

**Biogeography of Nonindigenous Species: From Description to Prediction**

by  
Deborah A. Reusser

A DISSERTATION

Submitted to

Oregon State University

in partial fulfillment of  
the requirements for the  
degree of

Doctor of Philosophy

Presented February 10, 2010

Commencement June 2010

**Committee in Charge:**

Dawn Wright, Major Adviser, Geography/Geosciences  
Julia Jones, IGERT Ecosystem Informatics Minor Adviser, Geography/Geosciences  
Mike Bailey, Computer Science Minor Adviser, Elec. Engr. & Computer Science  
Sylvia Yamada, Committee Member, Zoology  
Dominique Bachelet, Graduate Council Representative, Biological & Ecol. Engr.

AN ABSTRACT OF THE DISSERTATION OF

Deborah A. Reusser for the degree of Doctor of Philosophy in Geography presented on February 10, 2010.

Title: Biogeography of Nonindigenous Species: From Description to Prediction



Abstract approved: \_\_\_\_\_

Dawn J. Wright

Nonindigenous species are a major threat to the ecological integrity and biodiversity of marine and estuarine ecosystems. To become a successful invader, species must pass through four phases: (1) survive transport, (2) survive release, (3) establish a population, and (4) expand their range. To better understand these processes, an integrated framework was designed to capture life history characteristics, environmental preferences, dispersal mechanisms, and geographic distribution information for both native and nonindigenous marine and estuarine flora and fauna. Key aspects of this framework include: 1) consistent terminology; 2) translation of numerical habitat values and physiological requirements into classes; 3) development of classification schemas for natural history, environmental attributes, and geographic distributions; and 4) integration

of biotic attributes to allow database queries on single or multiple species across spatial scales. Species data for the North Pacific were collected from the literature, local surveys, and regional databases.

Ballast water discharges have been identified as a major source of species introductions. To predict the potential rate of invasion from ballast water, a linear invasion model predicting per capita invasion probabilities (PCIP) of new invaders was developed based on historic invasion rates and ballast discharge volumes for estuaries on the west coast of the United States. While the probability of invasion is likely to vary with ballast discharge values, organism concentrations in the ballast, and invasibility of individual ports, the PCIP provides a quantitative methodology for establishing protective ballast water discharge standards based on organism concentrations, the approach being used to regulate ballast water discharges both nationally and internationally.

Habitat or niche models can be used to predict a nonindigenous species' potential distribution in invaded areas over several spatial scales. The utility of non-parametric multiplicative regression (NPMR) was evaluated for predicting habitat- and estuary-scale distributions of native and nonindigenous species. Results indicate that NPMR generally performs well at both spatial scales and that distributions of nonindigenous species are predicted as well as those of native species. Development of approaches for regulating ballast water and identifying areas at risk through predictive modeling are useful tools in the management of the nonindigenous species threat.

©Copyright by Deborah A. Reusser  
February 10, 2010  
All Rights Reserved

Doctor of Philosophy dissertation of Deborah A. Reusser presented on February 10, 2010.

APPROVED:

Dawn J. Wright

Major Professor, representing Geography

Aaron Wolf

Chair of the Department of Geosciences

Sally Francis

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

---

Deborah A. Reusser, Author

## ACKNOWLEDGEMENTS

The author expresses sincere appreciation for the guidance, time and assistance provided by her committee members and for the forbearance and support of her family and friends during the completion of this dissertation.

## CONTRIBUTION OF AUTHORS

Dr. Henry Lee II assisted with data collection, analysis and editing of Chapters 2, 3 and 4.

Dr. Melanie Frazier assisted with statistical analyses in Chapters 2 and 3. Marla

Ranelletti, Dr. Jody Stecher, Katie Marko, Karen Ebert and Kathy Welch contributed to populating the Pacific Coast Ecosystem Information System (*PCEIS*). Rachel Nehmer

and Tad Larsen assisted in the implementation of the graphical user interface for *PCEIS*.

Dr. Gregory Ruiz and the Smithsonian Institution provided data for analysis in Chapter 3.

Dr. Jody Stecher, Dr. Melanie Frazier and Dr. Henry Lee II provided valuable comments and suggestions that improved the manuscripts in Chapters 2, 3 and 4.

