each individual would need to receive the appropriate payment. Estimating the total amount of monetary compensation necessary and then spending that amount on ecological restoration (value-to-cost scaling) could result in overcompensation or undercompensation.
- Individuals tend to prefer compensation sooner rather than later; estimates of compensation must take this preference into account. This is accomplished by discounting goods and services that will not be provided until some future time period.
- REA and HEA, both forms of service-to-service scaling, are designed to identify the level of ecological services that, if produced, would compensate the public for spill-related impacts. Both models are based on the economic model of public compensation, but they embody several key limitations:
    - REA addresses one service, or a bundle of services, all tied to a single population in fixed proportions. Compensation for any one service is said to have occurred if the discounted quantity of the service is not decreased following restoration compared to the baseline.
        - Key assumptions that are needed, but may or may not hold true in a given case, include:
            - “small” changes in service levels relative to regional supply; and
            - a constant level of baseline services over the relevant time period or an adjusted discount rate.

- HEA addresses all services flowing from an ecosystem simultaneously by evaluating a composite service metric (which may use a single service indicator—that is, put all the weight in one proxy for all services, effectively turning the HEA into a REA). Compensation is said to have occurred if the discounted quantity of services, measured at the composite level, is not decreased following restoration. Total compensation for a spill can be estimated as the sum of the compensatory requirement associated with a series of HEAs, each assessing a different habitat, if and only if the services provided by each habitat are strictly independent from one another. If services are not strictly independent, summation of requirements associated with a series of HEAs will likely lead to overcompensation.

  - In addition to the limitations imposed by REA, HEA requires that:

    - there is one service, or, if not, the people in the population affected by both injury and restoration need to have the same preferences;

    - a system of weights allows the various injured and restored services to be converted to a single unit. To be consistent with the economic model of compensation, that system of weights would be the ratio of the per-unit utilities associated with each ecological service.

- If both REA and HEA are used to estimate compensatory requirements for a single spill, the total compensatory requirement is less than the sum of the individual estimates unless all services evaluated are strictly independent of one another.

4 Resource Equivalency Analysis in Depth

Provided the conditions identified in Chapter 3 are met, REA estimates the scale of compensatory restoration in a manner consistent with the economic model of compensation. As previously noted, REA assumes that all services associated with an individual resource are provided in direct proportion to the amount of the resource. That is, the number of individuals in a population is an aggregator of the many different services provided by the population.

Under those circumstances, REA essentially reduces to computing the difference between two population projections (Figure 4-1), the only purely economic consideration being discounting. One projection is with the spill (including positive and negative effects of the response) and the effects of a specific restoration project. The other is a projection of the population under baseline conditions.

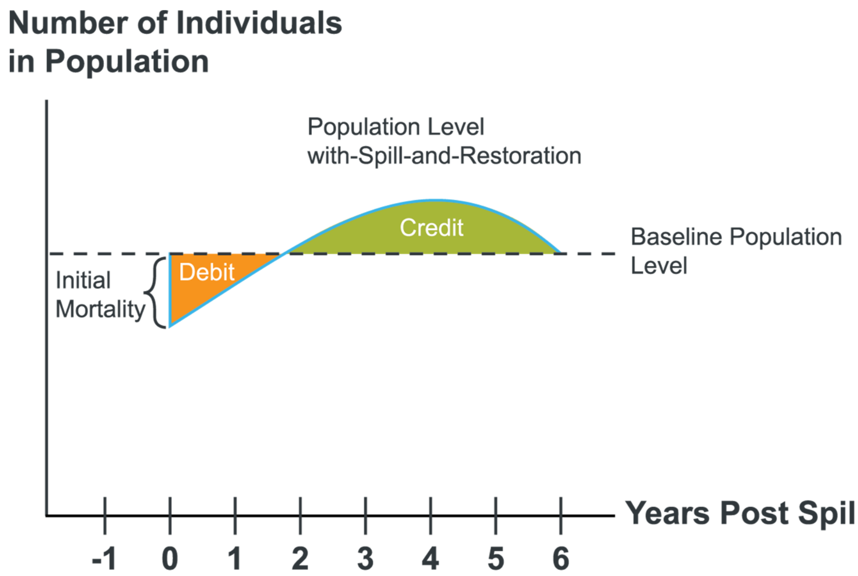


Figure 4-1—Graphic Representation of REA

In most situations, REA implementation should be a straightforward process of adapting existing population models to the REA framework and the specifics of the case at hand. The USFWS and state resource agencies have models that project future population levels for many species of birds and mammals. If there is no pre-existing model for a species, the required demographic parameters are generally available, the required computational models are well developed [starting with Leslie (1945)], and the computations are readily implemented in a fully transparent way in spreadsheets.

However, while the basic structure of REA is agreed upon, a variety of computational approaches have been implemented. For example, assessments conducted for the Chalk Point [(NOAA et al. (2002)]; Lake Barre [Penn and Tomasi (2002)]; *North Cape* [Sperduto et al. (1999, 2003)]; *Anitra* [New Jersey Department of Environmental Protection et al. (2004)]; *Star Eviva* [Skrabis (2004)]; *New Carissa* [Skrabis (2005)]; *Athos I* [NOAA et al. (2009)]; and Refugio Bay [CDFG et al. (2020)] oil spills each used REA calculation approaches that differ and occasionally conflict with one another. See Wakefield and McNutt (2008) for a more detailed discussion.

Importantly, the various REA *computational* methods have resulted in NRD liability estimates that differ by millions of dollars, even when all estimates of parameter inputs to the model, such as initial mortality, toxicological assumptions, biological assumptions, and ecological assumptions, are identical.

Thus, when restoration requirements for injuries to populations are assessed using REA, the key discussions are critically related to the mathematical structure of the population models underlying the REA and the ecological processes they are designed to reflect.

4.1 Implementing REA

The first data gathered in support of a REA analysis relate to the magnitude of the initial impact on the population. Several methods have been used to estimate the magnitude of initial impacts. The reader is referred to a report by ENTRIX (Annex C) for a review and assessment of different methods. All the methods seek to estimate the same parameters: the number of adults, juveniles, and young-of-the-year that were either killed or not created because of the spill.

In the next step, the demographic and ecological data necessary to project population levels into the future are assembled. The generally agreed-upon list of parameters for typical applications includes:

— annual survival rates by age class;

— maximum age attained by individuals in the population;

— proportion of the population that is female;

— age at first breeding among females;

— proportion of breeding-age females that attempt to reproduce annually;

— average number of offspring per reproductive attempt;

— an estimate of the recent population trend (increasing, decreasing, steady);

— for populations not increasing at their maximum intrinsic rate of growth, an understanding of what ecological factor(s) is/are limiting the population; and

— an understanding of the ecological mechanism(s) that would cause the with-spill-and-restoration population to return to and eventually exceed the baseline population over time.

The required information is available for most species through USFWS, state agencies, biological literature, and data repositories such as Birds of North America and Fishbase.[36] If data elements are missing or uncertain, they can usually be estimated by calibrating the parameters so that model predictions reflect recent population trends. Moreover, in our experience, as long as parameters are calibrated to reflect the observed population trend, even moderate variations in parameters result in relatively minor changes in estimated restoration requirements.

Once the necessary data have been assembled, the next step in REA is to specify the computational framework that will be used. There are three broad approaches: professional judgment of a recovery period without formal modeling; static arithmetic that does not consider ecological processes driving recovery; and population modeling. Two of these frameworks—professional judgment [Sperduto et al. (1999, 2003)] and static arithmetic [Skrabis (2005), Sperduto et al. (1999)]—are shortcuts that estimate debits and credits without formally projecting population levels. The third approach, population modeling, is best characterized as a “fully specified” REA, although exactly what specification is adopted can vary based on case circumstances and the biology of the resource.

Each method is briefly described below; readers are directed to the original references for more detail. The remainder of this section focuses on one-time mortality events; in situations where ongoing exposure occurs, the general approach can be adapted to apply.

4.1.1 Professional Judgment

The professional judgment approach to REA is exemplified by the Sperduto et al. (2003) treatment of gulls, cormorants, alcids, and gannets for the *North Cape* spill. No explicit population calculations were performed; professional judgment was used to specify debits and credits. This approach is often used when exposed populations are thought to be at carrying capacity pre-spill and likely to expand rapidly to fill all available habitat post-spill. In this case, recovery to baseline occurs within one breeding season independent of any spill-related restoration. If such rapid recovery occurs, one (or fewer) DBY of debit is associated with each mortality.

Restoration credit is then based on increasing the quantity or quality of available habitat under an assumption that the new habitat is rapidly filled by individuals that would not otherwise have existed. For instance, if converting a one-acre farm field to wetland increases mallard density in the area from zero to

---

36 https://birdsoftheworld.org/bow/home, www.fishbase.org.

two ducks per acre, two mallard-years are produced annually.[37] The REA credit associated with converting the field to wetland is then calculated as the present value of two duck-years produced annually over the life of the wetland.[38]

The professional judgment approach represents an expeditious path to compensation that may be applicable if the injured population is in a density-dependent equilibrium and is characterized as reaching adulthood rapidly, producing many offspring per reproductive cycle, and investing little energy in parental care. Ecologists often refer to these populations as having *r-selected* life strategies. The approach may be particularly applicable for waterfowl because density-dependent survival appears to govern population levels [Gunnarsson et al. (2013)].

In a slight modification to the professional judgment approach, CDFG et al. (2020) scaled pinniped and cetacean restoration by assuming that compensation for release-related mortalities could be scaled on a 1 to 1 basis (after discounting) to mortalities prevented through the funding of marine mammal response activities. This basically sets aside any nuanced differences between discounted animal-years arising from one mortality and those arising from one avoided mortality.

4.1.2 Static Arithmetic

The static arithmetic approach to REA is described by Skrabis (2005). In this method, populations of animals are not explicitly projected, and recovery mechanisms do not exist. This is not really an ecological model, but rather a way to use life history parameters to compute debits and credits without an underlying ecological justification behind the calculation.

The debit is estimated in two steps. First, the number of spill-related mortalities is multiplied by the average life expectancy of a population member. This is described as representing direct losses associated with individuals killed by the spill. Thus, if a spill occurs in 2019 (and that is the base year for the discounting) and kills 100 adult animals that on average would have lived for three more years, the total direct effect would be 291 discounted animal-years, as shown in Table 4-1.

Table 4-1—Direct Injury as Estimated via Static Arithmetic

Year	Loss of Animal Years	Discount Factor	Lost Discounted Animal Years
2019	100	1	100
2020	100	0.97	97
2021	100	0.94	94
			Total direct loss: 291

The second step is to add additional debit associated with the offspring these adults would have produced had they not died because of the spill (this is often referred to as production forgone associated with the F1 generation). Continuing the example, assume the 100 individuals killed represent 50 breeding pairs, each breeding pair produces on average 0.5 viable offspring annually, and on average the offspring live for four years. The additional debit associated with the F1 generation would be 278.9 discounted animal-years (Table 4-2). Thus, after accounting for the direct effect and the F1 generation, the total debit would be 569.9 discounted animal years.

---

37 Presumably, restoration projects valued using professional judgment increase carrying capacity and do not simply attract existing individuals from one area to another. In this expedited approach, the potential for landscape-level effects to increase the level of the target populations at sites in proximity to the restored site are ignored.

38 Note that the mallard compensation project would create other wetland services, which could be applied against other REA and potentially any HEA-based debit.

If the analyst opted to consider the F2 generation (i.e., the offspring of the F1 generation), the total debit would again increase. The 25 forgone offspring in the F1 generation would not produce 12.5 offspring of the F2 generation, losing four years of life expectancy.

Table 4-2—Indirect Injury Associated with the F1 Generation

Year	Animal-Years Lost Due to Progeny that Would Have Been Born in 2019	Animals-Years Lost Due to Progeny that Would Have Been Born in 2020	Animals-Years Lost Due to Progeny that Would Have Been Born in 2021	Total Animal-Years Lost in Any Given Year	Discounted Animal-Years Lost
2019	25	0	0	25	25.0
2020	25	25	0	50	48.5
2021	25	25	25	75	70.7
2022	25	25	25	75	68.6
2023	0	25	25	50	44.4
2024	0	0	25	25	21.6
2025	0	0	0	0	0.0
Total indirect injury associated with the F1 generation: 278.9					

Logical extrapolation of this thought process implies that offspring of offspring of the individuals indirectly lost due to the spill (i.e., the F3 generation) should also be included, and so on for all subsequent F_x generations, $x = 3,4,5, \ldots$. Since each initial mortality logically represents a debit that continues to accumulate indefinitely, one uses professional judgment to limit the calculations to a chosen number of generations (in practice usually one or two).[39]

REA credits in this framework are often based on the assumption that restoration projects will increase survival and/or productivity among a subset of the population. In parallel with the debit calculations, one estimates the number of individuals created/preserved annually and then multiplies that number by the (discounted value of) the average life expectancy of a population member. This calculation provides direct benefits. Credit associated with indirect benefits derives from the offspring of the directly created individuals and the average life expectancy of the offspring. Because the approach logically accumulates credits indefinitely, the credit calculation is truncated to one or two generations, often chosen to be the same as in the debit calculation.[40]

There are two primary limitations of the approach. First, the number of generations included in the calculation (1) has a major impact on the estimate of compensatory restoration and (2) is essentially arbitrary.[41] Determining the "correct" number can only be accomplished by comparing results to those generated by a population model, which obviates the need for the simplification. Second, by focusing on the individuals that died and what they would have done if the spill had not occurred, the approach does

39 The total bird-years lost increase with each additional generation included; it is possible that bird-years increase faster than the rate of discount, in which case the debit logically becomes infinite without a cutoff generation. This reveals the absence of ecological foundations of the approach.

40 Some have asserted that if the number of generations included on the debit and credit side is the same, how many to include is of little practical concern. However, REA results vary substantially according to the number of generations chosen. Simply choosing to include the same number of generations on the debit and credit side in no way increases the accuracy/reliability of this approach to scaling.

41 In theory, it appears that a static calculation that included a large (in the limit infinite) number of generations on the debit and credit side would approximate the level of restoration that would be required for a population not limited by any density-dependent mechanism, and so can grow without bound. For populations that are regulated by density-dependent mechanisms, there is no obvious rule of thumb.

not address how individuals that survived the spill actually respond to the post-spill environment and how that response dictates the evolution of the post-spill population level relative to baseline. Thus, the approach has no theoretical linkage to the question REA is attempting to address.

4.1.3 Population Modeling

The population modeling approach to REA is described in Appendix G of CDFG et al. (2004). The approach uses projection matrices to predict future population levels; debits and credits are a function of those projections.

4.1.3.1 An Example Projection Matrix

For a bird-specific model that recognizes two age classes (fledglings and adults) based on a post-breeding census, the model calculates the number of fledglings present at the time of census as follows. Let $A(t)$ be the number of adults at t, PB be the proportion of adults that breed, and PF be the proportion of animals that are female. Then, if CP is chicks fledged per breeding pair, the number of new fledges in year t, $F(t)$ is:

$$F(t) = [A(t) \times PB \times PF] \times CP \tag{3}$$

Letting SF be the rate of survival of fledges to become adults the next year and SA be the survival of adults, next year's adult population is:

$$A(t+1) = [F(t) \times SF] + [A(t) \times SA] \tag{4}$$

Given these equations and a starting population $A(0)$, the future population can be computed at any date. The components of the equations that are bold are determined in the model; the components that are not are parameters to be specified by the analyst and that must be estimated on some basis.

Equation (3) represents the production of fledglings in the census year. The first term, in brackets, is the number of occupied nest sites; this is the product of the number of adults, the proportion of adults that attempt to breed (some might be "floaters" available to take over nest sites should some adults die in a spill), and the proportion of the population that is female. Multiplying the number of nests by the number of chicks that successfully fledge per nest (CP) gives the number of fledges produced in year *t*.

Equation (4) then provides the number of adults in the next census year, t + 1. This is the sum of the number of fledges at *t* that survive through their first winter to adulthood, $F(t) \times SF$, and survival of adults from one year to the next, $A(t) \times SA$.

The basic model can be extended to multiple age classes and account for age of first breeding and age-specific breeding success and survival rates. Moreover, in the model above, survivorship (the proportion of the population that breeds) and fecundity do not depend on population levels. However, density-dependent survival and/or reproduction can be specified as appropriate, as can various limits or constraints, such as the number of available nest sites.

Given a starting population of adults in some base year, the model can be used iteratively to project the population through time. Simple spreadsheet computations can be used to do this in a transparent manner.

The model may be deterministic, with each of the key parameters specified as a single point estimate, or stochastic, with the parameters in any year drawn from a probability distribution of possibilities. Then, using statistical techniques such as Monte Carlo analysis, which is easily implemented through spreadsheet add-ons such as Crystal Ball or @Risk[42], one can simulate the population projections a number of times and compute the mean outcome (or any other statistic, such as 95th percentile) over a large number of runs. By specifying various forms of uncertainty in the future, one can simulate the implications of a variety of future contingencies for population outcomes. For example, effects of climate

[42] Crystal Ball and @Risk are used strictly as examples of spreadsheet add-ins that provide a specific statistical analysis. API documents do not endorse or require the purchase or use of proprietary products or services.

change on habitat availability can be included, as can extreme weather events during migration that may alter survival.

4.1.3.2 Using Projected Populations to Estimate Debits and Credits

Both baseline and with-spill-and-restoration population projections use the same equations and the same demographic parameter estimates in any given year. However, the with-spill-and-restoration projection differs from the baseline projection in three ways. First, any release-related reduction in young-of-year creation is subtracted from the number of young-of-year created under baseline conditions. Second, individuals aged 1 or greater lost to the release are incorporated by subtracting the number of adult mortalities from the total number of adults in the spill year under baseline conditions.[43] Third, the effect of restoration is incorporated into the with-spill-and-restoration projection by adding an appropriate increment to the number of young-of-year produced and/or the adult survival rate based on the specified effect of the restoration project.

Often, the population models used in REA assume the proportion of the population that breeds, annual survival rates, and fecundity rates are independent of population levels, suggesting there is no density dependence and therefore no natural recovery mechanism. While the absence of density dependence may accurately describe some species, other species may be regulated by density-dependent mechanisms. When appropriate, density dependence is incorporated into models by specifying a functional relationship between population levels and survival and/or productivity.

An approach that directly incorporates density-dependent effects is described in CDFG et al. (2004). In this assessment, multiple avian REAs are performed by asserting a specific form of density dependence, which the authors refer to as “single-generation stepwise replacement.” The approach assumes that “each year after the spill the juvenile age class will be entirely replaced. That is, despite the fact that some breeding adults have been removed from the population, the population produces the same number of juveniles post-spill as it would have under baseline conditions.” This could occur if, prior to the spill, a population was at carrying capacity perhaps limited by available nesting habitat or food.

In theory, the population modeling approach to REA can be used to identify the proper scale of restoration under most circumstances. The primary limitation of the population modeling approach is that, when populations are characterized as having a non-trivial risk of extinction in the near term, the model could underestimate the required level of compensatory restoration.[44]

4.2 Evaluating Alternative Approaches for Implementing REA

The simple approaches to REA (professional judgment and static arithmetic) are easy to apply, have few moving parts to negotiate, and, hence, could be used to reduce transaction costs. Population modeling is somewhat more complex, but still relatively easy to implement via simple spreadsheet programming.

The purpose of this evaluation is to investigate the ability of simple methods to predict the amount of restoration required compared with more complete models (i.e., to identify situations in which the simple approaches work well). We also seek to identify which assumptions in the simple methods lead to better approximations in the different situations studied. This will afford some insight into when one may be more comfortable with the simple approaches, and when it may be appropriate to incur the greater transaction costs associated with a somewhat more complex method.

---

[43] The age of adults lost to the spill is often assumed to follow the age distribution of adults in the population. However, this assumption can be modified if incident-specific conditions warrant.

[44] If the deterministic calculations outlined in Section 4.1.3.1 suggest the population will persist through time (and so credit will accumulate) but, in fact, the population goes extinct, restoration credits derived from future time periods would not be realized. Under this scenario, the population-based REA would underestimate the true level of restoration required.

Of course, evaluating the performance of a simple method applied to a complex environment requires a standard of comparison (i.e., it is necessary to know what the "true" result is so we can determine how well various approaches approximate that truth). We do not claim to know with certainty the "true" amount of compensatory restoration required when a spill impacts a complex, highly variable resource. Rather, we create a model that estimates the exact amount of restoration that would be required *if* a specific series of events occurred. Those events include a specific initial impact, a set of randomly selected shocks to future survival and productivity, and an assumption about the mechanisms that regulate the population. We then use the common approaches to REA to generate restoration estimates that correspond to that specific series of events and compare the exact restoration requirement associated with the specific series of events to the predicted restoration requirements. The difference between the exact and estimated restoration is defined as the error rate.

In this document, we report initial results along these lines. We have not yet tested the robustness of the alternative REA methods across a wide variety of environments. Thus, we recognize the limits of the conclusions that can be drawn based on our current comparisons.

4.2.1 Comparison of REA Approaches

The mathematical methods used to evaluate REA approaches are described in Section 4.2.1.1; those not interested in the details may find it convenient to skip to Section 4.2.1.2.

4.2.1.1 Calculating Exact Restoration Requirements

We use a model developed by USFWS [Melvin and Gibbs (1996)] to determine the exact amount of restoration that would be required in a specified environment. The model requires four inputs to project baseline population levels through time: the number of individuals in the population at the outset of the model run; annual survival (the probability that an individual lives to the next year); the proportion of the breeding-age population that attempts to breed; and reproductive success (young-of-year produced per breeding pair, or fecundity).

There are several age classes: young-of-year in year t, (Y_t) and age i adults alive at t ($A_{i,t}$), where the age index i runs from 1 to the maximum age (M). The analysis assumes half of each age class is female.

Equation (5) calculates the number of young-of-year in year t. This is the sum of the offspring of adults that breed. We let P_i be the proportion of age-i adults that breed and F_i be their fecundity (i.e., number of offspring per breeding female). Then, we have:

$$Y_t = \sum_i (P_i \times \frac{A_{i,t}}{2} \times F_i) \tag{5}$$

Equation (6) calculates the number of age $i+1$ adults in year $t+1$, based on the number of age i adults in year t and their survival rate (S_i) as:

$$A_{i+1,t+1} = A_{i,t} \times S_i \tag{6}$$

It should be understood that $A_{0,t} = Y_t$, that S_0 is the survival of young-of-year to become age-1 adults, and that $S_M = 0$.

These equations are applied sequentially to project the population size for all $t \geq 1$. In doing so, Monte Carlo methods are used to incorporate uncertainty in future conditions. To do this, adult survival, fledgling survival, and productivity are drawn from a distribution rather than being assigned a fixed value. Each simulation includes one baseline and one with-spill-and-restoration projection; these projections use the same randomly drawn demographic parameter(s) in any given year.[45]

The with-spill-and-restoration projection differs from the baseline projection in that the magnitude of the initial impact is incorporated. This is done by subtracting the estimated number of young-of-year killed by the spill from their number in the release year at baseline. Likewise, adults lost to the spill are subtracted from their number in the release year under baseline conditions.[46] Finally, the effects of restoration are incorporated by adding a user-specified increment to the number of young-of-year produced at a user-specified number of breeding sites during years in which restoration is in effect. The amount of restoration required is determined by adjusting the scale of restoration such that the baseline and with-spill-and-restoration populations provide the same number of DBYs.

For each environment evaluated, the exact amount of restoration is calculated for each of 5,000 different simulations. The distribution of outcomes across the 5,000 simulations is summarized as a mean and standard deviation.

4.2.1.2 Scenarios Evaluated

Scenarios were selected to represent the range of potential circumstances encountered by REA practitioners. The range includes populations regulated by density-independent and density-dependent mechanisms, and, for each mechanism, populations that were increasing, in steady state, and decreasing at the time of the release. In all cases, the extinction risk over 100 years is essentially nil.[47]

Table 4-3 contains parameters used in the assessment of populations not regulated by density dependence. Table 4-4 contains parameters used to assess density-dependent populations. In all cases, it is assumed that 10 young-of-year were not created in the spill year (year zero) and 20 adults were killed by the spill.

The professional judgment approach is not evaluated. We evaluate the static arithmetic approach including one or two generations and a deterministic population model.

In all cases, the assumed restoration project increases productivity at a specified number of breeding sites by 25 percent beginning in year 5 and lasting through year 9. The amount of restoration required is estimated as the number of breeding sites to be restored annually.

---

45 More types of stochasticity could be included, such as catastrophic environmental shocks or loss of management funds, and these could vary over time (e.g., increasing severity of storms or loss of habitat due to climate change). Such additional analyses are beyond the scope of this paper.

46 The age of adults lost to the spill is assumed to follow the estimated age-distribution of an equilibrium population.

47 Increasing and decreasing population levels may violate the requirement of a constant baseline service level through time if the population for NRDA analysis is a large fraction of all birds in the population valued by the public. NRDA REA applications assume this “smallness” and that the entire population is approximately constant. If the NRDA population is a large fraction of the total, and there are changes in the risk of extinction, this assumed constancy of values may not be not technically defensible. In this case, a modification of the discount rate may be in order; see Sections 3 and 7.

Table 4-3—Input Parameters for Density-independent Scenarios

	Increasing 1 % Annually	Steady State	Decreasing 1 % Annually
Pre-spill adult population	1,500	1,500	1,500
Fecundity (young-of-year per pair)[1]	1.27 (SD 0.25)	1.27 (SD 0.25)	1.27 (SD 0.25)
Survival (young-of-year to age 1)[2]	0.48	0.48	0.48
Adult survival[3]	0.74 (SD 0.060)	0.74 (SD 0.060)	0.74 (SD 0.060)
Proportion of age 1 breeding	0.662	0.539	0.42
Proportion of adults greater than age 1 breeding	1.0	1.0	1.0
Maximum age	12 years	12 years	12 years

SD = standard deviation

[1] The standard deviation is consistent with observed variability for piping plover populations in New England from 1994 to 2003.

[2] It is assumed that the relative deviation from mean survival is perfectly correlated for young-of-year age 1 and adults; therefore, a standard deviation is not necessary for the young-of-year to age 1 survival.

[3] The coefficient of variation is consistent with the findings described in Melvin and Gibbs (1996).

Table 4-4—Input Parameters for Density-dependent Scenarios

	Initial Population 25 % below Stable Level	Steady State	Initial Population 25 % above Stable Level
Approximate number of age 3+ adults at time of release	425	566	708
Fecundity (young-of-year per pair)	0.538 (SD 0.096)	0.538 (SD 0.096)	0.538 (SD 0.096)
Survival (fledges to age 3)	0.458	0.458	0.458
Annual survival (age 3+)	0.952 (SD 0.048)	0.952 (SD 0.048)	0.952 (SD 0.048)
Maximum age	30 years	30 years	30 years
Age at first breeding	6	6	6
Percentage of age 6+ breeding initially	100 percent	81 percent	57 percent
Available nest sites	145	145	145

SD = standard deviation

4.2.2 Results of Comparison of REA Approaches

Table 4-5 summarizes the estimated number of nests to protect using the various approaches to REA. The exact amount of restoration is calculated for 5,000 different event sequences; those estimates are summarized as a mean and standard deviation. These are directly comparable to the estimates generated by the deterministic population and two static arithmetic models.

Table 4-5—Predictive Power of REA Approaches

	Density-Independent			Density-Dependent		
REA Framework	**Increasing Population**	**Steady Population**	**Decreasing Population**	**Increasing Population**	**Steady Population**	**Decreasing Population**
Actual (mean, SD) sites to be restored	42 SD = 9	42 SD = 9	42 SD = 9	112 SD = 15	66 SD = 7	65 SD = 7
Population model estimate	41	41	41	95	62	62
Static 1 generation estimate	68	68	68	149	149	149
Static 2 generation estimate	60	60	60	131	131	131

4.2.3 Discussion of Comparison of REA Approaches

The population modeling approach to REA performs fairly well. This result is robust to assumptions regarding the magnitude of initial impact, the interval between the spill and the initiation of restoration, the duration of the restoration project, the overall population trend, and/or the existence or absence of density-dependent regulation mechanisms. In contrast, static approaches, as they are typically employed including a small number of generations, are consistently and significantly biased toward overestimates of restoration requirements. The one-generation model performs especially poorly.

This result is not unexpected. Some REA practitioners who use the static approach acknowledge that population modeling likely performs better for well-understood systems where extinction risks are minimal. They justify the use of a static approach as an ad hoc method to address uncertainty when extinction risks are non-trivial. This is because, if extinction occurs, the generation of credit ceases, and so a model that ignores extinction risks may predict accrual of credit that is not realized.

To investigate this, we evaluated the accuracy of static and population modeling-based REA approaches under non-trivial extinction risks. The purpose of this evaluation was to inform the REA practitioner as to the level of extinction risk that may cause population model-based results to diverge from the actual restoration requirements, setting aside the possibility for non-constant per-unit utility through time.

Figure 4-2 is a graphic representation of the predictive power of static and population modeling-based REA approaches under various extinction risks.[48] The analysis was based on the demographic parameters identified in the preceding tables; however, adult survival, age at first breeding, and/or productivity were set to trend downward through time until the necessary extinction risk was achieved. The vertical axis is the number of breeding sites restored annually; the horizontal axis reports the computed extinction risk.

[48] The results of this analysis are only applicable if the per-unit utility is constant even as a population approaches extinction. If, as is commonly observed, public preferences are characterized by diminishing marginal utility, the population approach is unlikely to under-predict restoration requirements even under high extinction risk scenarios.

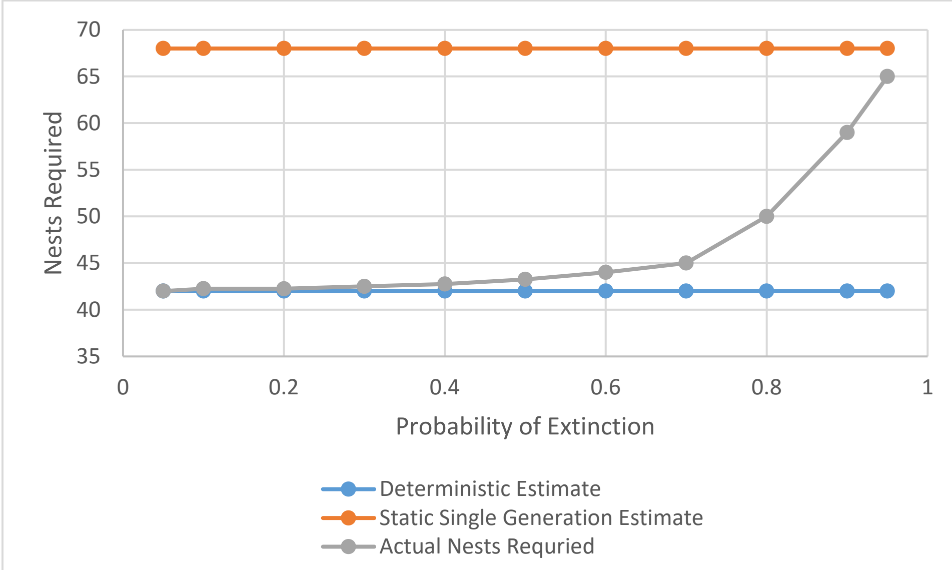


Figure 4-2—Predictive Power of REA Approaches as a Function of Extinction Risk

At moderate probabilities of extinction risk, the deterministic population approach to REA predicts restoration requirements well. It is only as the 100-year extinction risk approaches 70 percent that the population model begins to lose its predictive power.

An extinction risk as high as 70 percent would be extremely unusual. For context, the extinction risk of the endangered Atlantic population of piping plovers was near zero in 2010 [Plissner and Haig (2000)], while the extinction risk of the endangered northern spotted owl was estimated to be about 10 percent [Akcakaya (1998)]. Further, if extinction risks exceeded 70 percent, it is unlikely that economic assumptions justifying the use of REA would hold, as the marginal value of a population member would rise significantly as the population approached zero (see Section 3).

4.3 Common REA Issues and Sensitivities

The first commonly encountered REA issue relates to the selection of an overall REA approach. As long as affected populations will not expand rapidly to recover to their baseline level within one or two breeding cycles, one must choose between static arithmetic and a population modeling approach. Our analysis suggests that under nearly all realistic scenarios we have evaluated, the static approach to REA overestimates actual restoration requirements by as much as 50 percent. If injuries are minor and restoration is cost-effective, this overestimate may not be meaningful and the simpler approach is cost-effective. However, as was the case for marbled murrelets and common loons impacted during the *New Carissa* spill (USDOI 2006), the selection of static arithmetic to implement REA can result in liability overestimates that range from hundreds of thousands to millions of dollars.

The second commonly encountered REA issue relates to the assumption of small changes to current service levels and a relatively constant baseline service level through time. The assumption is often violated when species listed as threatened or endangered are impacted by a spill. This is because listed species often have populations that are small and trending downward. Under this scenario, bird-years provided in the future are likely to provide more utility than the bird-years lost to the spill and, as a result, REA tends to overestimate compensatory requirements. While REA might still be used, it would require an adjustment to the discount rate to reflect trends in service values. To the best of our knowledge, the potential magnitude of this error has not been quantified. A practical and defensible method awaits further research.

The third REA issue relates to the selection of demographic parameters. Specifically, any one set of demographic parameters implies a population trend (increasing, decreasing, or stable). Often, parameter bundles used in REA imply a population trend that is inconsistent with the observed population trend. For example, a bundle of demographic parameters may imply that a population is doubling every five years when, in fact, the population is observed to have been decreasing. In this case, the REA practitioner must use professional judgment to select a bundle of demographic parameters that is consistent with both the range of demographic parameter values reported in the literature and the observed population trend; we refer to this process as parameter calibration. Failure to calibrate parameters can result in overestimates or underestimates of restoration requirements, though these errors are unlikely to be significant. Of course, if certain parameters govern mechanisms of injury and restoration, scaling is more sensitive to selection of values for these. A range of values for scaling parameters, with values for others recalibrated to match trends, can be used to bound the range of effects.

A final REA issue, although not unique to REA, arises when a restoration project provides both effects on the subject resource of a particular REA, as well as other services. For example, a project that enhances nesting habitat by controlling invasive species may benefit other species or other habitat services. Certainly, if there is injury to similar habitat, the multiple benefits of the REA project should be credited against injuries.

4.4 REA in Cooperative versus Litigation Settings

As with all NRDAs, behind cooperative assessment or settlement negotiations lurks the prospect of litigation. Litigation of OPA NRDA cases is unusual. Even in the Deepwater Horizon case, the NRDA was settled. That said, the defensibility of methods is an issue in almost all negotiations with material damages at stake. From the perspective of a RP and/or grustee agency likely to be involved in multiple assessments, the establishment of precedents related to flawed methods may be increasingly problematic.

REA is not inherently indefensible. It uses well-developed, published methods for modeling individual populations. Estimation of life-history parameters employs established procedures. Further, some of the more difficult aspects of HEA regarding multiple services and people with heterogeneous preferences are absent in REA. The primary litigation issues for a REA likely relate to the issues described in Section 4.3.

4.5 REA Summary

REA is most commonly used to estimate compensatory restoration requirements for the loss of services tied to a single population (usually a particular species of bird, turtle, marine mammal, or fish) when the population has been injured disproportionately relative to its habitat.

Conceptually, REA is a three-step process:

1) The baseline population level is projected through time.
2) The population level given the spill and a restoration project is projected through time.
3) The size of the restoration project is adjusted until society experiences no net loss of discounted species-years.

There are three broad approaches to REA: professional judgment, static arithmetic, and population modeling. Two of these methods, professional judgment and static arithmetic, can be characterized as simplifications that attempt to directly estimate debits and credits without projecting baseline and “with-spill-and-restoration” population levels. The third approach, population modeling, actually estimates debits and credits as the difference between baseline and “with-spill-and-restoration” population projections.

The professional judgment approach is relatively accurate if the injured population is (1) in a density-dependent equilibrium and (2) characterized as having an r-selected life strategy. Under these circumstances, recovery of the injured population to baseline levels may be relatively rapid, and the professional judgment approach to REA may represent an expeditious path toward restoration. Under all other scenarios where REA is an appropriate tool, population modeling is the preferred REA approach, as static arithmetic tends to significantly overestimate compensatory requirements.

5 Habitat Equivalency Analysis in Depth

HEA is used to estimate compensatory restoration requirements under OPA when compensation for impacts to multiple services flowing from a single habitat is to be accomplished via habitat restoration.

At a conceptual level, HEA can be thought of as functioning very much like REA. However, rather than the number of individuals in a population acting as the aggregator of the many services provided by the population, HEA relies on a composite service index to function as an aggregator of the many services provided by an acre of habitat. If, in fact, all services provided by a habitat moved in fixed proportions to one another, the HEA is effectively a REA with the amount of resource denominated in acres. A HEA is also analytically a REA if a single services indicator is used to represent all services for both injury and restoration and the only available restoration is more acres of a similar habitat. A true REA for the indicator might look for cost-effective ways to restore the indicator itself, in addition to a habitat project.

When the regional composite service level with-spill-and-restoration is below baseline, a debit accumulates. When the regional composite service level with-spill-and-restoration exceeds baseline, a credit accumulates. The public is compensated when the discounted value of the debit (the orange area in Figure 5-1) is equal to the discounted value of the credit (the green area in Figure 5-1).

Within this conceptual framework, HEA can be thought of as a three-step process:

1) The regional baseline composite service level is estimated for all post-spill time periods.
2) The regional composite service level through time is re-estimated given the effects of the spill and the effects of a restoration project. This is represented by the solid blue line in Figure 5-1.
3) An iterative process is used to identify the size of the restoration project that, when implemented, ensures that society experiences no net loss of DSAYs.

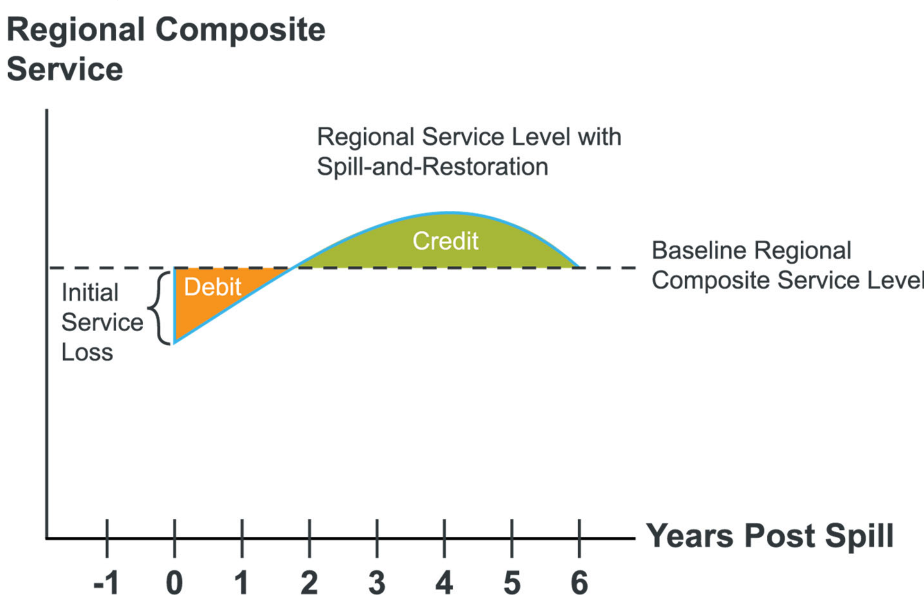


Figure 5-1—HEA Illustrated as the Regional Provision of a Composite Service

Figure 5-1 shows the composite service level for some habitat at a regional level (e.g., all composite wetland service in the region or all composite sandy shoreline service in the region). The regional aggregation may obscure an important difference between HEA and REA. In REA, one population is typically affected by both the spill and the restoration. In HEA, the impacted site and the compensatory restoration site are often at different physical locations. Thus, in a HEA, it is likely that (1) the compensatory restoration project provides somewhat different services, and (2) the effects of compensatory restoration do not actually speed recovery to baseline at the injured site.

It is because of the preceding facts that some practitioners prefer to visualize HEA as it was illustrated in Section 2. As discussed in Section 2 (and reproduced here as Figure 5-2), composite service flows from two discrete sites: the injured site and the restored site. Compensation is achieved when the debit associated with the injured site is equal to the credit associated with the restoration site.

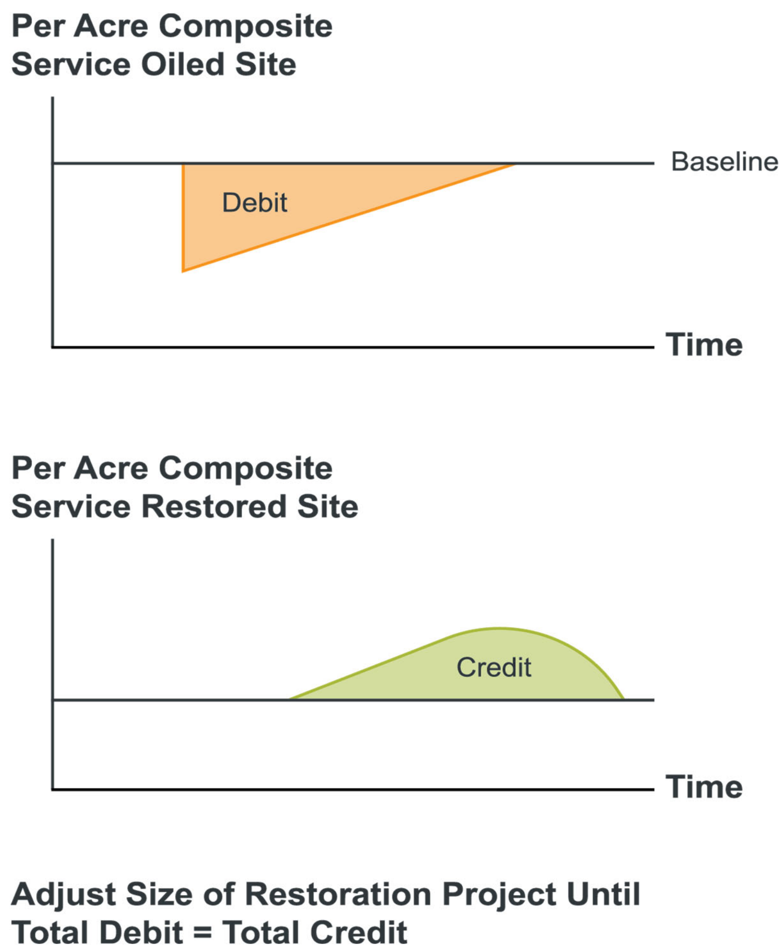


Figure 5-2—HEA Illustrated as Composite Service Flowing from Two Discrete Sites

Regardless of how the HEA model is visualized, the unit of analysis in HEA is a SAY, which is defined as the level of service provided by a base acre in one year. The injured and restored habitats are then judged relative to the base acre. For example, if the base acre is a pristine site, a degraded acre of habitat may generate 50 percent services (or 0.5 SAYs). SAYs occurring in the future are discounted to reflect the fact that society does not have the same value for a given service occurring in different time-periods. The resulting unit is a discounted service acre year (DSAY). That is, the DSAY is a composite measure of all the flood control, carbon sequestration, bird-watching opportunities, etc., provided by one acre of the base site.

Before discussing the actual HEA process, we believe it helpful to expand on two concepts: (1) HEA's reliance on a composite service index and (2) establishing the baseline relative to a base acre.

Composite Service Index: In Section 3, it was established that HEA would be most defensible if either (1) all the individual services flowing from the oiled habitat were injured in exactly the same proportion as would be provided by the restoration project or (2) only one individual service was thought to provide value. This obviates the need to aggregate the many services provided by a habitat into one composite, renders moot the issue of heterogeneous preferences for services among the humans being compensated, and simplifies the discount rate issue.

Unfortunately, rarely if ever do these conditions hold, and to limit HEA to such circumstances would greatly reduce its applicability. Thus, the primary HEA challenge is often dealing explicitly with multiple services changing in different ways at different locations and at different times. When thinking about the "multi-service" challenge, it can be helpful to distinguish between service indicators, individual services, and the composite service, and to assign terminology to differentiate between service indicators, individual services, and composite service:

— Because it often is not possible to measure a service directly, scientists measure habitat attributes (e.g., stem density or invertebrate biomass) thought to be indicative of the level of individual service(s) being produced; we call those measurable attributes "service indicators."

— The effects of a spill, response, and restoration on service indicators is based on how the indicators are produced in habitats. How these effects are predicted using measurements of oil concentrations and observations relative to reference areas is determined within the ecological production function (EPF).

— The equation (implicit or explicit) that converts service indicators into a service level estimate for each given individual service is a service production function.

— The equation (implicit or explicit) that converts levels of the individual services being provided into a composite service level is an aggregation method.

Figure 5-3 illustrates these relationships.

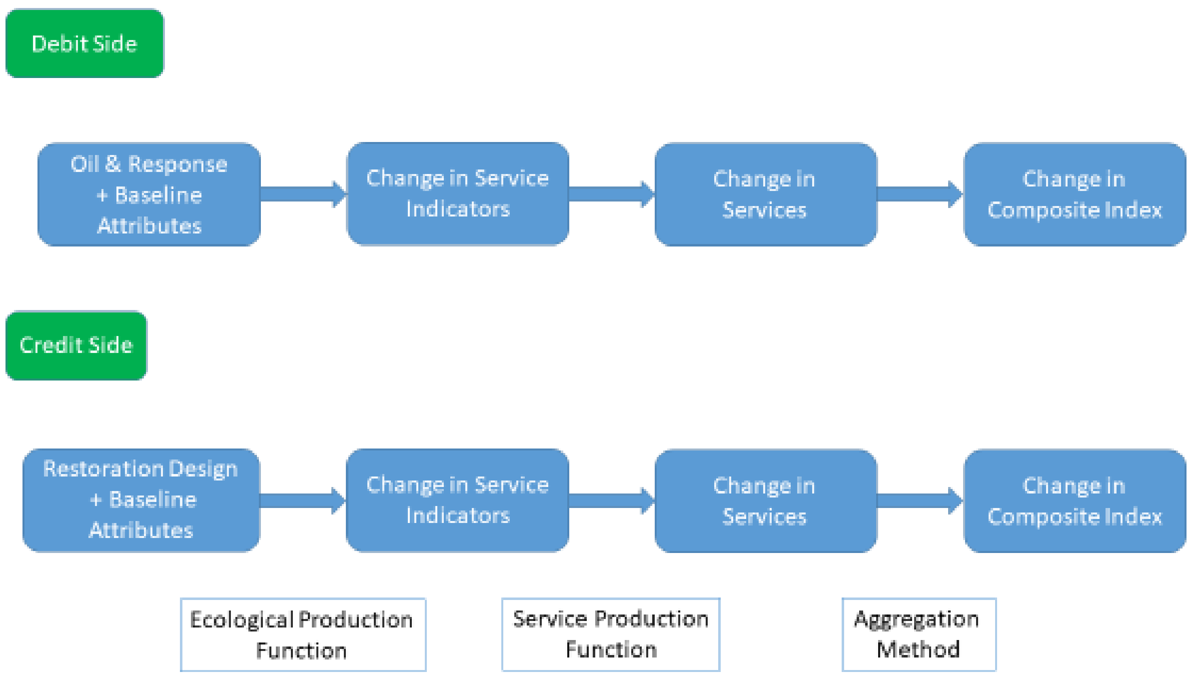


Figure 5-3—Relationships Between Service Indicators, Services, and the Composite Service

Baseline and the Base Acre: When HEA is visualized in terms of debits and credits flowing from two discrete sites (see Figure 5-2), the composite service level provided at each individual site is expressed relative to some chosen acre of habitat that serves as a basis for comparison. We call the habitat that serves as the basis for comparison the "base acre."

It does not matter what that base acre is; it could be an acre of "ideal" habitat, a nearby but un-impacted (i.e., reference) acre, or a pre-spill acre from the site that was actually injured. The choice is a matter of convenience and convention; HEA only requires that both the injured site and the restoration site be judged relative to the same base acre. For example, if the injured site in the year prior to the spill is the standard of comparison, the baseline service level at the injured site (the service level that would have been provided by each acre of the injured habitat "but for the spill") is likely to be defined as 100 percent at the outset of the spill. If that site would not have changed much over the relevant time period absent the spill, baseline would appear as a straight line, much as it does in Figure 5-2.

However, the baseline condition at an injured or restored site does not need to remain constant through time. To illustrate this idea, suppose the selected base acre is the affected site just prior to impact. However, due to the development of an adjacent site and associated runoff of sediments, the absolute level of service flowing from the impacted site under baseline conditions would have decreased linearly to 50 percent of the base acre by year 5. This decreasing baseline is illustrated in Figure 5-4. If a spill occurring in year 1 caused the level of the composite service flowing from the injured site to be reduced to 75 percent of the base acre until time period 3, at which time the service level returned to its baseline trajectory, the HEA debit, expressed on a per-acre basis, would appear as in Figure 5-4.[49]

[49] Would the situation illustrated in Figure 5-4 violate the HEA requirement of a constant baseline? Maybe, but not necessarily. The economics behind HEA requires the regional level of composite service to remain constant through time to ensure that the

Similar diagrams can be drawn for sites that, under baseline conditions, would generate more service through time or for sites that have annual or even seasonal service-level variation. The key is that the base acre is the standard against which both the injured site and the restoration site are judged.

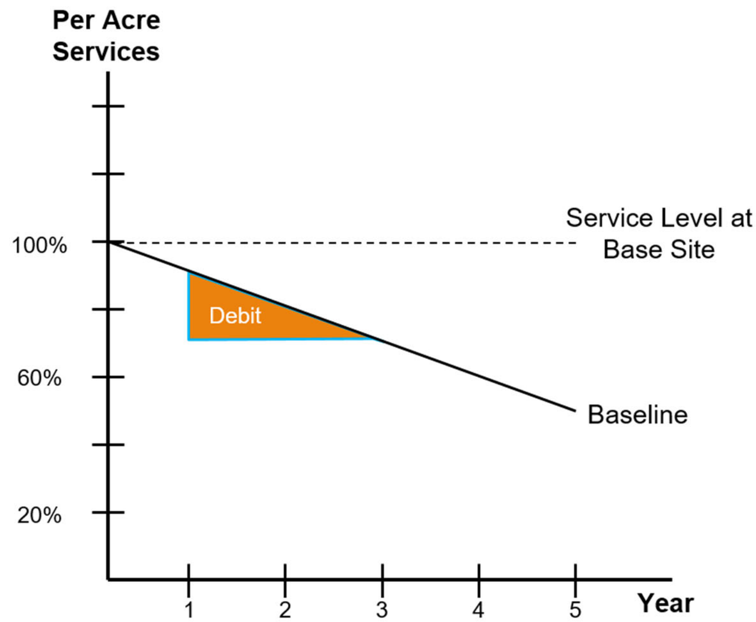


Figure 5-4—HEA Debit Relative to a Decreasing Baseline Service Level

It is common for HEA practitioners working in a collaborative setting to have disagreements related to the baseline and for those disagreements to cause damage estimates that diverge by hundreds of thousands or millions of dollars. In our experience, these disagreements tend to be conceptual, not factual. For example, an assessment team may agree that the baseline rate of shoreline erosion in an area is four feet per year. Disagreements may arise as to how that baseline erosion rate affects the credits associated with a beach nourishment project that increases the width of a beach that is currently 100 feet to 180 feet. Should credit for the project only be associated with the 20 years when the beach width exceeds 100 feet? No; credit would be generated so long as the beach is wider than it would otherwise have been, which would be about 45 years assuming the beach is backed by some barrier such as sheet piling, building foundations, or a road.[50]

In our experience, the best way to avoid (or, if necessary, resolve) these issues is to refer back to the underlying economic theory that defines the baseline as the condition that would prevail “but for” an action and to let fidelity to that concept guide the cooperative effort.

---

marginal value of composite service remains constant. If the composite service loss at the injured site is offset by an increase in composite service production elsewhere in the region, and/or if the affected site is small in the context of the region, the constant baseline assumption would not be violated. As previously noted, if the marginal value of services is not constant through time, the discount rate can be adjusted to accommodate this.

50 In this example, the nourished beach would again be 100 feet wide 20 years after nourishment. In that year, credit is correctly based on the difference between the beach width with the nourishment project (100 feet) and the beach width without the nourishment project (0 feet). Moving out to 30 years post-project, the nourished beach is 50 feet wide. Having run into the solid structure that backs the beach, the baseline beach width remains at 0 feet. Thus, 30 years post-project, credit is based on the extra 50 feet of beach width that exists relative to baseline.

5.1 Typically Available Information

The incident-specific information typically available to help HEA practitioners answer these questions as they relate to habitats that are generally two dimensional[51] is summarized in Table 5-1.[52] Some of these information sources (e.g., SCAT, habitat characterization, and sheen maps) are collected as part of the spill response[53], while other data are only likely to be available if NRDA practitioners collect them. As noted in Table 5-1, some of the data used in HEA characterize conditions that are ephemeral in nature, meaning they are transitory and do not persist for very long in time. Consequently, these types of data can only be obtained in the hours or days immediately following a spill; otherwise, the opportunity is lost. Additionally, data may be collected as a one-time event or may be generated by ongoing sampling efforts. Decisions regarding sampling frequency are usually based on the rate at which conditions are expected to change through time. Lastly, information obtained from baseline or background sampling, which is not addressed in this document, can also inform HEA.