## TABLE OF CONTENTS

	Page
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
TERMINOLOGY.....	2
SPECIES CLASSIFICATION AND CRITERIA.....	3
INVASION PROCESSES – VECTORS AND STAGES OF INVASION.....	10
<i>Vectors and Pathways</i> .....	11
<i>Survival</i> .....	17
<i>Population Establishment</i> .....	18
<i>Range Expansion – Patterns of Invasion</i> .....	19
ECOLOGICAL PROBLEM FORMULATION – RESEARCH OBJECTIVES.....	21
<i>Ecosystem Informatics for Marine Biogeography and Natural History</i> .....	22
<i>Management Strategy for Ballast Water</i> .....	23
<i>Ecological Niche Modeling for Prediction</i> .....	24
REFERENCES.....	25
 <b>CHAPTER 2: EVOLUTION OF NATURAL HISTORY INFORMATION IN THE 21<sup>ST</sup> CENTURY – DEVELOPING AN INTEGRATED FRAMEWORK FOR BIOLOGICAL AND GEOGRAPHIC DATA.....</b>	 <b>35</b>
ABSTRACT.....	36
INTRODUCTION.....	37
CLASSIFICATION SCHEMAS TO CAPTURE ENVIRONMENTAL AND NATURAL HISTORY ATTRIBUTES.....	40
<i>Temperature Classification Schema</i> .....	41

TABLE OF CONTENTS (CONTINUED)

	Page
<i>Wave Energy Classification Schema</i> .....	43
<b>HIERARCHICAL SCHEMAS FOR ENVIRONMENTAL AND NATURAL HISTORY</b>	
<b>ATTRIBUTES</b> .....	44
<i>Taxonomic Hierarchical Schema</i> .....	46
<i>Salinity Hierarchical Schema</i> .....	47
<i>Depth Hierarchical Schema</i> .....	53
<i>Reproductive Mode Hierarchical Schema</i> .....	57
<i>Geographic Hierarchical schemas</i> .....	57
<i>Biogeographic Hierarchical Schema</i> .....	59
<i>Land/Sea Connection Hierarchical Schema</i> .....	63
<b>POPULATING THE INFORMATION SYSTEM</b> .....	64
<b>CONCLUSIONS</b> .....	67
<b>ACKNOWLEDGEMENTS</b> .....	69
<b>REFERENCES</b> .....	70
<b>CHAPTER 3: PER CAPITA INVASION PROBABILITIES: A LINEAR MODEL</b>	
<b>TO PREDICT RATES OF INVASION VIA BALLAST WATER</b> .....	<b>75</b>
<b>ABSTRACT</b> .....	<b>76</b>
<b>INTRODUCTION</b> .....	<b>78</b>
<b>METHODS</b> .....	<b>84</b>
<i>Foreign Ballast Water Discharge Rates for Coastal Waterbodies and the</i>	
<i>Great Lakes</i> .....	85

TABLE OF CONTENTS (CONTINUED)

	Page
<i>Estimates of Organism Concentrations in Ballast Water</i> .....	86
<i>Estimates of Historical Invasion Rates</i> .....	91
UNCERTAINTIES IN HISTORICAL INVASION RATES AND SAFETY FACTORS .....	92
AMONG PORT PATTERNS OF INVASION RISK.....	94
ACROSS COAST PATTERNS OF INVASION RISK.....	99
ASSUMPTIONS AND LIMITATIONS.....	101
RECOMMENDATIONS/CONCLUSIONS .....	104
ACKNOWLEDGEMENTS.....	107
REFERENCES .....	108
<b>CHAPTER 4: PREDICTIONS FOR AN INVADED WORLD: A STRATEGY TO PREDICT THE DISTRIBUTION OF NATIVE AND NON-INDIGENOUS SPECIES AT MULTIPLE SCALES .....</b>	<b>111</b>
ABSTRACT .....	112
INTRODUCTION .....	113
METHODS.....	114
<i>Habitat (point) Scale</i> .....	115
<i>Estuary Scale</i> .....	117
<i>Geographic Variables and Datasets</i> .....	118
RESULTS AND DISCUSSION.....	119
CONCLUSIONS.....	124
ACKNOWLEDGEMENTS.....	124

## TABLE OF CONTENTS (CONTINUED)

	Page
REFERENCES .....	125
<b>CHAPTER 5: SUMMARY .....</b>	<b>126</b>
UNCERTAINTY .....	126
NICHE MODELING.....	127
CONCLUSIONS.....	129
<b>CHAPTER 6: BIBLIOGRAPHY .....</b>	<b>132</b>
<b>APPENDIX A: CALCULATION OF COASTAL PER CAPITA INVASION PROBABILITIES (R CODE).....</b>	<b>146</b>

## LIST OF FIGURES

Figure	Page
<b>Chapter 2</b>	
2-1: Taxonomic Hierarchy .....	46
2-2: Three level hierarchical schema for salinity.....	49
2-3: Number of species per benthic sample relative to the salinity .....	51
2-4: Salinity values from benthic samples.....	52
2-5: Hierarchical schema for depth.....	54
2-6: Number of chiton species in the Northeast Pacific by depth range.....	56
2-7: Hierarchical schema for reproduction .....	58
2-8: Total number of chiton species by MEOW ecoregion.....	61
2-9: Hierarchical schemas for marine biogeographic areas and land/sea connection...	62
2-10: Graphical interface for data entry.....	67
<b>Chapter 3</b>	
3-1: Hypothetical propagule supply dose-response curves.....	82
3-2: Distribution of per capita invasion probabilities .....	90
3-3: Risk diagrams illustrating the effect of three different safety factors .....	100
3-4: Risk diagrams based on less protective assumptions .....	102
<b>Chapter 4</b>	
4-1: Percentage of species falling into different classes of the AUC .....	122

## LIST OF TABLES

Table	Page
<b>Chapter 2</b>	
2-1: Temperature classes for marine/estuarine species.....	42
2-2: Wave energy classes for marine/estuarine species.....	43
2-3: Habitat Regimes for marine/estuarine species .....	66
2-4: Crosswalk of the biogeographic and bathymetric ranges of chitons in the Northeast Pacific .....	69
<b>Chapter 3</b>	
3-1: Summary of existing or proposed ballast water performance standards applicable to United States waters.....	80
3-2: Historical number of invaders ( $N_h$ ), foreign ballast discharge volumes ( $D_h$ ), and per capita invasion probabilities (PCIP) for the East, Gulf, and Pacific coasts of the United States. ....	87
3-3: PCIP <sub>CI</sub> based on upper 90%, 95%, 99%, and 99.9% confidence intervals .....	98
3-4: Major assumptions of the per capita invasion probability approach.....	106
<b>Chapter 5</b>	
5-1: Confusion matrix for niche models .....	127

## **Chapter 1: Introduction**

Aquatic ecosystems are undergoing a rapid revolution in species composition and ecosystem stability caused by the seemingly endless influx of nonindigenous species (Cohen & Carlton, 1998; Galil, 2000; Grosholz *et al.*, 2000). These non-native species represent one of the major threats to the ecological integrity and biodiversity of marine and estuarine ecosystems (Carlton, 1996a; Vitousek *et al.*, 1997; Ruiz *et al.*, 1999) and in lakes have been predicted to be one of the more important causes of biodiversity loss by 2100 (Sala *et al.*, 2000). Nonindigenous species are also the second most significant cause for declines in endangered or imperiled species (Lassuy, 1995; Wilcove, 1998).

Estuaries, in particular, appear highly vulnerable to invasions (Ruiz *et al.*, 1997; Cohen & Carlton, 1998; Wolff, 1999; Paavola *et al.*, 2005; Nehring, 2006). The classic example is the San Francisco Estuary where approximately one new nonindigenous species is predicted to arrive every 14 weeks (Cohen & Carlton, 1998). And while the San Francisco Estuary has been dubbed “the most invaded estuary in the world” (Cohen & Carlton, 1998), it is not the only highly invaded estuary (Wolff, 1999; Ruiz, 2000). The high vulnerability of estuaries may be related to the intensity of the introductions of foreign species (e.g., through ballast water discharges) or an innate property of estuarine communities (e.g., low biotic resistance). In any case, the future does not look better as estuarine invasion rates appear to be increasing (Cohen, 1998).

While the breadth of invasion biology is beyond the scope of this research, a general knowledge of the terminology, classification criteria, transport vectors, and stages of

invasion is needed to better understand the processes and patterns that describe the biogeography of nonindigenous species.