---

51 Until recently, HEAs under OPA had been limited to habitats that can generally be thought of as two-dimensional; e.g., sediments or shorelines. For the Refugio oil spill assessment, the trustees used HEA to estimate compensatory restoration for impacts to organisms that live in the water column [see Appendix H of CDFW et al. (2021)], which is best conceptualized as a three-dimensional habitat. Their logical construct relied on dividing the water column into depth profiles, estimating service loss of each depth, and then re-inserting the depth profiles into a "single water-column-wide" injury. While HEA applied to the water column is an interesting concept, the restoration requirement identified in the Refugio water column HEA deviates by more than an order of magnitude from the restoration requirement that is identified using more traditional REA methods. In light of this "first of its kind result," NRDA practitioners should be wary of ad hoc HEA modifications designed to address three-dimensional habitats such as the water column.

52 Recreational services provided by habitats are addressed outside HEA [see Section 2.1 and MacNair et al. (2022)].

53 Note that the detail and precision needed for spill response is often lower than is ideal for NRDA purposes. In many cases, dedicated NRDA documentation and augmentation of response field efforts can reduce uncertainty (and liability) in NRDA.

Table 5-1—Data Commonly Available for Use in HEA

Data Type	Typical Timing of Data Collection				Data Uses in HEA		
	Ephemeral	Non-ephemeral	One-time	Ongoing	Document Exposure	Quantify Injury	Estimate Recovery
Shoreline oiling information—location and description (aerial surveys and/or SCAT data)	X			X	X	X	X
Response effort documentation		X		X	X	X	
Shoreline habitat characterization		X	X			X	X
Visual observations of sheening with sediment disturbance	X			X	X	X	X
Visual observations of buried oil	X			X	X	X	X
Visual observations of mortality to benthic organisms	X			X	X	X	
Neat oil composition		X	X		X		
Polyaromatic hydrocarbon (PAH) concentrations in sediments—intertidal and subtidal		X		X	X	X	X
PAH concentrations in water column	X			X	X	X	X
PAH concentrations in soils (where applicable)		X		X	X	X	X
PAH concentrations in biota—typically shellfish or fish					X	X	X
Sediment toxicity—ideally with paired PAH data						X	X
Benthic community information (occasionally)						X	X
Data on vegetation (percent cover, stem height, discoloration)						X	X
Information on areas of primary restoration conducted during response						X	X

5.2 Estimating HEA Injury

When HEA is conceptualized as a composite service flowing from two discrete sites, the debit side of the HEA model can be thought of as addressing three questions.

1) How much habitat was exposed and to what degree (footprint)?
2) To what degree was the level of composite service flowing from the footprint reduced immediately following the spill (magnitude of initial service loss)?[54]
3) How will the post-spill composite service level change through time relative to baseline (recovery)?

5.2.1 Footprint

It is common for multiple habitat types to be exposed to oil and for the level of exposure to vary from completely covered with pooled liquid oil to small, scattered tarballs.

In practice, one cannot assess all habitats and all levels of exposure separately. As such, similar habitat types are typically grouped and exposure is aggregated into a few levels. Figure 5-5 and Figure 5-6 illustrate a grouping scheme typical of many OPA NRDAs.

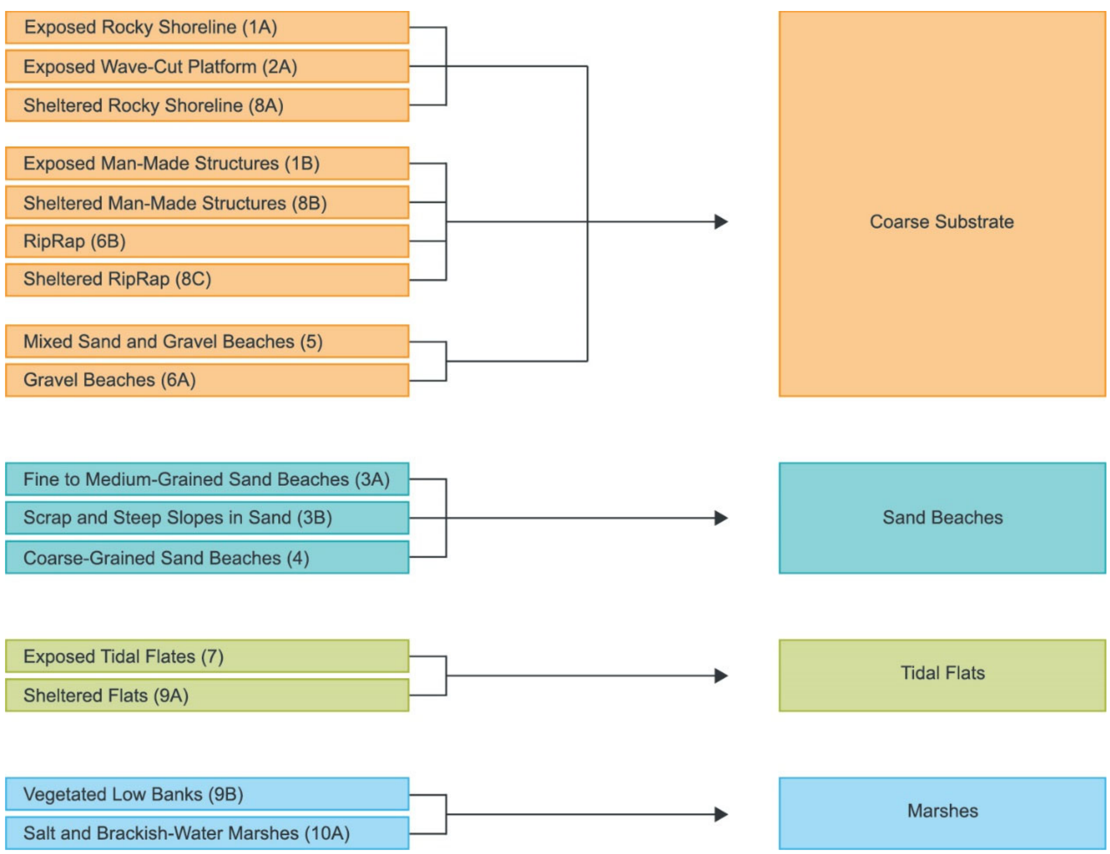


Figure 5-5—Example Grouping of Shoreline Habitat Types for NRDA Use

[54] We use the term "initial service loss" as implying service loss immediately after the spill. At contaminated sites, the initial loss dates back in time before present day, and the term "current service loss" is invoked, recognizing the need to "backcast" service losses to the time of initial onset of injury or December 1980, whichever is later.

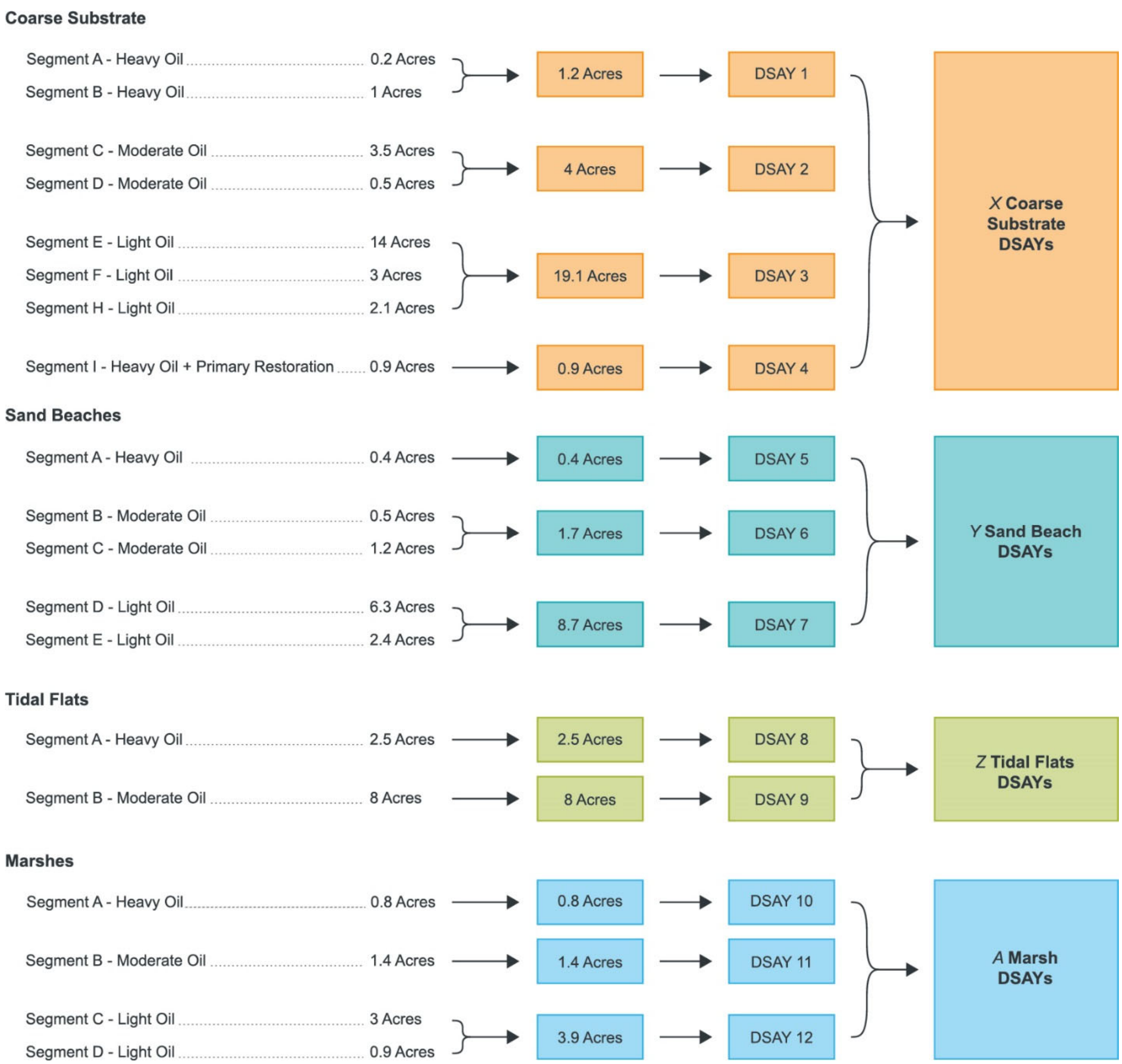


Figure 5-6—Schematic of a Typical OPA NRDA Conducted via HEA

5.2.1.1 Habitats

The process of grouping areas into habitats for assessment should be based on the following: If the array of services provided is the same and the individual services move proportionally to one another such that all that differs is the degree of composite service loss, geographic areas can be combined. For example, two beach areas (coarse and fine-grain sand) might reasonably be combined as long as there is no reason to believe the underlying services provided by the areas are differentially affected by oil.

As one increasingly departs from this construct, there is increasing need for weighting systems (formal or informal) to account for differential services and/or degrees of effect. As an example, in a HEA for benthic habitats, sandy bottom may provide fewer services per acre than rocky bottom to which algal communities attach. In this case, an RHV may be applied that reflects the notion that sandy areas are "worth" 10 percent as much as the rocky areas based on the relative productivity of the habitats. This RHV is then used to combine the areas into a single subtidal habitat.

— In theory, the RHVs will be based on public WTP; in practice, they are typically based on measures of productivity, diversity, or best professional judgment. As the bases for the RHVs become more tenuous (typically because more disparate areas are combined), it will become more reasonable to establish two habitats and conduct two HEAs.

— The approach of establishing new habitats when RHVs become overly tenuous only "works" if habitat-specific restoration is intended. If a single habitat-restoration project will be used to compensate for injuries to two or more habitats, as is often the case, an RHV will be required at some point. Under this circumstance, the decision to "lump or split areas" is largely a matter of convenience.

5.2.1.2 Exposure Levels

Records documenting visible oiling (e.g., slick maps, sheen maps, SCAT documents describing oiling on shorelines and/or vegetation, subsurface/buried bands of oil, and sheening of disturbed sediments) are typically used to establish exposure levels. Supplemental data used to fill in temporal or spatial gaps often include cleanup records, buried oil surveys, contaminant concentrations in environmental media (water, soil, sediments, biota), and chemical fingerprinting of the spilled oil.[55] In a recent marsh assessment, drone imagery collected several months after the spill was used to characterize marsh exposure by looking at the color and condition of the marsh vegetation in the images.

Degrees of shoreline habitat exposure are often categorized as very light, light, moderate, and heavy based largely on the degree of shoreline oiling as described on SCAT data sheets. Additional exposure categories may be created to address intertidal areas that were exposed to the oil before it stranded or areas that were only exposed to sheen or scattered tarballs.

Issues that often arise when assigning exposure levels during an assessment include:

— Data gaps: These may include shorelines that were never assessed and/or incomplete data on assessed shorelines (e.g., data sheets lacking specific information on location, oiling width, percent oil, oil description).

— Temporal changes in shoreline exposure: Integrating data collected on multiple days for the same shoreline.

— Variation within an impact category: For example, all marshes might have light oiling on vegetation, but only a subset may have sediments that sheen when disturbed. The practitioner would need to decide if this is best represented by two different levels of exposure.

— Inadequacy of data gathering methods: A common example relates to the potential for oil to sink. Submerged oil mapping is usually done for response purposes where the focus is on determining if submerged oil is recoverable. The data required to make this response determination may be insufficient for NRDA purposes.[56]

Once habitat grouping and exposure level issues have been resolved, calculating the number of exposed acres in each impact category can be accomplished using geographic information system or other mapping software.

---

[55] Comparison of chemical signatures (fingerprints) of the environmental (water, sediment, biota) samples to the neat oil sample can help determine if contamination in an area is spill-related.

[56] This data gap is often filled using conservative (tending to overstate injury) assumptions. A follow-up NRDA survey or modification of protocols to better address NRDA needs should be considered.

5.2.2 Magnitude of Initial Service Loss

Impairments to services arise through several possible routes: (1) physical fouling of habitats and/or organisms by oil, (2) toxicological effects on organisms from exposure to dissolved concentrations of oil, and (3) the physical and biological effects of response activities. The goal of estimating initial service loss is to integrate the immediate effects of these stressors on multiple services and then aggregate that into an estimate of composite service loss.[57]

The first two concerns that arise associated with this task are often (1) which individual services to consider and (2) how service indicators can be used to understand release-related changes in the provision of those individual services.

5.2.2.1 What Individual Services to Consider

Within the HEA construct, injuries are quantified as a change in the level of composite service provided by a habitat. The composite service is itself a function of many underlying individual ecosystem services. A reasonable first task when estimating a composite service loss, and one often overlooked in practice, is to specify the individual services to be considered in the HEA.

The question that often arises is whether to include all services provided by the habitat or only those that are impaired. To illustrate this issue, consider a sandy shoreline HEA where there are only two observations: (1) Inspections reveal that beach-dwelling invertebrates experienced nearly 100 percent mortality along heavily oiled shorelines, and (2) shorebird monitoring data reveal that area shorebirds experienced no adverse effects. Further, assume that HEA practitioners working in a cooperative setting agreed to assume that those individual services for which invertebrate biomass is the service indicator should be assigned a 100 percent initial loss, while those individual services for which shorebirds are an indicator should be assigned zero loss.

If each of the two service classes is given equal weight, there could be either a 50 percent loss (counting all services) or a 100 percent loss (counting only the impaired services). Which is correct? The answer depends on the proposed restoration action.

To understand the linkage between the two issues, suppose restoration results in an improvement (uplift) among both invertebrates and shorebirds. A HEA practitioner could either (1) adopt the 50 percent service loss and calculate credit based on the change in both invertebrate and shorebirds or (2) use 100 percent loss on the debit side and a 2:1 RHV for restoration, recognizing that each DSAY of restoration credit equates to (i.e., is worth) two DSAYs of debit. These are mathematically equivalent approaches.

Adopting the 100 percent service loss *without* the 2:1 restoration adjustment is an error, as the value of the lost composite service (which does not include shorebird-related services) is less than the value of the restored composite service (which does include shorebird-related services). This mismatch is a violation of a fundamental principle embedded in service-to-service scaling. In addition, judges hearing HEA-like

---

[57] At the outset, it is helpful to distinguish the task of quantifying service loss, with which we are concerned, from injury determination, which is a process step the trustees must complete. Injury determination is essentially a “yes/no” finding and typically goes forward as documentation that an injury occurred. Injury quantification requires estimating the amount of service loss relative to baseline conditions. Thus, the information needed for injury determination can simply be a shoreline oiling level or the observation of sheen in a marsh. In contrast, injury quantification requires information characterizing effects over a range of exposure categories, including baseline and establishing a causal link between the spill and the observed effect.

This distinction is drawn because there is a tendency among some practitioners to make an injury determination and then to leap directly to quantification without establishing the additional linkages necessary. For example, when evaluating the effects of response activities, injury determination may rely strictly on documents characterizing the spatial and temporal distribution of response effort. Then, there may follow an assignment of a service loss to all areas where response occurred, without any stated link between the amount and type of response activity and the assumed reduction in service. While this may be expedient, if it is made without considering if and how it diverges from disciplined practice, it can undermine both its own expediting intent and public confidence in the NRDA process.

arguments have identified such analysis as imparting to the trustees a “windfall” to which they are not entitled. Nonetheless, we note that this error is routinely made in HEAs.

5.2.2.2 How to Use Service Indicators

For the most part, one cannot go to the field and measure the amount of an individual service provided by a habitat. Thus, as a practical matter, one must rely on biophysical parameters that (1) can be measured and (2) are thought to reveal the capacity of a habitat to provide an individual service; we call these service indicators. For example, the food service a sediment habitat provides to fish might reasonably be represented by the biomass of sediment invertebrates per acre.

The selection of service indicators and the process of mapping indicators to individual ecosystem services is a key NRDA step. It is especially important that the practitioner clearly specifies the service at issue and makes sure that (1) the indicator(s) associated with that service are comprehensive, such that no important service is left out, and (2) if there are multiple indicators associated with a single service, they are treated as multiple lines of evidence and composited without double counting. That is, if one service is assigned two service indicators, it is important to make sure that the weight assigned to the service is not inadvertently doubled when changes in the composite service level are calculated as a function of changes in indicators. This “service accounting” task is a place where HEAs often go awry.

The method used to determine how service indicators are affected by an oil spill is also important. There are two basic approaches. One can predict the effects of exposure on the service indicators, or, in most cases, one can measure the effects directly. The former is a modeling approach and is often used when observation is difficult.[58] The latter is an observational approach, usually based on counts and/or measurements of organisms. A hybrid approach is sometimes used.

Finally, one must decide how the indicator actually relates to the provision of a service. For example, the degree to which exposure changes the blood chemistry of an organism can be both modeled and observed; the effect of a change in blood chemistry on the level of service provided by an organism is a separate question. Unless a practitioner can clearly address the second question as well as the first, it may be best to avoid the use of blood chemistry as a service indicator.

The indicators and approaches that are often used to establish injury in a cooperative setting include the following.

— *Visual oiling observations* such as the proportion of sediments, rocks, plant stems, and/or shellfish that are covered in oil are often viewed as providing information about the potential for, and extent of, mortality among organisms living in and on those surfaces. This is because the presence of sticky oil may result in the coating of biota resulting in mortality due to interference with respiration or photosynthesis. While narrow to moderate bands of oiling on vegetation generally do not kill the macro-flora, they may impact the algal communities living on the stems. Those algal communities may, in turn, be an important source of food for algal grazers. Often, exposure-related changes to these communities are evaluated by estimating the density of live organisms or live-to-dead ratios in impacted areas relative to unimpacted areas. In other instances, literature describing the relationship between visual oiling observations and mortality may be relied on.

  Because many of the easily observed organisms tend to be plants (primary producers) or shellfish (primary consumers), changes in the abundance and/or biomass of these organisms is often used to inform estimates of lost food services. Because these organisms provide micro-niches and cover for other organisms, they are also used to indicate changes in the level of habitat services provided.

— *Toxicity study data* with good reference samples have been used to evaluate the potential for toxicity-related injury. Endpoints can include acute and chronic mortality, as well as sub-lethal effects such as

58 For example, it is difficult to observe the effect of dissolved oil on larval fish.

reduced growth and reproduction. These data are often used to inform estimates of service loss because they provide insight into potential reductions in abundance and/or biomass of virtually any organism that can be studied in the lab.

However, toxicity tests generally require relatively large sample sizes and are expensive. To address this issue, toxicity testing is often conducted on a representative subsample. If the study is well designed, correlations can be drawn between measured or modeled polyaromatic hydrocarbon (PAH) concentrations and toxic effects, and applied to the larger data set. It should be understood that toxicity tests are often conducted on sensitive organisms; while a negative result (no effect) can be applied to less-sensitive organisms, a positive result (a change in physiological parameter) cannot be applied to other organisms.[59]

— *Health indicators* (plant discoloration, blood chemistry, lesion counts, DNA adducts) have, in the past decade, increasingly been forwarded by trustees as potentially useful indicator data. The rationale is that if a study identifies a correlation between exposure levels and the frequency of adverse health indicators, the relationship can be used to infer a sub-lethal effect of exposure, which is interpreted as indicative of a reduction in the services for which that species is an indicator. While high natural variability in many indicators can make studies costly and difficult to interpret, the primary challenge associated with this approach is in understanding the link between a change in the metric being measured and a change in the provision of any specific service (i.e., the service production function). For example, a HEA practitioner may establish a correlation between the frequency of DNA adducts and exposure to oil; whether and how that observation relates to a change in the level of an individual service provided by a habitat is an entirely different matter.

— *Community data* (species-specific density and diversity) can theoretically document changes in abundance, biomass, and benthic community structure (an injury as defined by the OPA NRDA regulations). However, high natural variability in many communities makes finding appropriate reference sites difficult, and the variability-driven need for large sample sizes often means the approach is not cost effective. This problem was acknowledged in the Dixon Bay NRDA [Finley et al. (1995)] and limits the use of community data in NRDAs. Moreover, while a reduction in the total density or biomass of organisms may be evidence of a reduction in the provision of the composite service, it is more common to observe a change in community structure. That is, we observe changes in species composition, but not a reduction in community density or biomass. In this case, it is often not clear how a change in community composition relates to a change in service provision.

— *Data on vegetation* are usually collected when vegetated habitats (e.g., marshes, mangroves, forested floodplains) are oiled. These data may include stem density, stem height, live-to-dead ratios, and documentation of physical oiling and plant stress (e.g., yellowing of leaves, early leaf drop). Data to develop diversity measures are also sometimes collected. The standard procedure is to collect these data within a standardized sampling frame (quadrat) typically encompassing an area of 1 square meter. Similar types of data have also been collected for oiled mangroves in larger sample plots.

Release-related changes in the density and/or biomass of vegetation are frequently used as an indicator of a wide range of individual services provided by vegetated habitat. This is because:

— plants are often the base of the food chain in the habitat;

— plant roots and stems often provide the micro-habitats other species rely on and an aesthetic value recreationists enjoy;

---

59 This is an example of the difficulties posed by adopting methods and protocols appropriate for ecological risk assessment (where a potential effect on a sensitive organisms is used to design remedial systems protective of all organisms) to a NRDA (where changes to a composite service should integrate actual impacts to all services).

  - plant structures stabilize soils, trap sediments, attenuate wave energy and storm surge, and facilitate water purification; and

  - plant and associated bacterial biophysical processes often mediate bio-geochemical cycling.

- *Concentrations of contaminants in sediment, soil, and water samples* have been used by NRDA practitioners to inform estimates of potential service reductions. Many chemistry-based approaches have been developed, the most basic of which compare concentrations of individual analytes (in the case of spill, PAHs or total PAH) to screening level benchmarks [Environmental Protection Agency (2003)]. Benchmarks typically identify "safe" concentrations below which toxic effects are not expected. Therefore, if PAH concentrations in exposed areas are below these benchmarks, a service reduction due to toxicity may not be a concern. However, concentrations above benchmarks cannot be used as conclusive evidence of service reductions.

Trustees and trustee contractors have published several methods for estimating the likelihood of injury related to toxicity when screening level thresholds have been exceeded. These methods include calculation of mean sediment quality guideline quotients [Long and McDonald (1998); Long et al. (2006)] and use of logistic regression models (LRM; LRM approach) [Field, Macdonald, Norton, Ingersoll et al. (2002); Field, Macdonald, Norton, Severn et al. (1999)].

5.2.2.3 Translating Injury into Initial Service Loss

Having observed some change in a set of service indicators, it is then necessary to specify a release-related change in composite service provision based on those service indicators. As noted earlier, if all services flowing from an injured habitat were reduced in exactly the same proportions as they are provided by restoration, HEA could be conducted by focusing on a single service tied to a single indicator. That is, one could do a REA on some easily observed indicator species, assign a service weight of 1 to this metric, and know that if service losses related to that injured species were addressed by restoration, all other service reductions would also be appropriately addressed. This is an approach often taken to simplify HEA when implemented in a cooperative setting.

However, such simplifications should be made in light of the three central issues associated with estimating the loss of a composite service:

- The first issue is how to use data such as visible exposure (e.g., oiling of shorelines), visible injury (e.g., dieback of oiled marsh vegetation), or amounts of contamination (e.g., PAH concentrations in sediments) to estimate changes from baseline among individual service indicators. We refer to this as understanding/developing the EPF.

- The second issue is how to predict the effect of many interrelated changes among indicators on the level of each individual service. If indicator measures are direct (e.g., a release-related change in the density and biomass of benthic organisms), this second step may be implicit. However, if the approach is indirect (e.g., a change in the frequency of liver lesions among resident mammals), an explicit service production function is required.

- The third issue relates to aggregation—how does a practitioner convert estimates of release-related changes across many individual services and translate those into a release-related change in the level of composite service provided by a habitat? We refer to this as the aggregation method.

Some of these translations may be straightforward. For example, the level of nutrient cycling service provided by benthos can most likely be reasonably related to the total biomass of benthos. Likewise, several wetland services can likely be assumed to be proportional to primary and/or secondary productivity. In contrast, some services bear a more complex relation to their indicator metrics. For example, with benthic organisms, it is common to observe a change in community structure (i.e., changes in relative abundance of species), but no reduction in density or biomass. Assumptions, for example, related to the level of food service provided to fish depend on the feeding preferences of prey species and whether they are species-specific or opportunistic feeders. Lastly, if the primary indicators are far-removed measures such as changes in blood chemistry or body burdens, estimates of release-related changes in composite service will be tenuous at best.

The logic of the preceding paragraph leads us to make one nearly universal recommendation: If data are being collected to support an eventual HEA, it is helpful to identify indicators as closely related to the service of interest as possible. For example, if a practitioner is interested in the effect of a spill on the quantity of nesting services provided by a forest to birds, it would be better to study the number of fledglings produced per unit of habitat than to simply count the number of birds observed in the area.

5.2.2.3.1 Initial Loss of Individual Services

There are many data sets and measurements that have been used by NRDA practitioners to measure spill-related reductions in a particular service. Below, we discuss three of the more commonly encountered approaches under OPA.

Practitioners translate some types of data, such as vegetative health, more or less directly to service losses. For example, in vegetated habitats, metrics such as percent cover, stem density, and stem height serve as proxies for primary production. Primary production is then used as an indicator for services that are dependent on the vegetation, such as shoreline stabilization, nutrient cycling, carbon export, and structural support for grazers. Therefore, if the metrics of productivity decrease by 20 percent relative to the baseline condition, a commensurate decrease in services tied to vegetation might be assumed. A recent example of such an approach is the Refugio Bay oil spill trustees' assessment of subtidal impacts, which was based on discoloration observed in seagrass [CDFG et al. (2020)].

Practitioners also often rely on toxicity tests to provide similar information on secondary productivity. If a toxicity test suggests acute mortality of 30 percent relative to baseline, that result may indicate that services associated with secondary producers, such as food for higher trophic levels and nutrient cycling, are lost to a commensurate degree. However, additional information such as field observations of biota or benthic community information might influence final service loss determinations. For example, if toxicity tests are conducted on species that are relatively sensitive to contaminants, a 30 percent service loss across an entire suite of species is very likely to be an overestimate. USFWS et al. (2015) outlined this approach as it relates to the assessment of instream habitats following the 2010 Enbridge Line 6B oil spill.

Finally, some practitioners have attempted to translate PAH concentrations in sediments to service losses using the results of the LRM method. As discussed in the previous section, the LRM method estimates the probability that a sample is acutely toxic. The LRM is a screening level assessment and is not a substitute for a direct effects assessment (toxicity testing). Some practitioners have applied a simplifying assumption that probability of toxicity equals percentage service losses. That is, a sample with a 20 percent probability of toxicity is assumed to represent a 20 percent service loss and a sample with a 60 percent probability of toxicity is assumed to represent a 60 percent service loss. Other non-linear transformations of this relationship might be used. While this approach for translating exposure and potential injury into service losses may be cost effective for use in some NRDAs, it is not technically sound without additional data to infer changes in ecologically relevant endpoints.

5.2.2.3.2 Integrating Service Losses into a Loss of the Composite Service

When the proportionality assumption does not hold, the release-related change to each individual service level must be weighted and combined to determine the change in the composite habitat service level. In practice, weights are applied to service metrics, not services, per se. Weighting can be implicit or explicit.

Implicit weighting occurs when the values associated with different services are not stated, yet effects on multiple services (or their metrics) are somehow integrated into an overall percent service loss. This implicit approach lacks transparency and often results in highly divergent service loss estimates. Trustees seem to place high implicit weight on the services that are most affected, while RPs place significant weight on the services that are not.

It is generally recommended that the practitioners conducting the HEA discuss the different services provided by a habitat and the fact that each service is receiving a relative weight even if those weights are not explicitly written down. In our experience, this listing of services and acknowledgement of the existence of weights is often sufficient to generate more consistent estimates of composite service reductions. For example, physical services of a habitat (e.g., erosion protection, flood control, perching sites for birds) are often not impacted by an oil spill. However, when services and weights are not discussed, trustees may assert 100 percent service loss to an intact and functioning, albeit somewhat impaired, ecosystem. This occurs less often when service loss determinations are preceded by a discussion of the full range of services injured and restored, and their relative contribution to overall habitat service provision.

If the trustees and RP cannot agree to a level of initial composite service loss, it may be appropriate and useful to quantify each individual service flow, explicitly determine the weight each service will receive, and compute a weighted average service loss. These weights should be based, to the best of the NRDA practitioners' ability, on the relative value of the service.

5.2.3 Recovery

First, we note that recovery to baseline does not necessarily mean that an injured site has returned to the exact biophysical conditions that prevailed prior to the spill. Recovery to baseline means that the site provides the same level of the composite service post-spill as it would have but for the spill (i.e., at baseline). This is critical to understand, as some HEA practitioners have asserted that recovery to baseline does not occur until the impacted habitat is returned to the state that would have existed but for the spill. This assertion, which would result in extended recovery periods, has no basis in either economic or ecological theory.

Recovery of the composite service level back to baseline should be estimated using the same individual service weights and considerations that went into the estimation of the initial impact. Among other considerations:

— Some types of service-oriented metric data, such as vegetative health, percent cover, stem density, and/or stem height, are often translated directly into measures of services. These same data, if collected over time, can be used to quantify service recovery. Either site-specific monitoring can be used to measure changes in the value of these parameters, or one can rely on literature describing how metrics recover following a spill or analogous perturbation.

— Changes in the density of primary consumers may also be directly linked to services. Recovery can be estimated by monitoring communities and comparing them to a control site.

If monitoring is relied upon, the frequency and length of monitoring is largely dependent upon the service being considered and may be linked to the physiology of resident species; seasonal patterns such as regrowth of vegetation in the spring; or reproductive and other demographic parameters. For example, depuration of PAHs from shellfish tissue is generally rapid after gross oil is removed and, to the extent PAH tissue concentrations provide insight into the recovery of a composite service, monitoring of tissue level concentrations over the weeks and months following the spill may be advisable. Other resources such as marsh vegetation might reasonably be monitored over months or years.

5.3 Estimating HEA Credit

HEA credits are typically estimated as DSAYs generated per acre of restoration. The ultimate size of the restoration project is then calculated by dividing the total DSAYs of debit by the DSAYs produced by an acre of restoration.[60]

The credit analysis requires an understanding of the EPF for production of service indicators. While we formulate HEA as if there is a single fixed design and the only issue is scaling, the selection of a design is important, as it affects both credits and costs. Having selected a project design, the credit side of the HEA model incorporates the answers to two questions:

1) To what degree is the level of composite service flowing from the footprint increased or decreased during and immediately following restoration (magnitude of initial service gain)?[61]
2) How will the post-restoration service level change through time (uplift)?

There are multiple restoration types: habitat creation, habitat rehabilitation, and habitat preservation; the information required to answer the three “credit questions” varies depending on the type of restoration being considered.

5.3.1 Habitat Creation or Rehabilitation

Habitat creation fundamentally changes an area from one habitat type to another. A common habitat creation project is the conversion of shallow open water to marsh using dredge spoil. Riparian areas and upland and lowland forests are also frequently created through planting of deforested areas. When oyster reefs, corals, and/or seagrass beds are impacted, new reefs/corals/beds are also often created through seeding and/or transplanting into subtidal areas. In slight contrast, habitat rehabilitation improves the functioning of an existing habitat; a classic example is removal of an invasive plant species from a sand dune habitat.

Critically, the credit associated with a created or rehabilitated habitat must reflect the *net* increase in ecological services. That is, if a corn field is to be converted to forested riparian habitat, the credit associated with that restoration action should reflect the development of the composite riparian habitat service *minus* the ecological services that would have been provided by the corn field had it not been converted. Unless it is viewed that corn fields and forests provide the same suite of services and only differ in the amount provided, this adjustment requires, either implicitly or explicitly, an RHV that allows the composite corn field service to be converted to a composite forested riparian service.

While the considerations associated with each site and habitat vary, the thought process behind HEA credit curves associated with habitat creation and/or rehabilitation are generally consistent. In the *Athos I* Damage Assessment and Restoration Plan *(*DARP) [NOAA et al. (2009)], the trustees described credit assumptions associated with two wetland sites. Their discussion is typical of the considerations, if not the level of detail, embodied in most HEA credit estimates.

In the text below, taken from the draft DARP, the reference site is a natural, fully functioning wetland.[62] One of the sites discussed, Mad Horse Creek, is best thought of as a wetland enhancement site, since

---

[60] The OPA NRD regulations specify that “to the extent practicable, when evaluating compensatory restoration actions, Trustees must consider compensatory restoration actions that provide services of the same type and quality, and of comparable value as those injured” (15 CFR § 990.53). This is generally referred to as “in-kind” restoration. However, should in-kind scaling prove impractical for some reason, “the scaling process will involve valuation of lost and replacement services,” whatever those replacement services may be (15 CFR § 990.53). When the habitat restored is not the same as the habitat injured, the process is referred to as “out-of-kind” restoration (see Chapter 5).

[61] We use the term “initial service loss” as implying service loss immediately after the spill. At contaminated sites, the initial loss dates back in time before the current time, and the term “current service loss” is invoked, recognizing the need to “backcast” service losses to the time of initial onset of injury or December 1980, whichever is later.

prior to project implementation, it was a poorly functioning wetland. The second project, Lardner's Point, is best thought of as a wetland creation site.

> *The Trustees relied on resource agency staff experience with creating wetlands in this region, data from other damage assessment cases, and information in the scientific literature. The Trustees assume that marsh construction for both projects will begin in 2009. Ecological services are expected to develop following a logistic model, reaching maximum service in fifteen years. For Mad Horse Creek, a baseline ecological service of 10 percent is used. This reflects the minimal level of service provided by the current area of Phragmites-dominated, disturbed wetlands. At Lardner's Point, a baseline ecological service of zero is used, reflecting the current state of the property, which is abandoned industrial upland, covered in invasive plants such as knotweed, with a steep riverbank.*
>
> *The maximum service level is estimated to be 85 percent, reflecting Trustee experience that restored marshes generally do not reach productivity levels associated with natural, fully functional marsh habitat. The project life span is estimated to be 50 years. Based on these inputs and using the three percent annual discount rate typically applied in HEA calculations, each restored acre at Mad Horse Creek provides a credit of 13.72 (discounted) service acre-years and each acre at Lardner's Point provides 15.56 (discounted) service acre-years* [NOAA et al. (2009)].

5.3.2 Habitat Acquisition

Habitat acquisition is the purchase and preservation of habitat to prevent its loss. Acquisition can include outright purchase of the property or purchase of a conservation easement. The deed or easement is typically transferred to a land management organization that will ensure habitat preservation. Because ecological services flow from the property in its natural state regardless of the ownership, the net service gain is contingent upon the (regional) baseline level of the composite service in question and the effect that preserving any one parcel will have on the future level of the composite service provided within the region (see Figure 5-1).

Critically, the credit associated with habitat preservation is not based on the service provided by the preserved parcel. From a technical perspective, the credit associated with habitat preservation is based on the net increase in the regional level of the composite service, if any, that occurs as the result of habitat preservation. If preservation of one parcel simply redirects development of one upland site to a different upland site that provides the region with the same level of service, the preservation would provide no credit. In fact, if the preserved habitat provides a lower composite service level than the acreage where the diverted development occurs, the regional composite service level would decrease as a result of the preservation. However, from a practical perspective, the redirection of development pressure is often ignored when HEA credits are estimated in a cooperative setting.

If the potential redirection of development is ignored, the analysis can take a very convenient form. The value of acquisition and preservation is calculated as:

$V = \sum_{t=1}^{T} \lambda * (1-\lambda)^{t-1} * L_t * 1/(1+r)^t$,

where λ is the development hazard rate that specifies the probability the parcel is developed at date t (given that it has not been developed before t), L_t is the habitat services lost if development occurs at date t, and r is the discount rate. L_t is calculated as:

---

[62] As discussed in Section 3.4.3, it appears as though the *Athos I* Trustees intended to use a natural, fully functioning wetland acre as the base site but in actuality used a severely impaired wetland acre as the base site.

$$L_t = \sum_{s=t}^{T} \frac{L}{(1+r)^s} = L * \left[[1-(1+r)^{-(T-t)}]/r\right],$$

where L is DSAYs of loss from development.[63]

5.4 Common HEA Pitfalls

This section summarizes several pitfalls commonly encountered during the HEA process and outlines strategies to help avoid them.

5.4.1 Incomplete Documentation Results in Overestimate of Footprint

Response data describing the extent and magnitude of oiling and search and collection of oiled wildlife are directly relevant to NRDA. In particular, data collected by SCAT are very useful for the NRDA *if* the data are well documented. SCAT data are usually the main source of oiling exposure information, and occasionally the only source. However, a common problem when implementing HEA is that SCAT data provide incomplete (or not representative) information regarding the extent and degree of oiling. Data gathered during the spill response are imperfect for use in the NRDA, for understandable reasons, since they are gathered to inform response, not estimate service loss. However, data deficiencies create uncertainty in the NRDA, which leads to conservative assumptions.

To collect the best possible data during the spill response and later in the NRDA phase, it is recommended that NRDA practitioners are on-site, coordinating closely with the spill responders. This can ensure that all responders tasked with shoreline surveys understand the information needed, that data sheets are filled out completely, and that data sheets are compiled and archived. In addition, coordination can often focus data collection on the most useful data and/or avoid collection of unnecessary or confounding data. NRDA teams should collect data in reference areas and in a range of representative conditions, rather than worst-case conditions only.

For example, following a spill that affected a highly dissected salt marsh, the RP used two airboats with Global Positioning System (GPS) units to document the full extent of the spill along all marsh edges in two weeks. In this case, it was documented that oiling extended only a few feet into the marsh. As such, 200 miles of shoreline oiling was associated with less than 100 acres of oiling.

5.4.2 Failure to Consider Uninjured Services

When weights are developed implicitly (and occasionally when they are developed explicitly), NRDA practitioners tend to place seemingly disproportionate emphasis on the services affected by a spill. The best way to avoid this pitfall is to be sure that all participants are cognizant of the weights they must assign (either implicitly or explicitly) and the fact that those weights should reflect public preferences, not personal bias.

This pitfall is clearly illustrated in the *Athos I* DARP [NOAA et al. (2009)], wherein the trustees assert that 100 percent of baseline services were initially lost from areas impacted by heavy or moderate shoreline oiling. This assertion cannot be supported when viewed in light of the shoreline injury report produced by the *Athos I* trustees, which lists baseline services provided by shoreline habitats. The list of services and an integration of trustee assessments of impacts on those services from documents in the administrative record for the spill are summarized in Table 5-2.

A 100 percent service loss in light of the observations summarized in Table 5-2 implies that the trustees placed zero weight on services such as primary production, runoff reduction, and flood control, which

---

[63] If the hazard rate is changing at a constant rate, this can easily be accommodated as a change in the discount rate. More complex adjustments can be included with additional simple spreadsheet programming.

were reported as having not been affected by the spill. It is difficult to defend the assertion that the public derives no value from these services.

While estimating the true liability overestimate associated with this error would require a system of weights for individual wetland services, the error is not inconsequential. For example, if all services are given equal weight, and noting that approximately 25 percent of the services provided by the shoreline were not impacted, estimates of shoreline-based compensatory restoration in the *Athos I* spill would be reduced by as much as $3 million, all else being equal.

Table 5-2—Summary of Services and Impacts Described by Trustees

Service	Function Described in DARP	Trustee Evaluation of Impacts
Primary production	Production of plant material that forms the base of the primary food web and the detrital food web. Much of salt marsh vascular plant production is exported to adjacent habitats as detritus.	The shoreline injury report states, "the spill occurred when the marshes were in senescence (not growing) and it was not possible to discern any significant impacts to marsh vegetation" when it began growing in spring 2005.
Habitat for biota	Marshes serve as physical habitat for organisms including birds, mammals, insects, fish, and invertebrates. The type and density of the vegetation is the primary determinant of species use.	See Primary Production discussion, which indicates no apparent impact.
Food web support	This encompasses the entire system, including invertebrates that are food for higher trophic levels that may spend limited time in the wetland.	The trustees assert that oil would have smothered most organisms within the oiled band.
Fish and shellfish production	Marsh edge and ponds are nursery areas for fish and shellfish. Dense shellfish provide microhabitat for a diverse assemblage of organisms that contribute to productivity and species composition.	In the final pre-assessment data report, measurements of tissue burdens are at levels below thresholds of concern.
Sediment shoreline stabilization	Marsh vegetation stabilizes the soil and prevents erosion during normal tides, wave action, or storm events.	See primary production discussion, which indicates no apparent impact.
Water filtration	This is the physical removal of particles and nutrients from water.	See primary production discussion, which indicates no apparent impact.
Nutrient removal transformation	Nutrients are converted to plant material, thereby reducing the occurrence of algal blooms and anoxic conditions in the bay.	See primary production discussion, which indicates no apparent impact.
Sediment /toxicant retention	Sediments and the toxicants bound to them are filtered in wetlands rather than being transported to the bay. Wetlands encourage redox reactions that can detoxify many compounds.	See primary production discussion, which indicates no apparent impact.
Soil development and biogeochemical cycling	The soil is a living system that converts chemicals from one form to another and supports the growth of higher plants through biogeochemical cycling and the breakdown of detritus.	In Appendix H of the shoreline injury report, a trustee contractor reports that little substrate penetration occurred. Also, see primary production discussion.
Storm surge protection	Wetland habitat is a buffer between the bay and other habitats. Vegetation absorbs wave energy and reduces impacts to inland habitat and property.	See primary production discussion for evidence that vegetation was unaffected.
Slow runoff from upland	Marsh surface absorbs runoff from upland; vegetation also slows flow, allowing more runoff to be absorbed .	See primary production discussion for evidence that vegetation was unaffected.

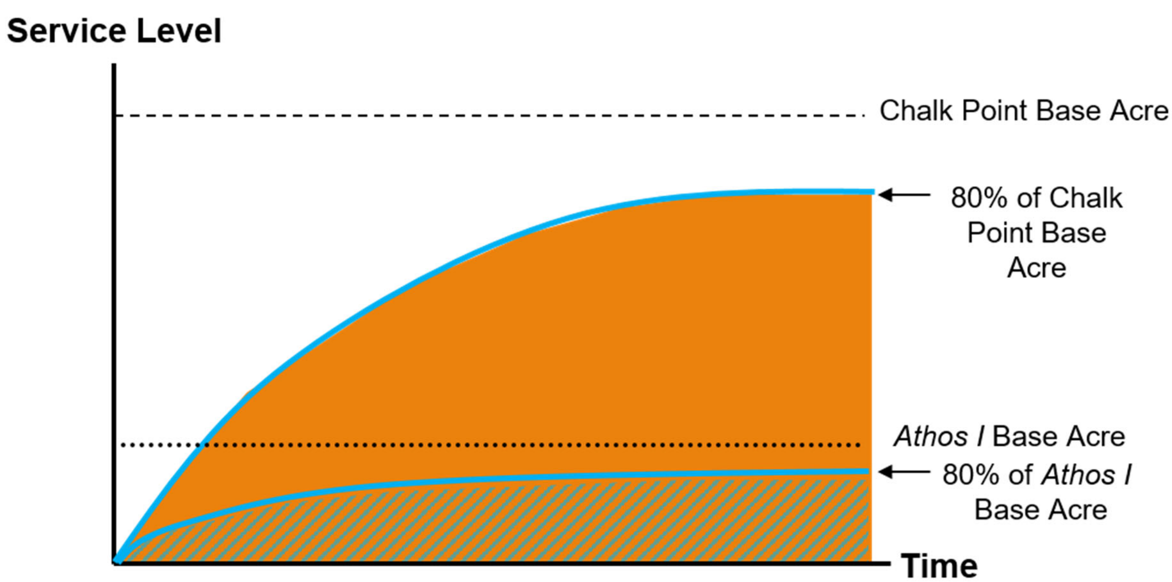
5.4.3 Inconsistent Treatment of the Base Acre

In HEA, variation in service levels is expressed as a "percent change" relative to some base acre. While it does not matter what the base acre is, HEA requires that both the injured site and the restoration site be judged relative to the same base acre.

This concept is commonly violated. For example, in the Chalk Point assessment, the base acre was a high-quality marsh with few background PAHs, few impacts from urbanization, and limited invasive vegetation. Using that high-quality marsh as a base, the trustees asserted that a created marsh site in a relatively pristine setting should eventually provide services equivalent to 80 percent of the base [NOAA et al. (2002)]. This is a reasonable assumption, and there is a degree of scientific consensus that a well-constructed created marsh in a relatively pristine area will provide somewhat less composite service than a naturally occurring, pristine marsh.

However, as often occurs in OPA NRDA assessments, the trustees in a subsequent case borrowed the 80 percent assumption from the Chalk Point assessment without making any adjustment for the facts surrounding the assessment. Specifically, the *Athos I* trustees selected a base acre from a degraded marsh located in an industrialized portion of the Delaware River adjacent to Philadelphia [NOAA et al. (2009)]. The base acre includes vast stands of an invasive plant (*Phragmites*) and is subject to urban runoff and combined sewer overflows. Noting that restoration projects related to the *Athos I* spill were created marsh sites in relatively pristine settings, it is inconsistent to assert that the *Athos I* habitat restoration would provide only 80 percent of the services provided by the base site because, in the *Athos I* case, the restoration should have been judged relative to a degraded (industrialized) base acre. As depicted in Figure 5-7, the proper credit calculation should have reflected a restoration site that, after a few years, provided a considerably greater level of the composite services than was provided by the *Athos I* base acre.

If the *Athos I* trustees built into their analysis the use of a degraded base acre when they transferred assumptions from the Chalk Point assessment, the asserted shoreline liability would likely be reduced by approximately $1 million, all else being equal.



The Chalk Point project increases services significantly (represented by the colored area). The similar Athos I *project is assumed to create fewer services (represented by hashed area) because the baseline does not reflect the correct/appropriate base acre.*

Figure 5-7—Similar Projects Generate Different Service Increments When Base Acres are Mismatched

5.4.4 Inconsistent Assumptions between Debit and Credit Calculations

HEA assessments sometimes embody assumptions that are inconsistent with one another. For example, it does not appear as though the trustees were consistent in their treatment of marshes in the *Athos I* spill [NOAA et al. (2009)]. In their response to comments on the *Athos I* DARP, the trustees state that *Phragmites*-dominated marshes provide a service level similar to that of wild rice and *Spartina* marshes. This assumption is employed when determining injuries to oiled *Phragmites* marshes. However, when justifying restoration projects, the trustees state that a degraded *Phragmites* marsh provides 10 percent of the services of a healthy *Spartina* marsh. If the trustees believed *Phragmites* marshes provide service levels similar to *Spartina* or wild rice marshes, compensatory restoration projects designed to convert *Phragmites* marshes to *Spartina* marshes provide little to no increase in services and should be rejected as compensatory restoration projects. Alternatively, if the trustees believed *Spartina* marshes provide 10 times greater services than do *Phragmites* marshes, the estimated injury associated with impacted *Phragmites* marshes should have been adjusted to incorporate the relatively low level of wetland service that was being produced under baseline conditions.

5.5 Sensitivities

HEA sensitivities, both on the debit side and credit side, are reviewed in this section.

5.5.1 Debit Side

On the debit side of the HEA model, there are three important variables: footprint (of impacted area), magnitude of initial service loss, and recovery. Any of the model parameters associated with these variables may be influential for NRDs, and this influence depends on the magnitude of the other variables.

5.5.1.1 Footprint

Regarding footprint, additional acreage increases the HEA debit and NRDs proportionally. The footprint in many spills is readily agreed to; acreages of different types of habitat are measured accurately, and RPs and trustees are working with similar data. This is not so straightforward when there are significant amounts of background oil, as is often the case in central California, Louisiana, and Texas, where natural seeps are common. In these cases, the footprint of the oil released will depend on the fingerprint of spilled versus naturally released oil. This was a significant issue in the Deepwater Horizon spill in the Gulf of Mexico and Refugio Bay oil spill in the Santa Barbara Channel, both areas with prolific natural seeps.

A second set of complications often arises when oiled habitats are dynamic and so oiling may be "missed." This can happen on exposed high-energy outer coast shorelines, especially sandy shorelines, where oil can be buried and/or "lifted" during subsequent high tides, leaving little evidence of its presence. In these circumstances, fate and transport modeling may be used to refine the footprint or, increasingly, it may be prudent to use drones or aircraft to rapidly and repeatedly document visible oiling.

5.5.1.2 Initial Injury

Initial injury estimates tend to be most sensitive to aggregation method (i.e., the method used to combine all of the individual services into a change in the level of composite service provision). When one or two services are heavily weighted and others are assigned weights of near zero, the choice of which individual services to assign non-zero weights to can and will be the primary factor determining damages.

In addition, as discussed earlier, response data and data collected by NRDA teams immediately after the spill may be focused in more heavily affected areas. In the absence of a more comprehensive random sampling design, there may arise considerable uncertainty about how to extrapolate measurements from more-exposed areas to those less exposed. This is an area where early focus on sampling for NRDA purposes rather than only relying on response efforts may be highly beneficial in reducing uncertainty.

Finally, in our experience, when laboratory bioassays are used as the primary means of informing service loss, HEA practitioners often make highly conservative assumptions regarding effects. For example, a bioassay may indicate that 30 percent of some sensitive organism and life stage die when exposed to a specific dissolved oil concentration for 48 hours. HEA practitioners have taken this observation and assigned a 30 percent initial loss of composite service because a single measurement from among many exceeded the specified level. This approach does not consider exposure duration, nor does it consider the wide range of sensitivities across species and life stages.

5.5.1.3 Recovery

Recovery depends critically on two factors. The first is the weight given to service indicators that may recover either very quickly or very slowly. Modest changes in weights for those with intermediate recovery times tend not to be material.

The second critical recovery factor is the shape of the recovery curve, as opposed to the duration of the recovery period. While it may take a long time for some service indicators to recover, there may also be significant early recovery followed by an extended period of slower recovery. Anchoring on the time to ultimate recovery and imposing a linear recovery curve will miss the potential early rapid-recovery period. This will lead to an overestimate of the debit.

5.5.1.4 Interplay between Debit Elements

The importance of either initial injury or recovery largely depends upon the magnitude of the other. For example, in Figure 5-8, the initial injury is relatively small. If the time to recovery is short, the debit is represented by the dark orange triangle. If the time to recovery is long, the debit increases to include both the dark orange and light orange triangles. However, the size of the light orange triangle (the increased debit associated with an extended recovery period) is not particularly large. In contrast, Figure 5-9 illustrates an initial injury that is relatively large. Again, if a long recovery is assumed rather than a short recovery, the total debit increases by the size of the light orange triangle. However, in Figure 5-9, the size of the light orange triangle is rather large.