### ***Terminology***

A variety of different adjectives such as non-native, alien, nonindigenous, and invasive are commonly used to describe species discovered in new environments outside of their native range. Similar to Elton's approach in his seminal book in 1958, titled *The Ecology of Invasions by Animals and Plants*, political proponents for the management of these species use the term "invasive species" as it implies an impending threat. However, Executive order 13112 defines the term *invasive species* to mean "a non-native species whose introduction does or is likely to cause economic or environmental harm or harm to human, animal, or plant health" (<http://www.invasivespecies.gov/>). Many nonindigenous species have not become invasive, but the potential is there for any nonindigenous species to become invasive given the right supporting environmental conditions. Within the context of this research, I refer to nonindigenous species as all species that are not native, including species that have become invasive.

Cryptogenic species are species that are not demonstrably native or nonindigenous (Carlton, 1996b). This classification is given to species that could be native or nonindigenous, but lack of information about the species makes it impossible to determine. For example, a cryptogenic species is a species that is found both in Japan and the West Coast of the United States, but it cannot be determined which way it was transported, or if it was transported at all. Species known to be introduced species whose

geographic origins or mechanism of introduction cannot be determined are not cryptogenic species (Carlton, 2009).

Cosmopolitan species have global geographic distributions within suitable habitats. Several species with cosmopolitan distributions (e. g. *Teredo navalis* considered nonindigenous in the Northeast Pacific) have been spread around the world on the bottoms of ships since man first set sail on the open sea. However, other species, (e. g. *Paracalanus parvus* that occurs in 24 of 27 oceanic regions (<http://copepodes.obs-banyuls.fr/en>)) have naturally occurring cosmopolitan distributions due to their long distance dispersal capabilities, as discussed under vectors and pathways.

Groups of species that appear similar morphologically and have been identified as a single species but are isolated reproductively are called species complexes. Analyses such as DNA sequencing is needed to distinguish between individual species (Westheide & Schmidt, 2003). Species complexes can sometimes be mistaken for a single “cosmopolitan” species due to the inability to distinguish between morphologically indistinguishable species without DNA sequencing (Klautau *et al.*, 1999).

### ***Species Classification and Criteria***

Identifying species as native, nonindigenous, or cryptogenic to a particular area is not a simple task and taxonomists do not always agree on the classification of a species. Take for example, the parasitic isopod, *Orthione griffenis*, discovered on the mud shrimp, *Upogebia pugettensis*, in Yaquina Bay about six years ago. *Orthione griffenis* was classified by Dr. John Chapman as an invader from Asia that traveled to the Northeast Pacific coast in the ballast water of ships

(<http://oregonstate.edu/dept/ncs/newsarch/2009/Feb09/isopod.html>). However, some taxonomists disagree with this classification (Chapman pers. comm.). A set of standardized criteria to evaluate whether a species is native, nonindigenous, or cryptogenic have been presented and utilized by invasion biology experts such as Chapman, Carlton, Sytsma, and Cohen in the United States (Chapman & Carlton, 1991; Chapman & Carlton, 1994; Sytsma *et al.*, 2004; Cohen *et al.*, 2005), and Boudouresque, Ribera, and Wolff in Europe (Boudouresque, 1994; Ribera & Boudouresque, 1995; Wolff, 2005). These authors all agree that determining a species classification requires high quality taxonomy to avoid confusion between distributions of closely related but different species. Sytsma *et al.*, (2004) indicated that application of these criteria required not only detailed information on their taxonomy, but also the biogeography, ecology and life history of the species. These criteria have been used in a multiple-criteria- weight-of-evidence approach to evaluate the classification of many species including *Crangonyx floridanus*, *Caecidotea racovitzai*, and *Asellus hilgendorfi* in the Northeast Pacific (Toft *et al.*, 2002). Lee *et al.*, (2008) evaluated the individual discriminatory power of these various classification criteria prior to their use in the classification of a suite of species collected through several U. S. EPA sampling programs. In the absence of historical records, they suggest this weight of evidence technique is the best approach to evaluate the classification of a species (Lee *et al.*, 2008a). These species classification criteria with their discriminatory power are discussed below.