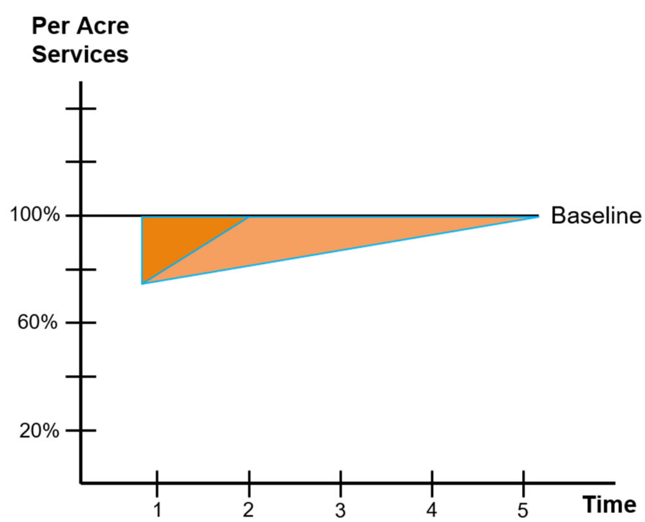


Figure 5-8—Increasing Recovery Time Results in a Modest Increase in Debit When the Initial Service Loss is Small

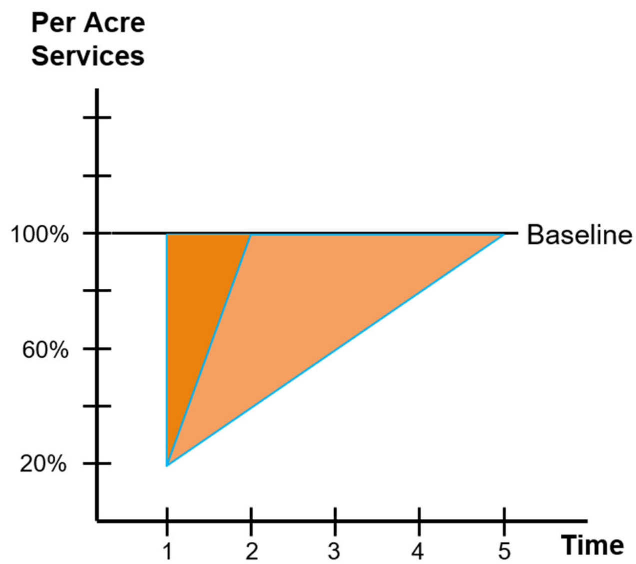


Figure 5-9—Increasing Recovery Time Results in a Large Increase in Debit When the Initial Service Loss is Large

In addition, the shape of the recovery curve also influences the relative importance of the initial injury estimate and the estimated time to recovery. As illustrated in Figure 5-10, if most services recover rapidly (curved line), even a large initial injury and an extended recovery period can result in a relatively small debit (represented by the dark orange triangle). However, if recovery is assumed to occur linearly, the same initial injury and time to recovery results in a much larger debit (represented by the dark orange and light orange triangle).

Finally, because discounting decreases the impact of future service losses, the farther into the future the service loss occurs, the less that service loss contributes to the total injury. For example, at a discount rate of 3 percent, the discounted value of a constant injury that extends 35 years into the future is only 20 percent greater than that same injury extending out over only 25 years, even though the time period is 40 percent greater. Thus, the near-term rate of recovery and shape of the recovery curve are much more important determinants of the total debit than is the total duration of recovery.

Estimating the sensitivity of NRDs to various parameters is readily addressed via Monte Carlo analysis. With this method, ranges of possible values for each parameter in the HEA are specified. A “draw” of each parameter from a probability distribution over this range provides one estimate of NRDs associated with the parameter combination in that draw. Doing this many times gives a distribution of NRDs. The data generated can then be used to identify which parameters are most influential to the overall result. This is primarily a tool to understand uncertainty and is discussed more in Section 7.

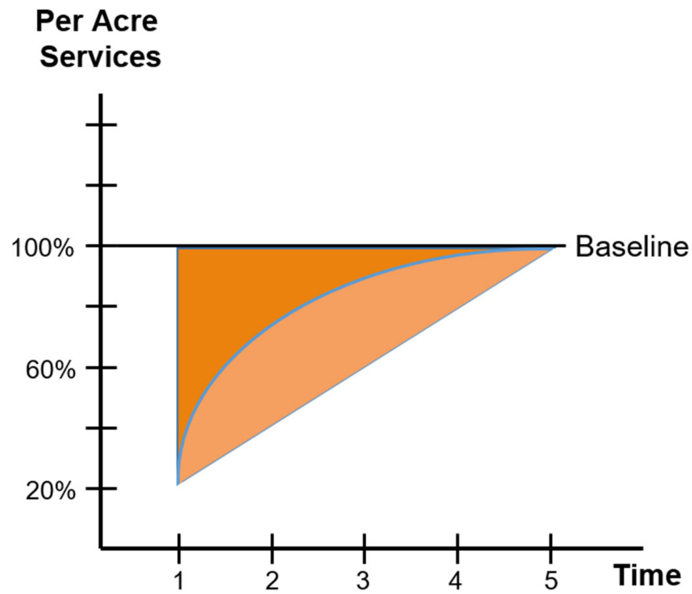


Figure 5-10—The Shape of the Recovery Curve Influences the Importance of Loss and Recovery Assumptions

5.5.2 Credit Side

Similar to the debit side, parameter estimates defining the speed at which service levels increase and the length of time over which they endure can be important; the relative degree to which any single parameter estimate is important depends on the values assigned to the other parameters. However, when it comes to restoration, assumptions regarding service flows are (1) fairly well defined by existing literature, experience, and precedence, and (2) generally not the key issue.

There are two important credit-side biological-physical and economic factors affecting ultimate NRD liability. The first are the weights given to different service indicators. As with recovery from initial injury, greater weights placed on either fast- or slow-developing indicators will shift the development of benefits over time and hence DSAYs of credit. The second is the RHV.

Of course, the per-unit cost of the identified restoration project scales directly to NRDs. The trustees are mandated to consider cost effectiveness when identifying restoration alternatives; it serves society at large if there is a diligent search for cost-effective restoration options. An example is the use of wetland terraces in southwest Louisiana.[64] In this part of the country, wetland terraces provide the same suite of services as a regular wetland creation project at approximately 10 percent of the cost. Creating wetland services via terracing rather than traditional wetland creation would reduce NRD liability by an order of magnitude, all else being equal.

An additional issue that can affect restoration costs is whether the project is a “real live project” or a “generic” project of unknown location and design used to monetize the debit. In the latter case, in determining a claim amount or cash-out settlement, trustees may wish to ensure that the restoration project ultimately selected is sufficiently funded. In this case, trustees often use the upper end of the range of restoration costs for the type of project selected. For example, in a recent case, the costs of a generic wetland creation project forwarded by the trustees included a large cost element for dealing with excavation and disposal of contaminated sediments should such a need arise. Noting that such an

---

[64] A wetland terrace is a sediment ridge constructed at marsh elevation using subtidal bottom sediments excavated from on-site.

occurrence is rare, a more appropriate method for addressing this uncertainty would be to base settlement on expected costs, perhaps coupled with an insurance or contingency clause related to contaminated sediments (see Section 7).

5.6 Cooperative versus Litigious HEA

The primary differences between a cooperative and litigious HEA would likely relate to the assumption of proportional injury and/or the creation of service weights, the technical defensibility of assumptions related to the magnitude of initial service loss and the speed of recovery, and assumptions related to per-unit restoration costs.

- While practitioners may find it cost-effective to turn a blind eye to an assumption that one proxy service accurately represents the composite service level or to the implicit derivation of service weights in a cooperative setting, it is likely that either of those approaches would be open to challenge in a litigious setting.

- In a cooperative setting, it may be cost-effective to accept certain assumptions regarding the magnitude of initial injury and recovery time, provided they do not result in significant increases in the cost of compensation. In a litigious setting, the acceptance of technically uncertain and/or unjustifiable assumptions may not be acceptable.

- In a cooperative setting, practitioners may find it cost-effective to accept poorly documented or undocumented restoration costing. In a litigious setting, it may be prudent to ensure that all cost elements are fully documented and defensible.

As discussed in Section 7, in a cooperative setting, a RP may find it cost-effective to accept an ad hoc treatment of risk and uncertainty. In a litigious setting, it may be prudent to ensure that risk and uncertainty are explicitly addressed in a theoretically rigorous manner.

A second class of defensibility issues related to HEA is the absence of a way to aggregate the preferences of the public when those preferences differ. As discussed in Section 3, there is no way to do this without invoking additional information that is typically not part of HEA. Other simplifying HEA assumptions, such as linearity of services in acres, whether or not service changes are appropriately "small," the distinction between a habitat's *capacity* to produce services and actual service provision, and the requirement that the value of service changes from injury and restoration be identical all would be sources of challenges of whether HEA as a method is reliable in the case at hand. We note that existing uses of HEA in court were for relatively simple cases, involving groundings and single-metric applications that were effectively REAs.[65]

5.7 HEA Summary

HEA addresses all services flowing from a habitat simultaneously by assessing a single representative/proxy or a composite service index. It is appropriately employed if:

- all services from the injured habitat are reduced in exactly the same proportion as they will be enhanced by the restoration project, thus allowing analysis based on any single service indicator, scaled by acres; or

- a system of weights is derived or agreed upon that allows the various injured and restored services to be converted to a single composite service.

[65] United States v. Great Lakes Dredge & Dock Co., 259 F.3d 1300 (11th Cir. 2001), cert. denied, 535 US 955 (2002).

While that system of weights could be based upon the individual preferences of the public (departing from HEA to the world of economics), in practice they tend to reflect the best professional judgment of NRDA practitioners.

Given the absence of a firm scientific basis for specifying service weights outside of economic analysis and the uncertainties inherent in estimating the magnitude of an initial service loss, service recovery, and increase in services via restoration, HEA is best thought of as a structure (i.e., a framework and language) for undertaking negotiations to resolve NRDA cases.

In our experience, the cooperative approach is most successful when data gaps are small or inconsequential. Material data gaps can be difficult to address after the fact, often resulting in either protracted negotiations or the adoption of highly conservative assumptions that result in large damage estimates relative to similar cases that are able to rely on information rather than assumption. That is, the timely collection of field data and a mutual agreement that assumptions embedded in an HEA be internally consistent can often be used to efficiently identify mutually agreeable outcomes. In addition, it can be helpful to have explicit discussions regarding all the services that flow from a habitat and the fact that those various services are weighted (either implicitly or explicitly) as they enter the composite service.

6 REA and HEA: A Comparison of Strengths, Weaknesses, and Appropriate Uses

Sections 3, 4, and 5 outline the basic constructs of REA and HEA and review the economic underpinnings of the models. The key conclusions reached in those sections are as follows.

Both models are economic in nature. By a series of assumptions, the economics “collapses” to allow a focus on the ecological services lost over time due to a spill and then gained over time due to restoration.

Both models are implemented in three steps:

1) The flow of services at baseline is identified.
2) The flow of ecological services given the spill and a restoration project is determined.
3) The size of the restoration project is adjusted until there is no net loss of the discounted present value of ecological services.

REA assesses impacts to a set of services that all move proportionally with a population of a single species (e.g., of bird or fish); HEA integrates the multiple services provided by a habitat and treats them as a single composite service.

Among other constraints, REA requires that:

— the ecological services are defined such that an increase in the service is desirable; and

— the spill and restoration result in “small” service level changes (so that marginal values are constant).

In addition to these constraints, HEA requires either:

— the suite of services flowing from the injured ecosystem to have been injured in exactly the same proportion as will be provided by the restoration project; or

— a system of weights that is used to translate the various injured and restored services to a single unit.

Given that both REA and HEA impose multiple constraints (i.e., they are only strictly applicable under a very limited set of circumstances), why use REA and HEA at all? Simply put, REA and HEA often provide a simple, technical basis for reaching a settlement.

In the context of a cooperative NRDA, REA and HEA can expedite the assessment process by providing a structure for organizing negotiations, facilitating the rapid identification of the nature and extent of agreements that need to be reached, and imparting to trustee agencies, the public, and RPs a measure of comfort regarding the technical defensibility of the cooperatively identified level of restoration. The expediting function is aided by a small but growing body of literature and case studies describing the application of both REA and HEA in the OPA context (see NOAA 2000 and the U.S. Department of Interior [DOI] Natural Resource Damage Assessment and Restoration rule).[66]

In the context of a litigious assessment, the conceptual underpinnings of HEA have been upheld in court cases related to seabed disturbance under the National Marine Sanctuaries Act.[67] While we are unaware of court cases explicitly upholding the use of REA, the general similarity of REA and HEA, the fact that REA applications are generally more consistent with theoretical requirements than are HEA applications, and the publication of scientific papers outlining the use of REA to assess compensatory requirements under OPA suggest its use may also be admissible in court, provided that generally accepted scientific methods for implementing REA are employed.[68]

Given that REA and HEA are increasingly being relied on to conduct OPA NRDA, is there a reason to select one over the other? There are no hard and fast rules that allow a practitioner to determine with precision whether REA, HEA, both, or neither are appropriate tools for estimating compensatory restoration requirements. However, by combining a general understanding of REA and HEA with a general understanding of the facts surrounding a spill, it should be possible to identify an assessment approach that balances the desire to ensure that appropriate compensatory restoration is identified with the need to avoid double-counting and minimize transaction costs.

The following exercise may provide some insight and guidance.

1) Make a list of all impacted resources. Usually, the list will contain:

   a) habitats (e.g., sandy shoreline, rocky shoreline, seagrass beds, coral beds, wetlands, water columns, offshore sediments, intertidal sediments);

   b) specific species (the list could be long and varied, but will usually include several species of birds, and may include sea turtles, fish, and marine mammals); and

   c) things people do (e.g., hunting, fishing, hiking, boating, going to the beach).

2) Impacts to “things people do” tend to be recreational issues, which are addressed separately. Remove them from the list, but do not forget possible overlap with impacts to species.

3) Assign all of the species to an injured habitat. If any species is not easily assigned to a habitat, it is a candidate for REA.

4) Now, focusing on any one habitat, list the services provided by the habitat and the species that reside therein.

---

[66] 43 CFR part 11 RIN 1090-AA97.

[67] United States v. Great Lakes Dredge & Dock Co., 259 F.3d 1300 (11th Cir. 2001), cert. denied, 535 US 955 (2002).

[68] For reasons discussed in Section 1 and Annex A, there remain a number of elements of REA and HEA open for challenge, and in general, litigation would require that the estimates generated by REA and HEA are consistent with full economic methods. If a REA or HEA approach were employed in a litigation setting, the consistency between the actual spill circumstances and the simplifying assumptions embedded in REA and HEA would need to be critically evaluated.

a) Were all services and inhabitants injured (and likely to be restored) approximately in a fixed ratio? If so, address via HEA. If not, proceed to the next question.

b) Is it cost-effective to conduct a service-weighted HEA that treats all services flowing from the habitat as a single composite service?[69] If so, address via HEA. If not, proceed to the next question.

c) Is it cost-effective to create a sufficient quantity of habitat such that even the most severely impacted service will be compensated? If so, use the most severely impacted service as a proxy for all services flowing from the habitat and address all services in an HEA. If not, proceed to Step 5.

5) If specific populations are removed, is it possible to address those populations using REA and then address the remaining services and inhabitants using HEA? If yes, do so. If not, proceed to Step 6.

6) Consider methods other than HEA and REA based on more explicitly economic approaches. The alternative method could include a modified version of REA or full economic methods.

7) Finally, after deciding which services can be addressed by REA or HEA, consider how the different injury assessments and restoration projects interact. Are injured services being omitted or double-counted? Are restoration credits being omitted or double-counted? How can the assessment strategy be adjusted to minimize omissions and double-counting in a cost-effective manner?

It is worth highlighting two ideas embedded in the above exercise.

— There is much to be gained by identifying, as soon as possible, the likely restoration projects associated with each potential REA and HEA; there is little to be gained by assessing injuries in a manner that does not facilitate the scaling of restoration.

— The goal of NRDA should be to compensate the public in a manner that minimizes overall NRDA costs (the cost of compensatory restoration plus the cost of assessment); in the long run, there is no other strategy that will make the public better off. With this goal in sight, it may be cost-effective to simply accept overcompensation. However, those errors can become future trustee expectations. Hence, there appears to be a critical role for RPs and trustees in ensuring that a high degree of technical rigor is maintained in OPA NRDA.

7 Emerging Issues

Section 7 focuses on several emerging issues related to the use of REA and HEA to conduct OPA NRDAs. In each case, we introduce the issue and discuss its implications. If the resolution to an issue is straightforward, but for some reason not currently adopted by the NRDA community (in our experience), we recommend a solution. When the resolution of an issue is less clear, we present what we believe to be a technically justifiable solution. For all issues, we advise the reader that Section 7 reflects Cardno's views and should not be interpreted as providing a consensus view of NRDA practitioners.

7.1 The Habitat-Based Resource Equivalency Method

In a recent paper in the journal *Environmental Management*, NRDA practitioners associated with federal trustee agencies (both NOAA and DOI) forwarded a new approach to implementing HEA [Baker et al. (2020)]. They call their method the Habitat-Based Resource Equivalency Method (HaBREM) and assert that there are several advantages of HaBREM relative to the traditional approach to HEA. Specifically, the

[69] To answer yes to this question, a RP must have some sense that the trustees may be willing to accept relative weight based on best professional judgment or survey. Developing litigation-proof weights would be cost-prohibitive except under the most extreme circumstances.

authors argue that HaBREM is more technically rigorous and defensible than HEA, reduces scope for "practitioner inference," and therefore will enhance cooperation and resolution of cases.

HaBREM implements HEA as a series of REAs applied to individual resources, each of which functions as a service indicator. For example, Baker et al. (2020) apply HaBREM to a hypothetical wetland assessment. They conduct six separate REAs: (1) above-ground vegetation biomass, (2) below-ground vegetation biomass, (3) white shrimp, (4) brown shrimp, (5) killifish, and (6) amphipods. Each REA identifies a different restoration requirement, ranging from 7.8 acres of wetland creation required to restore above-ground vegetation to 21.1 acres of wetland creation required to restore amphipods.[70]

To this point, HaBREM follows the same methodological approach as would a HEA, where the following are chosen as indicator metrics: (1) above-ground vegetation biomass, (2) below-ground vegetation biomass, (3) white shrimp, (4) brown shrimp, (5) killifish, and (6) amphipods.

HaBREM diverges from HEA in the following manner:

— A traditional multiple-service HEA focuses on the changes in the composite service where the level of composite service is a function of all chosen indicators. This construct allows over-restoration of some indicators to offset under-restoration of others. For example, if the six service indicators identified in the Baker et al. (2020) example were assigned equal weight, 12.25 acres of wetland creation would be identified as the restoration requirement. Four of the indicators would be over-restored (above-ground biomass, brown shrimp, white shrimp, and below-ground biomass), while two (killifish and amphipods) would be under-restored. The public is judged to be fully compensated because, in HEA, the windfall produced by the over-restored indicators offsets the shortfalls associated with the under-restored indicators.

— In contrast, HaBREM imposes a decision rule wherein the compensatory requirement is based on the indicator that requires the most restoration. In their example, amphipods require 21.1 acres of marsh creation, so 21.1 acres is identified as the compensatory requirement. In practical terms, HaBREM identifies the "neediest indicator" and ensures that it is fully restored. In doing so, the method ensures that all other indicators are over-restored.

Importantly, Baker et al.'s (2020) characterization of HaBREM is more than simply placing a weight of 1 on the neediest metric and just ignoring overcompensation of the other indicators. If this were so, the HaBREM calculations would just be dress-up clothes for a negotiation position. The HaBREM decision rule hinges on two technical assumptions: (1) indicators cannot substitute for one another in the production of the composite service and (2) the neediest indicator, whatever it may be, is always the indicator that limits the production of composite service. Under this set of assumptions, there is no overcompensation of the public as the over-restored metrics do not provide any additional services and so provide no value.

A thorough review of HaBREM is beyond the scope of this section. Here, we note the following points.

— The theoretical basis stated informally by Baker et al. (2020) does not in fact support the conclusions of HaBREM. However, an alternative, more carefully developed formulation can support HaBREM based on a special form of an EPF that implies that service indicators always appear in fixed proportions in a habitat with given features. This means that injury and restoration change the *densities* of the indicator species, but not their *mix*.

---

70 In the Baker et al. (2020) example, 7.8 acres of wetland creation would ensure no net loss of present value above-ground biomass, 8.9 acres would ensure no net loss of brown shrimp biomass, 10 acres would ensure no net loss of white shrimp, 10.8 acres would ensure no net loss of below-ground biomass, 14.9 acres would ensure no net loss of killifish, and 21.1 acres would ensure no net loss of amphipods.

HaBREM imposes EPF restrictions by *assumption*; the assumptions are not derived from underlying ecological or economic considerations. It is perfectly possible to make other EPF assumptions and reach different scaling conclusions, while still operating within the HaBREM multiple-REA framework. For example, if we assume the least needy indicator limits the production of composite service, one might identify 7.8 acres as the compensatory requirement. While all the other metrics would be under-restored, they are, by assumption, not needed to produce composite service.

If practitioners were so inclined, they could adopt all the assumptions of HaBREM and then recognize that restoration project design can be adjusted to shift the ratio at which the indicators are produced. That is, if restoring amphipods via a standard wetland creation approach requires almost twice as much acreage as is required by any other indicator, shift the project design to create relatively more amphipods. This more flexible approach could make use of HaBREM's simplifying assumptions while replacing the scaling decision rule of restoring the neediest metric with one that emphasizes cost-effective restoration design.

7.2 Landscape-Level Effects

REA is often used to address resources that are highly mobile throughout their life. Birds may migrate long distances, marine mammals may roam widely, and fish populations may integrate survival of eggs and larvae floating over large swaths of ocean. REA embodies no presumption that either injury or restoration takes place exclusively within the confines of the footprint of oiling. Restoration may best occur at a location distant from the spill site if that location is where the population is limited. For example, compensation for a ruddy duck injury that occurred in Maryland took the form of breeding habitat creation in the prairie pothole region of the upper Midwest. In such a circumstance, it is typical that collateral benefits (or harms) to the broader ecosystem are *not* included in the analysis. The creation, enhancement, or protection of pothole wetlands in the prairie landscape not only increases ruddy duck populations that migrate through Maryland, it increases the populations of other waterfowl and augments other habitat services at the restoration sites. A full economic and ecosystem perspective would seek to capture these corollary effects at the landscape level.

In a HEA, it is typical to confine the scaling analysis to effects on habitats within the footprint of the oiling. Further, when conceptualizing "services" in HEA, the focus is not always on the actual production of services to people, but rather on the *capacity* of the habitat to function and to support service provision. Actual service provision will depend on many other factors in the broader ecosystem. Thus, we have discussed the EPF underlying HEA as governing the production of service indicators in the injured and restored habitats. Actual production of valued services may depend on the biophysical contributions of the habitats, as well as "inputs" supplied by people.

Thus, provision of services to people often occurs at the landscape level of organization, not at the habitat level and certainly not at the level of an affected site. As such, the approach to HEA is generally to look at the ability of injured and restored habitats to contribute (supply) a positive set of functional outputs to the production of services without checking to see if this broader system exists and is able to accept and use (demand) the functions of the injured and restored habitats.

An ecosystem approach would look more broadly at the overall landscape to understand the interactions between the affected habitat site and other habitats and constraints on the system. Because realization of service values requires the mixing of services from natural resources and actions of human systems (both people and human-made capital), this broader ecosystem/landscape view includes both natural and social systems. This broader view was forwarded by the National Research Council in the aftermath of the Deepwater Horizon spill. Clearly, this is a much more elaborate vision than is embedded in most HEAs, and may be beyond the current state of both the biological and economic sciences.

However, we believe there is an intermediate path that can recognize the landscape perspective while remaining practical. This approach would recognize on-site as well as exported services from the injured and restored habitats. The relative value of the exported services may then depend on the landscape position of the injured or restored habitat.

For example, suppose an injured wetland exports energy to adjacent sediment habitats and that both are situated in a disturbed waterway with poor benthic habitat characteristics. In this case, a given supply of the wetland service "energy for adjacent sediments" would be less valuable than if the wetland were adjacent to a high-quality sediment habitat. Since restored habitats can be created with auspicious surroundings, this could be accounted for in an RHV that relates injured and restored wetlands and sediments.

Under some circumstances, oiling can have effects on populations that extend beyond the footprint of oil as well as more narrowly expressed effects at oiled sites. Under these circumstances, it may be efficient to address service reductions within the footprint using HEA, but excluding services tied to the population in question. REA would be used to address the impacts to the population beyond the footprint.

7.3 Discounting in HEA and REA

It is typical in NRDA to assume a constant 3 percent discount rate; the use of this rate has a nontrivial effect on restoration requirement estimates. Consider service acre years Q_t provided at 2 dates, time 0, the base year, and some future time t, with t > 0. When computing DSAYs, Q_0 contributes to DSAYs on a one-to-one basis, while Q_t contributes an amount $Q_t/(1+.03)^t$. If t is large, the same amount of service change has a much smaller impact on DSAYs than an impact in the base year. For example, after 50 years, 1 SAY is worth only 0.23 DSAYs. The "value" is reduced 62 percent because the public has to wait 50 years to receive the service. Damages that occur in the past are compounded forward at 3 percent interest; if the NRDA takes 7 years to complete, 1 SAY of injury in the year of the spill is equivalent to 1.23 DSAYs at the end of the NRDA process.[71]

In the following section, we summarize the reasons the 3 percent discount rate became a standard default value and discuss if and when that rate should be re-evaluated.

7.3.1 The Basis of the 3 Percent Discount Rate

The OPA NRDA regulations state that the discount rate should reflect a "riskless discount rate representing the consumer rate of time preference" (15 CFR § 990.53(d)(4)). However, the phrase "consumer rate of time preference" is ambiguous. Time preference with respect to what entity? Would it be the same for recreation as for ecological services? Is that rate the same for each ecological service across all cases?

NOAA (1999) provides additional specificity. The analysis looks at three lines of evidence, each directed at finding an average consumer's rate of time preference.

— The first line draws on Freeman (1993), which identifies (1) the real (taking out inflation) rate of return on Treasury bills with a three-month maturity (which is as safe an investment as one can find) and (2) the real after-tax rate of return on a portfolio of common stocks (which is considerably riskier). The former averaged zero percent from 1926 to 1978 and exceeded 2.5 percent in only a handful of years. The latter averaged 4.6 percent over this time period, but is not as indicative of the "riskless" discount rate the OPA regulations call for.

— The second line is the real rate of discount used by the DOI when evaluating government projects. This is based on the real return on three-month Treasury bills, which NOAA (1999) found averaged around 3 percent between 1984 and 1999.

[71] Discounting will have a greater effect as (a) the amount of time between the onset of injury and the completion of compensation increases and (b) the assumed discount rate increases. For example, if injuries date back to 1981 (which is often the case under CERCLA), 1 SAY of damage incurred in 1981 is equivalent to 3.26 DSAYs in 2021 if compounded forward at 3 percent. If compounded forward at 7 percent, which is sometimes recommended (e.g., by the Office of Management and Budget), 1 SAY of injury incurred in 1981 would be equivalent to 15 DSAYS in 2021. In some state cases not brought under the federal statutes, NRD claims have reached far back in time (to 1900 or earlier); at a 3 percent discount rate, 1 DSAY of injury in 1900 is worth about 35 DSAYs today, while at 7 percent, that increases to about 3,358 DSAYs.

— The third line is the real rate of growth in U.S. gross domestic product (GDP), which averaged about 3 percent between 1984 and 1999.

Dunford (2018) refined these analyses and updated them to address the period from 1981 to 2016; he concluded that 2 percent is a better estimate than 3 percent.

It is important to ask two questions. First, exactly what concepts do these lines of evidence represent and how are these concepts applicable to elements of a NRDA? Second, do the empirical measures actually pertain to the underlying concepts, and, if so, under what conditions?

Regarding the first question, all of the concepts relate to an average consumer's willingness to trade consumption of a bundle of market goods across time. This is applicable to NRDA to the extent that trades of human use or ecological services across time mirror trades of market goods across time. We shall return to this issue below, but in brief, they *are* applicable to service changes that are monetized via an economic valuation (typically recreation in oil spill NRDAs) and *are not* generally applicable to REA and HEA.

The answer to the second question is best illustrated with a model. Suppose there is only one aggregate consumption good. Pick two time periods named year 0 and year t, with $t > 0$. The representative (or average) consumer gains well-being (utility) of $u(M_0)$ from consumption at date 0 and $u(M_t)/(1+\delta)^t$ from consumption at date t. This utility function is a way to number the indifference curves presented in Section 3; it provides a numerical representation of preferences. Here, δ is called a pure rate of time preference (or utility rate of discount) and arises from impatience regarding the timing of well-being. Note this is a person's impatience, so δ is a behavioral parameter affecting choices of that individual. Note also that it is *assumed* to be a constant. The variable (δ) attaches to utility and captures anything the individual has preferences over, including both market goods and ecological services. For now, we assume it is just market goods, since this is the basis for the 3 percent rate.

Within this construct, consider a trade in consumption through time. This involves a reduction in consumption at date 0 by a small amount C and an increase in consumption at date t by another small amount B. Let m_t be the marginal utility of consumption (income) at date t—that is, the increase in well-being from a little more consumption. Utility is decreased at date 0 by $C \times m_0$ and is increased at date t by an amount $(B \times m_t)/(1+\delta)^t$. For this transfer of consumption to increase the person's well-being, it must be that the utility cost $C \times m_0$ is less than the utility benefit $(B \times m_t)/(1+\delta)^t$. To just break even, it must be that:

$$C = B\left\{\left[\frac{m_t}{m_0}\right] \div \ (1+\delta)^t\right\} \tag{7}$$

The term in brackets on the right-hand side of Equation (7) is the riskless "consumption rate of discount" or consumer rate of time preference for market goods. It is composed of two terms. The denominator is the utility rate of discount based on the rate of impatience. The numerator is the ratio of the marginal utility of consumption at date t to the marginal utility of consumption at date 0. Anything that alters this ratio modifies the discount rate. In particular, if the baseline amount of consumption M_t is growing or shrinking, this will affect the marginal utility of consumption as long as the utility function is not linear.

Assuming that the marginal utility of income decreases with increasing income (a small increase in income when poor provides more utility than that same increase when rich), the term in brackets is equal to 1 only if there is no change in income over time. In this case, the discount rate is just the rate of impatience. However, with a growing economy (an empirical fact) and a curved utility of consumption (estimated to be true on average), the discount rate is greater than the rate of impatience.

We can deconstruct the term $[m_t/m_0]$. Suppose the economy has grown at an average rate of g_t between date 0 and date t. Then m_t is falling as t is increasing. Let γ_t be a measure of the amount of curvature in the utility function between the points M_0 and M_t.[72] It can be shown that Equation (7) can be rewritten as:

$$C = B/\,(1 + \delta + \gamma_t g_t)^t \quad (8)$$

The consumption rate of discount is $r_t^C = \delta + \gamma_t g_t$. Note that there are time subscripts on the social rate of discount. The discount rate will be different depending on the length of time between date 0 and date t. This is called the term structure of interest rates. The rate will have a constant term structure if (1) the rate of growth of consumption is constant and (2) the utility function has the property that its curvature γ_t is constant.[73]

How does this relate to the evidence used by NOAA in 1999—the rate of growth of GDP and the rate of return on three-month Treasury bills? First, assume that consumption is a constant fraction of GDP so their rates of growth are the same. Then, if γ_t is a constant equal to 1 and the rate of impatience is zero, the consumption rate of discount equals the rate of GDP growth. That is, while the rate of GDP growth plays a role in the discount rate, GDP growth is a fairly weak basis for estimating the discount rate.

The 3-month Treasury bill argument is much stronger. Suppose the average consumer's choice of consumption over time maximizes the discounted value of utility (discounting at the rate of impatience), subject to a budget constraint that the discounted value of consumption cannot exceed the discounted value of wealth, and calculated using a competitive market rate of interest for borrowing and lending, r. Then it can be shown that consumption will be arranged such that:

$$\left[\frac{m_t}{m_0}\right]/(1+\delta)^t = 1/(1+r)^t \quad (9)$$

Recognizing the left-hand side of Equation (9) as the consumption rate of discount, if the rate of impatience is greater than the rate of return on savings, the individual will consume more in the current period as makes sense.

Now think about borrowing and lending in a capital market with banks acting as an intermediary between consumers and firms. A profit-maximizing firm will undertake an investment project only if its return in future profits is larger than the opportunity cost of the funds used to pay for the project—that is, the firm's cost of capital. The net return on the investment by the firm is the marginal product of capital (*MPK*). If the project is a safe one (i.e., riskless), competition means that that the opportunity cost of funds to the firm is the return to an alternative riskless investment. With both consumers and firms willing to transact trades at the rate of return on a safe investment, this will be the market equilibrium rate on safe investments, a potentially observable entity. A three-month U.S. Treasury bill is considered among the safest investments in the world. Subtracting the rate of inflation (the consumer price index) from the nominal (including inflation) return to three-month Treasury bills, one can calculate a real return. Using this as a measure of r, by Equation (9), this is equal to the consumption rate of discount. Thus, we have that:

$$r = MPK = \delta + \gamma_t g_t \quad (10)$$

This is a famous equation called the Ramsey Rule, derived by British mathematician Frank Ramsey in 1928.

Based on the Ramsey equation, in a market economy equilibrium $r = \delta + \gamma_t g_t$, where r is the real rate of return to safe investments[74], the real rate of return to three-month Treasury bills is a reasonable estimate

---

[72] Using calculus, $\gamma_t = -(M_t)\{u''(M_t)/u'(M_t)\}$, where $u'(M_t) = m_t$ is the first derivative and $u''(m_t)$ is the second derivative.

[73] This will be true if $u(M)$ has the special form $u(M) = M^{(1-\gamma)}/(1-\gamma)$, which is often assumed in analytical work.

[74] Ramsey derived his result as the outcome of an economy run by a social planner with an objective of maximizing $\sum_t w(M_t)/(1+d)^t$. Today, we would call this objective an SWF, with $w(M_t)$ a well-being function assigned by the planner to a person with income M_t and d an ethical parameter. See Adler (2019) for an introduction and overview of SWFs and papers in Portney and Weyant (1999), as well as Gollier and Hammitt (2014), for an application to long-term discounting (for instance, evaluating

of r. Thus, a reasonable estimate of the consumption rate of discount is the average real return on three-month Treasury bills over the time period when discounting takes place.

The NOAA period for the 3 percent rate was the average over the period 1984 to 1999. This was a time of fairly high interest rates. Rates in the past are observable. Hence, at least for past and near-future effects of an oil spill, a reasonable way to estimate the discount rate that is completely consistent with NOAA (1999) is to look at the recent real return to three-month Treasury bills. From 1981 to January 2021, which includes 9 months of a global pandemic, the average rate is 1 percent. From 1990 to 2021, it is 0.2 percent, and from 2000 to 2021, it is −0.6 percent. For projecting into the future, one might want to average over periods where there was a mix of both upswings and downswings. For example, the average rate from 1900 to 2006 was 1 percent [Gollier and Hammitt (2014)].

7.3.2 Is the 3 Percent Rate Applicable to OPA NRDA?

The 3 percent rate and the analysis on which it is based pertain to a discount rate for a bundle of market goods. This does not seem entirely applicable to ecological services. However, if/when services are monetized, intuition suggests the discount rate for market goods may be somewhat more applicable.

Again, we evaluate using a mathematical model. Let the average consumer's utility at date t be $u(M_t, Q_t)$, where Q_t is ecological services. The economic value of Q_t at date t is the marginal utility of the service, q_t, divided by the marginal utility of income, m_t. This division by m_t denominates the value in units of consumption; therefore, the consumption rate of discount $r = \delta + \gamma_t g_t$ would apply. This would be relevant to a typical OPA recreation analysis where the amount of service (trips to a recreation site) is multiplied by the economic value of a trip to that site.

But for HEA and REA, natural resource services are not monetized, so they are not translated into consumption equivalents. In this case, the public is trading off a loss of natural resource services in one time period due to injuries (ΔQ_0^I) for an increase in services from restoration in some other period (ΔQ_t^R). This appears to call for a different discount rate, called the ecological rate of interest [Gollier (2010); Gollier and Hammitt (2014); Hoel and Sterner (2007)].

Specifying the ecological discount rate for use in HEA and REA is a difficult matter in the general case. This is because deriving the Ramsey Rule in this more general setting requires an allowance for the potential that m_t depends on Q_t, and that q_t depends on M_t. In this general case, discount rates for either consumption of market goods or consumption of ecological services need to account for temporal rates of change among both income and services, as well as the rate of substitution between income and services in generating utility. If, for purposes of simplification, it is assumed that the overall utility function takes the additive form $u(M_t, Q_t) = \alpha(M_t) + w(Q_t)$, the marginal utility of consumption is a constant α, and the marginal utility of natural resource services, q_t, is independent of income, it is possible to re-derive the Ramsey equation for services and find that ecological rate of discount is:

$$r^E = \delta + \gamma_t^E g_t^E \tag{11}$$

In this equation, δ is the rate of impatience applied to overall utility $u(M_t, Q_t)$. The term γ_t^E relates to the curvature of the utility function for services and g_t^E is the rate of change of ecological services. The ecological discount rate r^E should be used in HEAs and REAs.

If a spill and restoration are of short duration relative to the time over which there is material change in baseline service levels, r^E is reasonably approximated as the rate of impatience δ. Economists have estimated δ to be on the order of 1 percent to 1.5 percent [Gollier and Hammitt (2014)].

climate change policies). If the planner respects individual preferences and sets $w(M_t) = u(M_t)$ and $d = \delta$, the Ramsey Rule can be derived as a competitive market equilibrium [see Blanchard and Fischer (1993)]; this is how we have interpreted it.

If a spill and restoration are not of short duration, the second term in Equation (11) is relevant. There is little readily available evidence about the parameter γ_t^E. However, some general observations can provide guidance.

— If there is a trend in the resource base that was injured, the rate of change needs to be considered, in addition to the impatience term. For example, consider a HEA conducted for a sediment habitat where the amount of composite sediment service in the region is increasing over time as point and non-point sources are controlled. In this instance, injury occurs when the resource base is relatively impoverished and restoration occurs when the resource base is relatively expansive. Because the average person is averse to inequality (as measured by γ_t^E), we need a larger future benefit to overcome the increase in inequality. As the aversion to inequality grows, so must the rate of return.

— Alternatively, consider a wetland HEA in the Gulf of Mexico, and note that wetlands are being lost at a material rate due to subsidence, erosion, and sea level rise. Now g_t^E may be negative and substantial. In this instance, the discount rate for a HEA may be very small or even negative. This is because, with declining marginal utility of service, a small reduction in service today is of relatively small importance. Future restoration will add to a degraded resource base, and will be relatively highly valued.

  — In other words, current injury robs from the resource rich and future restoration confers a benefit to the resource poor. Because the average consumer values equality, the time shift, in isolation, is a beneficial outcome. If this time-shifting effect is sufficiently strong, the discount rate is negative. The same would hold true for a REA conducted with a declining species population.

One important final note: Section 3, and many papers discussing the foundations of HEA and REA, note the requirement that the baseline is constant so that values of services are constant and can be canceled from the scaling equation. Under this simplifying assumption, we feel the current 3 percent discount rate is somewhat overestimated. The discussion of the ecological discount rate using the Ramsey Rule shows that even if the resource base is not constant, the effects of such changes can, perhaps, be embodied into the discount rate.

7.3.3 Discounting Over Long Time Periods

Some economists have argued that the market rate of interest approach outlined in Section 7.3.2 only applies to effects that occur within a single generation—that is, for time frames on the order of 30 to 50 years. In this scenario, individuals are trading enjoyment of services now versus later.

Some NRDA cases (even some addressed under OPA) deal with service changes over much longer time frames. As an example, the restoration action for injuries to marbled murrelets for the *M.V. Stuyvesant* spill in California was the purchase and preservation of a large expanse of old-growth redwood habitat, which has existed for hundreds of years and may continue to do so for many hundred more.

When discounting across generations, some economists argue that policy analysts should focus on an ethics-based approach. One frequent dictate from an ethical viewpoint is that the utility rate of discount (δ) in the Ramsey Rule should be equal to zero, thereby lowering the discount rate. The rationale is that this portion of the discount rate essentially penalizes a generation based on when it is born, and this “birthism” form of discrimination should not be admitted in public decision-making. A full discussion of long-term discounting is beyond scope of this document; instead, we point the reader to Gollier and Hammitt (2014).

7.4 The Treatment of Uncertainty in REA and HEA

The goal of NRDA is to compensate the public for injury to natural resources. Essentially every element of this task involves uncertainty arising from a lack of perfect information. Each uncertain element has an inherently unobservable true state “out in the real world.” The true states for each element combine in some unknown process to generate a true amount and form of compensation that, if implemented, would compensate each member of the public. In the absence of perfect information about each true element, there is a range of possible true states for each. The set of possibilities is astoundingly large.

The statutory basis for NRDAs, the regulatory guidance associated with them, and a collection of science and economics principles, methods, and findings developed over decades create a framework that eliminates many of these astoundingly numerous possibilities from consideration in practice. For example, compensating each person may require a set of private side payments in money or some other good or service in addition to resource restoration. These are not admissible under the NRDA statutes and so need not be considered. Further, the OPA NRDA regulations reflect a set of fairly standard economic principles and models, such as measuring values by WTP for changes in services, itself operationally tied to certain underlying concepts, such as rational preferences. Finally, science conceives of habitats, organisms, and ecosystems as operating in certain ways that point toward certain EPFs, preferential measurement of certain attributes over others, and interpretation of some indicators along somewhat consistent lines.

Imposing these limitations and constraints eliminates a large number of potentially true states that the NRDA process could conceivably identify. The basic problem is for practitioners to make a decision about each element of a NRDA from among the implementable possibilities. These decisions include the following:

— What are the injured resources that will be addressed in the NRDA?

— For each injured resource, what is the appropriate high-level approach to scaling? Will it be value-to-value, value-to-cost, HEA, or REA? This defines the possible scaling models, which are ways to combine more basic elements into an estimate of NRDs.

— Having chosen a high-level approach, what specific form will the scaling model take? Will a REA be based on professional judgment or population modeling? If HEA is chosen, what are the relevant services and service indicators? What are possible restoration actions and what specific EPF will be used? This choice makes specific the list of more basic elements (model parameters) that are needed, as well as the way those parameters will be combined to compute NRDs.

— Any specific scaling model will imply a set of parameters (basic elements) needed for its implementation. Each has a set of unknown possible true values. What methods will be used to select a value (or range of values) for each parameter, and what final decision (selection) is made? For example, a parameter might be the number of young fledged in the year after the spill. What data are available or will be collected, what statistical approach will be used, and what decision rule will be used that says, given these data, this is the estimated number of fledges?

— How will uncertainty itself be dealt with? Will uncertainty about a statistically estimated parameter be carried forward into a scaling model, or will the estimate, once made, be treated as known? If the former, how will the uncertainty around the estimated parameter be determined? Will it include only the chosen statistical model, or also alternative models? Will alternative scaling models be assessed and assigned a probability? Will a potential violation of an underlying assumption of HEA, such as whether the utility function is "too curved" relative to the possible range of effects such that approximation errors are large, be assessed and given a probability?

Ultimately, dealing with uncertainty in NRDA involves making a decision about the cost of compensatory restoration with imperfect information. A decision needs some rule that specifies if the information is X, then the decision is d. A good decision rule is one that minimizes assent costs but still leads, on average, to neither excessive overcompensation nor undercompensation of the public.

In the remainder of this section, we approach the question of uncertainty using methods in line with current practice and standard approaches.

7.4.1 Selecting a Damage Estimate from a Distribution of Potential Damages

Total damages are usually computed as NRDs for each injured resource (e.g., birds, marine mammals, sediments, shorelines, fish, and recreation) adjusted to adhere to service accounting principles, and then summed.

For example, suppose that NRDA practitioners have specified a probability distribution for every uncertain input value. Using these distributions, it is possible for the practitioner to estimate a distribution of total damages using Monte Carlo methods. These methods “draw” a value for every input from its underlying probability distribution. Based on this draw, the associated NRDs are computed.

This program of “draw elements and compute NRDs” is repeated a large number of times. The output will look something like Figure 7-1, where the width of each blue bar is a dollar range for total damages and each height is the relative frequency with which a particular value occurs.

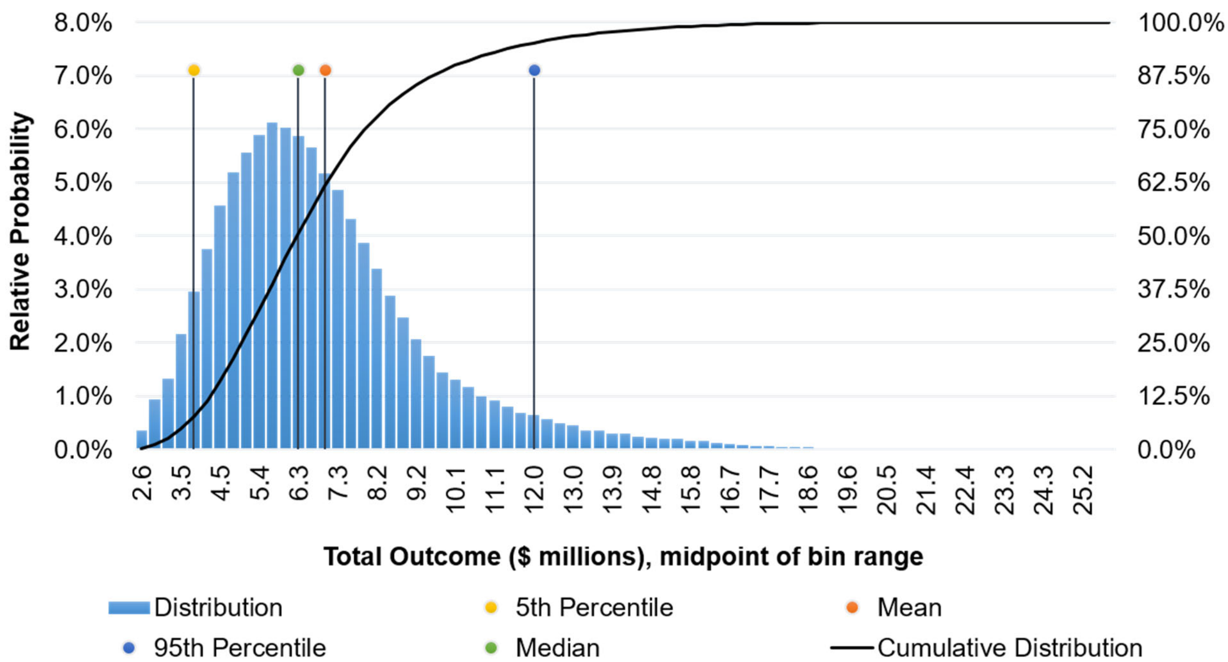


Figure 7-1—Monte Carlo Distribution of NRD Probabilities

Having arrived at the distribution of potential damages, which specific value should be chosen?

— If this same spill were repeated a large number of times, use of the mean would imply that, across all of the occurrences, the public would be properly compensated.

— If the mode were chosen, NRDA practitioners would be “right” more often than if another measure were selected.

— If the median were chosen, the assessment would overestimate compensation as often as it underestimates compensation.

One approach to identifying a preferred estimate is a decision analysis (DA) approach. Suppose the true value of damages is D. Based on the Monte Carlo distribution, true NRDs can take on any one of the possible values $(D_1, \ldots D_N)$ corresponding to the N bins of NRDs in Figure 7-1. Suppose a decision is made to pick d as the value of NRDs when, in fact, the true value is some other number D. In this situation, a decision error is made. The cost of a decision error is a loss of social value of $\mathcal{L}(d|D)$, which is positive if $d \neq D$ and zero if $d = D$. For example, the loss might be calculated as $\mathcal{L}(d|w) = |d - D|$, the absolute value of the difference between the estimated NRDs and the true NRDs. With absolute error loss, if true NRDs are $D =$ \$10 million and the trustees specify $d =$ \$8 million, the loss equals the amount of error: $\mathcal{L}(8|10) =$ \$2 million. If the estimate of NRDs is \$12 million, the same loss is incurred; this is called a symmetric loss function, as positive and negative errors are treated equally.

Alternatively, one may place an incrementally larger penalty on larger errors. Then one might specify $\mathcal{L}(d|D) = (d - D)^2$; this is called squared error loss. This is also a symmetric loss function, but now the penalty for being wrong increases rapidly as the amount of error increases.

Of course, one does not know the true value of NRDs, but Figure 7-1 provides the probabilities that the true value is any particular number D_n. Let $p(D_n)$ be the probability, taken from Figure 7-1, that true damages D fall into bin *n* and so equal n. The expected (or mean) loss from making a decision that NRDs equal some amount d is:[75]

$$E\{\mathcal{L}(d)\} = \sum_n \mathcal{L}(d|D_n)\, p(D_n)$$

A best decision is one that minimizes expected loss. It can be shown that if the penalty for being wrong about NRDs is absolute value loss, the best NRD decision is the median. However, if the penalty for being wrong is squared error loss, the best NRD decision is the mean.[76] Thus, the DA approach provides a basis for a decision rule about NRDs, and in the case of squared error loss, justifies the use of the mean. While these two loss functions are commonly used, the DA approach can also be used to address other loss functions that lead to estimates different than the mean or the median. For example, one could select an asymmetric function that associated a greater penalty with underestimation of NRDs (not fully compensating the public) than on overestimation (spending too much on restoration), which would appear to mimic the informal treatment of uncertainty often put forth by trustees.

The probabilities derived from Figure 7-1 represent the beliefs of the decision-maker about true NRDs, while the loss function reflects the decision-maker's motivations when making a decision. Both come into play in NRDA, and the above thought process provides a model of how one could incorporate beliefs, motivations, and decision-making as a way to understand decision-making under uncertainty.

7.4.2 Current Practices

As sensible and pragmatic as the preceding program is, it is rarely implemented in actual NRDAs, especially for oil spills. Instead, in our experience, trustees express a strong desire to address uncertainty informally by selecting a "reasonably conservative[77]" point estimate for each key input.

While common, this is not a sound practice, as a series of reasonably conservative assumptions embedded into a single HEA generally does not generate a reasonably conservative estimate of NRDs. For example, suppose a hypothetical HEA debit relies on three uncertain, independent parameters:

1) The footprint is between 20 and 35 acres.
2) The magnitude of initial service loss is based on biomass estimates that imply a mean service loss of 50 percent and a standard error (standard deviation of the estimated mean) of 5 percent.
3) Recovery is linear and will take between 20 and 35 years.

---

75 In general, if z is random and can take on the values z_i with probabilities p_i, the mean, or expected, value is $\sum_i z_i p_i$. If $g(z)$ is a function of z, the expected value of $g(z)$ is $\sum_i g(z_i) p_i$.

76 As an example, see DeGroot (1970).

77 A reasonably conservative assumption identifies as an input value one that (1) increases damages relative to the central tendency and (b) would not surprise anyone if, in fact, it were the true and correct value.

If based on a series of conservative assumptions, the HEA practitioners might assume a footprint of 33 acres, an initial loss of 58 percent (the ninety-fifth percentile), and a recovery period of 33 years. The debit associated with these assumptions would be 233 DSAYs. While this may appear, on its face, to be a set of fair expediting assumptions, Monte Carlo simulation reveals that the probability of a debit as large as 233 DSAYs is less than 1 in 1,000. The reason for the difference is the very low probability of getting all high draws or getting all low draws in any one event.

Simply stated, the reasonable worst-case approach should not be applied variable-by-variable. The result will likely be an overall NRD estimate at the extreme upper bound of its range. We highly recommended that when practitioners engage in a "reasonable worst-case" assessment framework, individual parameter values, whether they be expressed as point estimates or distributions, be based on best available information. If mutually agreed, any conservatism could be applied to the overall estimate of NRDs. Alternatively, it may be most expeditious to simply adhere to NOAA guidance on the treatment of uncertainty, which we review below.

7.4.2.1 NOAA Guidance on Uncertainty

The OPA NRDA regulations state [15 CFR § 990.53(d)(4)]:

> *When scaling a restoration action, trustees must evaluate the uncertainties associated with the projected consequences of the restoration action, and must discount all service quantities and/or values to the date the demand is presented to the responsible parties. Where feasible, trustees should use risk-adjusted measures of losses due to injury and of gains from the restoration action, in conjunction with a riskless discount rate representing the consumer rate of time preference. If the streams of losses and gains cannot be adequately adjusted for risks, then trustees may use a discount rate that incorporates a suitable risk adjustment to the riskless rate.*

Further insight is provided in NOAA (1999), which discusses methods for addressing uncertainty under OPA. These are not regulatory requirements, but they do indicate NOAA's interpretation of the regulations. The following can be gleaned from this document.