*Historical records of introduction.* This criterion provides the strongest evidence for a species to be classified as nonindigenous, especially for species that have been stocked. For example, detailed historical records exist where the Japanese oyster, *Crassostrea*

*gigas*, has been regularly stocked in the Northeast Pacific beginning in 1875 (Carlton, 1992) and the Northeast Atlantic beginning in 1966 to restore depleted native oysters (Wolff, 2005).

*Association with human vectors of introduction.* This criterion has been used for fouling species that could potentially have traveled on the hulls of ships. However, this criterion alone fails to discriminate between native and nonindigenous hard-bottom species so has very little discriminatory power.

*Fossil records.* A continuous fossil record for a species in a region provides strong evidence that the species is native to a region. However, this criterion has utility only for species that are likely to leave a fossil and is dependent on completeness of the fossil record for an area. Marine invertebrates with shells (e. g. clams) or skeletal structures (e. g. calcareous sponges and corals) provide an extensive geological record of their origin, range and migration patterns. For example fossil records indicate *Mya arenaria* went extinct in the Northeast Pacific in the late Tertiary period and was reintroduced accidentally into San Francisco Bay with a shipment of oysters in 1874 (Carlton, 1992).

*Insufficient natural dispersal mechanisms to account for observed distribution.* This criterion is useful to identify native species that are likely to be widespread either through active dispersal mechanisms (e.g., tuna) or passive dispersal mechanisms (e.g., many microbes, pelagic copepods). This criterion can be useful in identifying species that are likely naturally dispersed, and hence not nonindigenous. However, this criterion does not have high discriminatory power to identify a species as nonindigenous by itself.

*Recent appearance in a region where it has not been found previously.* This criterion provides reasonably strong evidence depending upon completeness of previous records, collections, and species size (e.g., stronger evidence for larger, well known taxa, such as fishes and crabs) that are likely to have been previously seen and recorded. However, for the less well studied taxonomic groups of very small animals, the discriminatory power of this criterion is much weaker. A good example of this criterion identifying a new invader is when the first blue crab, *Callinectes sapidus* was found in the harbor of Rochefort, in SW France in 1900 (Wolff, 2005). Known from the Northwest Atlantic, this species had no previous records of being found in the Northeast Atlantic prior to 1900. It is believed to have been transported across the Atlantic Ocean in the ballast water of ships.

*Discontinuous local distributions relative to similar endemic species.* A discontinuous distribution in a region may indicate a recent introduction, especially if the species is limited to ports or other areas of high invasion. For example, on the Northeast Pacific coast, if a species has been found in Puget Sound, Washington and Coos Bay, Oregon and no other estuaries in between, it could potentially be a nonindigenous species. This is particularly true if the species has broad dispersal potential, and other native species in the same genus can be found in estuaries all along the coast between Puget Sound, Washington and Coos Bay, Oregon. This criterion provides reasonably strong evidence depending upon the completeness of species records, sampling methodologies and collections. It also requires sufficient understanding of the natural history of the species to assess that the discontinuous distribution is not due to specific habitat or physiological limitations.

*Recent spread from one or a few locations to many locations within a region.* This criterion provides moderately strong evidence depending upon completeness of historical records and species collections. The recent spread should be evaluated to determine if it is related to recent changes in environmental conditions (e.g., El Niño) and whether ecologically similar native species are also showing a similar spread. In general, native species do not undergo sudden broad range expansions without altered environmental conditions to trigger the event. However, nonindigenous species have the potential to undergo a population explosion and rapid range expansion after initial population establishment, potentially because the new environment is lacking natural predators and/or barriers. The zebra mussel, *Dreissena polymorpha* is a good example. Once a population was established in the Great Lakes in 1988, it began to spread throughout the entire Mississippi river region displacing native mussels, because it had no natural predators to keep its expansion in check (Johnson *et al.*, 2006).

*Close associations with other introduced species.* Cohen *et al.* (2005) only used this criterion as evidence if the “association or dependency appeared to be obligate or near-obligate.” This criterion has been used to classify parasites and commensal species of stocked nonindigenous species such as *Mytilicola orientalis*, a parasite of *Crassostrea gigas*. *Mytilicola orientalis* is found where *Crassostrea gigas* has been stocked (Carlton, 1979).