1) Consistent with the regulations, the preferred approach is to account for uncertainty in the variables (estimated parameters or other NRDA elements) in damage calculations and develop a value for injuries or restoration gains that are treated as certain in each year of the analysis. A risk-free discount rate would be applied to these annual gains and losses.
2) Uncertainties about damages are to be evaluated explicitly, using ranges or probability distributions where possible. These can be developed by a variety of methods, including Monte Carlo techniques. It is not clear at what level of organization this is to be done, but the document mentions fairly high-level entities as "random variables" including injury, restoration benefits, and project scale. There is no guidance on how to select a value from the range or distribution of results.
3) If *variables* are represented as point estimates rather than with ranges or distributions, the expected value should be used as the point estimate.
    a) A reasonable inference based on the preceding is that if damages are expressed as ranges generated through some Monte Carlo-like process, practitioners should select some measure of central tendency. However, from a mathematical standpoint, a point estimate that relies strictly on expected values for each input may not be the same as the expected value generated by a Monte Carlo simulation.
4) Adjustments may be made to the expected value calculation to explicitly account for risk aversion.

7.4.2.1.1 Recommendations Regarding Ultimate Decisions

NOAA recommendations (1) and (2), in our opinion, basically call for the Monte Carlo approach. First, specify the probability distribution of each uncertain element. The idea of a range instead of a probability distribution seems (to us) to be a special case of a probability distribution where each value in the range is given equal probability.

Recommendation (2) has two parts. First, use the ranges or distributions to generate the "overall outcome distribution." This is the distribution in Figure 7-1, interpreting "overall outcome" to mean computed NRDs. The second part is an explicit answer to the question we posed above: Having arrived at the distribution of implementable NRDs in Figure 7-1, which should be chosen? We believe NOAA (1999) provides the answer: Choose the expected value.

We think this is a perfectly sensible way to estimate NRDs under uncertainty. If all practitioners involved in a cooperative assessment were guided by this principle, we believe assessment duration and costs would materially decrease.

7.4.2.1.2 Recommendation Regarding Individual Decisions

NOAA (1999) gives specific advice to the practitioner selecting point estimates for individual inputs: Choose the expected value for each point estimate.

Does this make sense from a technical perspective? The answer is yes, provided the underlying distribution is a reasonable representation of incident-specific uncertainty.

— When the probability distribution is viewed as the expected frequency of the possible outcomes from an event that can be repeated exactly the same each time, the following holds. The outcome of any individual trial may be above or below the expected value, but the arithmetic mean of multiple trials will converge to the expected value as the number of trials increases. In the NRDA context, consider a bioassay in which organisms are exposed to a dose of oil and fractional mortality is recorded. Since there is some random variation in the sensitivity of the test organisms, the outcome of the experiment will vary from trial to trial. But, if the same experiment is repeated many times using the same procedures, then the average of the outcomes will be close to the mortality associated with the mean sensitivity of the organisms. Reliance on the expected value makes perfect sense in these situations.

— Similarly, suppose the baseline value of some measurement is uncertain, but time-series data on that measurement across years is available. Let x_t be the measurement in year t and suppose there are T years of data available. It might be assumed that these data come from a fixed distribution of annual outcomes; thus, each year is a repetition of the process generating the data. Specifically, it might be assumed that $x_t = \mu + \varepsilon_t$, where μ is the true mean of baseline. As the number of years of data increases, the sample average converges to the true mean μ. Reliance on the expected value could make sense provided there is strong evidence that oil spills cause the sort of deviation observed in the data.

— In contrast, if the data come from a time-series and the years vary based on something other than pure randomness (i.e., El Nino events), and this is not accounted for in the analysis, this result does not hold. That is, when practitioners move away from the world of experiments and into the world of field data, reversion to the mean depends on untestable maintained assumptions. Inference in this world is dangerous and usually rests on ad hoc statistical analyses. What is needed in this circumstance is to control for the "other influences" or, at a minimum, to establish causality between the observed deviation and the spill. This requires formally considering the possibility that other factors might reasonably explain an observed deviation from the expected value.

Finally, it is noted that, when the distribution of possible damages is right-skewed as is generally expected,[78] the use of expected values will tend to result in higher damage estimates as uncertainty increases. This phenomenon tends to reduce the potential for undercompensation in highly uncertain situations.

7.4.2.2 Recommendation Regarding Risk Aversion

The OPA guidance mentions risk aversion, as well as the related concepts of *certainty equivalent approach* and *risk premia*. Risk aversion refers to a circumstance in which a decision-maker prefers a more certain outcome to a risky one. Risk aversion is closely related to the idea of wanting to smooth the time-path of services and avoid a high degree of inequality in services over time, as discussed in Section 7.3. Indeed, both derive from the curvature in the service value function.

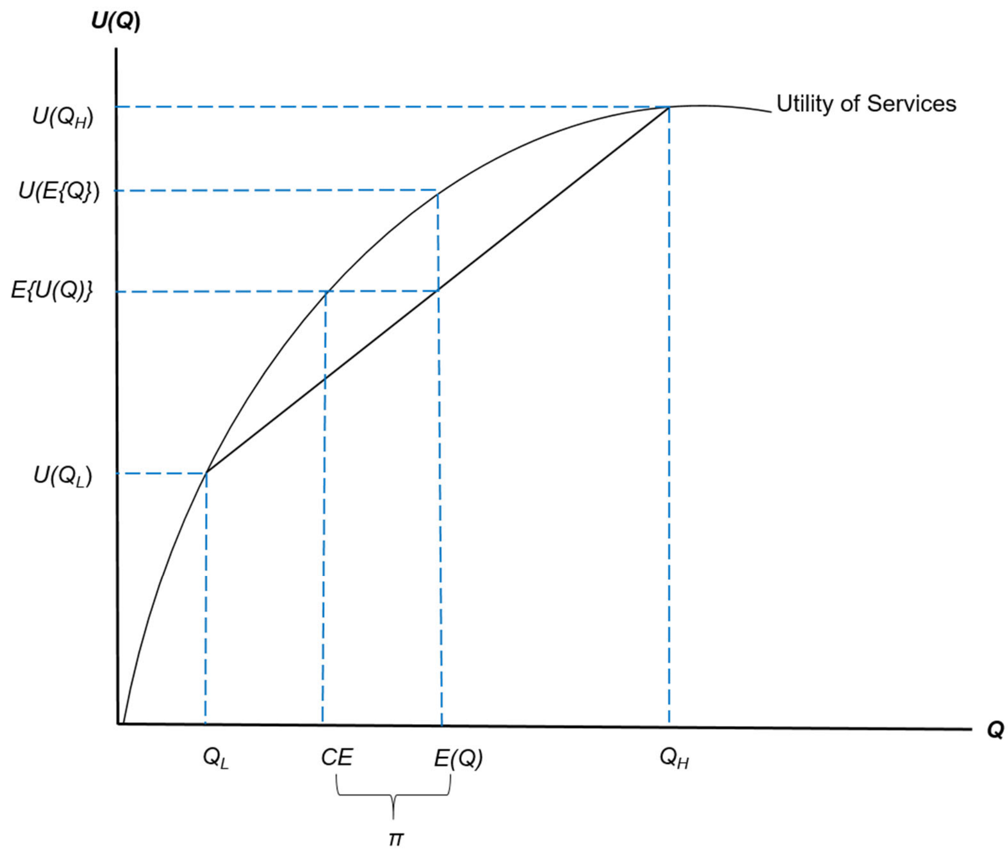


Figure 7-2—The Relationship Between Utility of Services and the Amount of Services in a Representative Time Period

Figure 7-2 describes the relationship between utility of services and the amount of services in a representative time period. Assume the utility function is the same across time periods and that the levels of services will change across periods, but note that the time subscript has been dropped for convenience.

---

[78] The distribution will often be right-skewed because of the multiplicative nature of REA and HEA, truncation of damages at zero, and the fact that environmental data tends to follow a lognormal distribution.

Services are uncertain; suppose they can take on one of two values, high (Q_H) or low (Q_L). The beliefs about these service levels are the probabilities p_H and p_L, respectively. Figure 7-2 plots the utility that is achieved if services turn out to be high or low as $U(Q_H)$ and $U(Q_H)$.

The utility function exhibits declining marginal utility of services, with the slope of the utility function steep when services are low and flat when services are high. The expected value of services in this period is $E(Q) = p_H Q_H + p_L Q_L$. The expected utility when services are uncertain is $E\{U(Q)\} = p_H U(Q_H) + p_L U(Q_L)$. These lie along the line joining $U(Q_L)$ and $U(Q_H)$; if low services are more likely, the expected value moves to the left of the center between $Q_L \text{ and } Q_L$, while if high services are more likely, it moves to the right of center.

If services were sure to be at the mean $E(Q)$, the sure value of utility is $U(E\{Q\})$. This lies above the expected utility when services are uncertain $E\{(U(Q)\}$. The representative individual is risk averse, meaning they get higher utility from the sure level of services V than the expected utility from a distribution of potential services with the expected value V.

If the practitioners were to focus on the mean but can, in special circumstances, consider risk aversion, how would they do this in a HEA or REA that does not enter utility (or value) into its computations? The relevant concept is the *certainty equivalent* amount of services. This is the *sure* level of services that has the same level of utility as the *expected* utility of the uncertain services. It is given by the service level CE in Figure 7-2. A related concept is the *risk premium*. This is shown as π in Figure 7-2; it is the difference between the expected value of services and the certainty equivalent. It is, in essence, a premium that must be applied to the uncertain situation to translate it into a certain equivalent.

We now relate this to discounting. If the representative person is facing a time path of random service levels, NRDA practitioners have a choice. They can translate each time period's level of services into its certainty equivalent (or deduct the risk premium from the expected value), which makes them no longer uncertain, and apply a riskless discount rate (such as the rate of return on three-month Treasury bills). Or, they can leave the uncertainty in the time path of services, but then they must use a risk-adjusted discount rate. The two formulations are equivalent in principle. The first is the preferred approach under the OPA regulations, but the second is admissible [15 CFR § 990.53(d)(4)] and often the approach taken in financial economics.

It is worth noting that if the utility function is linear, the marginal utility of services is constant, with several helpful results: (1) HEA and REA are exact, (2) there is no need to adjust discount rates for growing or shrinking levels of services at baseline, and (3) the representative person is risk neutral, the risk premium is zero, and the certainty equivalent is the same as the expected value. One can plug expected values of parameters into the HEA or REA equations, treat these as certain, and use a riskless discount rate equal to the rate of impatience.

Certainty equivalents and formal risk adjustments are rarely, if ever, applied in practice. However, the concept that the public requires additional compensation as a way to manage the risk of an underpayment is commonly included in practice in the form of ad hoc adjustments.

The 1999 NOAA guidance does allow for the use of contingency factors to address uncertainty. Essentially, this approach relies on expected values to estimate NRD liability, then adds some amount, based on professional judgment, to a settlement to decrease the possibility that the public might be undercompensated when restoration is implemented. This is an informal risk premium. As suggested earlier, if this contingency factor is applied to the overall damages approach, it is more reasonable than biasing individual point estimates.

7.4.2.2.1 Risk-Adjusted Discount Rates

Incorporating an adjustment for risk into the discount rate is noted as an option in the OPA guidance. If this option were chosen, it would be necessary to ensure that no risk adjustments, either formal or ad hoc, are incorporated into the underlying HEAs or REAs. To do both is to double-count the effect of uncertainty.

To make the correct adjustment, the various sources of uncertainty must be identified. Is it the rate of growth of services that is changing the baseline that is at issue? Or is it the amounts of debit and credit that are uncertain?

The issues are both complicated and technical [see Gollier and Hammitt (2014)] for a semi-technical overview. We can briefly summarize that if the debits and credits are certain and it is only the growth rate of services that is uncertain, one can adjust the risk-free rate to account for this source of uncertainty. This adjustment is complicated as it depends on the form of uncertainty (e.g., whether shocks to growth are correlated and persistent), but it can be accomplished with some assumptions. The need for risk premia arises when annual debits and credits are uncertain. There is an adjustment to the risk-free rate that can be made by adding a risk premium. This risk premium depends on how the debits or credits are correlated with potential swings in the size of the overall resource base—that is, with changes in baseline services. A service flow that swings opposite to swings in the resource base is very valuable.

These are important areas for further research and their importance is likely to increase in importance if climate change increases uncertainty.

7.4.2.2.2 Institutional Controls

Institution controls are mentioned in the OPA guidance as an approach that can be used to reduce uncertainty related to restoration credit. Often, the gains from restoration projects are uncertain because the “success” of the restoration project is unknown. For example, restoration may involve putting nesting platforms out for ospreys. If no ospreys occupy the platforms, no osprey services are gained.

Institutional controls, such as performance criteria and adaptive management, ensure that if restoration is not successful, the funds and expertise are in place to adjust the restoration project to ensure that the expected level of service gains (as measured by various field observations) are realized.

The use of institutional controls, particularly in a cooperative setting, can greatly reduce uncertainty associated with restoration outcomes.

7.4.2.2.3 Potential Double Counting of Uncertainties

The analyst must be careful not to apply multiple, overlapping factors to account for uncertainty. This commonly occurs in the cooperative NRD setting. For example, it would be incorrect to take the expected value of the distribution of restoration requirements, add a risk premium, and then add a contingency factor. Similarly, it would be incorrect to choose a high value from the range of potential debits, choose a low value from the range of potential credits, and then add a contingency factor.

Similar issues can arise regarding potential double-counting among performance criteria and the size of the restoration project. For example, if a NRDA practitioner explicitly reduced the expected credit of a project to address uncertainty, it would be incorrect to have performance criteria that ensure a high level of performance; one should either ensure very high performance using institutional controls and get very high credit, or lower the expected credit and not have stringent institutional controls. It is incorrect to ensure high performance and then scale restoration based on the assumption of low performance.

7.4.2.2.4 Final Thoughts on Risk and Uncertainty

In most cases, treating uncertainty in NRDAs in ways that more closely adhere to the trustee guidance will generally lower NRD liability estimates and increase the confidence of all parties involved that issues related to risk and uncertainty are being adequately resolved.

However, additional quantification will come with additional cost, both for the actual quantification and for negotiation. We note that it may not generally make sense to undertake a full probabilistic analysis of uncertainties for every NRDA. The analyst should balance the benefits and costs of more complex methods. We believe that use of ranges and distributions for NRDA elements, Monte Carlo, and expected values would go a long way toward increasing confidence in the process and reducing transactional costs.

7.5 Managing NRDA Costs: A Decision Approach

The NRDA process is not a linear one, marching through the steps of HEAs and REAs to arrive at scaled restoration, with a fixed array of data available. The process is iterative, with various elements chosen provisionally and revised or updated as new information becomes available. And what new information to obtain is itself a NRDA element to be decided upon. There is (a usually implicit) tradeoff that is made between: (1) a simple approach that conserves information and transaction costs, but may embody assumptions that do not apply in the case at hand, and therefore provide a biased result, and (2) a more complex approach that may be correct in principle, but costs more to implement and can raise its own uncertainties. These tradeoffs are made at various levels of the analysis, from choice of resources to assess to scaling model selection to parameter estimation. Often, the tradeoff analysis is informal and takes place in the background of cooperative NRDA negotiations.

One well-studied approach to these issues is decision analysis (DA). In the remainder of Section 7.5, the loss-minimizing approach presented in the previous section is applied to decisions related to the acquisition of information.

The OPA regulations state the following about methods and collection of information:

> *The additional cost of a more complex procedure must be reasonably related to the expected increase in the quantity and/or quality of relevant information provided by the more complex procedure. [15 FCR § 990.27(a)(2)]*

It is not at all clear how one should relate costs (measured in dollars) to the quantity and/or quality of information, neither of which is defined in the regulation.

The CERCLA regulations define a "reasonable assessment cost" [43 CFR 11 § 11.14(ee)] in a manner that is similar in spirit to the concepts in OPA, but more specific:

> *Reasonable cost means the amount that may be recovered for the cost of performing a damage assessment. Costs are reasonable when: … the anticipated increment of extra benefits in terms of the precision or accuracy of estimates obtained by using a more costly injury, quantification, or damage determination methodology are greater than the anticipated increment of extra costs of that methodology; and the anticipated cost of the assessment is expected to be less than the anticipated damage amount determined in the Injury Quantification, and Damage Determination phases.*

In this definition, the expected increase in the quantity and/or quality of information is linked to an expected *benefit* of increased precision or accuracy of estimates of uncertain elements of a NRDA. This is a useful step, as it is possible to measure such benefits in dollars, which are directly comparable to costs. The goal of this section is to outline how one might operationalize this idea.

7.5.1 Decisions Under Uncertainty and the Expected Value of Information

In the preceding section, Monte Carlo methods were used to generate a probability distribution for total damages based on the underlying uncertainties in scaling models and their input parameters. We also introduced the idea of a loss function $\mathcal{L}(d|D)$, which specifies the loss from estimating that total damages are equal to d when they are in fact equal to D. The shape of the loss function embodies the decision-makers' motivations in the NRDA and attitudes toward risk. A rational decision in the face of uncertainty about the true amount of total damages is to minimize the expected loss, where the averaging is across possible values for D. The probabilities attached to the possible true values of NRDs embody the decision-makers' beliefs about the various factors that govern NRDs, such as baseline, injury, recovery, and restoration credits.

In Section 7.4, we described $p(D_n)$ as the probability that true damages D equal D_n. We proposed that a rational way to make a decision about NRDs is to minimize the expected loss from restoration decision errors. That is, one should pick d to minimize:

$$E\{\mathcal{L}(d)\} = \sum\nolimits_n \mathcal{L}(d|D_n)\, p(D_n)$$

7.5.1.1 The Expected Value of Information

Suppose that fairly early in the NRDA, practitioners populate a Monte Carlo analysis using ranges of possible values for inputs based on preliminary spill information, literature values, NRDA precedents, and best professional judgment. Let $p^0 = (p_1^0, p_2^0, \ldots, p_N^0)$ be the preliminary probabilities attached to possible values of NRDs coming from that exercise. We will call these the "prior beliefs" as they exist prior to possible collection of additional data via site-specific studies.

Given these prior probabilities, the expected (or mean) loss from making a decision that NRDs equal some amount d is:

$$E^0\{\mathcal{L}(d)\} = \sum\nolimits_n \mathcal{L}(d|D_n)\,(p_n^0)$$

The estimate of NRDs that minimizes this expected loss is d^0.

For example, if the loss function is squared error loss, d^0 would be the mean of the preliminary Monte Carlo distribution. This is the best estimate of NRDs when you have the preliminary information associated with the prior beliefs. As noted in Section 7.4, squared error loss seems to comport with the NOAA (1999) guidance that one should typically select the mean of the distribution when estimating NRDs under OPA.

The minimum *amount* of expected loss we face when we undertake the preliminary analysis is $E^0\{\mathcal{L}(d^0)\}$. This is the best we can do with the preliminary information available; it is called the *risk* associated with the prior beliefs.

With squared error loss, the risk of the prior beliefs is just equal to the variance of the preliminary distribution. In this case, if $\bar{d}^0$ is the mean of the preliminary Monte Carlo, the risk associated with the prior information $R(p^0)$ is:

$$R(p^0) = \sum\nolimits_n (d_n - \bar{d}^0)^2\,(p_n^0)$$

Suppose one anticipates studying some uncertain NRDA elements. The probabilities of the various possible values of NRDs will change, depending on the study result. These are called *posterior beliefs*. Denote these by $p^1 = (p_1^1, p_2^1, \ldots, p_N^1)$. With the study results in hand, we minimize the new expected loss and get a new best estimate of NRDs, d^1. Given the new probabilities, making a best decision in the face of these gives rise to a new level of risk, $R(p^1)$.

Of course, before the studies are conducted, we do not know exactly what the results will be, so we must use our *expectations* of what the new probabilities will be—that is, the expected posterior beliefs. Two fundamental results of decision theory are:

Fact (1): The mean of the expected posterior beliefs is equal to the mean of the prior beliefs.

Fact (2): The variance of the posterior beliefs is generally smaller than the variance of the prior beliefs.

Fact (1) says that before conducting a study, you cannot expect that the data resulting from the study will change your current best guess about the unknown parameter, which is the prior mean. While this seems odd at first, this must be so—otherwise, the prior mean could not be your current best guess. Of course, the mean will change after the study has been conducted and the results analyzed.

Fact (2) is of great importance, as conducting an investigation is expected to reduce uncertainty, even though the mean (or best guess) is not expected to change. Recall that with squared error loss, the risk equals the variance. Fact (2) says that the beliefs we expect to hold after conducting a study result in computing a smaller variance of the uncertain parameter. Thus, with squared error loss, the expected risk falls since the risk equals the variance. This is the source of the value of information.

These two facts and their implications can be surprising to non-specialists, who often ask: How will my current best guess about the value of the parameter change if I conduct a study? These facts show that the answer is somewhat subtle. Fact (1) tells us that one must expect that, on average, a study will tell you nothing new (your best guess will not change). But, Fact (2) and its basis require that if the study is expected to estimate the true parameters precisely (the posterior variance is expected to be small because the sample size is large), it is also expected that the prior mean will be changing a lot (and the converse).

We can now define the expected value of information (EVOI) as:

$$EVOI = R(p^0) - E\{R(p^1)\}$$

The expectation in this expression is taken with respect to what the new probabilities will be. That is, we must have a probability distribution characterizing study results.

A full analysis is beyond the reach of this document. But in brief, a "study" in this setting is characterized by a *likelihood function,* which specifies the probability of getting study result x when the true value of what is being studied is w. This likelihood is a conditional probability, $\pi(x|w)$. Then, with the particular study result x^1 in hand, one can use a basic statistics result (Bayes' Theorem) to "invert" the likelihood function to get the posterior distribution. The posterior beliefs are the conditional probability that the true value of the NRDA element being estimated is w having collected the data x^1. We write this outcome as $p^1(w|x, p^0)$. The posterior beliefs depend on both the data and the prior beliefs. It is a weighted average of the two; if the prior beliefs are really certain (have a small variance), little weight is put on the new information from the study. Only a very powerful study would change beliefs that are well-informed. Conversely, if the prior beliefs are "we have no idea what the true value might be," then a great deal of weight is placed on the new data, and beliefs can be expected to change a lot. This is when the new information has its greatest value.

7.5.1.2 More Informative Studies

Of course, information can be collected in various ways, and study design is an important aspect of NRDA. We can relate study design to the value of information by defining study X to be more informative than another study X' if the EVOI of X is larger than the EVOI of X'. With squared error loss, we know that the EVOI is equal to the variance. Thus, in this case, one can rank alternative studies by their ability to reduce variance.[79]

---

[79] We know the mean using expected posteriors is the same as the mean using the priors (Fact 1), but the variance will shrink (Fact 2). Therefore, study X' will be less informative than X if X' is a mean-preserving spread of X. A mean-preserving spread is known to reduce the expected value of a convex function we are seeking to minimize, like a loss function. This insight links the value of information to many well-developed results in economics on the theory of decision-making under uncertainty.

This is true across study designs on a given topic and across topics that might be studied. In terms of overall NRDs from a Monte Carlo distribution of NRDs built up from many elements, which elements should be prioritized for study (e.g., baseline versus injury versus recovery) would depend on their expected ability to reduce the variance of the overall distribution of NRDs and their cost.

7.5.1.3 The Cost of Information

We are not as much interested in the gross value of information EVOI as we are in its net value, taking into account the cost of information. Let $C(X)$ be the cost of obtaining data from study X. The expected net value of information (ENVOI) is defined by:

$$ENVOI(X) = EVOI(YX) - C(X)$$

7.5.1.4 The Shape of the Cost Function for NRDA Studies

What might this cost look like? There are two sources of study costs in the NRDA process that may be relevant: (1) the cost of undertaking the study itself (e.g., study design, sampling, data management, analysis, quality assurance, reporting) and (2) negotiation costs regarding the implications of the study outcomes in the overall assessment. The first is likely to be increasing based on the information content of the experiment; more informative studies tend to be more expensive (e.g., increasing sample size). However, the second type of cost could decrease with more informative experiments. Imagine that the parties in a cooperative assessment have different prior beliefs and views of the prospects for a satisfactory study. Then, a weakly informative experiment could be unconvincing, and potential study deficiencies are then subject to considerable debate. A weakly informative experiment could be unconvincing, and potential study deficiencies are then subject to considerable debate. The results of a robust study design might be agreed to quickly. If we plotted the “amount of information” in the study along the horizontal axis and the costs on the vertical, the sum of these two sources of cost would make the total U-shaped, as depicted in Figure 7-3.

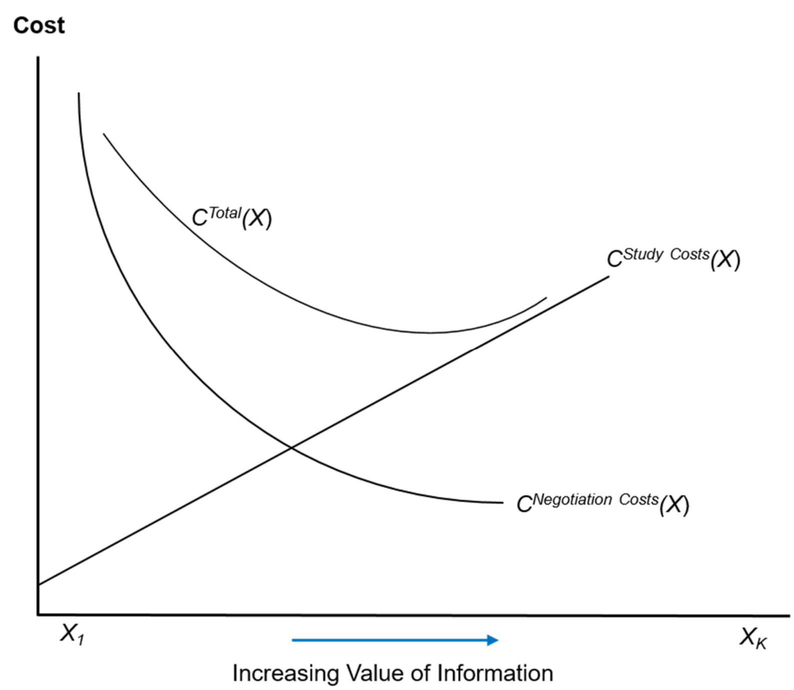


Figure 7-3—The Cost of Information

7.5.1.5 Optimizing Study Design

The optimization of a study design is to maximize the ENVOI by equating the incremental (marginal) EVOI from a more informative study to the incremental costs of conducting it. This is shown in Figure 7-4.

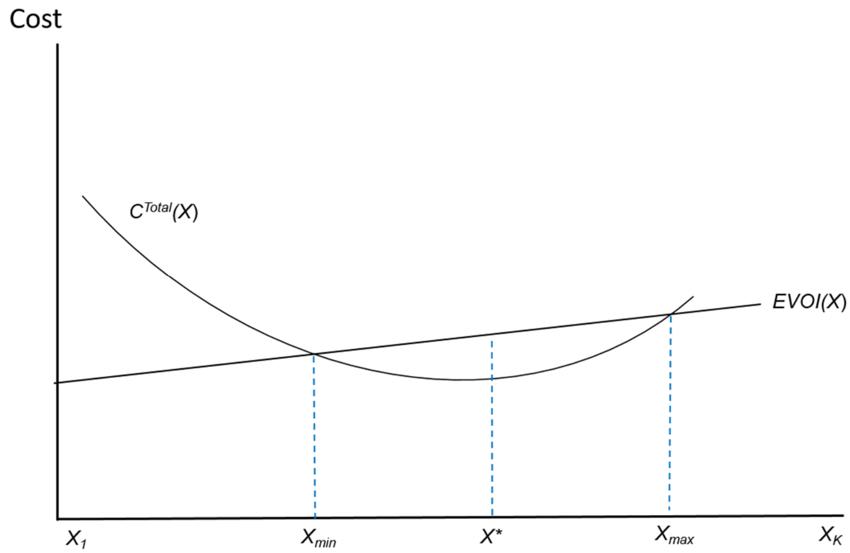


Figure 7-4—Optimizing a Study

Studies that are less informative than Y_{min} should not be undertaken, nor should studies more informative than Y_{max}, although without undertaking an analysis of costs versus benefits, a scientist may propose that Y_K is the best study. However, the cost of this study would not constitute a reasonable assessment cost. The optimized study is Y^*.

7.5.2 Further Implications of the Decision Approach

The basic approach of decision-making outlined above can be applied at the level of a single parameter in a scaling model. One can reasonably isolate one parameter if the others involve little or no uncertainty or have been agreed to. The uncertainty in the one parameter is then equivalent to the uncertainty in overall NRDs; the loss function was defined at this higher level.

In real cases, many more elements of the model would be uncertain, and they would combine with the certain elements to produce overall NRDs. There are several ways one could apply the above concepts.

One possibility is to carry the decision framework laid out here through to the more complex setting of multiple uncertain individual NRDA elements that interact to produce NRDs. Another possibility is to treat elements in groups surrounding higher-level topics; for example, the amount of initial injury or restoration credit in a given time period or the shape of a recovery curve. The decisions made at this higher level of aggregation could then be combined into an estimate of overall NRDs. This could be used to screen for which broad topics to work on by ranking them according to NRD variance reduction and then to decide how best to attack uncertainty at the topical level.

The discussion above is directed to parameters of scaling models, not the choice of the scaling model itself, as are most previous discussions of uncertainty in NRDA. The thought process outlined above can be modified to apply to the choice of scaling approach. An example is the paper by Byrd and Tomasi (2021) about recreation in NRDA. That paper examined whether to use a value-to-cost scaling approach or to spend more to measure the benefits of restoration projects and conduct value-to-value scaling. The difference hinges on a single uncertain parameter: the ratio of benefits of restoration to restoration cost. The model characterized when the simpler value-to-cost model was justified despite its associated bias. A similar analysis could be applied to different ways to implement a HEA or REA model or decide whether to use either of these instead of a value-to-value approach.

7.6 Climate Change

The effects of climate change are relevant to many contaminated sites, where injuries and restoration often stretch over many decades. Additionally, a rapid pace of change could mean that climate change issues will increasingly be a consideration for oil spill recoveries and longevity of restoration.

A paper by Rohr et al. (2013) addressed many of the scientific issues associated with climate change in NRDA. However, this paper stops short of analyzing issues related to restoration scaling. Here, we briefly discuss this issue. The basic issues have all been touched on already. Climate change just magnifies their importance.

1) **Change in the resource base.** Climate change may alter the resource base. If injury occurs when the resource base is larger and restoration occurs when the base is diminished, the value of a unit of future restoration will increase. This would call for a lowering of the discount rate.
2) **Uncertain annual debits and credits**. Climate change may increase uncertainty related to the level of future baseline services, as well as credits likely to be created in the future. This may call for the explicit treatment of uncertainty via calculation of certainty equivalents, which adjust probability distributions associated with gains or losses with their equivalent certain value. Alternatively, the assessment may embody a risk premia in discount rates.
3) **Interactions between income (market goods) and resource services**. If the resource base is changing at a large scale, it may no longer be a reasonable approximation to assume that the marginal utility of income and the marginal utility of services are independent. This complicates the specification of the ecological rate of discount for use in HEA and REA.
4) **Landscape-level scaling.** If climate is changing rapidly, the *location* of the best restoration may be more remote from the injury than is currently preferred; this brings landscape-level and social factors into the EPF.
5) **Out-of-Kind Restoration**. The benefits of maintaining a close nexus between injured and restored *resources* may fall relative to the need to identify creative ways to augment *services*. This places additional emphasis on EPFs. Such an approach might make use of demand-side indicators of value, such as proximity of resources to populations or certain types of resource attributes or services that engender value.
6) **Heterogeneous public preferences**. If, in the future, restoration projects are more frequently far from impacted areas, the people affected by credits may differ from those affected by injuries. This strains the assumptions embedded in both REA and HEA. With a better understanding of how individuals with different demographic characteristics value different services, it may be possible and appropriate to change service weighting when different human populations are affected.
7) **Adaptive management**. The timing of optimal restoration expenditures may change. Due to uncertain fluctuations in the resource base, it may be best to time restoration to offset downturns in resource availability. This may call for holding some fraction of restoration funds and expending them at places and times when negative shocks to resources have occurred.
8) **Learning**. When there is increased uncertainty and rapid change, it may be valuable to invest in information. This may augment the benefits of adaptive management. The expected value of future information may create incentives to delay restoration to a degree.

Some of these implications are technical modifications that pertain to implementation of a REA or HEA. Others call for an expansion of HEA and REA ideas and point toward adjustments that look like hybrid methodologies that integrate service-to-service and value-to-value approaches.

At a more practical level, climate change considerations are most likely to manifest through temporal adjustments reflected in either baseline trends or longevity of restoration benefits. Sea level rise, climatic shifts, and increased variability of extreme weather are perhaps the most visible indications of climate change likely to affect restoration scaling using REA and HEA. Recovery times for many common oil spill injuries are of short enough duration that climate-related changes in baseline are expected to be nominal. This is not so, however, for natural resources with slower recovery times that take longer to return to baseline. Climate-related changes may impede the ability of an injured resource to recover compared to recovery rates expected under conditions prevailing at the time of injury.

Such factors may simultaneously be acting to change the baseline service levels, as well. It may become increasingly difficult to establish a causal link between effects of oil on resources *versus* other baseline effects induced by climate change. It is common in oil spill NRDAs to establish baseline using historical data. But these data may not be appropriate to more recent trends. While this circumstance often arises, it will increasingly be an issue if the pace of change accelerates and past time trends become less applicable to baseline determinations.

Climate change may also impact the duration of restoration benefits. For example, with HEA the scale of compensatory restoration is determined not only by how quickly injured resources recover (e.g., the size of the debit), but also by how quickly the benefits of restoration are realized (time to maturity) and how long such benefits are maintained into the future (lifespan). Consider coastal marshes, for example, which must vertically accrete by means of sediment deposition and biomass accumulation in the face of sea level rise or risk conversion to open water. Predicting the lifespan of a restored marsh will increasingly require considering not only possible rates of sea level rise, but also the degree to which such habitats can effectively keep pace with sea level rise and continue providing services. Moreover, to continue our hypothetical example, once the restored marsh is fully converted to shallow open water or mud flat, how do the services of that new habitat compare to those of the injured marsh?

In practice, the need to address difficult questions fraught with uncertainty is likely to drive Trustees towards adopting simplifying assumptions, such as short lifespans for certain types of restoration, that avoid the additional complications posed by more meaningful consideration of climate-related effects.

7.7 A Better Framework for Cooperative Assessment?

It has been 36 years since Mazzotta, Opaluch, and Grigalunas (1994) and Unsworth and Bishop (1994) revolutionized compensatory restoration scaling by establishing the basic underpinnings of service-to-service scaling. Since that time, the OPA NRDA community has used HEA and REA to estimate compensatory requirements arising from oil spills on numerous occasions.

This raises the question: Can the community use our common past experiences to simplify and expedite future HEA and REA assessments under OPA? We believe the answer is yes—if both responsible parties and trustees act in good faith.

If the goal is to achieve such an outcome, we would propose the following guidelines.

— In many instances, OPA NRDAs are initiated during the emergency response. At that time, efforts should be made to ensure coordination and cooperation between trustee and RP NRDA representatives. From the outset, conceptual models, assessment methods, study work plans, data collection protocols, and data quality objectives should be transparent to all participants, and data should be shared in a timely manner. Wherever possible, consensus assessment methods should be identified and data collection protocols should be cooperatively developed with the express intent of supporting those consensus assessment methods.

— Shortly after the emergency response is over, trustees and RPs should exchange pre-assessment documents that include a good faith preliminary estimate of damages. This pre-assessment should include a conceptual framework and preliminary restoration scaling models that include technical details and ranges of each scaling model parameter.

  - Wherever practical, a Monte Carlo analysis should be used to estimate a range of damages and the key drivers of uncertainty for each resource or habitat assessed. This provides a basis for trustees and RPs to identify key sources of uncertainty; it also addresses the hesitancy many participants may feel when asked to stipulate parameter values from want of knowing how they interact with other decisions yet to be made, which has impeded early agreements in past cases.

  - If a resource or exposure pathway is, at that time, simply too uncertain to address, the resource/pathway should be identified, the concern should be articulated, and potential paths forward should be outlined.

- Given 36 years of experience by NRDA practitioners conducting this work, good faith technical efforts should often lead to overlapping damage estimates and the NRDA can end. If either party finds the preliminary estimates unacceptable, the Monte Carlo and associated diagnostics will show which parameters or other choices contribute to overall uncertainty. Those with very low impact on uncertainty can be agreed to. Those with moderate impact on uncertainty can be carried forward as a source of modest point spread between the parties for future reassessment. Those with large effects on uncertainty should be considered for additional information gathering as a next step.

- Absent some overarching concern, supplemental studies and data collection efforts should be limited to those that will refine an underlying input or assumption; in doing so, they will reduce the magnitude of damage estimate divergence in a manner that justifies the cost.

- After the initial exchange and discussion of pre-assessment documents, a timeline to assessment completion and corresponding milestones should be established and followed.

8 References

Adler, M. 2019. "Measuring Social Welfare: An Introduction," Oxford University Press, 2019, *Duke Law School Public Law & Legal Theory Series No. 2020-8*. Available at SSRN: https://ssrn.com/abstract=3541745.

Adler, M. and M. Fleurbaey. 2016. *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, 2016. Available at SSRN: https://ssrn.com/abstract=2816371.

Akcakaya, H. 1998. "Assessing human impact despite uncertainty: viability of the northern spotted owl metapopulations in the Northwestern US," *Biodiversity and Conservation* 7: 875–894.

Baker, M., A. Domanski, T. Hollweg, J. Murray, D. Lane, K. Skrabis, R. Taylor, T. Moore, and L. DiPinto. 2020. "Restoration scaling approaches to addressing ecological injury: the habitat-based resource equivalency method," *Environmental Management* 65: 161–177.

Blanchard, O. and S. Fischer. 1993. *Lectures on Macroeconomics*. The MIT Press, Cambridge, MA.

Brown, G. 1993. "Economics of Natural Resource Damage Assessment: A Critique," In, R. Kopp and V.K. Smith, Eds, *Valuing Natural Assets: The Economics of Natural Resource Damage Assessment*. Resources for the Future, Washington, D.C.

Byrd, H. and T. Tomasi. 2021. "Damages to Recreation from an Oil Spill: Why Current Assessments Typically are Wrong and What can be Done to Fix Them," *Proceedings of the International Oil Spill Conference*. Vol. 2021 (1).

California Department of Fish and Game (CDFG), *Stuyvesant/Humboldt Coast Oil Spill. Draft Damage Assessment and Restoration Plan*. California State Lands Commission, and United States Fish and Wildlife Service (USFWS), 2004. Available at: http://www.dfg.ca.gov/ospr/organizational/scientific/nrda/stuyvesant%20DARP/darp%20index.htm.

Refugio Beach Oil Spill Final Damage Assessment and Restoration Plan/Environmental Assessment, California Department of Fish and Wildlife (CDFW), California State Lands Commission, California Department of Parks and Recreation, Regents of the University of California, U.S. Department of the Interior, U.S. Fish and Wildlife Service, and National Oceanic and Atmospheric Administration, 2021. Available at: https://wildlife.ca.gov/ospr/nrda/refugio.

Deaton, A., and J. Muellbauer, 1980. “An Almost Ideal Demand System,” *The American Economic Review* 70 (3): 312-326.

DeGroot, M. 1970. *Optimal Statistical Decisions*, Wiley and Sons Inc.

Desvousges, W., N. Gard, H. Michael, and A. Chance, 2018. “Habitat and resource equivalency analysis: a critical assessment,” *Ecological Economics* 143: 74–89.

Dunford, R. 2018. “The Discount Rate for Assessing Intragenerational Natural Resource Damages,” *Journal of Natural Resources Policy Research* 8 (1-2): 89–109.

Dunford, R., T. Ginn, and W. Desvousges. 2004. “The use of habitat equivalency analysis in natural resource damage assessments,” *Ecological Economics* 48 (1): 49–70.

English, E., C. Peterson, and C. Voss. 2009. *Ecology and Economics of Compensatory Restoration*, NOAA Coastal Response Research Center, New Hampshire.

ENTRIX. 2009. *Assessment of beached bird modeling methods*, unpublished report submitted to American Petroleum Institute. 78 pp.

Procedures for the Derivation of Equilibrium Partitioning Sediment Benchmarks (ESBs) for the Protection of Benthic Organisms: PAH Mixtures. EPA-600-R-02-013. U.S. Environmental Protection Agency, Office of Research and Development, 2003, Washington D.C. 175 pp. Available at: http://www.epa.gov/nheerl/publications/files/PAHESB.pdf.

Field L., D. Macdonald, S. Norton, C. Ingersoll, C. Severn, D. Smorong, and R. Lindskoog, 2002. “Predicting amphipod toxicity from sediment chemistry using logistic regression models,” *Environmental Toxicology and Chemistry* 21 (9): 1993–2005.

Field L., D. Macdonald, S. Norton, C. Severn, and C. Ingersoll, 1999. “Evaluating sediment chemistry and toxicity data using logistic regression modeling,” *Environmental Toxicology and Chemistry* 18 (6): 1311–1322.

Finley, H., J. Hanifen, C. Piehler, L. Pace, J. Kern, M. Devany, T. Osborn, C. Scannell, B. Julius, and E. Rienharz, 1995. *Damage Assessment and Restoration Plan and Environmental Assessment, Wellhead Failure and Release*, Dixon Bay Louisiana, January 12-13, 1995. Available at: http://www.darrp.noaa.gov/southeast/index.html.

Flores, N., and J. Thatcher, 2002. “Money, who needs it? Natural resource damage assessments,” *Contemporary Economic Policy* 20: 171–178.

Freeman III, A.M., 1993. *The Measurement of Environmental and Resource Values: Theory and Methods. Washington, D.C.: Resources for the Future*.

Gollier. C., 2010. “Debating about the Discount Rate:The Basic Economic Ingredients,” *Perspektiven der Wirtschaftspolitik* 11(1): 38-55.

Gollier, C. and J. Hammitt, 2014. “The Long-Run Discount Rate Controversy,” *Annual Review of Resource Economics* 6 (1): 273-29.

Gunnarsson, G., J. Elmberg, H. Pöysä, P. Nummi, K. Sjöberg, L. Dessborn, and C. Arzel, 2013. “Density dependence in ducks: a review of the evidence,” *European Journal of Wildlife Resources* 59: 305–321.

Haddad, R. and B. Israel, 2017. “Natural Resource Damage Assessments: Is Cooperation a Good Thing?” *International Oil Spill Conference Proceedings* 2017 (1): 3091–3111.

Hausman, D., 2012. *Preference, Value, Choice, and Welfare*. Cambridge University Press, New York, NY.

Hicks, John, 1939. *The Foundations of Welfare Economics*. The Economic Journal. 49 (196): 696–712.

Hoel, M., and T. Sterner, 2007. “Discounting and relative prices,” Climatic Change 84: 265–280. Available at: https://doi.org/10.1007/s10584-007-9255-2.

Israel, B., B. Marston, and L. Daniel, 2017. *Natural Resource Damages: A Guide to Litigating and Resolving NRD Cases*, American Bar Association Publishing, Washington, D.C.

Jones, C., and K. Pease, 1997. “Restoration based measures in natural resource liability statutes,” *Contemporary Economic Policy* 15 (4).

Kaldor, N., 1939. “Welfare Propositions in Economics and Interpersonal Comparisons of Utility,” *The Economic Journal* 49 (195): 549–552.

Leslie, P.H., 1945. “The use of matrices in certain population mathematics,” *Biometrika* 33(3): 183–212.

Long, E., and D. MacDonald, 1998. “Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems,” *Human and Ecological Risk Assessment* 4 (5): 1019–1039.

Long, E., C. Ingersoll, and D. Macdonald, 2006. “Calculation and uses of mean sediment quality guideline quotients: a critical review,” *Environmental Science and Technology* / 40 (6): 1726 – 1736.

MacNair, D., G. Parsons, T. Tomasi, and H. Byrd, 2022. “Trip Equivalency for Economic Valuation in Recreation Demand Models: Implications for Compensatory Restoration and Benefits Transfer.” *Marine Resource Economics* (37):1-17.

Mazzotta, M., J. Opaluch, and T. Grigalunas, 1993. Restoration and Natural Resource Damage Assessment. Proceedings of 1993 International Oil Spill Conference, Tampa Bay, Florida.

Melvin, S., and J. Gibbs, 1994. Viability analysis for the Atlantic Coast population of piping plovers. Unpublished report to the U.S. Fish and Wildlife Service, Sudbury, Massachusetts. 16 pp.

Millennium Ecosystem Assessment. 2005. *Ecosystems and Human Well-being: Wetlands and Water Synthesis.* World Resources Institute, Washington, D.C.

National Oceanic and Atmospheric Administration (NOAA), 1999. *Discounting and the Treatment of Uncertainty in Natural Resource Damage Assessment—Technical Paper 99-1*, Damage Assessment and Restoration Program.

NOAA, 2000. *Habitat Equivalency Analysis: An Overview. Damage Assessment and Restoration Program*, National Oceanic and Atmospheric Administration, Silver Spring, MD, 23 pp.

NOAA, Maryland Department of Natural Resources, Maryland Department of the Environment, and USFWS, 2002. *Final Restoration Plan and Environmental Assessment for the April 7, 2000 Oil Spill at Chalk Point on the Patuxent River, Maryland.*

NOAA, USFWS, New Jersey Department of Environmental Protection, Delaware Department of Natural Resources and Environmental Control, and Pennsylvania Department of Natural Resources, 2009. *Draft Damage Assessment and Restoration Plan and Environmental Assessment for the November 26, 2004 M/T/* Athos *I Oil Spill on the Delaware River near the CITGO Refinery in Paulsboro, New Jersey*. Available at: http://www.darrp.noaa.gov/northeast/athos/pdf/DraftAthosDARP_Final_2.pdf.

National Research Council, 2005. *Valuing Ecosystem Services: Toward Better Environmental Decision-Making,* The National Academies Press, Washington, D.C. Available at: https://doi.org/10.17226/11139.

New Jersey Department of Environmental Protection, USFWS, and NOAA, 2004. *Draft Natural Resources Restoration Plan and NEPA Environmental Assessment for the Anitra Oil Spill of May 1996*.

Parsons, G., and A. Kang, 2010. "Compensatory restoration in a random utility model of recreation demand," *Contemporary Economic Policy* 28(4): 453–463.

Penn, T., and T. Tomasi, 2002. "Calculating resource restoration for an oil discharge in Lake Barre, Louisiana, USA," *Environmental Management* 29: 691–702.

Portney, P.R., and J. Weyant, eds., 1999. *Discounting and Intergenerational Equity*, Resources for the Future, Washington, D.C.

Plissner, J., and S. Haig, 2000. "Viability of piping plover, *Charadrius melodus,* metapopulations," *Biological Conservation* 92:163–173.

Rohr, J., P. Johnson, C. Hickey, R. Helm, A. Fritz, and S. Brasfield, 2013. "Implications of global climate change for natural resource damage assessment, restoration and rehabilitation," *Environmental Toxicology and Chemistry* 32: 93–101.

Skrabis, K., 2004. *Supporting Analysis of Damages: Lost Bird Years from the Star Evivva Oil Spill*. Available at: http://www.dnr.sc.gov/marine/pub/StarEvvivaFinal.pdf.

Skrabis, K., 2005. *Resource Equivalency Analysis for loons New Carissa Oil Spill*, February 1999. Available at: http://www.fws.gov/oregonfwo/Contaminants/Spills/NewCarissa/RestorationPlan/Documents/NewCarissaDARPAppendix7Loons.pdf.

Sperduto, M., S. Powers, and M. Donlan, 2003. "Scaling restoration actions to achieve quantitative enhancement of loon, seaduck, and other seabird populations," *Marine Ecology Progress Series* 264: 221–232.

Sperduto, M., H. Charles, M. Donlan, and S. Thompson, 1999. *Injury Quantification and Restoration Scaling for Marine Birds Killed as a Result of the North Cape Oil Spill*, Technical Bird Working Group, North Cape Oil Spill. March 25, 1999.

Tomasi, T., 2021a. "Considering Environmental Justice in Natural Resource Damage Assessment: Injury to Recreation Resources," *Integral NREE White Paper 2021-1*. https://www.integral-corp.com/wp-content/uploads/2021/12/Considering-EJ-in-NRDA-Injury-to-Recreation-Resources.pdf.

Tomasi, T., 2021b. "Considering Environmental Justice in Natural Resource Damage Assessment: Injury to Ecological Resources," *Integral NREE White Paper 2021-1*. https://www.integral-corp.com/wp-content/uploads/2021/12/Considering-EJ-in-NRDA-Injury-to-Ecological-Resources.pdf.

USFWS, 2020. *Migratory Bird Permits; Management of Conflicts Associated With Double-Crested Cormorants (Phalacrocorax auritus) Throughout the United States*, Available at: https://www.federalregister.gov/documents/2020/06/05/2020-11988/migratory-bird-permits-management-of-conflicts-associated-with-double-crested-cormorants.

USFWS and the United States Department of Agriculture, 2003. *Final Environmental Impact Statement: Double Crested Cormorant Management in the United States*. Available at: http://library.fws.gov/Bird_Publications/CormorantFEIS.pdf.

USFWS, Nottawaseppi Huron Band of the Potawatomi Tribe, and Match-E-Be-Nash-She-Wish Band of the Pottawatomi Indians, 2015. *Final Damage Assessment and Restoration Plan/Environmental Assessment for the July 25-26, 2010 Enbridge Line 6B Oil Discharges near Marshall, MI*. Available at: https://www.fws.gov/midwest/es/ec/nrda/MichiganEnbridge/pdf/FinalDARP_EA_EnbridgeOct2015.pdf.

Unsworth, R., and R. Bishop, 1994. "Assessing natural resource damages using environmental annuities," *Ecological Economics* 11: 35–41.

Varian, H.R. 1992. *Microeconomic Analysis*. 3rd Edition, W. W. Norton & Company, New York.

Wakefield, J.R., and A. Davis, 2017. "Assessing natural resource damages for commercially and recreationally harvested populations," *International Oil Spill Conference Proceedings* (May) 2017(1): 959–984.

Annex A

Comprehensive Bibliography

Relevant Literature

[1] API Standard 520, *Sizing, Selection, and Installation of Pressure-relieving Devices, Part 1—Sizing and Selection*, 9th Edition

[2] Allen, P.D., II, D.J. Chapman, and D. Lane, 2005. "Scaling environmental restoration to offset injury using Habitat Equivalency Analysis," Chapter 8 in *Economics and ecological risk assessment. Application to Watershed Management,* ed. R.F. Bruins and M.T. Herberling. CRC Press, Baton Rouge, Louisiana.

[3] Aquatic Technical Working Group, 2007. *Final Report Aquatic Injury Assessment M/T Athos I Oil Spill, Delaware River System.* June. Available at: http://www.darrp.noaa.gov/northeast/athos/admin.html.

[4] Baker, M., A. Domanski, T. Hollweg, J. Murray, D. Lane, K. Skrabis, R. Taylor, T. Moore, and L. DiPinto. Restoration scaling approaches to addressing ecological injury: the habitat-based resource equivalency method. *Environmental Management* 65: 161–177.

[5] Barnthouse, L.W., and R.G. Stahl. 2002. Quantifying natural resource injuries and ecological service reductions: challenges and opportunities. *Environmental Management* 30(1): 1–12.

[6] Bird and Wildlife Technical Working Group. 2007. *Final Report Bird and Wildlife Injury Assessment M/T Athos I Oil Spill, Delaware River System.* June. Available at: http://www.darrp.noaa.gov/northeast/athos/admin.html.

[7] Bruggeman, D.J., M.L. Jones, F. Lupi, and K.T. Scribner, 2005. "Landscape equivalency analysis: Methodology for estimating spatially explicit biodiversity credits," *Environmental Management* 36(4): 518–534.

[8] Byrd, H., E. English, N. Meade, and T. Tomasi, 2002. "Scaling the Washington Creek Restoration Project to the Chalk Point Oil Spill Diamondback Terrapin Injury," Available at: http://www.darrp.noaa.gov/northeast/chalk_point/admin.html.

[9] Cacela, D., J. Lipton, D. Beltman, J. Hansen, and R. Wolotira, 2005. "Associating ecosystem service losses with indicators of toxicity in habitat equivalency analysis," *Environmental Management* 35(3): 343–351.

[10] California Department of Fish and Game and United States Fish and Wildlife Service, 2008. Kure/Humboldt Bay Oil Spill Final Damage Assessment and Restoration Plan/Environmental Assessment. Available at: http://www.dfg.ca.gov/ospr/spill/nrda/nrda_kure.html.

[11] California Department of Fish and Game, California State Lands Commission, and United States Fish and Wildlife Service, 2004. *Stuyvesant/Humboldt Coast Oil Spill. Draft Damage Assessment and Restoration Plan*. Available at: http://www.dfg.ca.gov/ospr/organizational/scientific/nrda/stuyvesant%20DARP/darp%20index.htm.

[12] California Department of Fish and Game, California State Lands Commission, National Oceanic and Atmospheric Administration, United States Fish and Wildlife Service, National Park Service, and Bureau of Land Management. 2012. *Cosco Busan Oil Spill Final Damage Assessment and*

Restoration Plan/Environmental Assessment. Available at: https://wildlife.ca.gov/OSPR/NRDA/cosco-busan.