*Restriction to new or artificial environments.* This criterion alone does not have high discriminatory power and was excluded by Cohen *et al.* (2005). However, this criterion could be considered as additional evidence of a recent introduction if combined with

several other criteria such as a discontinuous distribution or recent appearance where the species has never been seen before.

*Conspecific with geographically isolated populations.* Globally disjoint distributions, as indicated by isolated populations in geographically separated oceans or on different continents, are suggestive that one of the populations was introduced. Using this criterion, when a Southwest Pacific barnacle, *Elminius modestus*, was discovered in the Netherlands in 1945, it was determined to be nonindigenous in the North Sea because it was found so far away from its known native environment of New Zealand (Wolff, 2005). While this criterion held true for *E. modestus*, due to the geographic distance between New Zealand and the Netherlands, in other places such as South America and Africa there could be another explanation. One alternative is the vicariance hypothesis that disjoint distributions result from the breakup of a previously continuous population due to plate tectonics or another geographical barrier. Another alternative is that the two populations represent sibling species that have yet to be recognized. The challenge with this criterion is that it is difficult to determine which way the species might have traveled if it is indeed an introduced species in one region. A date indicating when a species was first found in a location can be used in deciphering directionality in some cases.

*Type locality is geographically isolated from the region.* Lee *et al.*, (2008) included this as a new criterion arguing that the location of the type specimen in a geographically isolated area is essentially an application of the “Conspecific with geographically isolated populations” criterion. The location of the type specimen can be used as a simple filter to help identify possible nonindigenous species that need more detailed analysis. However,

Carlton (2009) indicates that a type specimen from another continent may indicate a species that was potentially misidentified through the application of European taxonomic records.

*Member of a nonindigenous taxonomic group.* This criterion has high discriminatory power, because membership in a family or genus that is only found in areas geographically remote from the local region suggests that one of the populations has been introduced (see Chapman, 1988).

*Ecological or physical adaptations dissimilar from endemic species.* Many introduced species are from climates where temperature ranges exceed those in the new location or where they escape parasites or diseases. Some introduced species tolerate temperatures, for instance, that do not exist in the new locations (Sytsma et al., 2004). While these differences in adaptation may help explain why a particular species is a successful invader, they do not have high discriminatory power to separate native versus non-native species.

*The species is absent from ecosystems with low invasion potential.* Current evidence indicates that some ecosystems do not contain many nonindigenous species, such as deep water (>200 m) habitats. Occurrence of a species in such ecosystems is suggestive of a native origin, while absence of the species from such ecosystems in itself has little discriminatory power.

*DNA analysis.* This criterion can provide strong evidence in identifying disjoint populations that are closely related (Coleman, 1996). For example, the copepod, *Eurytemora affinis*, was long thought to be native to San Francisco Bay, but genetic

analysis now indicates this species probably originated from the Northwest Atlantic coast and was transported to San Francisco Bay in ballast water (Lee, 1999). However, marine species with broad dispersal capabilities tend to have high levels of gene flow that make the genetic signal for population differentiation weak (Waples, 2004). In the case of the European snail, *Littorina littorea*, whose pelagic phase can be up to six weeks, there is disagreement with the conclusion from DNA analysis that it is native to the Atlantic coast of North America (Chapman *et al.*, 2008).

### ***Invasion processes – Vectors and Stages of Invasion***

The key elements of nonindigenous species becoming established in a new environment are the ability to arrive, survive, and thrive. The introduction, establishment and invasion of a species occurs in four discrete phases: 1) the potential invader is taken up by a transport mechanism (vector) and survives the transport; 2) the potential invader survives release into the new environment; 3) the invader establishes a self-sustaining population at or near the locality of initial release; and 4) for “invasive” species, the invader undergoes a population “explosion” and expands its range into new localities (Kolar, 2001; Sakai, 2001; Jeschke, 2005).