[13] California Department of Fish and Game, California State Lands Commission, and United States Fish and Wildlife Service, 2007. *Stuyvesant/Humboldt Coast Oil Spill Final Damage Assessment and Restoration Plan/Environmental Assessment*. Available at: ,http://www.dfg.ca.gov/ospr/spill/nrda/nrda_stuy.html.

[14] California Department of Fish and Game, National Oceanic and Atmospheric Administration, United States Fish and Wildlife Service and National Park Service, 2006. *S.S. Jacob Luckenbach and Associated Mystery Oil Spills Final Damage Assessment and Restoration Plan/Environmental Assessment.* November 1, 2006. Available at: http://www.dfg.ca.gov/ospr/spill/nrda/nrda_luckenbach.html.

[15] Carle, M., K. Benson, and J. Reinhardt, 2020. “Quantifying the benefits of estuarine habitat restoration in the Gulf of Mexico: an introduction to the theme section,” *Estuaries and Coasts* 43: 1680–1691.

[16] Chapman, D., N. Iadanza, and T. Penn, 1998. *Calculating Resource Compensation: An Application Of The Service-To-Service Approach to the Blackbird Mine Hazardous Waste Site*. Technical Paper 97-1. National Oceanic and Atmospheric Administration, Damage Assessment and Restoration Program, Silver Spring, MD.

[17] Cole, S., 2010. “Environmental Compensation using Resource Equivalency Analysis (REA) and Habitat Equivalency Analysis (HEA): is it Just for the Birds?” Department of Forest Economics, Swedish University of Agricultural Sciences, Umeå, Sweden.

[18] Commencement Bay Natural Resource Trustees, 2002. *Hylebos Waterway Natural Resource Damage Settlement Proposal Report. Review Draft.* Available at: http://www.darrp.noaa.gov/pacific/cbay/admin.html.

[19] Committee on the Effects of the Deepwater Horizon Mississippi Canyon Mc-252 Oil Spill on Ecosystem Services in the Gulf of Mexico, 2012. *Approaches for Ecosystem Services Valuation of the Gulf of Mexico after the Deepwater Horizon Oil Spill*. The National Academies Press, Washington D.C.

[20] Desvousgesa, W., N. Gard, H. Michael, and A. Chance, 2018. “Habitat and resource equivalency analysis: a critical assessment,” *Ecological Economics* 143: 74–89.

[21] Donlan, M., M. Sperduto, and C. Hebert, 2003. “Compensatory mitigation for injury to a threatened or endangered species: Scaling piping plover restoration,” *Marine Ecology Progress Series* 264: 213–219.

[22] Dunford, R.W., T.C. Ginn, and W.H. Desvousges, 2004. “The use of habitat equivalency analysis in natural resource damage assessments,” *Ecological Economics* 48: 49–70.

[23] Deepwater Horizon Trustees, 2016. *Deepwater Horizon Oil Spill: Final Programmatic Damage Assessment and Restoration Plan and Final Programmatic Environmental Impact Statement.* Available at: http://www.gulfspillrestoration.noaa.gov/restoration-planning/gulf-plan.

[24] Efroymson, R.A., J.P. Nicolette, G.W. Sutter II, 2003. *A Framework for Net Environmental Benefit Analysis for Remediation or* Restoration *of Petroleum-contaminated Sites.* ORNL/TM-2003/17. Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

[25] ENTRIX, 2009. *Assessment of beached bird modeling methods*, unpublished report submitted to American Petroleum Institute. 78 pp.

[26] Environmental Protection Agency, 2003. *Procedures for the Derivation of Equilibrium Partitioning Sediment Benchmarks (ESBs) for the Protection of Benthic Organisms: PAH Mixtures.* EPA-600-R-02-013. U.S. Environmental Protection Agency. Office of Research and Development. Washington D.C. 175 pg. Available at: http://www.epa.gov/nheerl/publications/files/PAHESB.pdf.

[27] Finch, B., 2019. "Phototoxicity Assessments of Field Sites in Barataria Bay, Louisiana, USA, and Heavily Weathered Macondo Crude Oil: 4 Years after the Deepwater Horizon Oil Spill," *Environ Toxicol Chem.* 38(8):1811-1819.

[28] Findlay, R., and M.J. Kealy, 2003. "Developing an ecological currency approach for identifying resource improvements," *Proceedings of Canadian Coastal Conference 2003.* Queen's University, Kingston, Ontario, Canada.

[29] Finley, H.W., J. Hanifen, C. Piehler, L. Pace, J. Kern, M. Devany, T. Osborn, C. Scannell, B. Julius, and E. Rienharz, 1995. *Damage Assessment and Restoration Plan and Environmental Assessment, Wellhead Failure and Release, Dixon Bay Louisiana, January 12-13, 1995*. Available at: http://www.darrp.noaa.gov/southeast/index.html.

[30] Flores, N.E., and J. Thacher, 2002. "Money, who needs it? Natural resource damage assessment," *Contemporary Economic Policy* 20: 171–178.

[31] Florida Department of Environmental Protection, National Oceanic and Atmospheric Administration, and Unites States Department of the Interior, 1997. *Damage Assessment and Restoration Plan/Environmental Assessment for the August 10, 1993 Tampa Bay Oil Spill. Volume I - Ecological Injuries*. Available at: http://www.darrp.noaa.gov/southeast/tampabay/admin.html.

[32] Fonseca, M.S., 1997. *Report for United States v. Fisher, et. al. Case No. 92-10027-CIVIL-DAVIS*, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, Beaufort Laboratory, Beaufort, North Carolina. Available at: http://www.darrp.noaa.gov/library/pdf/melfish3.pdf.

[33] Fonseca, M.S., B.E. Julius, and W.J. Kenworthy, 2000. "Integrating biology and economics in seagrass restoration: How much is enough and why?" *Ecological Engineering* 15: 227–237.

[34] Fonseca, M.S., W.J. Kenworthy, and G.W. Thayer, 1998. *Guidelines for the Conservation and Restoration of Seagrasses in the Unites States and Adjacent Waters, Appendix E*, National Oceanic and Atmospheric Administration, Coastal Ocean Office. Silver Spring, MD. Available at: http://shrimp.bea.nmfs.gov/library/digital.html.

[35] Ford, R.G., 2002. *Estimated Common Murre and Marbled Murrelet mortality resulting from the* T/V Command *Spill, Final Report*. Prepared for the Command Trustee Council. Available at: http://www.darrp.noaa.gov/southwest/command/pdf/cmd-r11.pdf.

[36] Ford, R.G., G.K. Himes Boor, and J.C. Ward, 2000. *Seabird Mortality Resulting from the M/V New Carissa Oil Spill Incident, February and March 1999.* Draft report prepared for U.S. Fish and Wildlife Service, Oregon Fish and Wildlife Office.

[37] Ford, R.G., G.K. Himes Boor, B.E. Sharp, and J.L. Casey, 2002. *Estimates of Bird Impacts Resulting for the M/V Kure/Humboldt Bay Oil Spill of November 5, 1997.* Final draft report prepared for California Department of Fish and Game, Office of Spill Prevention and Response.

[38] French McCay, D., and J. Jennings, 2002. *Final Report, Chalk Point Oil Spill of April 7, 2000 in Patuxent River, MD: Modeling of the Fates and Acute Biological Effects of the Spilled Oil on the Water Column.* Prepared for Research Planning Inc for Submission to NOAA Damage Assessment Center. ASA 00-148. April 2002. Available at: ,http://www.darrp.noaa.gov/northeast/chalk_point/admin.html.

[39] French, D., M. Reed, K. Joyko, S. Feng, H. Rines, S. Pavignano, T. Isaji, S. Puckett, A. Keller, F. French III, D. Gifford, J. McCue, G. Brown, E. MacDonald, J. Quirk, S. Natzke, R. Bishop, M. Welsh, M. Phillips, and B.S. Ingram. 1996. *The CERCLA Type A Natural Resource Damage Assessment Model for Coastal and Marine Environments (NRDAM/CME), Technical Documentation, Vol. I-Model Description.* Submitted to the Office of Environmental Policy and Compliance, U.S. Department of the Interior, Washington, D.C. April 1996, Contract No. 14-0001-91-C-11.

[40] Hampton, S., and M. Zafonte, 2005. "Calculating compensatory restoration in natural resource damage assessments: Recent experience in California," *Proceedings of the 2002 California World Oceans Conference*, Santa Barbara, California. American Society of Civil Engineers, Reston, Virginia. Available at: http://www.dfg.ca.gov/ospr/ organizational/scientific/nrda/CWO-REA-paper.pdf.

[41] Harrington, K., and T. Feather, 1996. *Evaluation of Environmental Investments Procedures: Overview Manual.* IWR Report 96-R-30. U.S. Army Corps of Engineers. Alexandria, Virginia. Available at: http://www.iwr.usace.army.mil/iwr/pdf/96r30.pdf.

[42] Industrial Economics Inc., 2004. *Final Lake Apopka Natural Resource Damage Assessment and Restoration Plan.* Report for U.S. Fish and Wildlife Service, Atlanta, Georgia. Available at: http://sjr.state.fl.us/programs/acq_restoration/s_water/lapopka/ pdfs/DARP.pdf.

[43] Jones, C.A., and K.A. Pease, 1997. "Restoration-based compensation measures in natural resource liability statutes," *Contemporary Economic Policy* 15: 11–122.

[44] Julius, B., 1997. *Report for U.S. v. Melvin A. Fisher et al.* National Oceanic and Atmospheric Administration Damage Assessment Center. Silver Spring, MD. Available at: http://www.darp.noaa.gov/library/pdf/melfish2.pdf.

[45] Julius, B., and T. Osborn, 1992. *Wetlands Compensation Analysis: Greenhill Oil Spill, Timablier Bay, Louisiana, September 29, 1992.* NOAA Damage Assessment Center, Silver Spring, MD.

[46] Julius, B.E., J.W. Iliff, C.M. Wahle, J.H. Hudson, and E.C. Zobrist, 1995. *Natural Resource Management Damage Assessment: M/V Miss Beholden Grounding Site, Western Sambo Reef, FKNMS, March 13, 1993*. NOAA, Damage Assessment Center, Silver Spring, MD.

[47] Kennedy, C., and S. Cheong, 2013. "Lost ecosystem services as a measure of oil spill damages: A conceptual analysis of the importance of baselines," *Journal of Environmental Management* 128: 43–51.

[48] King, D.M., 1997. "Comparing ecosystem services and values with illustrations for performing Habitat Equivalency Analysis," Report to the National Oceanic and Atmospheric Administration, Damage Assessment and Restoration Program, Silver Spring, MD. Available at: http://www.darrp.noaa.gov/library/pdf/kingpape.pdf.

[49] Kohler, K.E., and R.E. Dodge, 2006. "Visual HEA: Habitat equivalency analysis software to calculate compensatory restoration following natural resource injury," *Proceedings of the 10th International Coral Reef Symposium.* June–July. Okinawa, Japan.

[50] Lipton, J., D. Chapman, G. Koonce, and F. Rahel, 2004. "Damage Calculation for Aquatic Resources: Coeur d'Alene Basin Natural Resource Damage Assessment," Expert Report. *United States of America v. ASARCO Inc. et al.*, No CV96-01122-N-EJL. Prepared by Stratus Consulting Inc., Boulder, Colorado, for the U.S. Department of the Interior, Bureau of Land Management, U.S. Department of Agriculture, Forest Service, and Coeur d'Alene Tribe. August.

[51] Lipton, J., E. Özdemiroğlu, D. Chapman, and J. Peers, 2018. *Equivalency Methods for Environmental* Liability *Assessing Damage and Compensation Under the European* Environmental *Liability Directive.* Springer, Dordrecht.

[52] Lofton, J. and C. Zander, 1997. Report for *United States v. Melvin A Fisher et al.* United States District Court for the Southern District of Florida, Key West Division.

[53] Long, E.R., and D.D. MacDonald, 1998. "Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems," *Human and Ecological Risk Assessment* 4(5): 1019–1039.

[54] Louisiana Oil Spill Coordinator's Office, Louisiana Department of Environmental Quality, Louisiana Department of Natural Resources, Louisiana Department of Wildlife and Fisheries, National Oceanic and Atmospheric Administration, and United States Fish and Wildlife Service, 2000. *Final Damage* Assessment*/Restoration Plan and Environmental Assessment.* M/V Westchester *Crude Oil Discharge Lower Mississippi River, Louisiana.*

[55] Louisiana Oil Spill Coordinator's Office, Louisiana Department of Environmental Quality, Louisiana Department of Natural Resources, Louisiana Department of Wildlife and Fisheries, National Oceanic and Atmospheric Administration, and United States Fish and Wildlife Service, 2005. *Final Damage Assessment And Restoration Plan/Environmental Assessment for the September 22, 2002, Oil Spill at North Pass in the Mississippi River Delta, Plaquemines Parish, Louisiana.* May. Available at: http://www.darrp.noaa.gov/southeast/northpass/admin.html.

[56] Louisiana Oil Spill Coordinators Office, Louisiana Department of Environmental Quality, Louisiana Department of Natural Resources, Louisiana Department of Wildlife and Fisheries, National Oceanic and Atmospheric Administration, and United States Fish and Wildlife Service, 1999. *Damage Assessment and Restoration Plan Texaco Pipeline Inc. Crude Oil Discharge Lake Barre, Louisiana May 16, 1997*. Final version August 27, 1999. Available at: http://www.darrp.noaa.gov/southeast/lakebarre/admin.html.

[57] Louisiana Oil Spill Coordinators Office, Louisiana Department of Environmental Quality, Louisiana Department of Natural Resources, Louisiana Department of Wildlife and Fisheries, and National Oceanic and Atmospheric Administration, and United States Fish and Wildlife Service, 2005. *Final Damage Assessment and Restoration Plan and Environmental Assessment, Equinox Oil Company Crude Oil Discharge, Lake Grande Ecaille, Louisiana, September 22, 1998*. November 20. Available at: http://www.darrp.noaa.gov/southeast/index.html.

[58] Luxon, M., 2016. "Use of Habitat Equivalency Analysis as a tool for scaling ecosystem service losses and gains in Natural Resources Damages Assessments and beyond," Salish Sea Ecosystem Conference, Vancouver, British Columbia.

[59] Mazzota, M.J., J.J. Opaluch, and T.A. Grigalunas, 1994. "Natural resource damage assessment: The role of resource restoration," *Natural Resources Journal* 24: 41–59.

[60] McCay, D.P.F., and J.J. Rowe, 2003. "Habitat restoration as mitigation for lost production at multiple trophic levels," *Marine Ecology Progress Series* 264: 233–247.

[61] McCay, D.P.F., M. Gibson, and J.S. Cobb, 2003. "Scaling restoration of American lobsters: Combined demographic and discounting model for an exploited species," *Marine Ecology Progress Series* 264: 177–196.

[62] McCay, D.P.F., C.H. Peterson, J.T. DeAlteris, and J. Catena, 2003. "Restoration that targets function as opposed to structure: Replacing lost bivalve production and filtration," *Marine Ecology Progress Series* 264: 197–212.

[63] McCay, D.P.F., J.J. Rowe, N. Whittier, S. Sankaranarayanan, and D.S. Etkin, 2004. "Estimation of potential impacts and natural resource damages of oil," *Journal of Hazardous Materials* 107: 11–25.

[64] Michel, J. J. Hoff, K. Smith, M. Keiler A. Rizzo, and R. Ayella, 2002. *Injury to Wetlands Resulting From the Chalk Point Oil Spill*. Available at: http://www.darrp.noaa.gov/northeast/chalk_point/admin.html.

[65] Milon, J.W., and R.E. Dodge, 2001. "Applying habitat equivalency analysis for coral reef damage assessment and restoration," *Bulletin of Marine Science* 69 (2):975–988.

[66] National Oceanic and Atmospheric Administration (NOAA), 1995. *Habitat Equivalency Analysis: How Much Restoration is Enough?* Damage Assessment and Restoration Program, Silver Spring, MD. Available at: http://www.darrp.noaa.gove/library/pdf/heagenl.pdf.

[67] NOAA, 1997. *Natural Damage* Assessment *Guidance Document: Scaling Compensatory Restoration Actions.* Damage Assessment and Restoration Program, Silver Spring, MD. Available at: http://www.darrp.noaa.gov/library/pdf/scaling.pdf.

[68] NOAA, 1999. *Discounting and the Treatment of Uncertainty in Natural Resource Damage Assessment: Technical Paper 99-1.* Silver Spring, MD. Available at: ,http://www.darp.noaa.gov/library/pdf/doscpdf2.pdf.

[69] NOAA, 2000. *Habitat Equivalency Analysis: An Overview*. Damage Assessment and Restoration Program, Silver Spring, MD, 23 pp.

[70] NOAA, 2004. *Habitat Equivalency Analysis*. Available at: http://www.csc.noaa.gov/coastal/economics/habitatequ.htm.

[71] NOAA, 2006. *Habitat Equivalency Analysis: An Overview.* Damage Assessment and Restoration Program, Silver Spring, MD. Available at: http://www.darp.noaa.gov/library/pdf/heaoverv.pdf.

[72] NOAA, 2016. *Final Damage Assessment and Restoration Plan/Environmental Assessment for the Tank Barge DBL 152 Oil Spill.* Available at: https://casedocuments.darrp.noaa.gov/southeast/dbl152/pdf/Final_DARP_and_EA_for_DBL_152_Oil_Spill_w_Appendices_07_01_2016.pdf.

[73] NOAA, Maryland Department of Natural Resources, Maryland Department of the Environment, and U.S. Fish and Wildlife Service, 2002. *Final Restoration Plan and Environmental Assessment for the April 7, 2000 Oil Spill at Chalk Point on the Patuxent River, Maryland.*

[74] NOAA, Rhode Island Department of Environmental Management, U.S. Department of the Interior and U.S. Fish and Wildlife Service, 1999. *Restoration Plan and Environmental Assessment for the January 19, 1996 North Cape Oil Spill Revised Draft for Public Comment.*

[75] NOAA, U.S. Fish and Wildlife Service, Massachusetts Executive Office of Energy and Environmental Affairs, and Rhode Island Department of Environmental Management, 2014. *Final Programmatic Restoration Plan and Environmental Assessment for the Buzzards Bay Bouchard Barge-120 (B-120) Oil Spill Shoreline, Aquatic and Natural Resource Use Injuries, Massachusetts and Rhode Island.* Available at: https://casedocuments.darrp.noaa.gov/northeast/buzzard/pdf/B-120-Final-PRP-EA-and-FONSI-09-30-14.pdf.

[76] NOAA, U.S. Fish and Wildlife Service, New Jersey Department of Environmental Protection, Delaware Department of Natural Resources and Environmental Control, Pennsylvania Department of Natural Resources, 2009. *Final Damage Assessment and Restoration Plan and Environmental Assessment for the November 26, 2004* M/T/ Athos *I Oil Spill on the Delaware*

River Near the CITGO Refinery in Paulsboro, New Jersey. Available at: https://pub-data.diver.orr.noaa.gov/admin-record/6404/Athos_Final_RP.pdf.

[77] New Jersey Department of Environmental Protection, United States Fish and Wildlife Service, and NOAA, 2004. *Draft Natural Resources Restoration Plan and NEPA Environmental Assessment for the* Anitra *Oil Spill of May 1996.*

[78] Operational Science Advisory Team (OSAT 1), 2010. *Summary Report for Sub-sea and Sub-surface Oil and Dispersant Detection: Sampling and Monitoring*, Unified Area Command, New Orleans. Available at: http://www.dep.state.fl.us/deepwaterhorizon/files2/osat_report_17dec.pdf.

[79] Penn, T., and T. Tomasi, 2002. "Calculating resource restoration for an oil discharge in Lake Barre, Louisiana, USA," *Environmental Management* 29(5): 421–433.

[80] Peterson, C.H., M. Wong, M.F. Piehler, J.H. Grabowski, R.R. Twouldey, and M.S. Fonseca, 2007. *Estuarine Habitat Productivity Ratios at Multiple Trophic Levels.* Final Report to NOAA Office of Response and Restoration, Assessment and Restoration Division, Silver Spring, MD. 62 pp.

[81] Roach, B., and W.W. Wade, 2006. "Policy evaluation of natural resource injuries using habitat equivalency analysis," *Ecological Economics* 58: 421–433.

[82] Robilliard et al., 1997. "Ephemeral Data Collection Guidance Manual, with Emphasis on Oil Spill NRDAs," Proceedings of the 1997 International Oil Spill Conference. Fort Lauderdale, Florida.

[83] Scemama, P., and H. Levrel, 2015. "Using habitat equivalency analysis to assess the cost effectiveness of restoration outcomes in four institutional contexts," *Environmental Management* 57: 109–122.

[84] Shaw, W., and M. Wlodarz, 2012. "Ecosystems, ecological restoration, and economics: does habitat or resource equivalency analysis mean other economic valuation methods are not needed?" *AMBIO* 42: 628–643.

[85] Shellenbarger Jones, A., and M. Donlan, 2008. *Athos I Oil Spill Restoration Scaling Paper for Injuries to Non-Tributary Shorelines: Salt Marsh Restoration at Mad Horse Creek, New Jersey and Habitat Creation and Restoration at Lardner's Point, Pennsylvania.* Available at: http://www.darrp.noaa.gov/northeast/athos/admin.html.

[86] Shellenbarger Jones, A., and M. Donlan. 2008. *Athos I Oil Spill Restoration Scaling Paper for Tributary Injuries: Dam Removal and Riparian/In-stream Habitat Restoration on Darby Creek and Habitat Restoration at John Heinz National Wildlife Restoration*. Available at: http://www.darrp.noaa.gov/northeast/athos/admin.html.

[87] Shoreline Assessment Team, 2007. *Final Report Shoreline Injury Assessment M/T Athos I Oil Spill*. Available at: http://www.darrp.noaa.gov/northeast/athos/admin.html.

[88] Skrabis, K., 2004. *Supporting Analysis of Damages: Lost Bird Years from the Star Evivva Oil Spill*. Available at: http://www.dnr.sc.gov/marine/pub/StarEvvivaFinal.pdf.

[89] Skrabis, K., 2005. Resource Equivalency Analysis for Loons: *New Carissa Oil Spill, February 1999*. Available at: http://www.fws.gov/oregonfwo/Contaminants/Spills/NewCarissa/RestorationPlan/Documents/NewCarissaDARPAppendix7Loons.pdf.

[90] Snyder, J., and W. Desvousges, 2013. "Habitat and resource equivalency analyses in resource compensation and restoration in decision making," *Resources & Environment* 3 (2013–2014).

[91] Sperduto, M.B., S.P. Powers, and M. Donlan, 2003. "Scaling restoration to achieve quantitative enhancement of loon, seaduck, and other seabird populations," *Marine Ecology Progress Series* 264: 221–232.

[92] Sperduto, M., H. Charles, M. Donlan, and S. Thompson, 1999. *Injury Quantification and Restoration Scaling for Marine Birds Killed as a Result of the North Cape Oil Spill.*

[93] State and Federal Trustees, State of Montana, and U.S. Department of the Interior, 2017. *Final Programmatic Damage Assessment and Restoration Plan and Final Programmatic Environmental Assessment for the ExxonMobil Pipeline Company July 1, 2011 Yellowstone River Oil Spill.* Available at: https://dojmt.gov/wp-content/uploads/nrdp-yellowstone-restoration.pdf.

[94] Strange, E., H. Galbraith, S. Bickel, D. Mills, D. Beltman, and J. Lipton, 2002. "Determining ecological equivalence in service-to-service scaling of salt marsh restoration," *Environmental Management* 29(2): 290–300.

[95] Strange, E., D. Allen, D. Mills, and P. Paimondi, 2004. *Research on Estimating the Environmental Benefits of Restoration to Mitigate or Avoid Environmental Impacts Caused by California Power Plant Cooling Water Intake Structures.* California Energy Commission, PIER Energy-Related Environmental Research. Sacramento, California.

[96] Tetra Tech, Inc., 2006. *Draft Restoration Plan/Environmental Assessment for Morris J. Berman Oil Spill San Juan, Puerto Rico*. Prepared for National Oceanic and Atmospheric Administration, Puerto Rico Department of Natural and Environmental Resources, National Park Service. April 2006. Available at: http://www.darrp.noaa.gov/southeast/berman/admin.html.

[97] Tomasi, T., 2021a. "Considering Environmental Justice in Natural Resource Damage Assessment: Injury to Recreation Resources." Integral NREE White Paper 2021-1. Available at https://www.integral-corp.com/wp-content/uploads/2021/12/Considering-EJ-in-NRDA-Injury-to-Recreation-Resources.pdf.

[98] Tomasi, T., 2021b. "Considering Environmental Justice in Natural Resource Damage Assessment: Injury to Ecological Resources." Integral NREE White Paper 2021-2. Available at https://www.integral-corp.com/wp-content/uploads/2021/12/Considering-EJ-in-NRDA-Injury-to-Ecological-Resources.pdf.

[99] Tomasi, T., J. Wakefield, and H. Byrd, 2003. "Habitat Equivalency Analysis: Implications for Endangered Species," Proceedings of the 20"03 International Oil Spill Conference. Vancouver, British Columbia.

[100] Thur, S.M., 2007. "Refining the use of habitat equivalency analysis," *Environmental Management* 40: 161–170.

[101] U.S. Department of Commerce and NOAA, 1996. 15 CFR Part 990, *Natural Resource Damage Assessments*. January 5, 1996. 61 FR 440.

[102] U.S. Department of Interior, U.S. Fish and Wildlife Service (USFWS), 2020. *Migratory Bird Permits; Management of Conflicts Associated With Double-Crested Cormorants (Phalacrocorax auritus) Throughout the United States*. Available at: https://www.federalregister.gov/documents/2020/06/05/2020-11988/migratory-bird-permits-management-of-conflicts-associated-with-double-crested-cormorants.

[103] United States Department of the Interior, Bureau of Land Management, United States Fish and Wildlife Service, United States Department of Agriculture, United States Forest Service, State of Oregon, Department of Fish and Wildlife, Department of Environmental Quality, and Confederated Tribes of Siletz Indians of Oregon, Confederated Tribes of the Coos, Lower Umpqua and Siuslaw

Indians. 2006. *Final Damage Assessment and Restoration Plan & Environmental Assessment for the* M/V New Carissa *Oil Spill Oregon Coast.* January, 2006. Available at: http://www.fws.gov/oregonfwo/Contaminants/Spills/NewCarissa/RestorationPlan/default.asp.

[104] U.S. Fish and Wildlife Service, Nottawaseppi Huron Band of the Potawatomi Tribe, and Match-E-Be-Nash-She-Wish Band of the Pottawatomi Indians, 2015. *Final Damage Assessment and* Restoration *Plan/Environmental Assessment for the July 25-26, 2010 Enbridge Line 6B Oil Discharges near Marshall, MI.* Available at: https://www.fws.gov/midwest/es/ec/nrda/MichiganEnbridge/pdf/FinalDARP_EA_EnbridgeOct2015.pdf.

[105] Unsworth, R.E., and R. Bishop, 1994. "Assessing natural resource damages using environmental annuities," *Ecological Economics* 11: 35–41.

[106] Viehman, S., S. Thur, and G. Piniak, 2009. "Coral reef metrics and habitat equivalency analysis," *Ocean and Coastal Management* 52: 181–188.

[107] Wakefield, J.R., and A. Davis, 2017. "Assessing Natural Resource Damages For Commercially and Recreationally Harvested Populations," International Oil Spill Conference Proceedings 2017(1): 959–984.

[108] Wakefield, J., and McNutt, 2008. "An Ecological Framework for Resource Equivalency Analysis (REA)," Proceedings of the 2008 International Oil Spill Conference. Savannah, Georgia.

[109] Wakefield, J., T. Tomasi, and H. Byrd, 2003. "Predicting Spill-Induced Bird Mortality Using Beached Carcass Data," Proceedings of the 2003 International Oil Spill Conference. Vancouver, British Columbia.

[110] Wellman, E., A. Sutton-Grierab, M. Imholt, and A. Domanski, 2017. "Catching a wave? A case study on incorporating storm protection benefits into Habitat Equivalency Analysis," *Marine Policy*83: 118–125.

[111] Zafonte, M., and S. Hampton, 2007. "Exploring the welfare implications of resource equivalency analysis in natural resource damage assessments," *Ecological Economics* 61: 134–145.

[112] Zafonte, M., and S. Hampton. 2005. *Lost bird-years: Quantifying injuries from acute mortality events*. Proceedings of the 2005 International Oil Spill Conference. Miami, Florida. Available at: http://www.dfg.ca.gov/ospr/organizational/scientific/nrda/ IOSC_lost_bird-years_paper_FINAL.pdf.

[113] Zieman, J.C., 1997. *Report for United States v. Melvin A. Fisher et. al. Case No. 92-10027-CIVIL-DAVIS*. Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia. Available at: http://www.darrp.noaa.gov/library/pdf/melfish1.pdf.

Annex B

Quick Reference Summary of HEA and REA Models

REA and HEA are models designed to help identify the level of compensatory restoration that would ensure the public experiences no net loss of well-being due to a spill-related interim loss of ecological services. Both models are based on the economic theory of public compensation, but they embody several key assumptions that must exist if the results they generate are to be valid.

REA addresses services all tied to a single population. Conceptually, REA is a three-step process: (1) The baseline population level is projected through time; (2) the population level given the spill and a restoration project is projected through time; and (3) the size of the restoration project is adjusted until society experiences no net loss of discounted species years (Figure B-1).

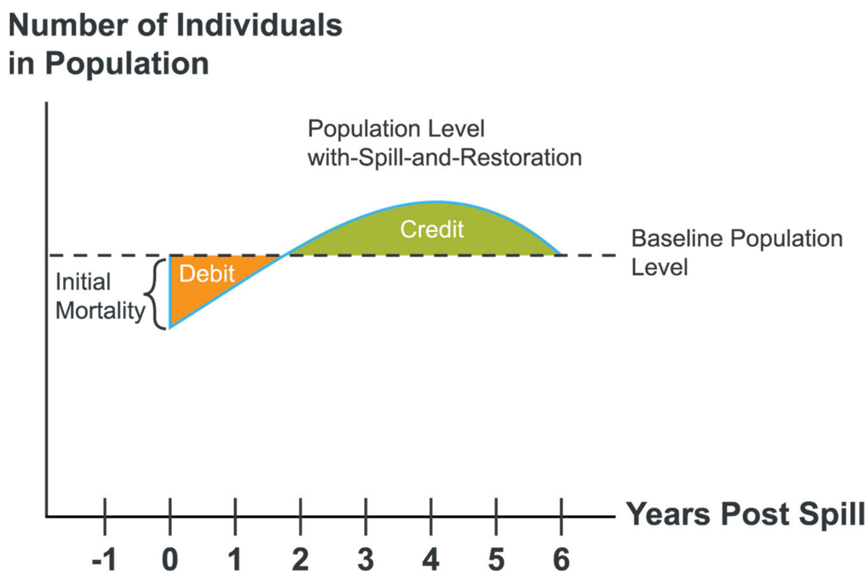


Figure B-1—Graphic Representation of REA

There are three broad approaches to REA. To be valid, all require that spill-related changes in population levels are "small" and that baseline population levels through time be relatively constant. The professional judgment approach may be useful if impacted populations expand rapidly in response to favorable conditions. Under all other scenarios, population modeling is preferred as the static approach is biased.

Once a REA approach has been selected, practitioners must remain cognizant of issues related to the calibration of demographic parameters, the implications of REA restoration projects that provide services, and increasing the target population. However, because actual NRDA liability is estimated as the per-unit cost of restoration multiplied by the compensatory requirement as estimated by REA, RPs may find it cost-effective to accept REA assumptions in a cooperative assessment even if they judge them to be unsound.

HEA addresses all services flowing from a habitat simultaneously via the use of a proxy or composite service. In addition to the requirements identified for REA, HEA often requires that all services flowing from a habitat be combined into a composite service; this requirement is challenging to implement in a technically defensible manner.

Conceptually, HEA is a three-step process: (1) The regional baseline composite service level is projected through time; (2) the composite service level given the spill and a restoration project is projected through time; and (3) the size of the restoration project is adjusted until society experiences no net loss of the composite or proxy service (Figure B-2).

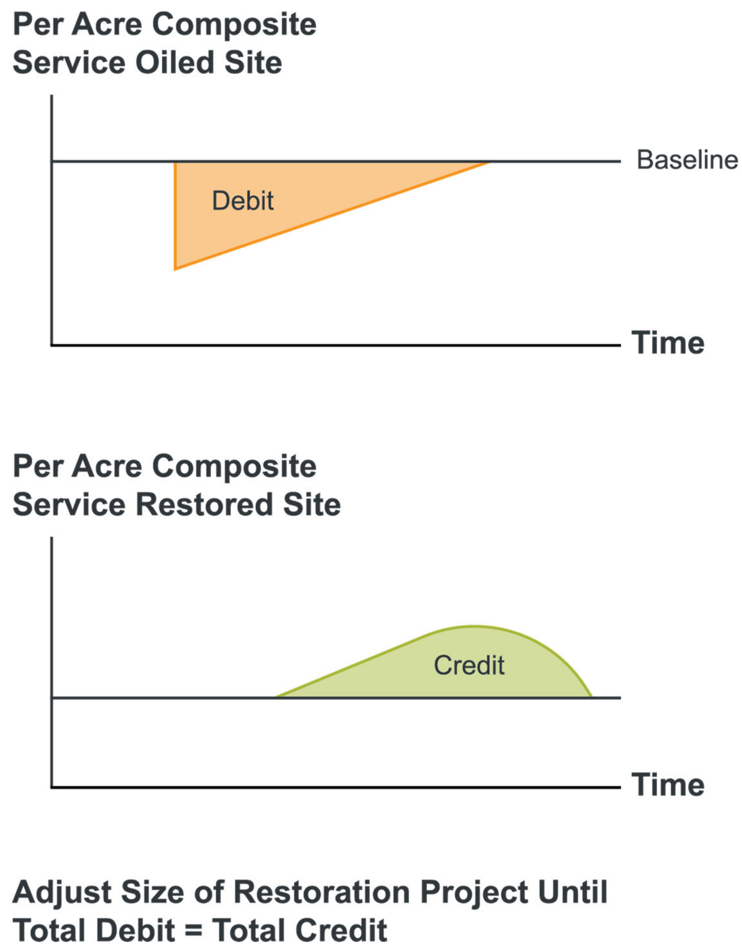


Figure B-2—Graphic Representation of HEA

In HEA, variations in the composite service level are expressed relative to a base (i.e., as a percent of the composite service provided by the base over one year). It is important that this base be common to both the debit and credit calculations. Because variation in a service level is difficult to measure, the service changes embedded in most HEAs are based on professional judgment. Given this fact, it can be difficult for a RP to significantly influence the estimation of these parameters using fact or science. However, the collection of field data and the requirement that assumptions be consistent with one another can often be used to identify the boundaries of a reasonable analysis. In addition, it can be helpful to have explicit discussions regarding all the services that flow from a habitat and the fact that those various services are weighted (either implicitly or explicitly) as they enter the composite service.

Once restoration is scaled, NRD liability is calculated by multiplying the per-unit cost of restoration by the compensatory requirement. As such, identification of cost-effective restoration is critical.

Selecting a model: There are no hard rules that allow a practitioner to identify the “best” model. However, by combining an understanding of REA and HEA with the spill facts, informed decisions are possible. The following exercise provides some guidance.

1) Make a list of all impacted resources. Usually, the list will include the following:
   - habitats (e.g., sandy shoreline, rocky shoreline, seagrass beds, coral beds, wetlands, water column, subtidal sediments, intertidal sediments);
   - specific species (the list could be long and varied, but will usually include several species of birds, and may include sea turtles, fish, and marine mammals); and
   - things people do (e.g., hunting, fishing, hiking, boating, going to the beach).
2) Impacts to “things people do” tend to be recreational issues that are addressed separately. Remove them from the list, but do not forget possible overlap with impacts to species.
3) Assign all of the species to an injured habitat. If any species is not easily assigned to a habitat, it is a candidate for REA.
4) Focusing on any one habitat, list the services provided by the habitat and the species that reside therein.
5) Were all services and inhabitants injured (and likely to be restored) approximately in a fixed ratio? If so, address via HEA. If not, proceed to the next question.
6) Is it cost-effective to conduct a service-weighted HEA that treats all services flowing from the habitat as a single composite service?[80] If so, address via HEA. If not, proceed to the next question.
7) Is it cost-effective to create a sufficient quantity of habitat such that even the most severely impacted service will be compensated? If so, use the most severely impacted service as a proxy for all services flowing from the habitat and address all services in a HEA. If not, proceed to Step 8).
8) If specific populations are removed, is it possible to address those populations using REA and then address the remaining services and inhabitants using HEA? If yes, do so. If not, proceed to Step 9).
9) Consider methods other than HEA and REA based on more explicitly economic approaches. The alternative method could include a modified version of REA, or full economic methods.
10) After deciding which services can be addressed by REA or HEA, consider how the different injury assessments and restoration projects interact. Are injured services being omitted or double counted? Are restoration credits being omitted or double counted? How can the assessment strategy be adjusted to minimize omissions and double counting in a cost-effective manner?

[80] To answer yes to this question, a RP must have some sense that the trustees may be willing to accept relative weights based on best professional judgment or survey. Developing litigation-proof weights would be cost-prohibitive except under the most extreme circumstances.

Annex C

ENTRIX Unpublished Report to API: Assessment of Beached Bird Modeling Methods

ENTRIX

ASSESSMENT OF BEACHED BIRD MODELING METHODS

Prepared for:

American Petroleum Institute
Washington, DC

Prepared by:

ENTRIX, Inc.
10 Corporate Circle, Suite 300
New Castle, DE 19720

January, 2009

TABLE OF CONTENTS

1.0 INTRODUCTION 1-1

2.0 ALTERNATIVE METHODS FOR ESTIMATING AVIAN MORTALITY 2-1

2.1 LITERATURE TRANSFER 2-1
2.2 SWEPT THROUGH CALCULATIONS 2-2
2.3 BEACHED BIRD MODELING 2-4
2.4 CHOOSING A MORTALITY ESTIMATING METHOD 2-6
2.5 DECISION MAKING WITHIN THE NRD FRAMEWORK 2-6

3.0 BEACHED BIRD MODELING MECHANICS 3-1

4.0 CRITICAL RESPONSE ACTIVITY AND POST-HOC STUDIES 4-1

4.1 SEARCH PATTERN 4-1
4.2 VECTOR OF CARCASS COLLECTIONS 4-5
4.3 PHYSICAL REMOVAL 4-6
4.4 SEARCH EFFICIENCY 4-7
4.5 BACKGROUND DEPOSITION 4-9
4.6 SINKING 4-10

5.0 ESTIMATING MORTALITY USING THE BBM 5-1

5.1 STEPWISE PARALLEL APPROACH TO BBM 5-1
5.2 CRITICAL BBM ASSUMPTIONS AND DATA THAT CAN BE USED TO EVALUATE THEM 5-2

6.0 BBM SENSITIVITY EXPERIMENTS AND POTENTIAL REFINEMENTS 6-1

6.1 EXPERIMENTS AND APPROACHES 6-1

7.0 BBM SUMMARY AND RECOMMENDATIONS 7-1

7.1 SWEPT THROUGH OR BBM MORTALITY ESTIMATION 7-1
7.2 EMERGENCY RESPONSE STAFFING 7-1
7.3 MANAGING SEARCH EFFORT 7-2
7.4 GATHERING INPUT DATA SETS 7-3
7.5 GUIDANCE ON BBM SET-UP AND IMPLEMENTATION 7-4

8.0 WORKS CITED 8-1

LIST OF TABLES

TABLE 2.2.1-1. NRDAM/CME ASSUMPTIONS REGARDING THE JOINT PROBABILITY OF AT RISK BIRDS BECOMING OILED AND EXPERIENCING MORTALITY. 2-4
TABLE 3.1-1. DATA INCLUDED IN A BBM SEARCH PATTERN AND VECTOR OF CARCASS COLLECTIONS 3-1
TABLE 3.1-2. THE PROBABILITY THAT A CARCASS WILL PERSIST A GIVEN NUMBER OF DAYS AFTER IT IS DEPOSITED. 3-2
TABLE 3.1-3. THE PROBABILITY THAT A CARCASS DEPOSITED AT AN UNKNOWN TIME WITHIN A SEARCH INTERVAL WILL PERSIST UNTIL THE SEARCH 3-2
TABLE 3.1-4. TOTAL CARCASS DEPOSITION RATES 3-4
TABLE 3.1-5. SPILL-RELATED CARCASS DEPOSITION RATES 3-4
TABLE 4.4-1. SINGLE SEARCHER EFFICIENCIES ESTIMATED IN KURE STUDY 4-8
TABLE 4.4-2. SEARCH EFFICIENCIES FOR FOOT BASED SEARCHES ON LARGE COBBLE/BOULDER SHORELINES. 4-8
TABLE 6.1.1-1. THE RELATIONSHIP BETWEEN SEARCH FREQUENCY AND EXPECTED ERROR. 6-2

LIST OF FIGURES

FIGURE 2.4-1. SELECTING A BIRD MORTALITY ASSESSMENT METHOD. 2-8
FIGURE 6.1.2-1 BIAS ASSOCIATED WITH OMITTED SEARCH. 6-3
FIGURE 6.1.3-1. BIAS ASSOCIATED WITH MISS-SPECIFIED SEARCH EFFICIENCY. 6-4
FIGURE 6.1.4-1. BIAS ASSOCIATED WITH MISS-SPECIFIED PERSISTENCE RATES. 6-5
FIGURE 6.1.5-1. ASSESSMENT OF ALTERNATIVE SCHEMES TO ADDRESS PARTIAL SEARCH. 6-7
FIGURE 6.1.6-1 MODELING BIRDS COLLECTED BETWEEN SEARCHES. 6-8
FIGURE 6.1.8-1 USER SPECIFIED PATTERN OF DEPOSITION THROUGH TIME. 6-10
FIGURE 6.1.8-2 ALTERNATIVE METHODS FOR ADDRESSING TEMPORAL PATTERNS OF CARCASS DEPOSITION. 6-11
FIGURE 6.1.11-1 MORTALITY SCENARIOS WITH OUTLIER MULTIPLIERS 6-15

LIST OF APPENDICES

APPENDIX A: ALTERNATIVE BBM MODELS

APPENDIX B: DETAILED EXPERIMENTAL METHODS

APPENDIX C: QUICK REFERENCE AID TO BEACHED BIRD MODELING

1.0 Introduction

Potential avian impacts are often a key liability component of Natural Resource Damage Assessment (NRD) under the Oil Pollution Act of 1990 (OPA). The foundation of those liability estimates is the estimate of total acute avian mortality. Indeed, many Trustee agencies, including the United States Fish and Wildlife Service (USFWS), and the California Office of Spill Prevention and Response (OSPR), currently translate acute mortality estimates directly into injury and restoration projects using predetermined conversion factors based on "standard" Resource Equivalency Analysis (REA) (See Sperduto et al. 2003 for USFWS methods, and CDFG 2004 for OSPR). As such, liability levels are largely established once total acute mortality is estimated.

The methods used to estimate total acute mortality following an oil spill fall into three broad categories: literature transfer (RIDEM 1998), Beached Bird Modeling (BBM) (Ford 1987), and swept through calculations (French et al. 1996). Among the three, literature transfer requires the least amount of data but also generates the most uncertain results. Beached Bird Modeling is the most data intensive method, but is generally thought to generate the most reliable results. Swept through calculations are intermediate in both data requirements and uncertainty.

This paper has three purposes. The first is to provide NRD practitioners with an overview of methods commonly used to asses acute avian mortality following oil spills. The second purpose is to provide NRD practitioners with an understanding of the BBM, the data required to support BBM modeling, uncertainties associated with BBM results, and steps that can be taken to reduce those uncertainties. Finally, we outline methods designed to help Trustee and responsible party representatives cooperatively identify optimal strategies throughout the BBM process.

The BBM has been selected for in depth analysis for several reasons:

- The BBM has been the assessment tool of choice for most recent west coast spills (Kure, New Carissa, Stuyvesant, Luckenback),
- The BBM was used recently used on a major east coast spill (Bouchard Barge-120 in Buzzards Bay MA),
- USFWS recently conducted training to facilitate wider use of the BBM on the east coast, and

BBM results can be sensitive subtle variations in assumptions and modeling approach. These sensitivities have not been addressed in the literature.

Our goal is to facilitate an understanding of the input data required by the BBM, the uncertainty associated with the BBM, and how spill response and post-hoc activities can be used to address those uncertainties.

The main body of this reported is divided into an additional six sections:

- Section 2: Alternative methods for estimating acute avian mortality;
- Section 3: BBM mechanics;
- Section 4: Critical response activity and post-hoc studies in support of BBM;
- Section 5: Estimating mortality using the BBM;
- Section 6: BBM sensitivity; and
- Section 7: Summary and BBM recommendations based on expected benefits and costs.

Appendix C is a series of short tables designed to help the NRD responder make BBM related decisions in an emergency response situation.

2.0 Alternative Methods for Estimating Avian Mortality

This section describes the three alternative approaches for estimating acute avian mortality following an oil spill: literature transfer, swept through calculations, and beached bird modeling.

If the NRD practitioner becomes engaged in the NRD process following the spill response, he or she will be limited to approaches that can be supported by the data collected during the spill response. However, if the NRD practitioner is involved during the response, he or she may have the opportunity to evaluate the spill environment, select an appropriate mortality estimation approach based upon the spill environment, and ensure that required data are collected. Sections 2.1 through 2.4 provide insight into the process of selecting an assessment approach. Section 2.5 identifies and discusses some of the dynamics associated with decision making within the NRD construct and identifies the roles of Trustee and responsible party representatives

2.1 Literature Transfer

Literature transfer uses mortality estimates published in the literature to estimate mortality levels for the spill in question. It is the only mortality estimation method that can be implemented without any incident specific data.

Literature transfer can be as simple as stating "the average acute avian mortality following an oil spill is 500 birds. Therefore, the acute avian mortality associated with this spill will be 500 birds." Typically however, literature transfer is somewhat more sophisticated. A rule of thumb may be applied such as "total mortality is ten times the number of birds collected." At other times, the transfer may attempt to control for factors thought to influence avian mortality. For example, one might estimate mortality associated with a spill in California as the average mortality associated with previous spills in California.

The logical progression of literature transfer would use regression analysis to identify a mathematical function that predicts acute avian mortality following a spill. In such an analysis, data from past oil spills would be used to predict acute avian mortality as a function of explanatory variables such as volume of oil spilled, water body type, season, etc. Unfortunately researchers have not identified robust explanatory variables. Burger (1993) reports:

> "A statistical analysis of 45 oil spills shows a weak…correlation between spill volume and number of seabirds killed. This relationship cannot be used to predict mortality and loses its significance if one extreme case is omitted. The data show the wide variance in mortality in spills of all sizes. A loose 'rule-of-thumb' that is often used in poorly documented spills is that the overall mortality is ten times the actual body count. There is no justification for this notion. The mean estimate is 4–5 times the body count, but each spill should be examined independently."

Ford *et al.* (No date A) updated this analysis to include additional spills and to control for type of habitat impacted (estuarine versus open ocean) and the latitude at which the spill occurred. While Ford's expanded analysis did identify variables with some explanatory power, his conclusion was similar to Burger's findings. The regression-based relationship was "unlikely to be very accurate in any specific incident" and actual mortality relative to predicted mortality may vary by more than an order of magnitude.

In concluding the paper, Ford *et al.* (No Date A) noted that Piatt and Ford (1996) analyzed several experiments in which authors released marked seabirds into the ocean following oil spills and determined how many of those were subsequently recovered. Results ranged from 0 to 47 percent, with an average of 17 percent recovered, e.g., one in five birds was recovered. This observation supports Burger's recommendation (1993) that if literature transfer is the selected estimation method, total acute mortality may be estimated as four to five times the body count.

2.2 Swept Through Calculations

Swept through calculations estimate avian mortality as the product of:

- The number of birds at risk of exposure to oil;
- The proportion of at-risk birds that become oiled; and
- The mortality rate among oiled birds.

Swept through calculations can be performed with limited spill-related data. Other than release location, release time, and an estimate of the spill volume, no incident specific data is required. Under conditions of very limited spill related data, a generic set of swept through calculations can be thought of as occurring in 5 steps.

a. Hydrodynamic modeling is used to simulate the movement of oil as it “sweeps” across surface waters and shorelines.
b. Surface waters and shorelines swept by oil are divided into cells and the habitat type of each cell is identified. Habitat types might include ocean, sandy shoreline, and wetland.
c. Pre-existing biological databases are used to assign a number of birds to each cell. For example, the model might assign one common eider to each ocean cell, two mallards and two willets to each wetland cell, and ten least sandpipers to each shoreline cell. These birds are identified as “at risk” birds.
d. The proportion of at risk birds that become oiled is estimated. This parameter may be estimated as a function of the volume of oil in a cell, the amount of time oil remains in a cell, the characteristics of the oil that was spilled, and/or species-specific behaviors.
e. A mortality rate is applied to all oiled birds. These estimates are often guild-specific because hypothermia-induced mortality appears to be directly related to the proportion of time birds spend on the water and may increase at higher latitudes.

To determine the reliability of their swept through algorithm, French-McCay and Rowe (2004) assessed the correlation between avian mortality estimates generated by their swept through algorithm and avian mortality estimates based on field observations. They reported that, absent any field data, swept-through estimates can diverge significantly from estimates generated by field based methods. However, their analysis also suggested that, when field data are used to augment swept-through calculations, the divergence can be decreased to a factor of 2 or 3.

2.2.1 Augmenting Swept Through Calculations

Confidence in the results of swept through calculations is increased when field data are used to calibrate the number of birds at risk, the probability of an at risk bird becoming oiled, and the probability of spill-related mortality among oiled birds. The text below briefly outlines methods that can be used to accomplish that goal. Any data collection protocol implemented during a spill response would need to be tailored to best address the specific circumstances and logistical constraints associated with the spill.

2.2.1.1 Estimating the Number of Birds at Risk

As described above, the number of birds at risk is a function of the amount of habitat swept by oil and an assumed bird density for each habitat.

The amount of habitat swept by oil is typically determined via hydrodynamic modeling. While the physics underlying these models is fairly well understood, uncertainty associated with the spill release scenario and the effects of random events introduce error into hydrodynamic simulations. Therefore, hydrodynamic modeling should be calibrated to any data describing slick and shoreline oiling observations. If sufficiently detailed, these observations can be used as a substitute for hydrodynamic modeling.

Absent spill-specific field data describing bird densities, swept through calculations rely on average bird density estimates reported in the literature; often these averages are aggregated to state or regional levels. French-McCay and Rowe (2004) noted that avian abundances are extremely variable in space and time and so reliance on literature-based average densities introduces a high degree of uncertainty into swept through calculations. This uncertainty is greatly reduced if the data necessary to estimate bird densities are collected during the spill response.

The NRD practitioner should be aware that the data required to estimate bird densities are not generally collected during a spill response. Therefore, the necessary observations will likely only be available if specific protocols for their collection were distributed and implemented during the response. The protocols can be land based, air based, or boat based. Land based protocols require observers to be distributed randomly throughout the spill area. Each day observers record their location, the surrounding habitat, viewing distance, and the number of birds they observed by species, and the time at which they were observed. There is a large literature on conducting aerial and boat based surveys to estimate the number of birds in an area, Henkel et al. 2007 provides a relevant overview.

2.2.1.2 Estimating the proportion of at risk birds that get oiled

Data describing the proportion of at risk birds that, on average, actually get oiled is sparse and highly variable. This variability is likely due to the movement of birds in space and time. As such, reliance on literature-based estimates introduces a high degree of uncertainty. This uncertainty is exacerbated when the species of concern tend to flock or raft. For flocking and rafting species, it may be inappropriate to assign an average rate of oiling calculated over many spills to a specific spill where the entire flock or raft will likely either encounter, or avoid oil.

The uncertainty associated with the proportion of birds at risk that get oiled can be reduced if the data necessary to estimate the proportion of at risk birds actually oiled is collected during the spill response. Again, the required observations are not typically collected during an oil spill response and will likely only be available if a NRD practitioner established protocols for their collection during the response. These protocols can be identical to those used to estimate bird density with the notable exception that observers are also required to note the degree of oiling, i.e. no visible oiling, trace, light, moderate, heavy, for each bird they observe. This requirement makes the use of air and boat based surveys more challenging.

2.2.1.3 Estimating mortality rates among oiled birds

Among the parameter estimates required to complete a swept through calculation, estimates of mortality rates given oiling may be the most uncertain and are the most difficult to address via data collection.

One rule of thumb often asserted is that “an oiled bird is a dead bird”. There is no justification for this notion. Mortality rates among some species may approach 100 percent under certain circumstances. These very high rates would likely be associated with small birds that spend the majority of their time on the water in cold climates (high latitudes). The Natural Resource Damage Assessment Model for Coastal and Marine Environments (NRDAM/CME), which was developed for the US Department of Interior (USDOI), reports the joint probability that at risk birds will become oiled and experience spill related mortality (Table 2.2.1-1) (French *et al.* 1996). However, in a discussion of the data underlying their assumptions, French *et al.* report “documentation of the probability of oiling and the mortality of wildlife which have been oiled is not readily available" and mortality estimates for Piscivorous Raptors, Aerial Seabirds, Waders, Shorebirds, and Terrestrial birds are based on best professional judgment alone.

Table 2.2.1-1. NRDAM/CME assumptions regarding the joint probability of at risk birds becoming oiled and experiencing mortality.

Behavioral Group	Examples	Assumed Probability of Oiling and Mortality
Dabbling Waterfowl	Mallards, Geese, Swans	99%
Nearshore Aerial Divers	Kingfishers	35%
Piscivorous Raptors	Eagles, Osprey	35%
Surface Seabirds	Eiders, Alcids, Cormorants	99%
Aerial Seabirds	Kittiwakes, Terns, Pelicans	5%
Wading and Shore Birds	Plovers, Dunlins, Herons	35%
Terrestrial Birds	Hawks, Ravens, Owls	0.1%

Since the 1996 publication of NRDAM/CME, ENTRIX has compiled additional data regarding the probability of mortality given exposure. These data lead us to believe that the values published in French *et al.* 1996 tend to overestimate the probability of mortality for several behavioral groups. Our research suggests that mortality rates among oiled aerial birds, terrestrial birds, shorebirds and wading birds are in the one to ten percent range even among moderately and heavily oiled individuals. However, this is an area of ongoing research for ENTRIX and, at this time, it is most appropriate to simply identify the parameter as being highly uncertain.

Under certain circumstances, the NRD practitioner may be able to collect data that will allow estimation of mortality rates given oiling. However, the process is labor intensive and can not be completed for at least one year post spill. If an impacted species is already banded and intensely monitored, it may be possible to identify oiled and un-oiled birds during the spill and then use return rates in the following year(s) to estimate survival rates among oiled and un-oiled birds. This was the case for kittiwakes following the *Braer* spill and two species of terns following the *Bouchard B-120* spill. Similarly, if nearly all members of a small, easily censused sub-population with high site fidelity are exposed to oil, it may be possible to monitor mortality rates in that sub-population. This was the case for piping plovers exposed to oil in the *Bouchard B-120* spill and western snowy plovers following the *New Carissa*. Absent those unusual circumstances, it is extremely difficult to measure this parameter.

2.3 Beached Bird Modeling

BBM is the most data intensive mortality estimation approach. The approach estimates total mortality as a function of the dead and injured birds collected during the spill response, the intensity of bird collection effort, physical processes that naturally add and remove carcasses from a shoreline, and sinking rates. In total, six input data sets are required:

1. The mainland search pattern describes the spatial distribution, frequency, and intensity of shoreline searches;
2. The vector of carcass collections describes each carcass collected and the circumstances surrounding its collection;
3. The physical removal function describes the rate at which carcasses are removed from the shoreline by scavengers, burying, or rewash;

4. The search efficiency function describes the proportion of carcasses that, if present during a search, are likely to be found;
5. The background deposition rate describes the number of birds that would have died and been deposited on the shorelines had there not been a spill; and
6. The sinking function describes the probability that an oiled carcass sank prior to being deposited on a shoreline.

While the mainland search pattern and vector of carcass collections must be documented during the spill, the remaining data sets are generally obtained via post-spill field studies, non-spill related programs, or from the existing literature.

There are five steps involved in BBM calculations:

1. The number of birds collected on any given search is converted to a rate of carcass deposition in the days prior to the search. This rate is often reported as daily carcass depositions per km. The conversion controls for the length of the search (often reported in km), the amount of time elapsed since the last time the area was searched (often reported in days), the number of carcasses removed[1] from the search area prior to a search, and the number of carcasses that, though present, are not found by searchers (imperfect search).
2. The bird deposition rate is extrapolated to areas for which there are no search data. This extrapolation adjusts for incomplete search effort and results in an estimate of total carcass deposition.
3. Total deposition is divided into spill-related deposition and background deposition by applying a background deposition rate.
4. Spill related deposition is converted to spill-related mortality by adjusting for the portion of carcasses believed to have sunk before being deposited on a shoreline (sinking).
5. The commonly reported "multiplier" is calculated as total spill-related mortality ÷ the total number of carcasses collected.

Ford (1991) showed, and the experiments reported in Section 6.0 of this report confirm, that BBM can be an unbiased estimator of total deposition if a spill includes a large number of spill-related carcasses deposited randomly or uniformly in space and time. If these conditions hold, the level of uncertainty associated with any BBM mortality estimate is a function of the precision of the six input data sets and the level of consistency between the assumptions made by the BBM practitioner and the facts of the spill.

That is, the technical underpinnings of the BBM are sound. If a large number of carcasses are deposited randomly or uniformly in space and time, the BBM can produce an unbiased estimate of total carcass deposition. However, for this to happen, the 6 input data sets must accurately represent the circumstances surrounding the event and the assumptions selected by the BBM practitioner must mimic spill events. If data sets are incomplete and or a series of assumptions "protective of the resource" are employed, effects can compound resulting in significant overestimates of mortality.

Section 5 of this paper discusses methods that can be used to maximize the precision of the six input data sets. Section 6 describes methods to minimize the divergence between BBM assumptions and the facts of the spill. Section 7 summarizes and provides recommendations for the BBM practitioner. The recommendations contained in each of these sections are designed to minimize the need for assumptions and the resulting tendency to overestimate avian impacts by applying a series of assumptions that are "protective of the resource."

---

[1] Removal can occur via scavenging, burying, or rewashing and sinking during subsequent high tides.

2.4 Choosing a Mortality Estimating Method

The preceding text describes the three methods that can be used to estimate total acute avian mortality associated with an oil spill: literature transfer, swept through calculations, and BBM. Each method has strengths and weaknesses and each may be appropriate under specific circumstances.

The primary advantage to literature transfer is that it can be conducted with limited site-specific data. While data commonly collected during a spill response such as the location of the spill, or the volume of oil spilled are often used to qualitatively adjust mortality estimates, the validity of those adjustments appears questionable. Authors who have investigated the approach note that the best literature based approach may be to simply multiply the number of carcasses collected by four or five.

Swept through calculations can be performed given only a release location, a release time, and an estimate of the spill volume. However, when performed with these minimal data requirements, swept through mortality estimates are highly uncertain. Confidence in swept through estimates can be increased by using incident specific data describing the movement of oil, the number of birds at risk of getting oiled, and the probability of at risk birds getting oiled, and the probability of oiled birds dieing. French-McCay and Rowe (2004) report that when incident specific data is incorporated into the calculations, the difference between swept through mortality estimates and those generated by more reliable field based methods tend to diverge by a factor of two or three.

BBM is the most data intensive mortality estimation approach. It requires data describing all avian search efforts that occurred during the spill response as well as the details surrounding the collection of each carcass. Data describing the rate at which carcasses are naturally removed from the shoreline via physical processes, the ability of searchers to find carcasses, rates of background bird deposition, and the probability that carcasses sank prior to being deposited on shorelines are also required. These data are generally obtained via post-spill field studies or from the literature. When the six required data sets accurately describe the physical processes that prevailed during the spill, the BBM has been shown to be an unbiased mortality estimator.

Figure 1 is a flow diagram designed to help the NRD practitioner select a method for assessing avian mortality. We identify several decision points where RPs must determine if an outcome is acceptable. The "acceptability" of various outcomes is specific to each individual RP as each RP may consider multiple criteria such as total expected NRD and response cost, Trustee preferences, and the level of response cooperation.

That said, from a cost benefit perspective, literature transfer is generally best applied to spills where extremely high levels of uncertainty are acceptable or in circumstances where all parties can reach a rapid agreement. If reduced uncertainty is preferable, and/or agreement can not be reached, swept through calculations or BBM will be required.

When choosing between swept through and BBM, if field data to support estimates of the number of birds at risk, the probability of oiling among at risk birds, and a mortality rate among oiled birds can be collected, swept through modeling may be preferred because of its generally lower transaction costs. In the absence of field data to support these parameter estimates, swept through calculations may still be preferable if a mortality estimate accurate to within a factor of 2 or 3 is acceptable, or if repeat shoreline searches are impractical for logistical reasons. BBM is best employed when a relatively low level of uncertainty is required and when potential NRD liability justifies the incremental transaction costs; that break point is generally around $1,000,000 in NRD liability (See Section 7.1)

2.5 Decision Making Within the NRD Framework

In the few days following a major spill, it is not always easy to assemble all stake-holders (state, tribal and federal Trustees as well as responsible party representatives) to identify near-term assessment activities that will make best use of scarce resources to reduce future assessment uncertainties and transactions costs. To address that issue NOAA, in conjunction with several P&I clubs, has suggested the cooperative development of bird injury assessment guidelines. We strongly encourage industry, state, and federal participation in such an effort should it come to pass. A set of common ideas and guiding principles established prior to a spill could significantly improve the quality and efficiency of response activity.

Absent some consensus regarding guiding principles, ENTRIX believes the following ideas may prove helpful.

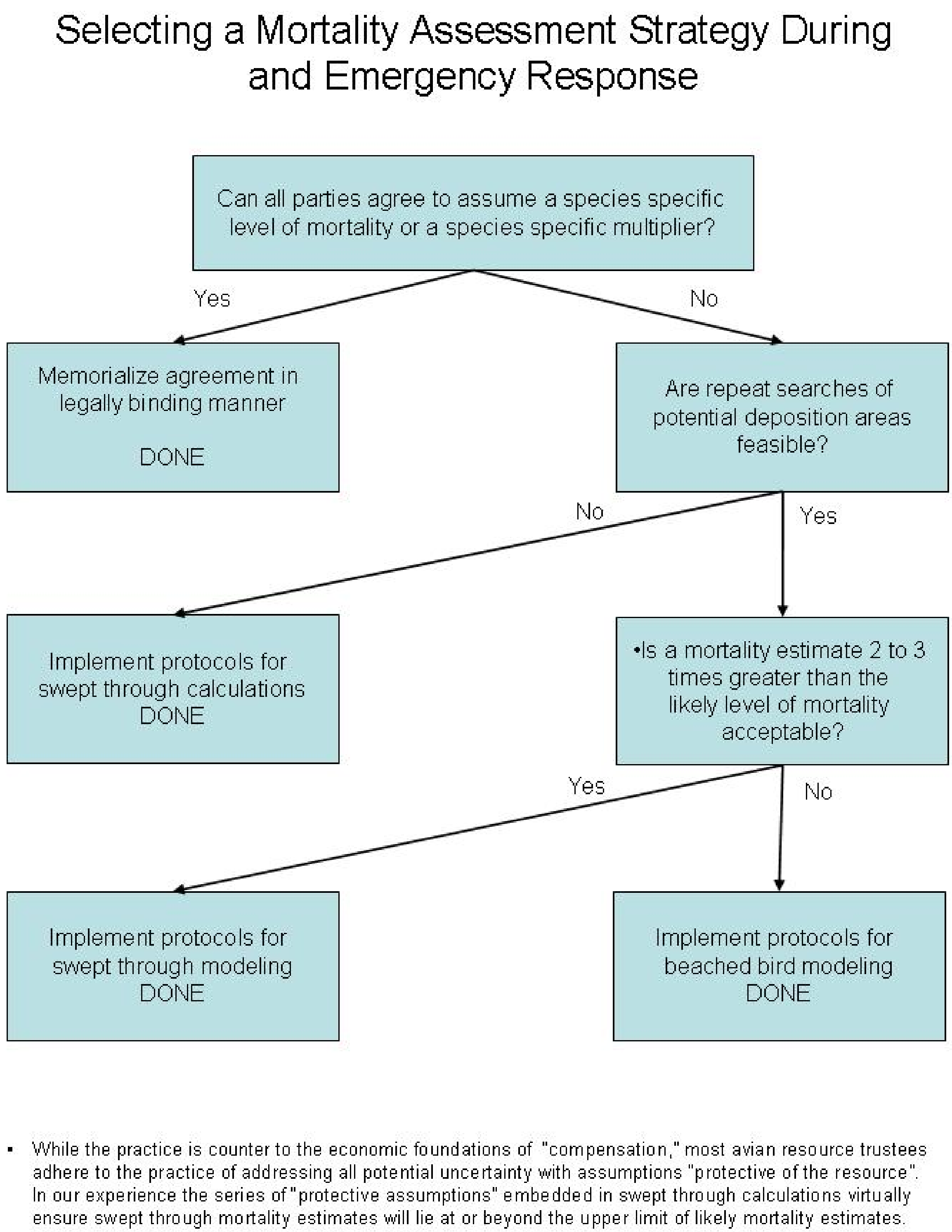
When working within the context of a cooperative NRD assessment, it is recognized that the Trustees have legal responsibility for the assessment. Further, the degree of actual cooperation or collaboration allowed an RP in an assessment varies considerably from case to case. To the extent possible, the RP should attempt to exert influence on the process to achieve the following goals.

- Ensure that recommendations for activities are being made by individuals experienced in the assessment of avian impacts related to oil spills. Efforts should be made to ensure that both the Trustees and the responsible party have at least one experienced NRD bird modeler on-site.
- Encourage the team to plan first. It is not efficient to collect data without having chosen a mortality assessment method, nor is it efficient to select an assessment method if it cannot be supported by sufficient data.
- If possible, conduct field operations jointly. Having a Trustee representative and an RP representative paired in tasks tends to decrease future transaction costs, especially if there is joint documentation of outcomes/observations.
- Encourage open discussion and accommodate, to the maximum extent possible, the concerns of all stakeholders. Adding extra documentation procedures, clarifying a method, and/or developing an understanding of how and why the data will eventually be used may slow operations initially but will reduce assessment uncertainty in the long run.
- Conduct modeling in a stepwise parallel manner (See Section 5.1).

If an NRD assessment is not cooperative,

- Legal constraints associated with the handling of birds protected under the migratory bird treaty act prohibit an RP from independently assembling the data sets necessary for a BBM. Unless the spill area is quite small, even monitoring trustee activities will be impossible.
- Trustee organizations are encouraged to use the decision tree in Figure 2.4-1 to determine if BBM modeling is the optimal assessment strategy. If BBM is selected as the assessment method, Trustees are encouraged to consider the ideas outlined herein as they assemble data sets and conduct modeling.
- RPs could collect data sets that could be used to augment/verify independent swept through modeling; these include bird density, the proportion of birds oiled, and, if possible, the mortality rate among oiled birds. If the independent Trustee assessment is ultimately acceptable to the responsible party, the data need not be used. However, if the assessment is not acceptable and the RP has not collected data, few options remain.

Figure 2.4-1. Selecting a Bird Mortality Assessment Method.



3.0 Beached Bird Modeling Mechanics

In this section, beached bird modeling mechanics are illustrated using a tabular example. In Appendix A, an algebraically motivated BBM and a simulation based approach to BBM are described. Importantly, under many common conditions, the mortality estimates generated by each of the approaches are similar given similar assumptions.

Table 3.1-1 includes the information typically contained in the mainland search pattern and the vector of carcasses collected.

For illustrative purposes, assume all spill-related deposition occurred on 4 km of shoreline which has been broken into a northern and southern segment. For simplicity, also assume that a storm removed all background carcasses from the shoreline the day prior to the onset of spill-related deposition. Finally carcasses are removed from the shoreline when found.

Table 3.1-1. Data Included in a BBM Search Pattern And Vector of Carcass Collections

	Northern Segment 2 km long	Southern Segment 2 km long
Day 0: Onset of spill deposition		
Day 1	Search N1 4 carcasses found Search interval is 1 day	
Day 2		
Day 3		Search S1 1 carcass found Search interval 3 days
Day 4		
Day 5		
Day 6		
Day 7	Search N2 2 carcasses found Search interval is 5 days	

Assume physical processes remove 25 percent of carcasses present each day, search efficiency is 25%, the daily rate of background bird deposition is 0.5 carcasses per km, and 10 percent of spill-related carcasses sink before being deposited on a shoreline. Tables 3.1-2 and 3.1-3 translate the 25 percent carcass removal rate into two key functions: the probability a carcass will persist a given number of days following its deposition and the probability that a carcass deposited at an unknown time within a search interval will persist until the search occurs.

Table 3.1-2. The Probability that a Carcass Will Persist a Given Number of Days After it is Deposited.		
Days Since Deposition	**Probability of Persistence**	**Formula**
0	1.00	0.75^0
1	0.75	0.75^1
2	0.56	0.75^2
3	0.42	0.75^3
4	0.32	0.75^4
5	0.24	0.75^5
6	0.18	0.75^6

Table 3.1-3. The Probability that a Carcass Deposited at an Unknown Time Within a Search Interval Will Persist Until the Search		
Search Interval in Days	**Probability of Persistence**	**Calculated As**
0	1.00	1
1	0.88	(1+0.75)÷2
2	0.77	(1+0.75+0.56)÷3
3	0.68	(1+0.75+0.56+ 0.42)÷4
4	0.61	(1+0.75+0.56+0.42+0.32)÷5
5	0.55	(1+0.75+0.56+0.42+0.32+0.24)÷6
6	0.50	(1+0.75+0.56+0.42+0.32+0.24+0.18)÷7

Step one in the BBM estimation converts the number of carcasses collected on each search into carcass deposition rates.

NORTHERN SECTION

During Search N1, 4 carcasses were collected. The 25 percent search efficiency implies 16 carcasses were deposited and persisted until search N1. That is, given a 25 percent collection efficiency, the collection of 4 carcasses implies 16 carcasses must have been present during the search.

The probability that a carcass deposited at an unknown time within a 1 day search interval will persist until the search is 0.88 (Table 3.1-3). Hence for each carcass that persisted, the BBM calculates that 1.14 were deposited (calculated as 1 carcasses deposited ÷ 0.88 carcasses that persist).

The number of carcasses deposited since the spill is 18.29 (calculated as 16 carcasses persist X 1.14 depositions per carcass that persisted until the search).

This yields a carcass deposition rate of 9.15 carcasses per km per day and it applies to the 2 km of the Northern segment on day zero (Table 3.1-4).

During search N2, 2 carcasses were collected. However, not all carcasses collected are associated with deposition that occurred after search N1. Some of the carcasses collected on search N2 were present but not found on search N1, they persisted to N2 and were then found. We refer to these as holdover carcasses.

Because the BBM is calculating deposition that occurred in the prior search interval, adjustments are made to control for holdover carcasses. Recall that, during search N1, 16 carcasses were present and 4 were collected. Hence, 12 carcasses remained on the Northern segment after search N1. Six days later, when search N2 occurs, approximately 2.16 of those 12 holdover carcasses remain (Calculated as 12 carcasses X 0.18 probability of persistence for 6 days where 0.18 is taken from Table 3.1-2). Of those 2.16 carcasses, 0.54 are expected to have been collected on search N2 (calculated as 2.16 carcasses X 25 percent search efficiency).

The 0.54 holdovers are removed and the BBM estimates the deposition rate for the days between search N1 and N2 based on the collection of 1.46 carcasses (calculated as 2 carcasses collected – 0.54 holdover carcasses).

Coupling the 1.46 carcasses with 25 percent search efficiency implies 5.84 carcasses were deposited in the 6 days between search N1 and N2 and persisted until search N2.

Note the probability that a carcass deposited at an unknown time within a 6 day search interval will persist until the search is 0.5 (Table 3.1-3). Hence for each carcass that persisted, 2.0 were deposited (calculated as 1 carcasses deposited ÷ 0.5 carcasses that persist).

The number of carcasses deposited is 11.68 (calculated as 5.84 carcasses persist X 2.0 depositions per carcass that persists).

This yields a carcass deposition rate of 0.97 carcasses per km per day and it applies to the 2 km of the Northern segment on days 1 through 6 (Table 3.1-4).

SOUTHERN SECTION

On Search S1, 1 carcass was collected.

The 25 percent search efficiency implies 4 carcasses were deposited and persisted until search S1.

The probability that a carcass deposited at an unknown time within a 3 day search interval will persist until the search is 0.68 (Table 3.1-3). Hence for each carcass that persisted, 1.47 were deposited (calculated as 1 carcass deposited ÷ 0.68 carcasses that persist)

The number of carcasses deposited is 5.88 (calculated as 4 carcasses that persist X 1.47 depositions per carcass that persist).

This yields a carcass deposition rate of 0.98 carcasses per km per day and it applies to the 2 km of the Northern segment on days 0 through 2 (Table 3.1-4).

Step 2 is to extrapolate deposition rates to segment-day combinations for which no search data applies. In the example above, deposition data is not available for the Southern segment between days 3 and 6. While the data can be extrapolated in a variety of ways[2], in this example, the rate of 0.98 birds per km per day (the rate that prevailed on the southern segment earlier in the spill) is applied to the Southern segment on days 3 through 6. Table 3.1-4 reports daily deposition rates for all days and segments.

[2] In the past we have also used deposition rates estimated from nearby segments while controlling for shoreline oiling, orientation, and complexity.

Table 3.1-4. Total Carcass Deposition Rates		
	Northern Segment **2 km long**	**Southern Segment** **2 km long** **(rates estimated via extrapolation in italics)**
Day 0: Onset of spill deposition	9.15 carcasses per km	0.98 carcasses per km
Day 1	0.97 carcasses per km	0.98 carcasses per km
Day 2	0.97 carcasses per km	0.98 carcasses per km
Day 3	0.97 carcasses per km	*0.98 carcasses per km*
Day 4	0.97 carcasses per km	*0.98 carcasses per km*
Day 5	0.97 carcasses per km	*0.98 carcasses per km*
Day 6	0.97 carcasses per km	*0.98 carcasses per km*
Day 7	Not Available	Not Available

Step 3 is to convert total carcass deposition rates into a rate of spill-related carcass deposition by subtracting background birds. In this example, 0.5 background carcasses are deposited per km per day. Thus, for the 2 km stretch of beach, 1 carcass is deposited per day. Subtracting 1 carcass per day from the deposition rates reported in Table 3.1-4, yields the spill related daily deposition rates (Table 3.1-5)

Table 3.1-5. Spill-Related Carcass Deposition Rates		
	Northern Segment **2 km long**	**Southern Segment** **2 km long**
Day 0: Onset of spill deposition	8.15 carcasses per km	0 carcasses per km
Day 1	0 carcasses per km	0 carcasses per km
Day 2	0 carcasses per km	0 carcasses per km
Day 3	0 carcasses per km	0 carcasses per km
Day 4	0 carcasses per km	0 carcasses per km
Day 5	0 carcasses per km	0 carcasses per km
Day 6	0 carcasses per km	0 carcasses per km
Day 7	Not Available	Not Available

Step 4 is to convert the 8.15 spill-related depositions into spill-related mortality by adjusting for the proportion of carcasses that sank before being deposited on a shoreline. In our example, 10% of carcasses sank before being deposited. <u>This implies about 9-spill related mortalities.</u>

Step 5 is to calculate the commonly reported "multiplier". Taking the 9 spill-related mortalities ÷ 7 carcasses collected from the northern and southern segments, implies 1.3 spill-related mortalities per carcass collected. This would commonly be reported as a multiplier of 1.3

4.0 Critical Response Activity and Post-Hoc Studies

In our experience, the largest single source of BBM uncertainty relates to incomplete or poorly documented input data sets. As such, one of the most important activities for NRD practitioners during a response is documentation and data set construction. Failure to accomplish this goal often leads to assumptions that are "protective of the resource", a series of which can lead to significant mortality overestimation.

In this section we identify and discuss response activities, response documentation, and post-response field work that can be used to reduce BBM uncertainty. The discussion is organized around the six BBM data inputs: search pattern, vector of carcass collection, physical removal, search efficiency, background deposition, and sinking. Before discussing each of these topics in detail, some important terms are defined below.

Our recommendations are limited to "typical spills." A "typical spill" is defined to have impacted no more than 300 miles of mainland U.S. shorelines. Searching these shorelines does not entail unusual logistical challenges related to weather, access, or staging. Spill related deposition is limited to about 1 month after the spill.

The shoreline where potentially spill related carcass deposition may occur is defined as the geographic extent of deposition. It is critical that the shoreline searches be conducted throughout the entire geographic extent of potential deposition (up-coast, down-coast, and inland).

Generally, the geographic extent of deposition will be divided into bird search segments. These are relatively short stretches of shoreline, a few km at most, that can be searched by a single search team in their entirety, in ½ days or less. Unified command will also be dividing the shoreline into response segments. If possible, the NRD team should adopt the unified command response segments. If not possible, special attention must be paid to ensure that the existence of two segmentation systems does not cause confusion among bird responders or other response personnel.

The NRD team will only be able to direct the search effort of teams under their control. These teams should be instructed to always search bird segments in their entirety; partial segment searches require additional post-response analysis and so should be minimized. The NRD team will not be able to direct SCAT activity, inspection teams, clean-up crews or live wildlife recovery teams. However, close coordination with unified command can help minimize the occurrence of partial search as most of these teams will be instructed to perform their jobs over entire response segments.

The amount of time between repeat searches of a bird segment is referred to as search interval. The NRD responder should strive to keep search interval to less than 10 days; 3 to 5 days may be better. If the NRD responder has a-priori knowledge or expectations that the rate of physical removal may be very high, shorter search intervals should be targeted.

Finally, physical removal rates, background deposition rates, and sinking rates are specific to both location and time. These rates are typically determined using data collected during post-spill response field studies or via computer simulation. Given the sensitivity of these rates to the conditions that prevail during the spill, the NRD responder should carefully consider any opportunity to collect data during the spill that will facilitate the estimate of these rates at a later time.

The remainder of this section uses the terms and ideas identified above to discuss the creation of the six necessary BBM data sets.

4.1 Search Pattern

The longer the search intervals, the greater the probability that a carcass will be removed by some natural physical process (scavenging, burying, rewash and subsequent sinking) before it is collected. Therefore, when areas are searched less frequently, multipliers are increased. Similarly, as the proportion of unsearched shoreline increases, multipliers will increase. Therefore, failure to document search effort that does occur biases the BBM towards mortality overestimation.

A typical spill response includes multiple efforts that, when participants are properly trained and their efforts documented, represent bird carcass searches. These include Shoreline Clean-up Assessment Team (SCAT) efforts, sign-off inspections, clean-up crew activity, live wildlife recovery efforts, NRD specific carcass searches, and search efforts contributed by interested citizens and local organizations unrelated to the spill response. We refer to the complete listing of all bird search efforts as the search pattern.

In our experience, the construction of a high quality search pattern for a spill that covers 200 to 300 miles of shoreline requires 3 dedicated personnel during the emergency phase of the response and 1 or 2 persons during the operational phase. These individuals do not participate in other types of field activities. They:

- Establish bird search segments and direct NRD carcass recovery teams;
- Train search groups and organize public search efforts;
- Collect, review and organize documentation from all sources of search on a daily basis;
- Ensure that each carcass collected can be associated with the search upon which it was collected;
- Use GIS to map all searches and carcass recovery locations; and
- Coordinate with the environmental unit of the unified command.

4.1.1 Training Bird Searchers

For beached bird modeling purposes, bird searchers do not need be trained in ornithology or bird identification. They only need to document the precise stretch of shoreline being searched, the date and time of the search, the number of persons in their group, and the number of carcasses observed. If no carcasses are observed the teams must note the absence of carcasses. Finally, if carcasses are observed, searchers need to be trained in the proper procedure for collecting the carcass, documenting the collection, and turning the carcass over to USFWS.

During past spills, NRD practitioners have trained others by disseminating bird search requirements through the unified command via the environmental unit and by participating in the mandatory tail gate health and safety briefings for SCAT, inspection teams, clean-up crew supervisors, and live bird collection teams.

4.1.1.1 Working With SCAT and Sign-Off Inspection Teams

SCAT and sign-off inspection teams are already collecting all of the documentation necessary to describe a bird carcass search. The teams need only record the number of bird carcasses observed, or document that no carcasses were observed, and then collect any observed carcasses.

SCAT and inspection team members are generally comfortable performing the extra requirements of a bird search. The only commonly encountered concern is slowing down the SCAT and inspection effort. If this is a concern, an NRD member can be embedded in each team to handle carcass collections. Alternatively, teams can be asked to bag carcasses, label them so that they can be unambiguously linked to the documentation they are filling out, and relocate them to the super-tidal area of the shoreline. The SCAT or inspection team can then notify USFWS to make a collection via phone if a central wildlife collection number has been established. If not, the NRD team will be made aware of the carcass when they review the documentation that evening and the NRD team can coordinate carcass recovery.

4.1.1.2 Working With Live Wildlife Recovery Groups and Clean-Up Crews

Unlike SCAT and inspection teams, these teams may not be filling out forms that contain all of the documentation necessary for BBM modeling. In the past we have found it helpful to provide these groups (or their supervisors) with a form that prompts them to record: the GPS point where they begin and end their efforts; the date, start time and end time; the number of persons in the group; their mode of search (boat, foot, ATV, other); and the number of carcasses they observed or document that no carcasses were observed. It has also proved helpful to have a supervisor provide his or her name, their contact information during the response, and a contact number where they can likely be reached post-response.

These teams should already be collecting carcasses they observe and turning them over to USFWS. If they are not, an NRD member can be embedded in each team/crew to handle carcass collections. Alternatively, teams can be asked to bag carcasses, label them so that they can be unambiguously linked to the search documentation, and relocate them to the super-tidal area of the shoreline. The group or crew leader can then notify USFWS for subsequent collection.

Remember, the primary function of wildlife recovery groups and clean-up crews is not documentation. As such, it is often helpful to have an NRD responder meet with each group as they muster in and muster out each day to distribute forms for the upcoming day, review forms from the day just completed, and generally ensure that the process is running smoothly.

4.1.1.3 Working with the Public and Groups Unrelated To the Spill

When spills occur in developed areas, environmental groups, local animal control, and the public in general will be conducting wildlife searches. In a recent major spill, individuals unrelated to the spill response parties/organizations collected the vast majority of carcasses despite daily searches conducted by over a dozen NRD carcass search teams and extensive SCAT and inspection activity. Failure to document this search effort tends to bias the BBM towards overestimation.

First and foremost, when working with groups unrelated to the spill response, health and safety is the major concern. By having those groups document their efforts and report them to the NRD team, those groups have become, on some level, part of the spill response. Work through unified command to ensure that any and all health and safety issues are addressed (typically USCG will provide a health and safety briefing) and never allow non-spill related personnel to be present in areas where oil poses any threat to their physical well being.

Only after ensuring the health and safety of these individuals and confirming that all administrative questions have been addressed, should non-spill related personal be provided forms and instruction. Forms, necessary documentation, and carcass collection procedures are identical to that described for the wildlife recovery and clean-up crews. If convenient, it may also be beneficial to train these groups to conduct bird carcass searches in a systematic manner to ensure maximum search efficiency.

Again, the primary function of these public groups is not to provide spill documentation. It is often helpful to have an NRD responder meet with each group regularly to distribute forms for the upcoming time period, review recently completed forms, and generally ensure that the process is running smoothly.

Finally, while each spill is unique, it has been our general experience that non-spill personnel conducting bird carcass surveys "self-identify." Many are identified by reviewing the vector of carcass collections and noting individuals who repeatedly bring wildlife to the rehab center. These individuals can be contacted and taught to document their activity. Often, the wildlife rehabilitation veterinarians will have more volunteers than they can use; these individuals can be taught to conduct carcass searches and help with search documentation. It is also common for bird and wildlife groups to have pre-existing wildlife monitoring programs in place (for example plover monitoring, seabird stranding monitors, private refuge monitoring). These groups are generally pleased to provide documentation of their efforts to NRD responders.

4.1.1.4 NRD Specific Carcass Search Teams

This is the only group under the direct control of the NRD practitioner. These teams are deployed with the same documentation forms and carcass collection protocols as all other groups. These teams should be instructed to conduct searches in a systematic manner that maximizes search efficiency. Generally this entails one searcher searching the water line and the other searching the wrack line as they walk a shoreline. If the searchers must return to the starting point (round-trip search), they switch positions when they begin their return trip

4.1.2 Documenting Bird Search Effort

Search effort documentation occurs in two steps. The first step is to have the searchers document their searches and the carcasses they observe, if any. Each search should have its own documentation form (do not record multiple searches on a single form). Each search documentation form should contain at a minimum:

- The name of the group leader/supervisor;
- Contact information for the group leader/supervisor both during the spill response and post spill response;
- The number of individuals in the group;
- The group's purpose (SCAT, inspection, carcass search);
- Date, start time, end time;
- GPS coordinates for the start of the search and the end of the search
- Notice as to whether this was a one way search or a round trip search;
- The number of bird carcasses observed. If no carcasses are observed state "no carcasses observed"; and
- The fate of each carcass (e.g. collected and turned over to USFWS where USFWS assigned the carcass bird ID number D-107)

Inevitably, the rate of carcass encounters will drop as the response continues. When this occurs, documentation in the field becomes lax. It is critical that all searchers understand that a search on which no carcasses were observed provides as much information as a search on which 10 carcasses were recovered. Therefore, all searches are to be documented.

The second step of search documentation is the collection of all forms in a central location and the real time interpretation of that information. It may be helpful to have one NRD individual responsible for obtaining a copy of each SCAT form at the end of every day and bringing them to the central wildlife location. Similarly, one NRD individual should be assigned to collect forms from clean-up supervisors, one for collection from inspection teams, one for collection from NRD carcass searchers and one for collection from any other resource deployed to the field.

Each form is assigned a unique ID number (Remember, many groups may be collecting and cataloging forms so identify forms unambiguously such as Wildlife Search 1). Each evening those forms should be reviewed to ensure that all information is properly recorded and that every carcass observed has been collected and can be identified in the carcass collection vector. If data quality problems are identified, the NRD team can work with the group during the next day's tailgate safety briefing to ensure that the previous day's activity is properly documented and to prevent the duplication of the problem during future searches.

Every effort should be made to create a daily GIS map that documents each search, what type of group conducted it, and what was found. This mapping exercise can be used to ensure that all areas are receiving the appropriate level of search effort, to help identify trends in carcass deposition rates, and to identify the existence of non-spill related searches[3].

4.1.3 Distributing Bird Search Effort through Space and Time

The ability of the NRD practitioner to distribute search effort through space and time will have a direct bearing on the level of uncertainty associated with extrapolation of BBM results to date-location combinations for which there is no search information.

Extrapolation uncertainty is eliminated if all shorelines where potential deposition may be occurring are completely searched at least once a week. For relatively localized spills (deposition occurring over no more than 300 km of shoreline) and given sufficient resources, this may be possible. It has been our experience that the extra response costs associated with complete search effort are considerably less than the transaction costs associated with identifying reasonable extrapolation methods. If complete coverage is sought, it is critical to conduct searches at the geographic extremes of potential bird deposition and, each time a potentially spill-related carcass is identified at those extremes[4], to further expand the geographic extent of the search effort.

When assessing spills where deposition occurs over large areas or when resources are limited, it may be preferable to use stratified random sampling to select the bird segments that will be repeatedly searched. It is critical to sample throughout the entire geographic extent of potential deposition. It is also noted that the only resources at the direction of the NRD practitioner are the NRD carcass search teams. As such, the strain placed on that resource will be greater under a stratified random sampling procedure. Also note that, under a stratified sampling protocol, SCAT effort, inspection teams, wildlife recovery efforts, and public search is still likely to occur on the designated bird search segments. Documentation of these efforts is still critical.

4.2 Vector of Carcass Collections

The vector of carcass collections includes a description of every carcass collected during the spill response. That description should include at a minimum a unique bird identification number, the GPS coordinates of the collection location, and a link to the document that describes the search on which the carcass was collected. Ideally, documentation will also include a description of carcass oiling if any, a description of any carcass scavenging, photo-documentation, and an oiled feather sample collected for potential fingerprint analysis.

A NRD team member should be assigned the task of continuously updating the vector of carcass collections. This individual will also be responsible for collecting oiled feather samples and ensuring that appropriate chain of custody is established for those feathers. This team member will normally be stationed outside the central carcass collection area and, under ideal circumstances, will be working with USFWS law enforcement as they catalogue the carcasses for legal purposes.

Every carcass should be linked to a search documentation form and a GPS point describing its collection location. If, as is often the case, carcasses are brought in by non-spill related personnel, the NRD team member logging in the carcass should help the collector fill out the appropriate search form and identify an approximate collection location (Google Earth™ is an excellent tool for this). Under these

[3] Non-spill related searches are likely occurring in areas where carcasses are repeatedly collected but no documented search is occurring.

[4] Noting that un-oiled carcasses may be related to the spill, this determination is best made if the NRD practitioner can determine if the rate of carcass collections is elevated above background.

circumstances, the team member logging in carcasses is responsible for ensuring that the search form is filed at the central location and, if the public is being encouraged to conduct wildlife searches, the collector be directed to a place where he or she can receive appropriate health and safety and documentation training.

It is critical to note that USFWS law enforcement will be collecting data for their specific purposes and they should not be relied upon to provide NRD spill documentation. It is entirely possible that, at any point, all carcasses and all USFWS law enforcement documentation will become classified due to a criminal investigation and the NRD team will not be able to access those carcasses or records for years.

In our experience, two NRD personal should be dedicated to creating the vector of carcass collections during the emergency phase of the response, and one individual is needed during the operational phase.

4.3 Physical Removal

Carcasses can be removed from shorelines by natural physical processes including scavenging, burial, and rewash followed by sinking. The rate of physical removal is specific to time and location with rates varying from nearly 70 or 80 percent removal per day in some areas to no more than 5 percent over an entire week in other areas. Within any one geographic area, the rate has been shown to be dependant upon the size of the carcass (large carcasses tending to be removed at a slower rate), the condition of the carcass (fresh domestic carcasses being removed more rapidly than aged wild birds), and the habitat (carcasses in open areas tending to be removed more rapidly that carcasses in cover).

Given this level of variability, studies are generally conducted to estimate a site-specific rate of physical removal. In low energy areas such as bays and estuaries, these studies generally focus on scavenging only. In high energy environments, burying and rewash are also considered. The following text outlines procedures and considerations for a scavenging only study. If burying and rewash are to be simultaneously considered, the protocol would need to be modified such that carcasses are fitted with radio transmitters and floated to shore.

4.3.1 Generic Protocol and Considerations for a Scavenging Study

Carcass removal rates have been shown to be a function of the size and condition of the carcass. Therefore, it is best to use spill birds in the experiments. They do not necessarily need to be related tot eh incident being assessed and can often be obtained from USFWS[5]. To the extent practical, carcasses should mimic the size and species distribution associated with the spill being studied.

If carcasses from previous spills are used, each should be cleaned, stitched shut if the body cavity has been opened, and weighed at a logistics center. Carcasses are transported to shoreline access points where field personnel distribute them to predetermined locations based on a stratified random sampling design.

Wooden blocks with placards are placed beneath each carcass. These blocks serve two purposes. First, blocks float away in the case of an extreme high tide or wave action; this indicates re-floating rather than scavenging as the cause of carcass removal. Carcasses associated with blocks that can not be relocated would not be included in the scavenging data. Second, the placards reduce human interference by asking the reader to return any carcass they may have disturbed to its original position.

[5] USFWS often collects and archives carcasses for use in criminal investigations. Once that investigation is complete, they may be available for use in other studies.

The presence/absence and condition of each carcass is monitored at least one time per day; morning and evening checks may be necessary if, during the spill response, bird segments were searched more than once a day. If possible, monitoring will be conducted from a distance. Segments are monitored until all carcasses have been removed or until the data spans the greatest search interval observed in the search pattern. After the experiment, any carcass remains and wooden blocks are recovered and properly disposed of.

When fleshing out the preceding generic protocol, the NRD practitioner should be aware of the following potential constraints:

- Season and weather are likely to influence carcass removal rates. The proposed study should control for seasonal variation by collecting data on or around the spill anniversary.
- Carcass size may be an important factor in determining scavenging rate. Prior to being distributed, carcasses should be weighed. This allows the practitioner to estimate the relationship between carcass size and removal rate.
- Undocumented removal by humans and pets may be limited during a spill due to area closures. To evaluate and control for human interference during the experiment, a subset of carcasses may be continually monitored. These carcasses are not part of the scavenging rate determination.
- Carcass removal rates are likely to vary on a micro-environmental level. Placing few carcasses throughout the area of potential deposition via a stratified random distribution scheme, while logistically challenging, may provide more representative data, than placing many carcasses in a small area.
- Interstate transport and possession of bird carcasses may require a Federal Migratory Bird Special Permit.
- Carcasses may be classified as hazardous material. If so, proper permits for carcass distribution, handling, and disposal are required.

Depending upon logistical challenges and the complexity of the study, the design, implementation and analysis of a scavenging study can range in cost from $15,000 to $250,000. The majority of these costs are a function of factors beyond the control of the researchers. These factors include the difficulty in obtaining carcasses, travel costs, and the number of habitats to be assessed. One cost component that can be controlled by the researcher is the length of the study. Ideally, scavenging studies should continue until all carcasses are removed from the shoreline or one day longer than the longest search interval identified in the search pattern. As a rule of thumb, extending a scavenging study one additional day will increase total costs by about 5 percent.

It may be possible to conduct physical removal studies during the spill response. This would require the NRD team placing radio-tags on recently deposited carcasses and tracking their disposition until they are naturally removed from the system. This option poses many advantages including the simultaneous evaluation of all forms of physical removal during the actual spill and the use of actual spill-related carcasses.

4.4 Search Efficiency

Not all carcasses present on a shoreline during a search are found. Birds in the wrack line, small birds hidden in the rocks, and even birds on open sand can be hard to find. Bird size and coloration, as well as shoreline substrate will affect the proportion of birds present during a search that will actually be found. The mode of search (foot, boat, ATV, automobile), the rate of search, and the number of searchers also influence searcher efficiency.

Search efficiency can be estimated via literature transfer, during post-spill response studies, or during the actual spill response. Because existing search efficiency studies are often sufficiently detailed to control for the factors known to influence search efficiency, literature transfer may be the preferable method of search efficiency estimation. Literature transfer may not be a viable option if the habitat being searched is sufficiently different from previously studied habitats.

ENTRIX is aware of several studies that estimate search efficiency for shoreline areas. An unpublished paper by Ford *et al.* (no date B) summarizes the existing literature:

- Detection rates for King Eiders ranged from 44% to 94% on St Paul Island;
- On a sunny day and an unobstructed beach, one of nine carcasses were observed by a search team; and
- An average of 1 in 5 birds was missed by a trained searcher on an easy-to-search beach in the Orkneys.

That paper goes on to describe the Kure bird search efficiency study which was conducted along the Pacific coast in northern California's Humboldt County. The design of the Kure study controlled for several important variables including carcass size, carcass coloration, habitat type (large cobble and boulder, sand with extensive driftwood deposition, and wetland) and search mode (foot, truck, ATV). The search efficiencies reported in Table 4.4-1 are reported in Ford et al. (no date B). They were calculated to reflect the efficiency of a single searcher.

Table 4.4-1. Single Searcher Efficiencies Estimated in Kure Study.

	Very Small Carcass (cowbirds 12 to 20 cm) search efficiency (percent)	Larger Carcass (As large as cormorants) search efficiency (percent)
Sandy Beach (ATV)	12.5	43.9
Sandy Beach (Pick-up Truck)	3.1	40.9
Rocky Beach (On Foot)	27.9	55.3
Wetland (On Foot)	24.0	42.3

In 2004, ENTRIX conducted a more detailed analysis of the Kure search data and estimated foot-based search efficiencies for three carcass size categories and three search group sizes (Table 4.4-2).

Table 4.4-2. Search Efficiencies for Foot Based Searches on Large Cobble/Boulder Shorelines.

Carcass Size	1 Searcher (Percent)	2 Searchers (Percent)	3+ Searchers (Percent)
Small (>500g)	27	32	37
Medium (500 to 1599g)	55	67	73
Large (1600g or more)	80	91	96

If the NRD team opts to conduct a post-spill search efficiency study, the study should rely on searchers who actually participated in carcass searches during the response. These searchers should be instructed to search using protocols, modes, and rates actually employed during the spill. IN addition, the shorelines selected for the study should include those actually searched during the spill. It may also be prudent to control for season by conducting the study on a spill anniversary.

Typically researchers assemble test carcasses prior to the experiment. These carcasses are banded with low visibility bands and the size and coloration of each carcass is recorded. Next, researchers distribute these carcasses throughout the study area and record the precise location of each carcass[6]. After all carcasses are distributed, searchers are instructed to search the area at rates similar to those observed during the spill response and record the identification number and location of all carcasses observed. Ideally, only one search group is in the study area at a time and the study design is sufficiently rich to estimate search efficiency for all the unique combinations of search mode, habitat, carcass size, and number of searchers.

At the end of each experiment, researchers collect all carcasses and identify those carcasses which have been removed from the area by scavengers. The removal of these carcasses by scavengers is controlled for when search efficiencies are estimated.

It may be possible to estimate search efficiency during a spill-response by implementing the same experimental design outlined above. It may also be possible to implement an experimental protocol that eliminates potential study bias (the possibility that searchers increase their search intensity when they know they are being studied). This would be done by distributing tagged carcasses along bird segments just prior to a search without the searchers knowing that information. Researchers would again need to control for potential scavenging by recovering all carcasses not collected by the search team following the search. The individual constructing the vector of carcass collections and search pattern would also need to be informed of the study so they can remove those collections from the spill-related databases.

Depending upon logistical challenges and the complexity of the study, the design, implementation, and analysis of a search efficiency study can range in cost from $50,000 to $150,000. The majority of these costs are a function of factors beyond the control of the researchers. These factors include the difficulty in obtaining carcasses, travel costs, the number of habitats to be assessed, and the number of searchers who need to be evaluated.

4.5 Background Deposition

Carcasses unrelated to the spill (background birds) need to be netted out of total deposition estimates to determine the number of spill-related depositions. Ford (2006) published an analysis suggesting that the absence of physical oiling may not be a reliable indicator of spill-related mortality. Similarly, the presence of oil on a carcass does not necessarily imply the carcass experienced spill-related mortality. Therefore, the presence or absence of physical oiling on a carcass is generally not relied upon when adjusting for background deposition. Instead, existing beached bird monitoring efforts (SeaNet on the East Coast and Beach Watch monitors on the West Coast, among others) are used to estimate rates of background bird deposition.

Unfortunately, beach monitoring studies suggest that background rates of deposition are extremely variable from year to year and season to season. Hence, the application of an historical average rate of deposition introduces considerable uncertainty. Several statistical approaches are available to control for year-to-year variation. This is particularly true if data sets include time series information from areas both within and outside the geographic extent of deposition and if background deposition rates estimated for each area are correlated with one another.

In many cases, beach monitoring studies do not estimate a rate of background bird deposition. Instead, they report the number of carcasses encountered per mile of search. The use of background bird collection rates is complicated by a physical effect we call "build up." That is, beach monitoring data is generally collected about once a month. Therefore, the rate of carcass encounters is based on all

[6] If the habitat is such that easily identifiable trails are left, efforts to mask these trails may be necessary.

carcasses that have been deposited and persisted over the previous month. Such rates may be indicative of carcass encounters the first time a bird segment is searched. However, the rate of background carcass collections on subsequent searches will reflect only those carcasses held over from the previous search and background birds deposited since the previous search (i.e. there is limited carcass build up). Therefore, the rate of background bird collections on subsequent searches is expected to be lower than rates reported in beach monitoring studies.

It may be possible to gain insight into background deposition rates by going to shorelines outside the geographic extent of deposition and repeatedly searching and collecting carcasses from those beaches. Note that repeat searches are necessary. This is again because carcass "build up" results in relatively high background collection rates on first searches and lower subsequent search rates because subsequent search rates reflect primarily background birds deposited since the previous search

In addition to the estimation of a background bird deposition rates, the identification of specific background birds can be critical. This is the case when the collection of one or two carcasses at distant locations implies hundreds of additional spill mortalities after rates are extrapolated to unsearched shorelines. In such cases, uncertainty may be reduced by chemical "fingerprinting" of the oil to determine if the oil on the bird is spill-related. If possible, necropsy of these birds may also help determine their spill-relatedness.

4.6 Sinking

The proportion of carcasses that sink is conceptually dependent upon two factors: *buoyancy through time* and *time at sea*.

Buoyancy through time is reported in the scientific literature to be a function of carcass condition and environmental conditions:

- Ford *et al.* (1996) notes observational data suggesting fresh carcasses may float longer than carcasses that have been frozen;
- Nearly all scientists conducting sinking studies point out the intuitive assumption that intact carcasses will float longer than those which have been shot, punctured, scavenged or otherwise damaged;
- Studies conducted by Ford suggest that carcasses adrift in relatively sheltered areas float longer than those in open ocean conditions, and that this difference is statistically significant; and
- Weise (2003) points out that those carcasses likely to encounter fixed objects or debris are likely to sink more rapidly than those in relatively open water.

Time at sea is a function of the "time course of death" and physical forces acting upon the carcass.

- The time course of death describes where and when birds actually die at sea. Because the "time course of death" is generally impossible to know, the time course of oiling has often been used as a proxy (Ford et al. 1996). This simplification does not incorporate the active locomotion of oiled birds prior to death.
- The primary physical forces acting upon a carcass are winds, tides, and currents.

Buoyancy through time is fairly well understood. Ford *et al.* (1996) estimated sinking rates for bird carcasses under actual at-sea conditions by utilizing fresh carcasses that were oiled to varying degrees and fitted with radio-transmitters prior to release. One set of experiments was conducted in the relatively calm waters of Prince William Sound, the other in the more turbulent waters of the Gulf of Alaska.

Only 1 of 107 (less than 1%) of carcasses released into the Gulf of Alaska sank within 3 days; and only 5 of 107 (less than 5%) sank within 6 days. In three of the four releases in Prince Williams Sound, no birds sank until 6, 9, and 11 days had elapsed. In the fourth release, carcasses first sank 3 days after being released. In each of the four releases, only 1 or 2 carcasses (approximately 5-10%) sank before being deposited on shorelines (the releases continued until all birds had been deposited or sank, lasting 16-32 days). Median persistence rates were 15 to 20 days.

In three out of four releases in the Gulf of Alaska, no carcasses sank within six days of being released. For the fourth release, approximately 0% and 10% of carcasses sank within three days (depending on species) and 10% and 30% of carcasses sank within 6 days. Median persistence rates were 7 to 18 days for the Gulf of Alaska.

Since the literature that describes the rate at which carcasses sink is well developed and consistent, the primary BBM task is to estimate the time course of death. If data describing the timing and location of bird mortality has not been collected during the response, this input must be approximated. If the location of concentrations of un-oiled seabirds were recorded during the spill, this is accomplished by simulating the movement of the oil slick and determining the time and location at which birds and oil intersect. If data describing the location of birds has not been collected, it may be assumed the birds were uniformly distributed throughout the area swept by oil.

Once the time course of death is approximated, there are at least three oceanographic models that can estimate *time at sea*. They are COSIM (ENTRIX), SIMAP (Applied Science Associates), and the publicly available Type-A. These hydrodynamic models estimate the proportion of “bird carcasses” released at a specific time and location that have not been deposited on a shoreline by day X.

Note that underlying all time course of death analyses is the assumption that birds stop swimming (die) at locations where they are oiled. Coupling the observation that many birds live for some time after becoming oiled with the observation that oiled/injured birds typically move towards land implies that the proportion of birds dying near shore is likely underestimated. Sinking, therefore, is overestimated.

Though we are not aware of its ever having been implemented, the question of active movement of live birds could be addressed by fitting live oiled birds encountered at sea with radio transmitters and tracking their movement. Individuals have also attempted to address the question using hydrodynamic models to backcast the passive transport of a carcass from its time and location of deposition to its location at an assumed time of death (i.e. 24 or 48 hours after oil swept the area).

5.0 Estimating Mortality Using the BBM

As noted in Section 3, the various approaches to BBM (tabular, algebraic, and simulation modeling) generate similar mortality estimates when they employ similar assumptions. However, apparently minor variations in assumptions can cause mortality estimates to more than double. The goal of the BBM practitioner is to ensure that assumptions match the circumstances of the spill as closely as possible. This is best achieved by having two experienced modeling teams estimate mortality in a parallel stepwise manner. Section 5 first describes a parallel stepwise approach to BBM modeling that we have successfully employed in the past, and then identifies critical BBM assumptions and the considerations and data that can be used to inform those assumptions.

5.1 Stepwise Parallel Approach to BBM

Ford (1991) showed that the BBM is an unbiased estimator of total deposition under specific conditions. When those conditions hold, the level of uncertainty associated with any BBM mortality estimate is a function of the precision of the six input data sets and the level of consistency between the assumptions made by the BBM practitioner and the facts of the spill. In our experience, variation in assumptions alone can result in mortality estimates that diverge by more than 100 percent.

This section describes an approach to BBM designed to minimize the divergence between BBM assumptions and the facts of the spill. We refer to this approach as stepwise parallel BBM. In our experience, the incremental costs associated with a stepwise parallel BBM assessment may range up to $150,000 for a typical spill. Given the complexity of beached bird modeling and the associated potential for both simple error and divergent interpretations of facts, the stepwise parallel approach is likely to prove cost effective when expected bird liability exceeds $500,000.

In stepwise parallel BBM, two independent BBM modelers assess spill-related mortality. These independent efforts are calibrated in a stepwise manner to identify areas of divergence, clarify and reach consensus where information has been misinterpreted, and evaluate uncertainty where multiple interpretations are possible.

When compared to the more traditional approach in which a single BBM practitioner evaluates mortality and a second group reviews a report, the stepwise parallel approach offers several advantages. First, the parallel nature of the efforts generates two independent evaluations of the data describing the spill and the appropriate translation of those facts into BBM assumptions. If, via independent assessments, both BBM practitioners interpret facts similarly, it is likely that little uncertainty exists. If the BBM practitioners interpret facts differently, the stepwise calibration facilitates the elimination of simple error and the identification and quantification of true uncertainty where such uncertainty exists. The end result of this approach will either be a consensus mortality estimate or a range of mortality estimates where the source(s) of divergence are well understood.

This parallel approach is most efficiently implemented in a stepwise fashion.

Beached bird models, regardless of the specific version, require the same data related to search and carcass collection. Thus, the cooperative development of these data sets not only reduces uncertainty, it allows the modelers to develop constructive dialogue as they identify data gaps and develop data gap solutions. It also allows the BBM practitioners to develop a common vocabulary, and common bird search segments.

Once the search pattern and carcass collection vector have been finalized, the BBM practitioners review the available data related to physical removal, search efficiency, sinking and background deposition. They then develop recommendations regarding potential field studies and/or data analysis. The goal of this exercise is the cooperative development of functions describing each of these processes.

When data set construction is complete, daily spill related deposition for a single bird segment is estimated and results are compared. As was previously noted, the existing BBM models generate similar results (within about 5 percent) when they use identical data and assumptions. If divergent results are observed, it suggests that mistakes have been made or the modelers have interpreted the facts of the spill differently. If the former, mistakes can be corrected. If the later, the source of uncertainty can be identified and its effect quantified.

Next, daily spill related deposition is estimated for all dates and locations for which search data exists. Again, if the assumptions of the BBM practitioners are similar, results should be within about 5 percent. If results differ significantly, this calibration step allows for the correction of errors and the identification of true sources of uncertainty.

Once daily spill related deposition is estimated for all dates and locations for which search data exists, rates are extrapolated to the remaining dates and locations. If bird segments are sparsely distributed throughout the spill area, this extrapolation may need to consider whether or not the collection of carcasses from areas geographically removed from the spill imply deposition over a large geographic extent or whether those birds were likely background. This can be a major source of uncertainty which may be reduced via oil fingerprinting or necropsy.

The background rate of deposition is removed from the estimate of total deposition. This subtraction may consider whether or not live oiled birds collected during the spill could be background birds, and whether or not spill-related deposition was randomly distributed through space or time.

Finally, total spill-related mortality is estimated by adding in the number of carcasses that may have sunk prior to their deposition.

5.2 Critical BBM Assumptions and Data That Can Be Used To Evaluate Them

This section outlines a series of critical BBM assumptions and decisions that must be made by the BBM practitioners. In our experience errors and divergent interpretations of spill facts can lead BBM modelers to make assumptions that generate mortality estimates diverging by more than 100 percent. Once assumptions are compared and purged of errors, true uncertainty associate with a well-documented spill generally results in mortality estimates that diverge by no more than 30 percent.

5.2.1 Segmentation Scheme

During the spill response, bird search segments will have been established. Inevitably, the search pattern will reveal that some of these segments were incompletely searched during the spill response. We are aware of two potential methods for dealing with this occurrence. The simple approach adjusts search efficiency to reflect the partial search. For example, if search efficiency is assumed to be 60 percent and only 1/3 of a bird segment was searched, the BBM be implemented assuming that, the entire segment was searched but only 20% of the birds were collected during the search. The more complex approach is to divide the original bird segment. This results in more bird segments to model but each segment is always searched in its entirety.

Unfortunately the two approaches may generate different mortality estimates. To see this, consider a situation where a homeowner searches her 100 yards of waterfront daily and finds 1 carcass. The homeowners 100 yards are part of a 2,000 yard bird segment. For simplicity assume 100 percent search efficiency and no physical removal. By the simple approach, the collection implies deposition of 20 carcasses over the 2,000 m bird segment. The more complex approach implies 1 bird deposition and relies on the search data from the other 1,900 m to determine deposition on those 1,900 m.

5.2.2 Treatment of Birds Collected Between Searches

When carcass deposition occurs along a developed coast, carcasses will often be collected by non-spill personnel during events that can not be described as searches. These birds are defined as having been collected between searches. ENTRIX has encountered three approaches for modeling of these carcasses. The simplest approach is to remove them from the vector of carcass collections, estimate mortality, and then add the collections to the mortality estimate. A slightly more complex approach is to add these collections in after applying the average multiplier associated with all carcasses collected on searches.

The final and most involved approach is to calculate the probability that the carcass would have been collected during a subsequent documented search had it not been collected during the undocumented search. This probability is used to calculate the expected number of carcasses collected absent undocumented search effort. The beached bird model is then run on the expected number of carcass collections. For example, a carcass found 1 day before a document search, with a scavenging rate of 0.5/day and search efficiency of 0.5, would be treated as 0.25 carcass collected on the day of the documented search (1 x 0.5 x 0.5 = 0.25). If two additional birds were actually collected during the documented search, the beached bird model would be run as if the search resulted in the collection of 2.25 birds.

5.2.3 Treatment of Long Search Intervals

As previously noted, longer search intervals are associated with larger multipliers. In addition, longer search intervals are associated with increased uncertainty. This is because functions describing rates of physical removal become increasingly uncertain as they are estimated over longer time periods.

BBM modelers have noted that if carcass deposition and/or bird search effort are randomly distributed through space and time, this uncertainty can be reduced by discarding information related to longer search intervals and using extrapolation to estimate deposition rates for the days preceding those searches. We refer to that as a "short-BBM." However, if search effort is biased towards areas of high bird deposition[7] the short-BBM will be biased towards overestimation.

In previous assessments, some BBM modelers have simply asserted randomness and discarded data related to searches preceded by long search intervals.

5.2.4 Distribution of Carcass Deposition through Time

BBM practitioners commonly assume that the rate of bird deposition is constant within any search interval. In reality bird deposition is often observed to increase in the days following a spill, peak, and drop off. Other times, birds are deposited in pulses. Ford (1996) showed that, if pulses are sufficiency rapid and search intervals sufficiently short, the assumption of constant deposition generates an unbiased mortality estimate. However, when deposition increases, peaks, and decreases the assumption of uniform deposition may lead to biased mortality estimates.

It has been suggested that, rather than assuming uniform deposition, BBM models be set up to incorporate trends in deposition rates that manifest themselves in collections data. For example, if the rate of carcass collections on subsequent searches increases 10% per day for 1 week and then decreases at 10% per day for the next week, this pattern can be incorporated into the BBM. Similarly, it may be possible to use the

[7] Search effort may be biased toward areas of higher deposition if bird deposition is correlated with oil deposition (this is because SCAT and inspection effort is generally biased towards oiled areas) or if public and live wildlife search is directed toward known areas of species deposition.

traditional BBM assumption to generate trends in deposition rates and then to integrate those trends into a modified BBM in a two step process.

5.2.5 Distribution of Carcass Deposition through Space

When complete search has not been achieved, it may be necessary to assign a deposition rate from a searched segment to another, unsearched segment. Any extrapolation needs to control for the distribution of carcass depositions through space.

There are many statistical approaches for extrapolation and each has its own strengths and weaknesses. The critical point is that any decisions regarding extrapolations be based on estimated bird deposition rates not collection rates. To see this, begin by defining *carcass collection rates* as the number of bird carcasses collected per mile of search. In contrast, *deposition rates* are the number of birds deposited per mile of shoreline over the course of the entire spill. Note that the BBM is designed to translate *carcass collection rates* into *deposition rates* by accounting for things like search frequency.

Next, assume two beaches (X and Y) each 1 mile long:

- X is searched once a day every day for a week. The *deposition rate* is 2 birds per day.
- Y is searched once at the end of the week. The *deposition rate* is 1 bird per day.
- Search efficiency on each beach is 100 percent
- There is no scavenging on either beach.

Because search efficiency is 100 percent and there is no scavenging, we know we would collect 2 birds during each search of Beach X; 14 total collections during 7 miles of search. We also know we would collect 7 birds during the 1 search of beach Y; 7 total collections during 1 miles of search.

Beach X's *carcass collection rate* is 2 birds per mile of search (calculated as 14 birds collected ÷ 7 miles of search). The *carcass collection rate* on beach Y would be 7 birds per mile of search (calculated as 7 birds collected ÷ 1 mile of search).

In this example then, *carcass collection rates* are inversely related to *carcass deposition rates*. Note that Beach X has the higher *deposition rate* (2 birds per day) but the lower *carcass collection rate* (1 bird per mile of search). In fact, depending on the actual search frequency and scavenging rates, *carcass collection rates* and *carcass deposition rates* can be roughly equivalent, directly proportional, or inversely proportional.

Once carcass deposition rates have been estimated for available date-location combinations, the practitioner can use statistical methods to determine if extrapolation should control for factors such as the degree of shoreline oiling, shoreline orientation, or physical proximity to areas with deposition rates.

5.2.6 Methods for Netting out Background Deposition

The BBM practitioner faces two challenges when controlling for background. First, they must estimate the rate at which background carcasses were likely collected during first searches and then estimate the rate of background bird deposition during the spill time period. The specific details of this effort vary greatly and are largely a function of the available data set. As previously noted, background deposition can range from nearly zero to levels where background birds may represent the majority of collections. Therefore, few generalizations can be made.

The second challenged faced by the BBM practitioner is to identify methods that can be used to adjust for the presence of background birds. These methods must consider the following factors.

Not all oiled carcasses are related to the spill and not all un-oiled carcasses are un-related to the spill.

Due to carcass build up, the rate of collections on first searches will exceed the rate of collections on subsequent searches (Section 4.5)

Because bird/search specific multipliers commonly vary from less than 1 to over 100 for any one incident, the method used to adjust for background can have very large effects on estimated mortality. That is, it matters if we assume all birds collected during a spill response are equally likely to be background or if we note that birds with specific characteristic (live oiled birds, birds oiled by product not related to the spill) may be more or less likely to be spill related.

5.2.7 Treatment of Outlier Multipliers

The BBM estimates a unique multiplier for every bird collected during a spill response. It is not uncommon for those multipliers to be narrowly distributed around some mean with the exception of one or two multipliers that are obvious outliers. For example, Ford et al. (2001) reports that, for the New Carissa spill, a shoreline at the southern extent of potential spill-related deposition "received relatively little search effort, the one marbled murrelet found there was estimated to represent 120 marbled murrelet depositions." The average multiplier for the other 25 marbled murrelet collected during the response was 17.8. Given this scenario, the New Carissa Trustees and RP agreed to apply an average multiplier to the Marbled Murrelet in question rather than the multiplier calculated by the BBM.

In theory this decision was made because the assertion of 120 murrelet mortalities at this one point in space was unlikely in light of data describing murrelet densities and other spill specific data.

The appropriate treatment of these outlier multipliers is theoretically challenging. When all assumptions are met, the BBM is unbiased in the sense that, given a sufficiently large data set (or a sufficiently large number of spills), BBM estimates will approach the true level of deposition. Sufficiently large is defined such that random patterns of deposition in space and time approach uniformity. However, when sample sizes are smaller, the BBM estimate generated for any single spill will not necessarily approach the true level of deposition.

The BBM practitioner is then left to determine if their data set is sufficiently large and, if not, what, if anything can be done to address the issue.

6.0 BBM Sensitivity Experiments and Potential Refinements

In this section we assess the sensitivity of the BBM to several common occurrences and to several potential refinements. The results of the sensitivity experiments form the foundation of our BBM recommendations reported in Section 7. A total of eleven analyses were conducted. They include:

1. Variation in search frequency;
2. A failure to document search effort;
3. Miss-estimates of search efficiency rates;
4. Miss-estimates of persistence rates;
5. Alternative methods for incorporating partial segment searches;
6. Carcass collections that can not be associated with a search;
7. Treatment of long search intervals;
8. Non-constant deposition in time;
9. Non-constant deposition in space;
10. Alternative approaches to addressing background birds; and
11. A potential method to address outlier multipliers.

Detailed methods are reported in Appendix B.

6.1 Experiments and Approaches

The following describes the results of our 11 experiments and approaches

6.1.1 The Effect of Increased Search Frequency

One of the key questions faced by the BBM practitioner is at what frequency should segments be searched. As search frequency increases, BBM uncertainty decreases. However, the cost of mobilizing a bird search team to a spill likely range from $1,500 to $3,000 per day and that team is likely to cover no more than 5 miles of segment in any given day. As such, reducing uncertainty by increasing search frequency comes at a cost.

The purpose of this experiment is to provide the BBM practitioner guidance as they decide how frequently to search segments. To do this we estimated the expected error (the average of the absolute value of the difference between the true number of depositions and the estimated number of depositions) associated with the BBM as search frequency decreases from once daily to once every 10days.

For each search frequency, we created 50 search patterns; we required each segment to be searched on day 29 and held the probability of a search on other days to be uniform in space and in time. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in space and time, average daily persistence was 75 percent and average search efficiency was 75 percent.

The tabular BBM was used to estimate mortality for each of the 1,000 mortality scenarios associated with each of the 4 search frequencies. The error for any 1 mortality scenario is calculated as the absolute value of the difference between the true number of carcass depositions and the number of carcass depositions estimated by the tabular BBM. The expected error is the average taken over the 1,000 mortality scenarios.

Importantly, for all search frequencies, the tabular BBM generates an unbiased deposition estimate. That is, the estimated deposition averaged over the 1,000 mortality scenarios was equal to the true deposition. However, as search frequency decreases the expected error increases (Table 6.2.1-1).

It is also important to note that, as search efficiencies and or persistence rates decrease, the expected error increases. Hence, when a large portion of avian liability is likely to be associated with relatively small bodied birds (marbled murrelets and or plovers), which tend to be associated with lower search efficiency and persistence, 3 to 7 searched per week may be justified. If the primary source of avian liability is likely to be associated with larger birds (loons, pelicans) there is little reason to increase search frequency to more than once per week.

Table 6.1.1-1. The Relationship Between Search Frequency and Expected Error.

Daily Search	4%
One search every 2 days	7%
One search every 6 days	12%
One search every 10 days	15%

6.1.2 Failure to Document Search Effort

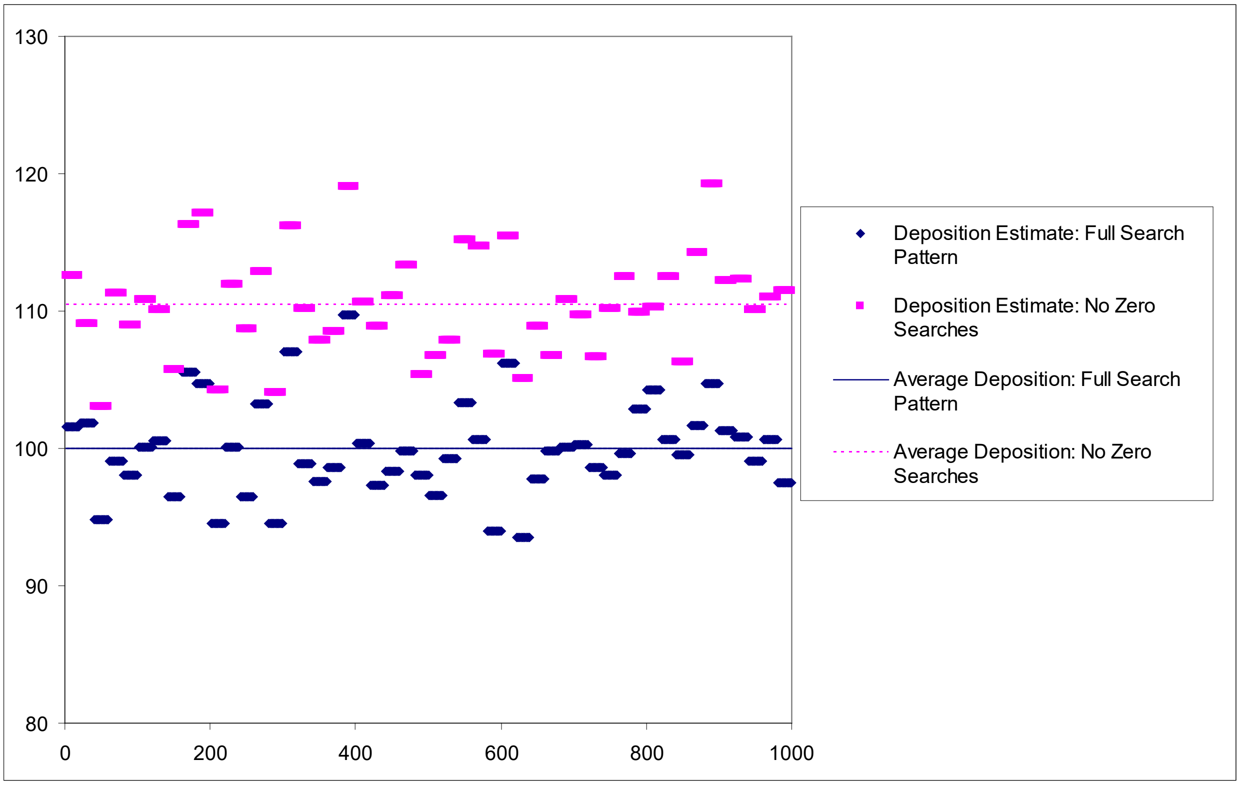
During a spill response it is common for wildlife search effort to go undocumented. This occurs most often when live bird recovery teams, SCAT, public groups, and or inspection teams fail to properly document searches on which no birds were found. This failure to document arises via one of two mechanisms. If documentation of a search exists but the absence of carcasses is not recorded, it is often asserted that the team may not have been looking for oiled wildlife and so their effort does not constitute search. Wildlife search also goes undocumented when wildlife search teams fail to document search efforts entirely because "they didn't see anything to report."

To investigate the sensitivity of the BBM to a failure to document search effort, we created 50 search patterns. We required each segment to be searched on day 29 and allocated 20 searches uniformly through time. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in space and time, average daily persistence was 75 percent and average search efficiency was 75 percent.

The tabular BBM was used to estimate mortality for each mortality scenario using the full data set. Each of the 50 blue dashes in Figure 6.2.2-1 is the average of the 20 mortality scenarios associated with a given search pattern. As expected, when operating on the full data set the BBM is an unbiased estimator (the average deposition estimate is 100). However, when searches on which zero birds were found are omitted from the data set, the BBM overestimates mortality by approximately 11 percent (Figure 6.2.2-1).

This overestimate of 11 percent reflects a failure to document approximately 22 percent of search effort. In previous spill assessments all search effort associated with SCAT teams, inspection teams, and live bird recovery teams has been omitted from the BBM assessment. This level of omission is likely to result in deposition estimates that may double the true level of deposition.

Figure 6.1.2-1 Bias Associated with Omitted Search.



Our approach for reducing the number of omitted searches is to have a dedicated BBM liaison on-site. That person coordinates closely with all SCAT, inspection teams, clean-up crews, and public groups to ensure that all searches are properly documented including searches on which no birds are observed. For a 1 month spill, the incremental cost of a dedicated BBM liaison is likely $15,000 to $35,000.

Assuming even the modest number of omissions identified in our experiment, and assuming the liaison can reduce the number of omissions by 50 percent, the presence of a BBM liaison is likely to reduce total bird related costs (assessment costs plus restoration costs) for spills where avian restoration costs are expected to exceed $700,000.

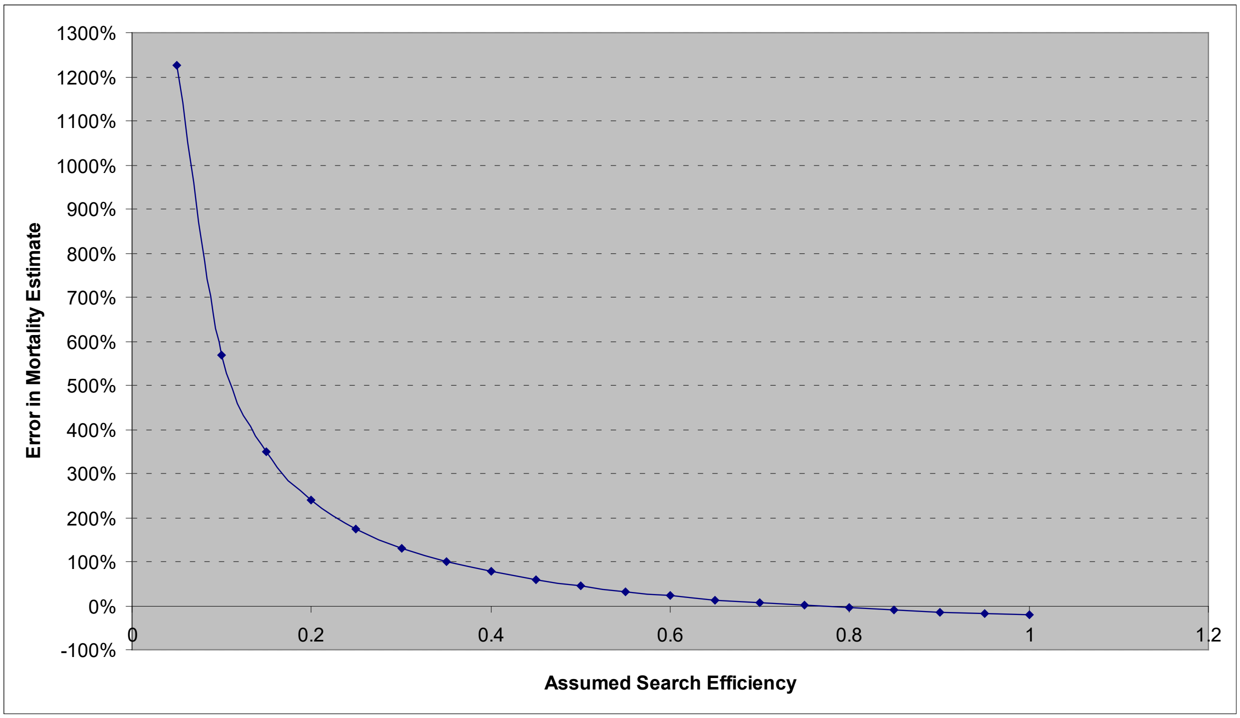
6.1.3 Misspecifying Search Efficiency Rates

Search efficiency rates can be estimated during a spill response, during a post-spill experiment, or values estimated for other spills can be utilized in lieu of site-specific experiments. The cost of a well implemented search efficiency study and the costs of the subsequent data analysis are likely to range from $50,000 to $150,000. These costs are largely avoided if existing search efficiency estimates are used. However, search efficiency rates can vary across sites and personnel and the reduced assessment costs associated with the use of existing search efficiency data increases the level of uncertainty associated with the final mortality estimate.

To investigate the sensitivity of the BBM to miss-specified search efficiency rates we created 50 search patterns, we required each segment to be searched on day 29 and allocated the remaining 20 searches assuming a uniform probability in space and time. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in space and time, average daily persistence was 75 percent and average search efficiency was 75 percent.

The tabular BBM was used to estimate deposition for each mortality scenario using a range of search efficiencies. When the true search efficiency is employed (in this case 75 percent), the BBM generates an unbiased estimate of deposition. If search efficiency is misspecified as 100 percent, the model underestimates total mortality by about 22 percent. When search efficiency is underestimated, the BBM overestimates deposition and the magnitude of this bias increases as the assumed search efficiency drops below 50 percent. When the assumed search efficiency is 50 percent, deposition is overestimated by 32 percent, when the assumed search efficiency drops to 40 percent; deposition is overestimated by 77 percent (Figure 6.2.3-1).

Figure 6.1.3-1. Bias Associated With Miss-Specified Search Efficiency.



Two additional facts should be noted in association with search efficiency. First, as the frequency of search increases, the error associated with misspecified search efficiency decreases. Our results are associated with an average of 1 search every 6 days. When search frequency is doubled, the effect of misspecifying search efficiency is approximately halved. Second, when compared to the effect of misspecifying persistence rates, the effect of misspecified search efficiency is relatively minor (Section 6.2.3)

The cost of designing, implementing, and analyzing a well designed search efficiency study is likely to range from $50,000 to $150,000. Noting that, for large carcasses (loons, eiders, geese) search efficiency rates are generally believed to exceed 50 percent per search, if potential bird liability is associated with larger birds, and if segments were searched at least once a week, a site specific search efficiency study may not be warranted. If the primary source of liability is associated with smaller birds, and if the species of small birds have high restoration costs as is likely the case for any west coast spill, a site specific search efficiency study may be warranted.

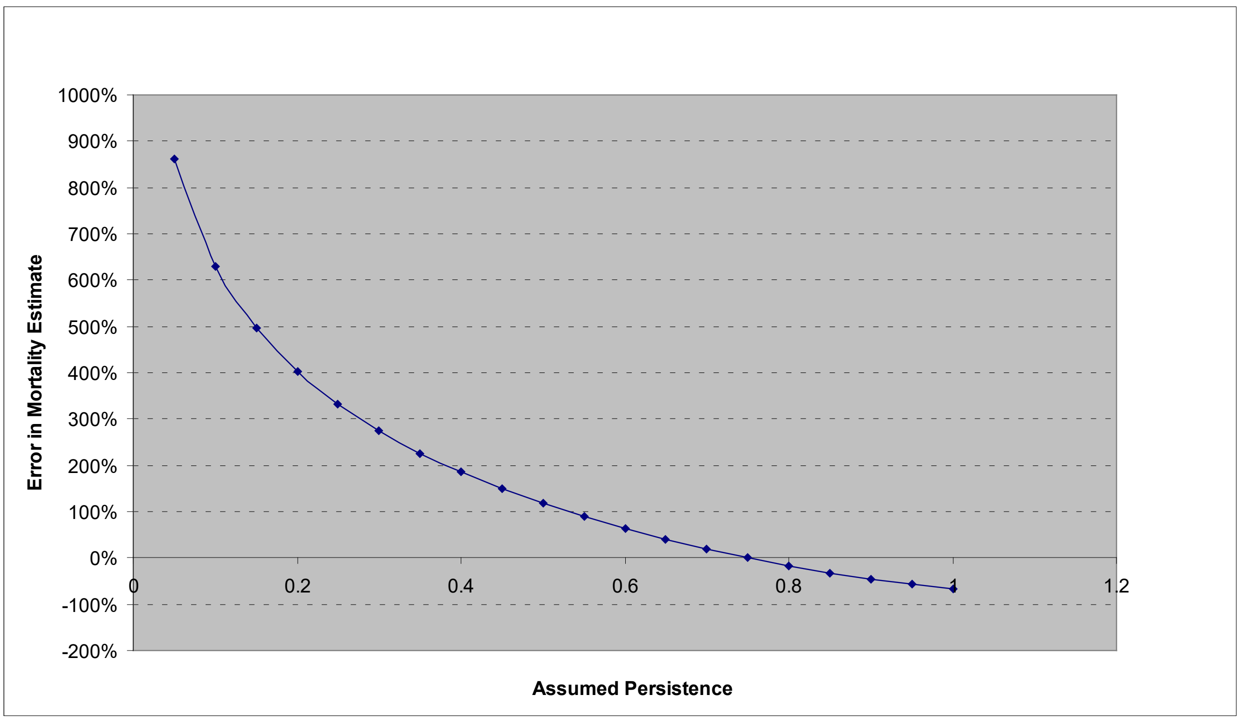
6.1.4 Misspecifying Persistence Rates

Persistence rates can be estimated during a spill response, during a post-spill experiment, or by using values estimated for other spills. The cost of a well implemented persistence study and post-spill analysis is likely to range from $150,000 to $250,000. These costs are largely avoided if existing persistence estimates are used. However, the level of uncertainty associate with transferring persistence rates from one site to another is significant.

To investigate the sensitivity of the BBM to misspecified persistence rates we created 10 search patterns, we required each segment to be searched on day 29 and allocated 20 searches assuming uniform probability in space and time. One hundred mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in space and time, average daily persistence was 75 percent and average search efficiency was 75 percent.

The tabular BBM was used to estimate deposition for each mortality scenario using a range of persistence rates (Figure 6.2.4-1). When the true persistence is employed (in this case 75 percent), the BBM generates an unbiased estimate of deposition. If persistence is misspecified as 100%, the model underestimates total mortality by about 66 percent. When persistence is underestimated, the BBM overestimates deposition and the magnitude of this error increases dramatically as the assumed persistence deceases. If true persistence is 75 percent and the rate is misspecified as 50 percent, the BBM deposition estimate is more than double the true level of deposition. When persistence is assumed to be 40 percent, the BBM deposition estimate is nearly triple the true level.

Figure 6.1.4-1. Bias Associated With Miss-Specified Persistence Rates.



The cost of designing, implementing, and analyzing a well designed persistence study is likely to range from $150,000 to $250,000. Given the observation that persistence rates can vary greatly even at the micro-spatial scale, the uncertainty reduction associated with a well designed persistence study is considerable. Even if avian liability is primarily associated with large carcasses which tend to have high persistence rates, the value of uncertainty reduction associated with a persistence study is likely to exceed the study cost in cases where avian liability is likely to exceed $1.25 million. If the primary source of liability is associated with smaller birds, and if the species of small birds have high restoration costs as is likely the case for any west coast spill, a site specific persistence is likely justifiable if total avian liability is likely to exceed $250,000.

6.1.5 Segmentation Scheme

During spill responses it is likely that a pre-established search segment will be searched but not in its entirety; we refer to this as a partial search. We are aware of two potential methods for dealing with partial search. The simple approach adjusts search efficiency to reflect the partial search. For example, if search efficiency is assumed to be 60 percent and only 1/3 of a bird segment was searched, the BBM is implemented while assuming that the entire segment were searched but only 20% of the birds were collected during the search. The more complex approach is to sub-divide the original bird segment such that all segments, when searched, are searched in their entirety. We refer to this as re-segmenting. Re-segmenting results in more bird segments to model but each is modeled in a theoretically consistent manner.

To investigate the sensitivity of the BBM to the two approaches, we created 50 search patterns. Unlike other simulations, we used only four search segments[8]. All searches on segment 1 were partial searches that cover either the northern half or the southern half of the segment. Searches on the other three segments were all complete. We required that all segments be searched in their entirety on day 29 and allocated 20 searches randomly in time with an average of 40 percent of the searches occurring on segment 1 and the remaining 60 percent of searches having an equal probability of occurring on segments 2, 3, or 4.

Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in time with an average of 40 percent of the depositions occurring on segment 1 and the remaining 60 percent of depositions randomly distributed to segments 2, 3, or 4. Average daily persistence was 75 percent and average search efficiency was 75 percent.

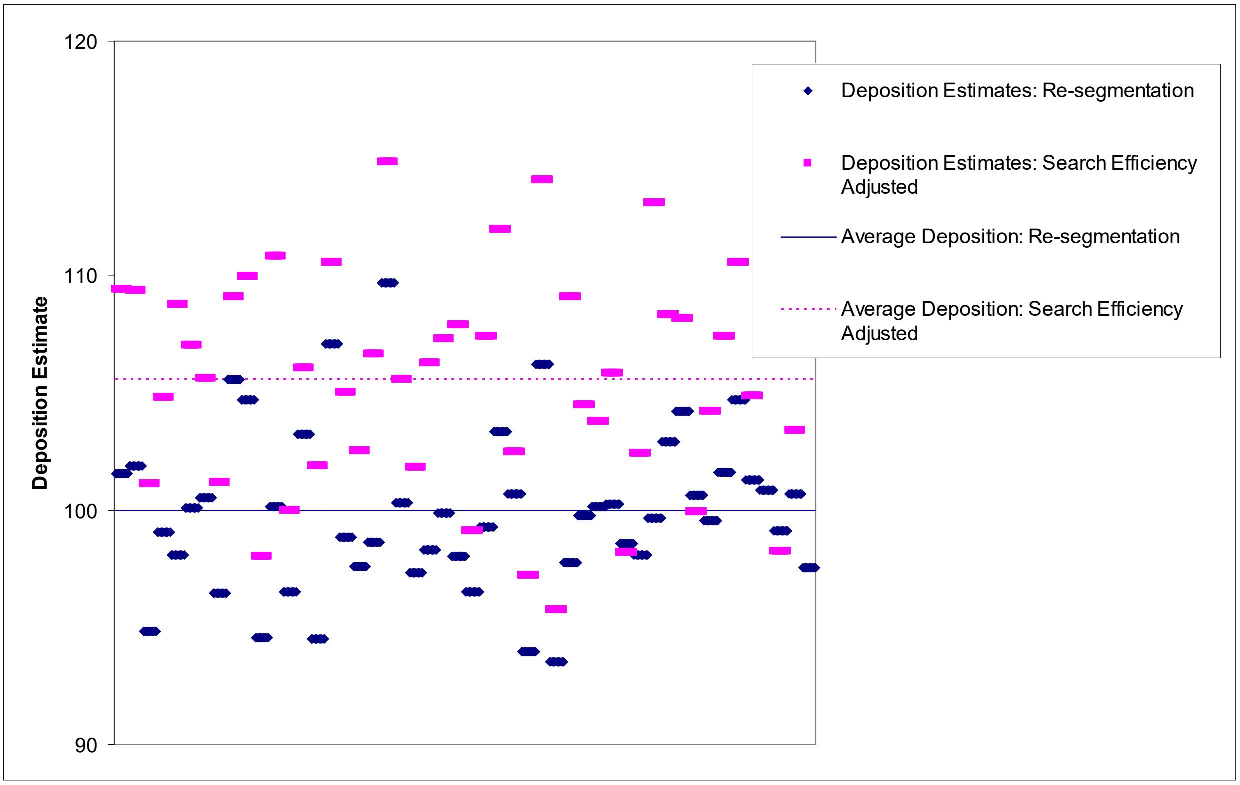
The tabular BBM was used to estimate mortality for each mortality scenario using the two alternative approaches. Each of the 50 blue dashes in Figure 6.2.5-1 is the average of the 20 mortality scenarios associated with a given "re-segmented search pattern." When re-segmenting is used to adjust for partial search the BBM proves to be an unbiased estimator (the average deposition estimate is equal to the true level of deposition). In contrast, the approach of adjusting search efficiency to reflect partial search biases results towards overestimation (Figure 6.2.5-1). In our simulation example, the magnitude of that bias is about 5 percent.

In our experience "re-segmenting" a spill of average complexity (75 original segments converted into 150 refined segments) is a fairly straight forward task that requires only moderate coordination among interested parties. The additional complexity is carried throughout the BBM analysis and is likely to increase total assessment costs by $5,000 to $10,000.

The effect of re-segmentation on the final deposition estimate is dependant upon the number, spatial distribution, and temporal distribution of partial searches. The results of our experiment suggest that the value of the bias reduction associated with re-segmenting is likely to exceed the cost of doing so when total expected avian liability is likely to exceed $200,000.

[8] The refinement was implemented to simplify our computer programming. It in no way alters the sign or approximate magnitude of the result.

Figure 6.1.5-1. Assessment of Alternative Schemes to Address Partial Search.



6.1.6 Carcass Collections That Can Not Be Associated With a Search

When carcass deposition occurs along a developed coast, carcasses will often be collected by non-spill personnel during events that can not be described as searches. These carcasses are defined as having been collected between searches. Similarly, spill related personnel may collect carcasses during a search but fail to document the search effort. This also results in carcasses being identified as having been collected between searches.

ENTRIX has encountered three approaches for the modeling of carcasses collected between searches. The simplest approach is to remove them from the vector of carcass collections, estimate deposition, and then add the number of carcasses collected between searches to the deposition estimate. A slightly more complex approach is to add these collections to the total deposition estimate only after applying the average multiplier associated with all carcasses collected on searches. The final and most involved approach is to calculate the probability that the carcass would have been collected during a subsequent documented search had it not been collected during the undocumented search. This probability is then used to calculate the expected number of carcasses that would have been collected absent undocumented search effort and BBM is run on the expected number of carcass collections.

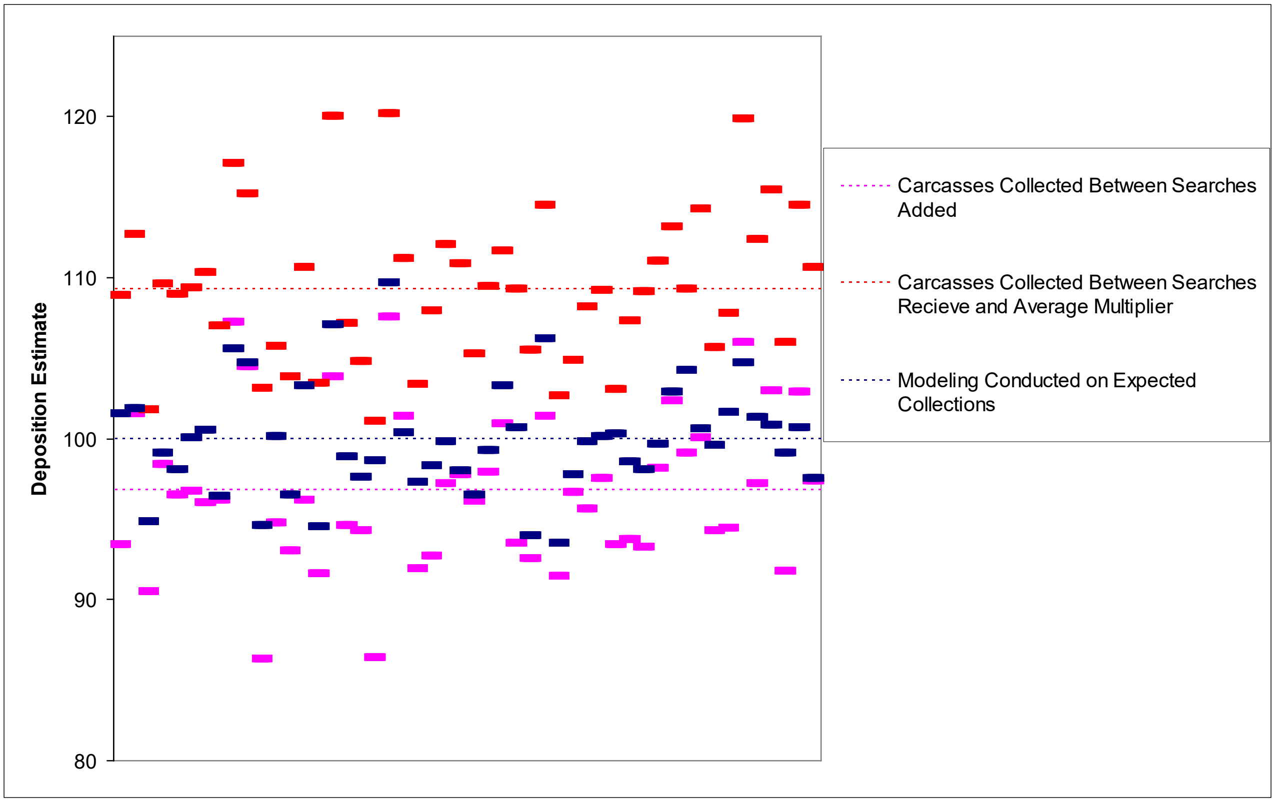
To investigate the three alternative approaches we created 50 search patterns; we required each segment to be searched on day 29 and allocated the remaining searches uniformly through time. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in space and time, average daily persistence was 75 percent and average search efficiency was 75 percent.

We modified the simulation algorithm to randomly select 5 segment days on which no search occurred. Any carcass present on one of these five segment days is assigned the final fate “collected between searches" and could not be found on any subsequent search. The remainder of the simulation algorithm was unchanged.

The tabular BBM was used to estimate deposition using the standard carcass collection vector and total deposition was adjusted for the carcasses collected between searches by adding the number of collections that occurred between searches to the deposition estimate (Approach 1). We also estimated deposition by adding the number of carcasses collected between searches to the BBM deposition estimate but only after applying the average multiplier to them (Approach 2). Finally, we estimated total deposition by calculating the expected number of carcasses that would have been collected absent undocumented search effort and ran the BBM on the expected number of carcass collections.

Figure 6.2.6-1 reports deposition estimates generated by the three alternative approaches. Applying the average multiplier biases the BBM toward overestimation by about 9 percent. In this simulation, the expected value approach and simply adding the number of carcasses collected between searches both perform well. The expected value approach being unbiased; adding 1 to total deposition for each carcass collected between a search underestimating deposition by just over 3 percent. We do note that as scavenging rates and or search intervals increase, we would expect the downward bias associated with simply adding the number of carcasses collected between searches to be exacerbated.

Figure 6.1.6-1 Modeling Birds Collected Between Searches.



In our experience, the expected collections approach is complicated by the fact that, outside of ENTRIX, most BBM practitioners have not developed automated algorithms for its implementation. As such they are required to carry out additional calculations and carry them throughout the BBM analysis. This may increase total assessment costs by $10,000 to $20,000.

If a data set has an unusually large number of birds collected between searches, relatively low persistence rates (below 50 percent per day), or relatively long search intervals (greater than 1 week), the expected value approach may be warranted. The expected value approach is also recommended if bird related liability is likely to exceed $5 million. Absent these conditions, simply adding the number of carcasses collected between searches to the total deposition estimate generated by a tabular BBM using a standard vector of carcass collections may be optimal. Under no circumstances should an average multiplier be applied to carcasses collected between searches.

6.1.7 Treatment of Long Search Intervals

Longer search intervals are associated with larger multipliers. In addition, longer search intervals are associated with increased uncertainty. This is because functions describing rates of physical removal become increasingly uncertain as they are projected over longer time periods. BBM modelers have suggested that uncertainty may be reduced by discarding information related to longer search intervals and instead using rates from nearby segments to estimate deposition rates for the days preceding long search intervals via spatial extrapolation. We refer to this approach as a "short-BBM." While ENTRIX believes the discarding of data is generally inappropriate and that preferred methods exist for addressing long search intervals, it can be demonstrated that, if either carcass deposition or search effort are randomly distributed in space, the short BBM will generate an unbiased deposition estimate provided the number of searches and or depositions is sufficiently large.

Importantly, if search effort is biased towards areas of high bird deposition the short-BBM will be biased towards overestimation. In our experience, both carcasses and search effort tend to be positively correlated with shoreline oiling and so the short-BBM would not be appropriate. However, other BBM practitioners have not identified a correlation between search effort and carcass deposition rates and, absent conclusive statistical evidence or a conclusive modeling test, chosen to implement a short-BBM.

ENTRIX has developed a modeling test that can conclusively identify bias in the short-BBM provided the correlation between deposition rates and search frequency is very strong. To do this we construct a short-BBM and a "test-BBM." We arbitrarily identify a maximum allowable search interval (in this case 5 days) and estimate mortality using a short-BBM. The test-BBM varies somewhat from the short BBM in that it acknowledges that the results of the search at the end of the long interval represents information upon which one can generate an upper bound estimate of the deposition that occurred during the 5 previous days. This is done by assuming that all collections made during a search that followed a long interval represent deposition occurring over only the previous 5 days. Rates within those 5 days are estimated by assuming zero birds were present at the beginning of the five day interval. For days beyond the 5 day test interval, rates are estimated via extrapolation. If mortality estimates generated via the test-BBM are less than those generated by the short-BBM search effort and or deposition are not random through space and time and the short-BBM will, unambiguously, result in an overestimate of deposition.

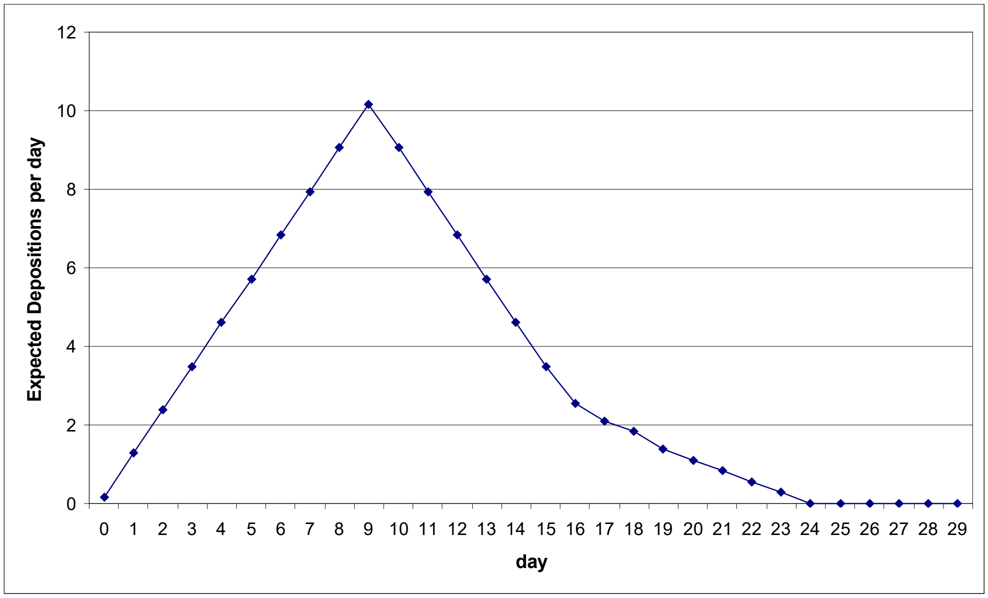
The cost of constructing, implementing, and explaining a test-BBM is likely to range from $1,000 to $5,000. It is recommended that, if a BBM practitioner conducts a short-BBM it always be tested. It is further noted that use of either a short BBM or a test BBM is unusual both in theory and in practice. Under most circumstances, the two step estimation process outlined in Section 6.1.8 is the preferred method for dealing with any issues related to long search intervals.

6.1.8 Distribution of Carcass Deposition through Time

The traditional BBM assumes the carcasses are deposited at a uniform rate during any search interval. In reality, deposition in most spills, increases, peaks, and then decreases through time. It may be possible to increase the accuracy of the traditional BBM by estimating deposition in two steps. First, a traditional BBM is used to predict daily deposition rates specific to each segment. These results are average across segments to generate an average deposition trend through time. The trend is then assumed in a second BBM run that is used to estimate the level of deposition.

To investigate the approach we created 50 search patterns, we required each segment to be searched on day 29 and allocated the remaining 20 searches uniformly through time. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, average daily persistence was 75 percent and average search efficiency was 75 percent. The pattern of carcass deposition through time is specified in Figure 6.1.8-1.

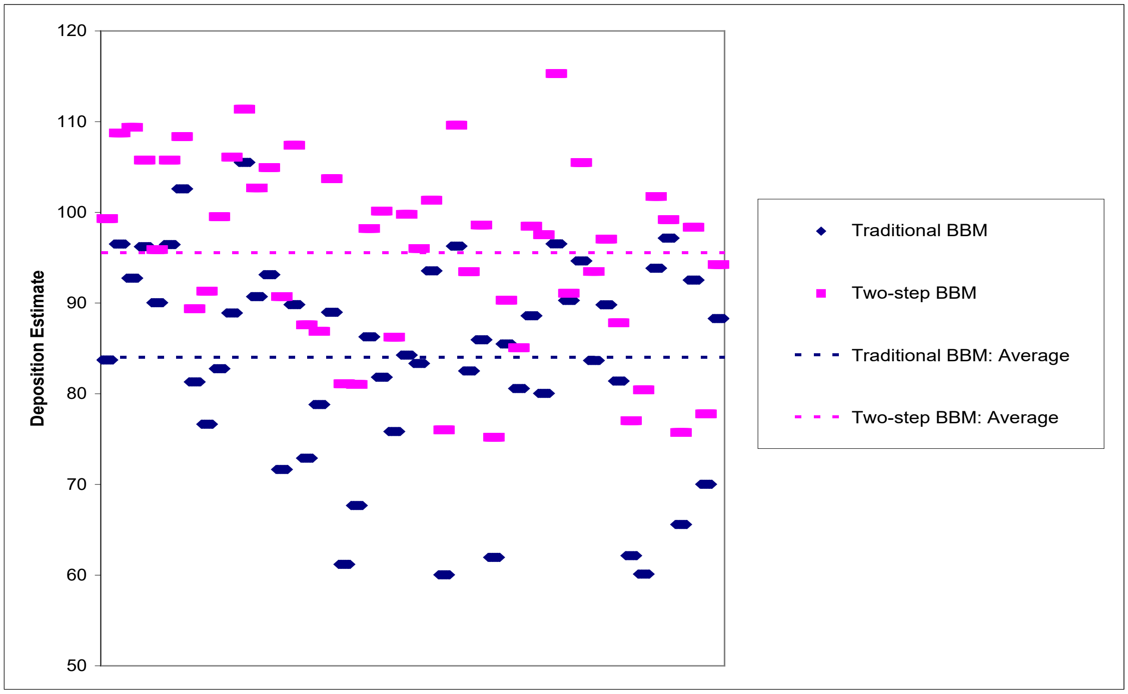
Figure 6.1.8-1 User Specified Pattern of Deposition Through Time.



Because the traditional BBM assumes a constant rate of deposition within any search interval but the rate of deposition was not constant, the traditional BBM generates a biased mortality estimate (Figure 6.1.8-2). However, the trend in daily deposition predicted by the traditional BBM does approximate the true deposition trend more accurately than the assumption of constant deposition. When that trend is used as an input to the second BBM run, the level of bias is decreased (Figure 6.1.8-2).

It is important to note that the sign and magnitude of potential bias is related to the actual pattern of deposition in time relative to the search pattern (that is, non-constant deposition can result in over or underestimates depending on the specific circumstance). However, the observation that two-step estimation may reduce bias given non-constant deposition rates is not dependant upon actual pattern of deposition.

Figure 6.1.8-2 Alternative Methods for Addressing Temporal Patterns of Carcass Deposition.



6.1.9 Distribution of Carcass Deposition in Space

If shoreline segments within the geographic range of potential avian deposition were not searched at all, or if they were not searched with sufficient frequency, the BBM practitioner may choose to extrapolate deposition rates from other segments to the unsearched areas. In the past BBM practitioners have done this by applying the average deposition rate from all searched segments to the unsearched areas. This approach is unbiased if and only if either search effort or bird carcass deposition is randomly distributed in space.

In contrast, if search frequency and *carcass deposition rates* are positively correlated[9] then any extrapolation of rates would need to adjust for that relationship. Recently a BBM practitioner suggested that "If deposition and search frequency are correlated there will be a positive relationship between the number of searches conducted on a segment and *deposition rate* (emphasis added)." However, the practitioner did not go on to test whether "there is a positive relationship between search frequency and *deposition rate*. Instead, they tested whether or not there is a positive relationship between search frequency and *carcass collection rates*. When no relationship was found, an average *carcass deposition rate* was applied to the less frequently searched areas.

To understand the potential technical flaw note that *carcass collection rates* are defined as the number of bird carcasses collected on a given segment divided by the total miles of search on the segment. In contrast, *deposition rates* are the number of birds deposited divided by segment length over the course of the entire spill. It is therefore possible that a segment could have a very high *carcass deposition rate* relative to other segments but because it was searched frequently is would have very many miles of search and therefore a relatively low *carcass collection rate*.

9 Search effort and bird deposition may be positively correlated if bird searches are drawn to areas where they have found carcasses in the past as is often the case with live bird search teams and or if the degree of shoreline oiling and the rate of carcass deposition are positively correlated.

Noting that the BBM is designed to translate *carcass collection rates* into *carcass deposition rates*, a more sensitive and theoretically justifiable analysis would test for a correlation between search frequency and estimated *deposition rates*.

To test this idea we created 50 search patterns applied to two segments, we required each segment to be searched on day 29 and allocated 9 additional searches to the first segment and 4 additional searches to the second segment. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 30. On average, twenty carcasses were deposited on segment 1 and 10 were deposited on segment 2. In both cases carcasses were deposited randomly in time. Average daily persistence was 75 percent and average search efficiency was 75 percent.

The results of our simulation suggest that, over the 1,000 mortality scenarios, the average *rate of carcass collection*, assuming each segment is 1 mile in length, was between 0.7 and 0.8 carcasses per mile of search for each segment. At the same time, the tabular BBM properly estimated a *carcass deposition rate* for Segment 1 of about 0.66 birds per mile per day and a *carcass deposition rate* for Segment 2 of 0.33 birds per mile per day. That is, relative carcass collection rates are not good indicators of relative deposition rates; BBM results are.

Given the results of our simulation, we conclude that, when assessing potential correlations between *carcass deposition rates* and search frequency (or any other variable of interest), the analysis should be performed on estimated *carcass deposition rates* not *carcass collection rates*.

ENTRIX also notes that other parameters (shoreline orientation, proximity to the epicenter of the spill, proximity to known areas of high bird density, degree of shoreline oiling) may be correlated with deposition rates. As such, it is prudent to conduct a regression analysis wherein the set of segment specific estimated deposition rates is the dependant variable and potential explanatory variables include, at a minimum, those listed above. The results of that regression analysis should be considered prior to any spatial extrapolation.

6.1.10 Methods for Netting out Background

One method used to adjust for background bird collection is to assume all collections are equally likely to be spill-related. Under that assumption the BBM practitioner:

1. Estimates the number of background birds collected by combining first search background bird collection rates, subsequent search background bird collection rates, and data describing the amount of first search and subsequent search.
2. Calculates the proportion of bird collections that were background as the total number of background birds collected ÷ the total number of birds collected
3. Estimates total spill related deposition as total deposition multiplied by the proportion of birds that were background.

This approach is recommended if and only if all birds have been assigned multipliers that are of similar magnitudes and if estimated deposition rates are above expected background deposition rates for all segments.

However, during many spill responses shorelines thought to be near or beyond the geographic extent of potential spill-related bird deposition are searched periodically. This “shoreline scoping” is done to ensure that the wildlife search effort is covering the appropriate geographic range. Because these shorelines are searched only periodically, two generalizations can be made. First, the longer intervals between searches suggest that, all else equal, the multipliers for any bird collected during “shoreline scoping” will tend to be disproportionately large. Second, and again related to the longer interval between searches, the rate at which background birds are collected (number of background birds per mile of search effort in the segment) will be elevated.

Under these circumstances, adjusting for background based on the assumption that all collections are equally likely to be spill-related, will bias results toward overestimation. This source of bias is exacerbated if spill-related deposition rates for geographically remote shorelines are overestimated (as would be the case if the subsequent search background rate were not adjusted to reflect relatively long intervals) and then those results are extrapolated to other unsearched shorelines.

While the magnitude of potential bias is incident-specific, Ford et al. (2001) reports the following for the New Carissa spill. Because a shoreline at the southern extent of potential spill-related deposition "received relatively little search effort, the one marbled murrelet found there was estimated to represent 120 marbled murrelet depositions." The average multiplier for the other 25 marbled murrelets collected during the response was 17.8. If it is assumed that all collections were equally likely to represent background deposition, spill related mortality would be 541. In contrast, if the marbled murrelet with the 120 multiplier was actually a background bird, total marbled murrelet mortality would be 445. This range represents several million dollars of liability given typical marbled murrelet restoration costs.

Below we report a theoretically appropriate method to adjust for the collection of background carcasses.

- Use beach monitoring data to estimate background deposition rates (not the rate at which background birds were collected).
- Combine search efficiency data, scavenging data, and the search schedule to predict the number of background birds likely to have been collected on each search during the spill.
- For each search, subtract the expected number of background collections from the actual number of birds collected on that search. This calculation generates a vector of spill-related collections.
- Finally, run the BBM on the vector of spill related collections.

The incremental costs associated with the approach outlined above are likely to range from $5,000 to $10,000. However, there may exist practical challenges related to the conversion of beach monitoring data to estimates of background carcass collections. Given these challenges the BBM practitioner may instead be limited to

- Using beach monitoring data to estimate first search collection rates and subsequent search collection rates.
- Ensuring that all shorelines where the expected number of background collections exceeds or is equal to the actual number of collections are assigned zero spill related depositions.
- Identifying any bird with a disproportionately large multiplier and ensure that oil fingerprinting, necropsy, and any other methods that can be brought to bear, are used to determine the spill-relatedness of the collection.
- For the remaining birds, estimating total spill related deposition as total deposition multiplied by 1 minus the proportion of carcass collections that were likely background.

The incremental cost associated with this background approach is broken into two categories: assessment costs and laboratory expenses. Assessment costs on a complex spill may range up to $5,000. Laboratory expenses including costs for Trustee and RP representatives to review and assess results may range up to $2,500 per fingerprint or necropsy.

Given the relatively modest assessment costs, it is recommended that, when data supports the estimation of background deposition rates, these rates be used to adjust for background. If background deposition rates can not be estimated, all segments should be screened to see if the number of carcass collections exceeded the expected number of carcass collections. Fingerprinting and necropsy may be appropriate when per bird restoration costs exceed $100 and or the bird specific multipliers are unusually large.

6.1.11 Methods to Address Outlier Multipliers

The BBM estimates a unique multiplier for every bird collected during a spill response. It is not uncommon for those multipliers to be narrowly distributed around some mean with the exception of one or two multipliers that are obvious outliers.

The appropriate treatment of these outlier multipliers is theoretically challenging. When assumptions are met (sufficiently large sample size such that random events are uniformly distributed) the BBM is an unbiased estimator of deposition. This argues against making adjustments based on outlier multipliers because any systematic adjustment will bias the model towards underestimation.

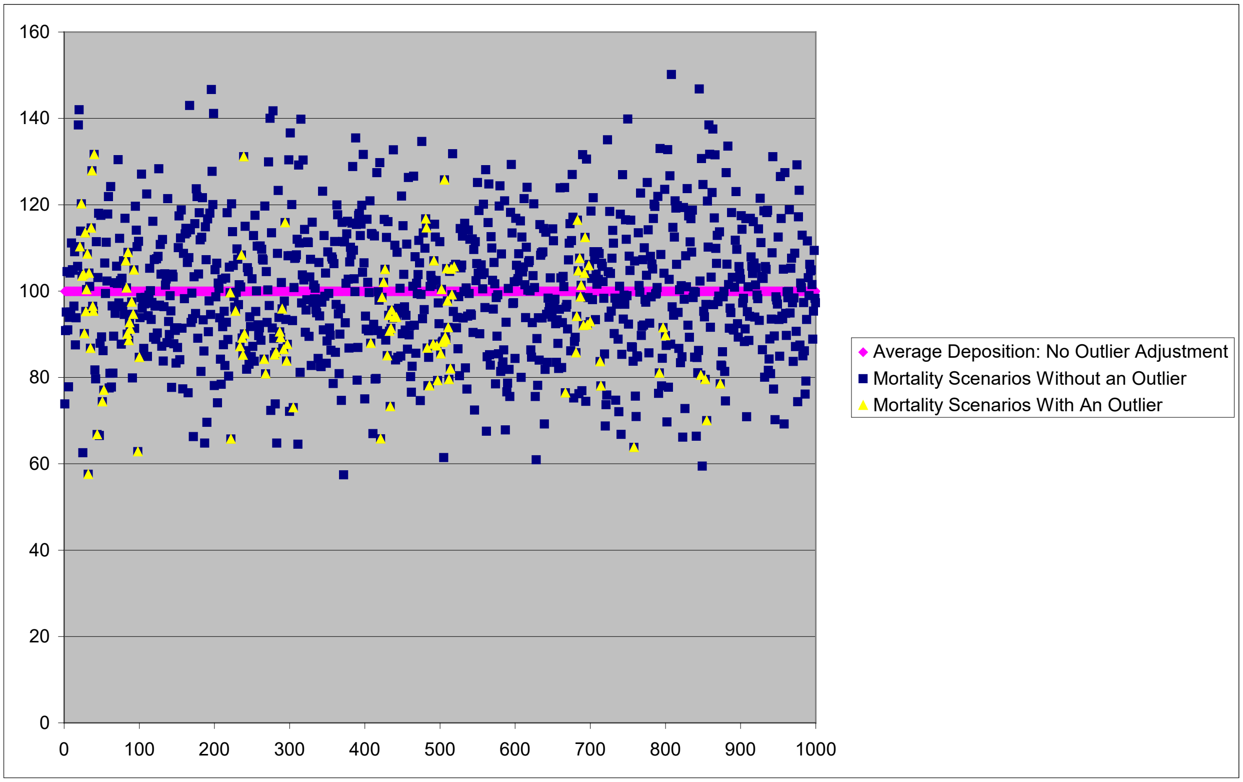
However, for any given mortality scenario, random variation can result in a deposition estimate that exceeds the true deposition by a considerable amount (±50 percent). ***If*** these large exceedances are driven by outlier multipliers that are generally absent among the mortality scenarios that are closer to the true level of deposition, a systematic identification and elimination of outliers may significantly increase the precision of the BBM estimates. It may be appropriate to trade some small level of bias for a large increase in precision if the goal of a BBM assessment is to accurately estimate deposition for the spill at hand.

To investigate the effect of adjusting outlier multipliers we created 50 search patterns, we required each segment to be searched on day 29 and allocated the remaining 20 searches uniformly through time. Twenty mortality scenarios were associated with each search pattern. For each mortality scenario the number of carcass depositions was 100, carcasses were distributed randomly in space and time, average daily persistence was 75 percent and average search efficiency was 75 percent.

Next, the tabular BBM was used to estimate deposition for each mortality scenario. As expected this estimate was unbiased. To investigate the effect of controlling for outlier multipliers we reviewed each of the 1,000 mortality scenarios one a time. We calculated the average multiplier for the mortality scenario and then identified any multiplier greater than 3 times the average. This multiplier was adjusted downward to the average.

Among the 1,000 mortality scenarios, approximately 20 percent contained a multiplier greater than 3 times the average for that mortality scenario. Figure 6.2.11-1 identifies those mortality scenarios which were adjusted. Noting that the outlier adjustment scheme is equally likely to identify a mortality scenario for which the BBM underestimates deposition as it is to identify a mortality scenario where deposition is overestimated, the scheme does not increase the precision of the BBM but does bias it towards underestimation. As such the systematic adjustment of outlier multipliers when not supported by exogenous data is not recommended.

Figure 6.1.11-1 Mortality Scenarios with Outlier Multipliers



7.0 BBM Summary and Recommendations

This section provides data to help the NRD responder determine if the incremental costs associated with a BBM assessment relative to a swept through assessment are justified. Assuming a BBM assessment is selected, this section goes on to identify the decision points typically encountered during a BBM assessment. These decision points are broken into 4 groups, mobilizing BBM personnel during an emergency response, managing search effort, decisions related to input data sets, and BBM set-up. For each decision point, we summarize the alternatives, their costs, and their likely benefits.

Appendix C contains a series of tables each relating to a particular phase of the BBM assessment. Within each table key issues are summarized and decision making tips are outlined. The purpose of this appendix is to function as a quick reference aid to the practitioner who may be in a position of making rapid decisions during an emergency response.

7.1 Swept Through or BBM Mortality Estimation

If all interested parties can not agree to a species specific level of mortality or species specific mortality literature transfer is not a viable assessment alternative. As such, most NRD practitioners must choose between swept through calculations and BBM methods.

- If repeat searches of shorelines where deposition may occur are not practical, swept through modeling must be implemented.
- If repeat searches are possible, the NRD practitioner must determine if the incremental costs associated with BBM are justified by the reduction in uncertainty associated with BBM relative to swept through calculations. In the unusual circumstance that the number of birds at risk of being oiled, the proportion of at risk birds that get oiled, **and** the proportion of oiled birds that experience spill-related mortality can be accurately estimated, swept through calculations may be the optimal assessment strategy due to lower assessment costs. This circumstance is likely to manifest itself only if birds in the area have been previously tagged and the degree of oiling, if any, can be determined for each of the tagged birds.
- If circumstances do not facilitate an accurate estimate of the proportion of birds that experience spill related mortality, swept through calculations are only accurate to within a factor of 2 or 3 (French-McCay and Rowe 2004). Our research suggests the expected error associated with a well implemented BBM analysis may be approximately 15 percent (Section 6.2.1). The incremental assessment cost of moving from “accurate to within a factor of 2 or 3” to an expected error of 15 percent is likely to range from $250,000 to $750,000 dollars.

Under the paradigm that, in the absence of data, assumptions that are “protective of the resource” will be employed in any swept through calculations, BBM appears to be the optimal avian assessment strategy if the proportion of oiled birds experiencing mortality can not be accurately determined and bird related liability is likely to exceed $1,000,000.

7.2 Emergency Response Staffing

During the emergency phase of an oil spill response it may be advantageous to have 1 to 3 experienced BBM modelers on-site.

1. A BBM coordinator is responsible for allocating bird search effort. This includes identifying the geographic extent of bird search effort, dividing the area into bird search segments, and determining if a stratified random sampling scheme or a complete shoreline sampling scheme will be implemented. This individual also determines the frequency with which each segments is to

be searched. These efforts are designed to minimize uncertainty associated with extrapolation of results through space and time.

2. A BBM data manager is responsible for ensuring that all data related to bird search effort and the collection of birds is fully documented and cross linked. Ideally, the data manager creates daily bird search maps. At a minimum these report the last time each bird segment was searched, the results of the last search, and the total number of birds collected from each segment to date. These maps are designed to assist the BBM coordinator and to ensure that data is efficiently organized for post spill analysis.
3. A BBM liaison is responsible for integrating the BBM data gathering effort into other facets of the response. They ensure that SCAT and Inspection teams are documenting their efforts such that they will be accepted as bird search effort. They also coordinate with live bird recovery teams, clean up crews, and the general public to ensure that efforts put forth by those groups are fully documented and qualify as bird search. The purpose is to avoid under-reporting of bird search effort and minimize the occurrence of birds collected between searches.

The cost associated with mobilizing a specialist ranges from $7,500 to $25,000 per specialist per week. Typically all three specialists may be required for the first 7 to 14 days following a spill. As the response moves into its operational phase the number of on-site BBM specialists is often reduced.

Noting the relatively high level of uncertainty associated with extrapolating results in space and time as well as the bias associated with the omission of searches (11 percent in our simulations Section 6.1.2), and improper treatment of partial searches (5 percent in our simulations Section 6.1.5), it is recommended that at least one BBM specialist be on-site when ephemeral data is being collected for the purposes of BBM. If the expected level of bird liability is likely to exceed $500,000, the benefit associated with mobilizing a specialist for each position will likely exceed the incremental assessment costs.

7.3 Managing Search Effort

During a response, the BBM practitioner will need to determine the geographic extent of search effort, the search frequency, and whether search will be complete or based on a stratified random sampling scheme. The benefit of increasing search frequency and the completeness of search is a reduction in uncertainty and the avoidance of assessment costs related to the identification of appropriate modeling and extrapolation methods. However, this benefit comes at the cost of mobilizing addition bird search teams at a cost of $1,500 to $5,000 per team per day. The purpose of Section 7.3 is to assist the BBM practitioner in managing search effort.

7.3.1 Geographic Extent of Bird Search

In our experience, search effort is often missing near the margins of the area of potential deposition. It is therefore often the case that deposition rates are extrapolated up-coast, down-coast, and inland not because birds were collected in those areas but because there was not sufficient search effort to demonstrate that birds were unlikely to have been deposited in those areas. Addressing these issues are often theoretically challenging and so “conservative assumptions” are used.

To avoid this issue, a search team can be assigned the task of repeatedly searching segments just beyond the geographic extent of potential deposition. If these teams encounter elevated levels of bird deposition and or oiled birds, the geographic extent of potential deposition is redefined to include these segments (and they must continue to be searched). New segments beyond the newly defined geographic extent of potential deposition are then identified and repeatedly searched.

For a spill with a 1 month bird search duration the cost of deploying a single team to constantly “search the edges” is likely to range from $45,000 to $100,000. In our experience issues related to the extrapolation of BBM results to and beyond the geographic extent of potential deposition are difficult to resolve (Sections 6.2.7, 6.2.8, and 6.2.9) and generally result in a 10 to 20 percent divergence in mortality estimates associated with most likely assumptions relative to assumptions that are "protective of the resource." As such, the value of uncertainty reduction associated with a search team dedicated to searching the edges is likely to exceed the incremental cost if total bird liability is likely to exceed $1,000,000.

7.3.2 Selecting Search Frequency

The results of the experiment described in Section 6.1.1 suggest that the BBM is unbiased over all reasonable search intervals. However, the level of uncertainty increases as search frequency decreases.

Given average search rates, a 100 mile geographic extent of deposition, and 1 month of bird search, it would typically cost $120,000 to $250,000 to completely search the area with a frequency of once every 6 days. Reducing average intervals to 3 days would double those costs and increasing intervals to 12 days would halve them.

Comparing those costs to the levels of uncertainty associated with various search frequencies (Section 6.1.1), under most circumstances it is reasonable to target search intervals of about 6 days. If very small birds that are costly to restore are likely to form significant proportion of total bird liability, as would be the case on nearly all west coast spills, a targeted interval of 3 days or less may be more appropriate.

7.3.3 Complete Search or Stratified Random Segments

A stratified random sampling algorithm is appropriate for spills that occur in remote coastal areas (Alaska), when the geographic extent of potential deposition exceeds 200 to 300 miles, and or if total expected bird liability is under $2 million. Absent those conditions, the cost savings associated with reduced search effort are likely to be exceeded by the incremental assessment costs associated with extrapolation and the costs associated with assumptions that are "protective of the resource" (Sections 6.1.7, 6.1.8, and 6.1.9) and lead to mortality overestimates.

7.4 Gathering Input Data Sets

Each BBM assessment requires 4 input data sets: persistence, search efficiency, background rates, and sinking rates. While it may be advantageous from both a cost effectiveness standpoint and an accuracy standpoint to gather these data sets during the response, practical constraints limit our ability to do so. As such, and with the exception of background rates, we evaluate the trade offs between various post-spill approaches for assembling these data sets. The goal is to assist the BBM practitioner as they proceed through the process.

7.4.1 Physical Removal of Carcasses

The cost of designing, implementing, and analyzing a well designed carcass persistence study is likely to range from $150,000 to $250,000. Given the observation that persistence rates can vary greatly even at the micro-spatial scale and the extreme sensitivity of the BBM to assumed persistence rates (Section 6.1.4) the value of uncertainty reduction associated with a persistence study is likely to exceed the incremental costs in cases where avian liability exceeds $1.25 million. If the primary source of liability is associated with smaller birds, and if the species of small birds have high restoration costs as is likely the case for any west coast spill, a site specific persistence study is warranted provided total avian liability is likely to exceed $250,000.

7.4.2 Search Efficeicny

The cost of designing, implementing, and analyzing a well designed search efficiency study is likely to range from $50,000 to $150,000. Noting that for large carcasses (loons, eiders, geese) search efficiency rates are generally believed to exceed 50 percent per search, if potential bird liability is associated with larger birds, and if segments were searched at least once a week, a site specific search efficiency study may not be warranted; the Kure search efficiency study is sufficiently well designed to facilitate search efficiency estimates for moderate to large birds over a range of terrain, search modes, and search team configuration. If the primary source of liability is associated with smaller birds, and if the species of small birds have high restoration costs, a site specific search efficiency study may be warranted due to the sensitivity of BBM results to minor miss-specifications when true search efficiency is relatively low (Section 6.1.3).

7.4.3 Background Rates

Rates of background bird deposition are hyper-variable in both space and time.

In ideal circumstances a beach monitoring group will be assessing shorelines within the geographic extent of potential deposition just prior to the spill. Under these circumstances the site specific data should be used to estimate rates of background bird deposition. If site specific data is not available, data from nearby shorelines monitored just prior to and during the spill response is the next best alternative. The availability of this data is often easily determined via an internet search or by contacting local universities.

If shoreline monitoring groups are not collecting data, it is recommended that a search team be deployed to repeatedly search a series of shorelines located an appropriate distance from the spill for the express purpose of estimating background bird deposition rates. This team should not collect carcasses but should instead tag birds and leave them in place. The cost of deploying a search team to an area outside the spill to collect background deposition data for two weeks would likely range from $21,000 to $42,000.

Noting that background birds have exceeded 20 percent of total spill related collections in at least one major spill, deployment of a search team to an area outside the spill is recommended if no shoreline monitoring data exists and if total bird related liability is likely to exceed $0.5 million.

7.4.4 Sinking Rates

The proportion of carcasses that sink is dependent upon two factors: *buoyancy through time* and *time at sea*. *Buoyancy through time* is well studies and reported in the literature (Section 4.6); carcasses generally float for a week or more. *Time at sea* is incident specific and depends upon the distance between the oil and the shore, the animals behavior once it becomes oiled, and local drift patterns.

For near shore spills where nearly all oil comes ashore within a week's time, it should be possible to agree that sinking is not an issue. For offshore spills and or spills where oil is not rapidly driven to shorelines trajectory modeling will be required. Noting that a primary source of uncertainty, the movement of live oiled birds toward shorelines, is difficult to incorporate into this sort of analysis, large expenditures designed to increase the precision of the trajectory modeling are not recommended.

7.5 Guidance on BBM Set-Up and Implementation

As noted in Section 3.0, the various approaches to BBM (tabular, algebraic, and simulation modeling) generate similar mortality estimates when they employ similar assumptions. However, apparently minor variations in assumptions can cause mortality estimates to diverge by more than 100 percent. The purpose of Section 7.5 is to provide guidance to the BBM practitioner as they generate BBM results.

7.5.1 Parrallel Stepwise Modeling

In parallel stepwise BBM, two independent BBM modelers assess spill-related mortality (Section 5.1). These independent efforts are calibrated in a stepwise manner to identify areas of divergence, clarify and reach consensus where information has been misinterpreted, and evaluate uncertainty where multiple interpretations are possible.

Depending on the level of scrutiny that would be applied to any Trustee generated mortality estimate, the incremental costs associated with a parallel stepwise approach could range from nearly zero to $150,000 for a moderately complex spill.

In our experience, variation in assumptions alone can result in mortality estimates that diverge by more than 100 percent. As such, parallel stepwise modeling is recommended in any circumstance where bird liability is likely to exceed $500,000.

7.5.2 Guidance on Model Set-Up

It is common for BBM modelers to discover patterns in data that suggest systematic search occurred but was undocumented. Given the bias associated with omitted searches (Section 6.1.2) it is recommended that attempts to document such search is made. In the past we have phoned individuals who consistently turn in oiled birds but who do not appear as participants on any documented search effort. We have also phoned residents and workers known to have been in the area. Such efforts are likely to cost $1,000 to $2,500.

It is common for birds to have been collected by bird searchers but to not be able to link those birds to specific searches. It is also common for birds to have been collected between searches. When persistence rates are near 75 percent per day, it may be acceptable to simply add these birds to the total mortality estimate (Section 6.1.6). If persistence rates are lower, it may be appropriate to conduct a BBM analysis on expected collections. The application of an average multiplier to these birds may lead to an overestimate of mortality.

It is common for segments to be incompletely searched. The theoretically appropriate method for addressing partial search is to create smaller segments such that, when segments are searched they are searched in their entirety. The value of the reduction in bias associated with re-segmenting is likely to exceed the cost of doing so when total expected avian liability is likely to exceed $400,000 (Section 6.1.5).

It is common for BBM practitioners to diverge in their approach for dealing with long search intervals when depositions rates may not be uniform in time. The preferred approach to dealing with this occurrence is to integrate the temporal pattern of deposition via a two step estimation process (Sections 6.1.7, 6.1.8, and 6.1.9). Some practitioners may suggest discarding all data related to long search intervals and replacing it via extrapolation. While this is not a recommended approach, it should, at the very least, be ground truthed against a test-BBM.

It is common to extrapolate deposition rates to areas that were not repeatedly searched during the spill. If search and or deposition are believed to have been random in space, it is appropriate to apply an average rate to unsearched areas. Absent this circumstance, it is recommended that regression analysis be used to identify variables that explain variation in deposition rates and that these variables be considered when deposition rates are extrapolated (Section 6.1.9).

It is common for some BBM practitioners to employ simplified methods for adjusting total deposition rates for the existence of background deposition (Section 6.1.10). It is recommended that any segment where the rate of carcass collection approximates the expected rate of background bird collection be assigned zero spill-related deposition and be removed from the analysis. It is also recommended that the BBM practitioner identify any bird with a disproportionately large multiplier and ensure that oil fingerprinting, necropsy, and any other methods that can be brought to bear, are used to determine the spill-relatedness of the collection.

It is common for bird specific multiplier to be narrowly distributed around some mean with the exception of one or two multipliers that are obvious outliers. Systematic adjustments to the outlier multipliers are not recommended (Section 6.1.11). However, exogenous data may warrant adjustment on a case specific basis.

Finally, every assessment is unique. Extensive methodological modifications and assumptions may be required on any given spill. As such, estimates generated by well intended BBM practitioners can diverge significantly. This divergence is most easily reduced by taking appropriate actions in the days and weeks following the spill, by meticulously documenting response efforts, and by implementing a stepwise parallel modeling approach.

8.0 Works Cited

Burger, A.E. 1991. Experiments to improve the assessment of mortality of oiled seabirds. Report to Environment Canada, Environmental Protection Service, Vancouver, B.C., Canada.

Burger, A.E. 1993. Estimating the mortality of seabirds following oil spills: Effects of spill volume. Marine Pollution Bulletin 26(3): 140-143.

Butler, R.G. and P Lukasiewicz. 1979. A field study of the effect of crude oil on Herring Gull (*larus argentatus*) Chick Growth. Auk 96:809-812.

Butler, R.G., A. Harfenist, F.A. Leighton, and D.B. Peakall. 1988. Impact of sub-lethal oil and emulsion exposure on the reproductive success of Leach's Storm-Petrels: short and long term effects. Journal of Applied Ecology 25, 125-143.

California Department of Fish and Game (CDFG), California State Lands Commission, and unites States Fish and Wildlife Service. 2004. *Stuyvesant/Humboldt* Coast Oil Spill. Draft Damage Assessment and Restoration Plan. Mat 2004. http://www.dfg.ca.gov/ospr/organizational/scientific/nrda/stuyvesant%20DARP/darp%20index.htm

Ford ,R.G., G.W. Page and H.R. Carter. (1987). Estimating mortality of seabirds from oil spills. In *Proceedings of the 1987 oil spill conference*, pp. 747-751. American Petroleum Institute, Washington, DC, USA.

Ford, R.G. Casey, J.L., Hewitt, C.H., Lewis, D.B., Varoujean, D.H., Warrick, D.R., and Williams, W.A. 1991. Seabird mortality resulting from the Nestucca oil spill incident, winter 1988-89. Report for the Washington Department of Wildlife. Ecological Consulting Inc., Portland Oregon, USA.

Ford, R.G., M.L. Bonnell, and D.H. Varoujean, G.W. Page, H.R. Carter, B.E. Sharp, D. Heinemann, and J.L. Casey. 1996. Total Direct Mortality of Seabirds from the Exxon Valdez Oil Spill. American Fisheries Society Symposium 18: 684-711.

Ford, R.G. G.K. Himes Boor, and J.C. Ward. 2001. Seabird Mortality resulting from the M/V New Carissa Oil Spill Incident February and March 1999. Prepared for U.S. Fish and Wildlife and Oregon Fish and Wildlife Office. Available at http://www.fws.gov/oregonFWO/Contaminants/Spills/NewCarissa/RestorationPlan/Documents/NewCarissaDARPAppendix4SeabirdMortalityReport.pdf

Ford, R.G. 2006. Using Beached Bird Monitoring Data for Seabird Damage Assessment: The Importance of Search Interval. Marine Ornithology 34: 91-98.

Ford, R.G. and M.A. Zafonte. 2008. "Scavenging of Seabird Carcasses at Two Oil Spill Sites in California and Oregon. Available at http://www.acjv.org/Marine_Bird_page/Working_Group_Materials/Oil_Spill/Ford%20&%20Zafonte%20Scavenging%20MS.pdf

Ford, R.G., N.A. Strom, R.G. Casey-Ford, and H.R. Carter. No date. Using Historical Oil Spill Data to Predict Seabird Mortality from Small Oil Spills. Accepted by the Marine Pollution Bulletin. Available at http://www.acjv.org/Marine_Bird_page/Working_Group_Materials/Oil_Spill/Ford%20et%20al.%20SmallSpills.pdf

Ford, R.G., B.E. Sharp, and M. Zafonti. No Date B. The efficiency of searchers recovering seabirds and waterfowl killed in the 1997 M/V Kure Oil Spill in Northern California. Report available at http://www.dfg.ca.gov/ospr/report/ssep/final_reports/search_efficiency.pdf

French, D., M. Reed, K. Joyko, S. Feng, H. Rines, S. Pavignano, T. Isaji, S. Puckett, A. Keller, F. French III, D. Gifford, J. McCue, G. Brown, E. MacDonald, J. Quirk, S. Natzke, R. Bishop, M. Welsh,

M. Phillips and B.S. Ingram. 1996. The CERCLA type A natural resource damage assessment model for coastal and marine environments (NRDAM/CME), Technical Documentation, Vol. I-Model Description. Final report, submitted to the Office of Environmental Policy and Compliance, U.S. Dept. of the Interior, Washington, D.C. April. 1996, Contract No. 14-0001-91-C-11.

French McCay, D. and J.J Rowe. 2004. Evaluation of Bird Impacts in Historical Oil Spill Cases Using the SIMAP Oil Spill Model. Applied Science Associates, Inc. Available at http://www.appsci.com/publications/pdf/McCay/FrenchMcCay_AMOP04.pdf

Henkel, L. A., R.G. Ford, W.B> Tyler, and J.N. Davis. 2007. Comparison of aerial and boat-based survey methods for marbled murrelets Brachyramphus Marmoratus and other marine birds. Marine Ornithology 35: 145-151.

Piatt, J. E, and R. G. Ford. 1996. How many seabirds were killed by the Exxon Valdez oil spill? Pages 712-719 in Exxon Valdez Oil Spill Symposium Proceedings (S. D. Rice, R. B. Spies, D. A. Wolfe, and B. A. Wright, Eds.). American Fisheries Society Symposium 18, Bethesda, Maryland.

Rhode Island Department of Environmental Management (RIDEM), National Oceanic and Atmospheric Administration, and U.S. Fish and Wildlife Service. 1998. Draft Restoration Plan and Environmental Assessment for the January 19, 1996 North Cape Oil Spill.

Sperduto, M., Powers, SP., Donlan, M. 2003. Scaling restoration actions to achieve quantitative enhancement of loon, seaduck, and other seabird populations. Marine Ecology Progress Series 264:221–232.

Walton, P., Turner, C.M.R., Austin, G., Burns, M.D., and Monaghan, P. 1997. Sub-lethal effects of an oil spill pollution incident on breeding Kittiwakes Rissa tridactyla. Marine Ecology Progress Series 155:261-268.

Wiese, F.K 2003. Sinking Rates of Dead Birds: Improving Estimates of Seabird Mortality Due to Oiling. Marine Ornithology 31: 65-70.

APPENDIX A

1.0 Appendix A: Alternative BBM Models

The following describes two approaches to BBM.

1.1 Algebraic Approach

This stylized version of the beached bird model might be applied if spill-related carcass deposition is limited to a single, uniform coastline. For illustrative purposes, assume that some portion of the coast is searched multiple times; the remainder is never searched. When located, bird carcasses are catalogued by time and location and removed from the area.

Let t denote time and $t = 0$ be the onset of deposition. Let t_i be the time when search i occurs. Define the time interval between t_i and t_{i-1} as time interval i. Then i indexes both searches and their preceding time intervals.

Let C_i be the number of carcasses collected on search i, and $\boldsymbol{C} = (C_1,\ldots,C_n)$ where n is the total number of searches. Similarly, let M_i be the multiplier applied to the carcasses collected on search i, and $\boldsymbol{M} = (M_1,\ldots,M_n)$. Then $C_i{*}M_i$, is an estimate of the total number of carcasses deposited on the shoreline during time interval i, and total deposition from the beached bird model is estimated as:

$$\boldsymbol{T}^{BB} = \boldsymbol{C}\bullet\boldsymbol{M} = \sum_{i=1}^{n} C_i M_i .$$

The beached bird model estimates the multiplier, the vector $\boldsymbol{M}$. Each element of $\boldsymbol{M}$ incorporates incomplete search effort, removal of carcasses by scavengers, imperfect search effort, and carcass holdover. Carcass holdover refers to the possibility that a carcass may have been deposited prior to interval i, neither scavenged nor found on previous searches, and then found on search i.

1.1.1 Incomplete Search

For simplicity, assume the affected coastline is homogeneous and has a constant rate of carcass deposition. Let P be the inverse of the proportion of coastline that is searched. Then, given estimated mortality in the area searched, multiplication by P expands this to the entire coastline. In practical applications with heterogeneous shorelines, more complex extrapolation schemes may be employed. In the past, we have extrapolated while controlling for variables such as proximity to oiled shorelines, proximity to other searched segments, and shoreline complexity.

1.1.2 Imperfect Search

Let F be the inverse of the proportion of carcasses that are present during a search and are found. For simplicity, we assume there is a constant "search efficiency" across searches and types of carcasses. In practice, search efficiency will vary across searches due to different numbers of searchers, different modes of search (e.g. foot, motorized, boat, etc.), different types of beach substrate (boulders, cobble, sand), or beach width. In addition, search efficiency typically varies by carcass size and coloration.

1.1.3 Physical Removal

The proportion of carcasses removed by scavengers, rewash, or burying is defined as a (probability) rate that may vary across time. Similarly, the deposition of carcasses may vary across time. Let d_t be the number of carcasses deposited at time t, and λ_t be the probability of physical removal at t. Define:

$$R_{t_i - t} = \prod_{s=t}^{s=t_i} 1 - \lambda_s \ ;$$

this is the number of carcasses deposited at t that persist until t_i. Then the proportion of carcasses deposited during interval i that are on the beach to be found during search i is:

$$E_i = \left(\frac{\sum_{t=t_{(i-1)}}^{t_i} d_t \times R_{(t_i - t)}}{\sum_{t=t_{(i-1)}}^{t_i} d_t} \right) .$$

Further define the inverse of this proportion as $S_i = 1/E_i$.

1.1.4 Holdovers

Finally, let H_i be the number of carcasses collected during search i that were not deposited during time interval i. H_i assumes non-zero values if search effort is imperfect, *i.e.,* a carcass might be deposited in interval j, neither scavenged nor found during search j, and then found on search i.

1.1.4.1 Estimating the Vector *M*

Ford (1996) notes that, the number of carcasses deposited during each time step (d_t) cannot be known with precision. In addition, the proportion of carcasses removed by scavengers would vary depending on the pattern of deposition. However, Ford also notes that, when computed over many observations, the expected value of S_i assuming deposition at a single, randomly selected step within a time interval is equal to the expected value of S_i assuming a constant rate of deposition throughout the time interval (Ford 1991).

Total mortality associated with the time interval prior to search i (T_i) is estimated by:

$$T_i = (C_i - H_i) \times P \times S_i \times F \ .$$

Since:

$$T_i = C_i M_i \, ,$$

we can equate these expressions for T_i and solve for M_i as:

$$M_i = P \times F \times S_i - \left(\frac{H_i \times P \times F \times S_i}{C_i} \right) .$$

1.1.5 Estimating Total Spill Related Deposition

Recalling that total deposition is estimated as:

$$\boldsymbol{T^{BB} = C \bullet M}$$

total spill related mortality is:

$$\frac{(C \bullet M) - B}{(1 - D)}$$

where B is the number of background birds deposited over the course of the spill and D is the proportion of spill-related carcasses that sunk before being deposited on the shoreline.

1.2 Simulation Approach

Unlike the tabular and algebraic models, the simulation approach does not estimate search-specific multipliers. Rather, it identifies total mortality estimates that are consistent with the number of carcasses collected on the known search pattern (carcass collection vector) by simulating the fates of individual carcasses.

The first step in the simulation approach is to generate multiple “mortality scenarios”. A mortality scenario is defined as a specific number of carcasses, $\hat{T}^k$.Each carcass is assigned a fate based upon four exogenously determined parameters: the proportion of carcasses that sink prior to being deposited, the proportion of the shoreline that is unsearched, a probabilistic physical removal function, and the probability that a carcass is found given it is present on a shoreline during a search. Potential fates are: sunk, deposited but not on a searched shoreline, scavenged, found on search i, or never found. These fates are used to determine the carcass collection vector (*i.e.,* the number of carcasses collected on each search i) associated with the mortality scenario.

The second step is to identify all those mortality scenarios with carcass collection vectors identical to the carcass collection vector from the spill. These form the set of “consistent mortality scenarios.” We estimate total mortality as the average number of carcasses associated with each of the consistent mortality scenarios.

1.2.1.1 Generation of Individual Mortality Scenarios

Each mortality scenario assumes a specified number of carcasses. Each carcass is assigned a fate using exogenously derived probabilities. This is done in a stepwise fashion.

First, each carcass is assigned a positive random number less than or equal to 1. If that random number is less than or equal to the proportion of bird carcasses that sink, the carcass is assigned a final fate “sunk;” otherwise the carcass is labeled “deposited.”

Next, each carcass labeled “deposited” is assigned a second positive random number less than or equal to 1. If that random number is less than or equal to the proportion of the shoreline that was unsearched, the carcass is assigned the final fate “deposited but not on a searched shoreline;” otherwise the carcass is labeled “deposited on a searched shoreline.”

Next, each carcass described as “deposited on a searched shoreline” is randomly assigned a unique hour of deposition sometime between the spill and the final search. Because each carcass has been assigned a specific hour of deposition and the time of each search that occurred during the spill response is known, the length of time between the assigned deposition and the first actual search during which the carcass might be found is known. The carcass persistence function $R(t,t_i)$ is used to calculate the probability that the carcass is removed by scavengers prior to search i. A random number less than or equal to 1 is drawn. If the random number is less than or equal to the probability of removal by scavengers, the carcass is assigned the final fate “scavenged in interval i”.

If the carcass is not scavenged in interval i, another random number is assigned. If the random number is less than or equal to the probability that a carcass is found given it is present on a shoreline during a search, the carcass is assigned the final fate “found on search i.” Alternatively, the carcass is labeled “missed on search i".

For any carcass labeled "missed on search *i*", the process is repeated for search *(i+1)* and the carcass could be assigned the final fate: "scavenged in interval *(i+1)*", "found on search *(i+1)*, or "missed on search *(i+1)"*.

The process continues until each carcass in the mortality scenario is assigned a final fate. Potential fates are: sunk, deposited but not on a searched shoreline, scavenged, found on search *i*, or never found.

1.2.2 Identifying Consistent Mortality Scenarios and Estimating Mortality

Letting $\hat{C}_i^k$ be the number of carcasses from the mortality scenario $\hat{T}^k$ assigned the fate "found on search *i*" for all the searches i = 1,…,n if this collection vector is identical to the actual carcass collection vector, then $\hat{T}^k$ is judged to be consistent with spill observations. That is, we say that $\hat{T}^k$ is a consistent mortality estimate if $(\hat{C}_1^k,\ldots,\hat{C}_n^k) = (C_1,\ldots,C_n)$. Let Γ be the index set of consistent mortality estimates, i.e. $\Gamma = \{k|\hat{T}^k \text{ is consistent}\}$.

Mortality scenarios, covering a wide range of total mortality specifications, are generated until Γ is large enough, (*i.e.,* until some appropriate number of consistent mortality estimates is identified). Let *K* be the number of consistent mortality scenarios in Γ. This procedure generates *K* consistent mortality estimates, conditional on the actual carcass collection vector and the exogenous probabilities specified. The expected total mortality, $\hat{T}$, is the average of the consistent scenarios:

$$\hat{T} = {}^{1}\!/_{K} \sum_{k\in\Gamma} \hat{T}^k$$

Expected total mortality is converted to expected spill-related mortality by subtracting the number of background birds deposited while spill-related deposition was ongoing.

APPENDIX B

2.0 Appendix B: Detailed D Experimental Methods

For the many of the analyses we conduct, we rely on three different computer tools. We refer to the first as the search generator. It creates search patterns with user-specified attributes. In this paper, potential searches can occur beginning day zero (the day of the spill) and on each of the 29 days thereafter. Each search can occur on one of five segments. The relative probability of a search occurring on any given day and on any given segment is user-specified. The user also specifies a target number of searches. The computer then randomly generates search patterns with those attributes. Table B-1.0-1 is a search pattern where the relative probability of a search on any given segment is equal, and where the relative probability of a search on any day less than 29 is equal. On day 29 we specified that all segments should be searched and our target for the total number of searches was 20. Note that on this particular search pattern the computer actually identified 25 searches.

Table B-1.0-1. Sample Search Pattern

Day	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
0		X			
1			X		
2	X				X
3					
4			X	X	
5	X	X			
6					
7					
8					
9					
10			X		
11					
12					X
13					X
14	X				
15					
16			X		
17				X	
18		X			
19	X				
20				X	
21					
22					
23					
24			X		
25					
26					X
27				X	
28					
29	X	X	X	X	X
Total Searches	5	4	6	5	5

The second computer tool we rely on simulates bird depositions and collections given a specific search pattern; we refer to the tool as a mortality scenario generator. Each mortality scenario includes a search pattern, and a user-specified number of carcasses each deposited at a specific location and time (locations and times are randomly determined but designed to conform to user-specified distributions). Each carcass associated with the mortality scenario is assigned a final fate using exogenously specified persistence and search efficiency rates. Potential final fates includes scavenged in interval i, found on search i, or never found.

When complete, the mortality scenario generator creates a data package that includes

- A search pattern;
- A vector of carcass collections (the number of carcasses collected on each search);
- The target persistence rate;
- The target search efficiency;
- The true number of carcasses deposited; and
- The targeted segment specific daily deposition rate.

Table B-1.0-2 reports the carcass collection vector associated with a hypothetical mortality scenario.

Table B-1.0-2. Data package associated with each mortality scenario.

Day	Segment 1 Birds	Segment 2 Birds Found	Segment 3 Birds	Segment 4 birds Found	Segment 5 Birds Found
0		1			
1			2		
2	2				2
3					
4			3	5	
5	5	3			
6					
7					
8					
9					
10			1		
11					
12					3
13					0
14	3				
15					
16			5		
17				9	
18		7			
19	0				
20				0	
21					
22					
23					
24			1		
25					

Table B-1.0-2. Data package associated with each mortality scenario (continued).					
Day	**Segment 1 Birds**	**Segment 2 Birds Found**	**Segment 3 Birds**	**Segment 4 birds Found**	**Segment 5 Birds Found**
26					6
27				8	
28					
29	5	4	2	1	1
Total Collections	15	15	14	23	12
Total Depositions 300, carcass deposition uniform in space and time, average daily persistence rate 75 percent, average search efficiency 75%					

The third computer tool we use is the tabular BBM. This BBM imports the data associated with a single mortality scenario and generates a total deposition estimate. It is then possible to compare the number of carcass depositions estimated by the BBM to the true number of depositions. By assessing thousands of mortality scenarios, we are able to assess both the accuracy and precision of the tabular BBM under various assumptions.

It is important to note that the tabular BBM was used as a matter of convenience. The results of these eleven assessments, had they been generated via the algebraic model or the simulation model rather than the tabular model, would not differ in a material sense.

An @Risk® model was used to generate sets of randomly drawn search patterns. In our simulations we have 5 segments and each of these segments can be searched on the day of the spill and for up to 29 days thereafter (30 days of potential search in total). The user specifies,

- A target total number of searches,
- The relative probability that a search occurs on a specific segment, and
- The relative probability of a search on any given day.

A search pattern is generated in 4 steps. First, the two relative probabilities are multiplied to estimate the relative probability of a search on any segment/day combination. For example, the modeler may decide that 60 percent of search effort should occur on segment 1 and that the remaining 40 percent should be allocated uniformly to the remaining 4 segments. Similarly, they may decide that 3.33 percent of the searches occur on any of the 30 days where search may occur. Given these exogenous variables, the relative probability of a segment 1 search on day zero is 1.99 percent (estimated as 60 percent multiplied by 3.33 percent). The relative probability of a search on segment 2 day zero is 0.33 percent (estimated as 10 percent multiplied by 3.33 percent). This relative probability is estimated for each segment day combination.

Next, the absolute probability of a search is estimated for each segment day. This is done by multiplying the relative probability of a search on any given segment/day by the target total number of searches. Continuing on with the example from the prior paragraph, if the target number of searches is 20, the absolute probability of a segment 1 search on day zero is 39.8 percent (calculated as 1.99 percent relative probability of a search multiplied by 20 searches), whereas the absolute probability of a search on segment 2 search on day zero is 9.9 percent (calculated as 0.33 percent relative probability of a search multiplied by 20 searches). This exercise results in a specific absolute probability of a search for each of the 150 potential segment/day combinations (i.e. 30 days of potential search and 5 segments implies 150 segment day combinations).

In the third step, a random number between zero and 1 is assigned to each segment/day combination. If the random number is less than the absolute probability of a search, a search is assumed to occur on that segment day. Finishing the example, if the random number for segment 1 day zero is 0.35, then a search is assigned to segment 1, day zero. If the random number assigned to segment 2 day zero is 0.11, no search is assigned to segment 2 day zero.

Finally, all segments are assumed to be searched on the last day of the simulation. This is done because, in practice, searches are typically conducted until spill related carcass deposition ends. That is, if search effort is halted before bird deposition ends, all models will be biased towards underestimation unless methods are developed to estimate deposition in the unsearched time period..

A single search schedule is shown below in Table 6.0-1. Multiple search patterns were often generated for each analysis.

Mortality scenarios were also generated by an @Risk® model. Each mortality scenario includes a previously generated search pattern, and a user-specified number of carcasses each deposited at a specific location and time (locations and times are randomly determined but designed to conform to user-specified distributions). Each carcass associated with the mortality scenario is assigned a fate using exogenously specified persistence and search efficiency rates. Potential final fates includes scavenged in interval *i*, found on search *i,* or never found.

To begin, the mortality scenario generator selects a previously constructed search pattern. The user specifies:

- The number of carcasses in the mortality scenario;
- The relative probability a carcass is deposited on any 1 segment;
- The relative probability a carcass is deposited on any given day;
- The probability a carcass will persist from one day to the next (persistence rate); and
- The probability a carcass is found provided it is present during a search (search efficiency).

In our analysis all carcasses are deposited on one of the five segments (i.e. sinking is zero). This assumption simplifies the analysis without loss of generality. Next, each carcass is assigned a positive random number less than or equal to 1. That number is used to assign each carcass to a segment based on a discrete distribution between 0 and 1 and the user-specified relative probability of deposition on any given segment.

Next, each carcass is assigned a day of deposition; this day may include day zero. To do this each carcass is assigned another positive random number less than or equal to 1. That number is used to assign each carcass to a day of deposition based on a discrete distribution between 0 and 1 and the user-specified relative probability of deposition on any given day.

Because each carcass has been assigned a specific segment and day of deposition, and because the date and location of each search is known, the length of time between deposition and the first actual search during which the carcass might be found is also known. The carcass persistence function $R(t,t_i)$ is used to calculate the probability that the carcass is removed by scavengers prior to search *i*. A random number less than or equal to 1 is drawn. If the random number is less than or equal to the probability of removal by scavengers the carcass is assigned the final fate "scavenged in interval *i*".

If the carcass is not scavenged in interval *i* another random number is drawn. If the random number is less than or equal to the probability that a carcass is found given it is present on a shoreline during a search, the carcass is assigned the final fate "found on search *i*." Alternatively, the carcass is labeled "missed on search *i*.

For any carcass labeled "missed on search *i*" the process is repeated for search *(i+1)* and the carcass could be assigned the final fate "scavenged in interval *(i+1)*", "found on search *(i+1)*, or it could be labeled "missed on search *(i+1)*. The process continues until the carcass is assigned a final fate. Potential fates are scavenged in interval *i*, found on search *i*, or never found.

When complete, the mortality scenario generator creates a data package that includes

- A search pattern;
- A vector of carcass collections (the number of carcasses collected on each search);
- The target persistence rate;
- The target search efficiency;
- The true number of carcasses deposited; and
- The true segment specific daily deposition rate.

This data package is used to assess the performance of the tabular BBM. Inputs to the tabular BBM include the search pattern, the vector of carcass collections, the target persistence rate, and the target search efficiency. The tabular BBM outputs (total estimated deposition and segment specific daily deposition rates) are compared to the true number of carcass depositions and the true segment specific daily deposition rates.

The tabular BBM is described in Section 3.1.1 and implemented in @Risk®.

APPENDIX C

3.0 Appendix C: Quick Reference Aid to Beached Bird Modeling

This appendix is intended to function as a quick reference aid for the practitioner who may be in a position of making rapid decisions during an emergency response. The appendix contains a series of tables each relating to a particular phase of the BBM assessment. Within each table key issues are summarized and decision making tips are outlined. The single most important tip is to ensure that a person experienced with beached bird modeling is participating in the decision making processes.

Table C-1.0-1. Quick reference staffing guide

Position	Duty	Number of Participants	On Site Week 1	On Site Week 2	On Site Week 3	On Site Week 4
BBM Coordinator	Define Geographic extent of potential deposition. Identify shorelines to be searched. Determine target search frequency.	1	Y	Y	Y	Y
BBM Data Manager	Assure all carcass collections and all search efforts are fully documented. Provide real-time mapping and carcass collections data to support coordinator's decision making.	1	Y	Y	Y	
BBM Liaison	Work with SCAT, live bird recovery team, and public groups to ensure that all search efforts and documentation are suitable for use in BBM analysis.	1	Y	Y		
Bird Search Teams	Search designated shorelines for bird carcasses. Collect observed bird carcasses. Document all search effort.	Sufficient to achieve coverage and frequency identified by BBM coordinator.	Y	Y	Y	Y
Geographic Extent team	Conduct searches at the geographic edges of the spill to ensure that the entire area of potential deposition is addressed.	2	Y	Y		
Background Team	Identify all existing sources of potential background bird collection rates. Conduct searches specifically intended to determine background collection rates in nearby areas that did not experience spill related deposition.	2			Y	

Table C-1.0-2. Quick reference to potential bird search effort.		
Group	**Normal Activity**	**Required Modification**
SCAT/Inspection Teams	Surveying shorelines and documenting areas searched an oil observed	Add carcass observation or document absence of carcass on existing forms. Collect or report carcasses for collection when observed
Live Wildlife Collection Team	Searching for live oiled wildlife and responding to reports of live oiled wildlife	Document the start and endpoints of any searches. Record carcasses observations or absence of carcasses. Collect or report carcasses for collection when observed.
Public/Ongoing Bird Monitoring Efforts	Public frequently searches shoreline stretches during spills and groups such as Audubon frequently have ongoing bird monitoring efforts in spill areas	Train participants to search the areas they frequent for carcasses. Document the start and endpoints of any searches. Record carcasses observations or absence of carcasses. Collect or report carcasses for collection when observed.
Carcass Search Teams	Searching shorelines for carcasses	None

Table C-1.0-3. Quick reference for creating a vector of carcass collections.	
For Each dead bird collected	**For each live bird collected**
Collection date	Collection date
Collection location (latitude and longitude)	Collection location (latitude and longitude)
Species	Species
Degree of oiling if any	Degree of oiling if any
Degree of scavenging if any	Feather sample identification number if any
Feather sample identification number if any	Picture identification number if any
Picture identification number if any	Unambiguous link to search effort database
Unambiguous link to search effort database	Final fate (died, released to wild, released to protective care)
USFWS law enforcement may, at any time, stop sharing and allowing access to information. It is necessary that NRD has its own database and database construction methods.	

Table C-1.0-4. Quick reference for construction of non-ephemeral data sets.		
Data Set	**Description**	**Keys**
Physical Removal of Carcasses	Rate at which carcasses are removed from the shoreline.	Rates vary on a micro spatial level. Rates appear to be systematically related to carcass size and carcass condition with small fresh birds removed more rapidly. These studies have been conducted several times and existing protocols should be considered.
Search Efficiency	Proportion of carcasses found during a shoreline search.	Rates shown to vary by substrate, carcass size, and the number of searches in a group. Estimates based on existing literature may be possible. If a site specific study is undertaken, these studies have been conducted several times and existing protocols should be considered.
Background Rates	Rate at which carcasses would have been collected from the shoreline had there not been a spill.	Several groups collect data describing rates of background bird collection. Rates tend to show great seasonal and inter-annual variation. First search rates are likely elevated relative to subsequent search rates. Rates may be augmented with oil fingerprint analysis.
Sinking Rates	Rate at which bird carcasses sink	Buoyancy of carcasses through time is well understood. The time at sea for each carcass must be estimated by the modeler. This is often achieved via hydrodynamic modeling which does not capture active movement of oiled birds.

Table C-1.0-5. Quick reference for BBM set-up and implementation.	
Topic	**Recommendation**
Working within the Trustee/RP framework	Parallel Stepwise modeling
Identification of undocumented search	Often patterns in the data suggest someone was systematically searching shorelines though there is no documentation. It is often possible to identify this activity via phone interviews and investigation. Incorporation of this search effort often reduces significantly avian mortality estimates.
Birds collected between searches	If persistence rates are above 75% per day, add in as 1. f persistence rates are lower, conduct BBM on expected collections.
Incomplete segment searches	Re-segment so that all searches represent complete segment searches.
Methods to address long search intervals	Two step BBM analysis that directly integrates temporal pattern of deposition is recommended
Geographic extrapolation	Use of regression analysis to identify appropriate extrapolation factors is recommended
Addressing background bird collections	If actual collection rate is less than or equal to expected background bird collection rate, assign zero deposition to the segment. Carcasses with unusually large multipliers should be assessed individually.
Outlier multipliers	Absent exogenous data that may warrant unique treatment of an outlier, no adjustments are recommended.
Unanticipated Occurrences	Every assessment is unique. The goal of BBM is to estimate the number of birds that experienced acute, spill-related mortality. Assessments that rely on a series of assumptions each one being "protective of the resource" tend to significantly overestimate that number.

American
Petroleum
Institute

200 Massachusetts Avenue, NW
Suite 1100
Washington, DC 20001-5571
USA

202-682-8000

Phone Orders:	1-800-854-7179	(Toll-free in the U.S. and Canada)
	303-397-7956	(Local and International)
Fax Orders:	303-397-2740	

Information about API publications, programs and services is available on the web at www.api.org.