Carlton (1996b) identifies and evaluates six interrelated processes which mediate invasion success to provide a better understanding of why species become established when and where they do. These six processes are: 1) changes in donor regions (i.e., environmental changes in a donor region that cause increases in species densities making more species available for transport); 2) new donor regions (i.e., different species available for transport); 3) changes in recipient regions (i.e., environmental changes that

make the recipient region more inhabitable); 4) invasion windows (i.e., changes that create colonizing conditions that may not have existed before); 5) stochastic inoculation events (i.e., release of an unusually large number of a single species increasing potential for establishment); and 6) dispersal vector changes (i.e., changes in vector size, speed and quality) (Carlton, 1996b). Each of these processes is important in determining invasion success, and even potentially predicting when and where invaders are likely to invade. The processes which dictate *when* an invader will invade are beyond the scope of this research. However, Carlton (1996b) does suggest a simple predictive technique for answering *what* and *where* called “the weed theory: invasive species are likely to continue to invade elsewhere if corridors are available and conditions permit.” In other words, species that have shown a capability of invading new areas have a high probability of invading yet more areas if the vectors exist. The first step towards utilizing this predictive technique is identifying and recording what and where nonindigenous species are currently and what the primary transport vectors might have been.

### *Vectors and Pathways*

Pathways leading to the current range and distribution of a species are either natural or anthropogenic. While I am most interested in the anthropogenic or human mediated vectors that transport species to new areas, it is also important to understand the extent to which species may have been able to travel naturally. Natural ranges for marine and estuarine species are defined by dispersal ability, environmental preference, and natural barriers (Watts *et al.*, 1998; Cox & Moore, 2005). Across the diverse breadth of marine species, a variety of dispersal mechanisms exist (Grantham *et al.*, 2003). In addition, dispersal is strongly influenced by currents that transport larvae short to long distances

and therefore directly affect natural range size and expansion for marine populations (Scheltema & Williams, 1983). Another reproductive strategy is species that brood or release juveniles as small adults such as peracaridan crustaceans, including amphipods, isopods, and tanaids. Species with this kind of dispersal scheme are not well suited to survive a transoceanic voyage by themselves, but they can be transported long distances through rafting and other means such as ballast water, hull fouling, or as hitchhikers in shipping material (Watts *et al.*, 1998; Grantham *et al.*, 2003; Thiel, 2003). Likewise, most species with lecithotrophic larvae would probably not be able to survive a long voyage because they are dispersed with a very short-lived yolk-sac that provides nutrients. Once the nutrients in the yolk-sac have been exhausted, the larvae must find a place to settle and metamorphose to the next stage. Therefore, the only kind of larvae that might survive a transoceanic voyage on their own are the planktotrophic larvae. Some types of planktotrophic larvae are longer lived and can feed off plankton when traveling great distances because they have feeding and digestive structures. These features give some planktotrophic larvae the capacity to disperse great distances during the larval stage, but large numbers suffer mortality from predation, food shortages, and inappropriate substrates available at the end of a journey. Planktotrophic larvae that are known to travel great distances are called teleplanic larvae. For example, the larvae of the gastropod, *Philippia radiata* and the bivalve, *Streptopinna saccata* have been found in the open ocean of the tropical Pacific and have a very wide distribution throughout the islands of the Indo-Pacific (Scheltema & Williams, 1983).

Although, oceanographic currents are responsible for the natural long-distance dispersal of a few marine species such as those with teleplanic larvae, many larvae live for so short

a time that there are comparatively few examples of east-west linkages between shelf faunas. In general, wide, island-free oceans act as effective barriers to natural dispersal across oceans (Cox & Moore, 2005). The Line Islands are located 2,500 kilometers south of Hawaii in the Central Pacific, and the 5,400 km expanse of deep ocean between the Line Islands and the Clipperton Islands off Western Mexico constitute the “East Pacific Barrier”, the single largest oceanic barrier in the world (Scheltema, 1988; Collin, 2003). And, while other natural dispersal mechanisms such as wind, birds, rafting, and stochastic events have been suggested as potential pathways leading to the current distribution of a few species (Scheltema & Williams, 1983; Watts *et al.*, 1998), the greatest opportunities for a species to travel great distances over broad barriers are by anthropogenic vectors. Human mediated vectors are discussed below.

Vectors that transport a nonindigenous species directly from a native region to a new location are considered primary vectors, while those vectors that disperse the species from the landing location to new areas within the invaded region are considered secondary vectors (Minchin *et al.*, 2009). It is important to understand the primary vector or vectors responsible for the transport of species across these large geographic barriers in order to identify management strategies to mitigate the influx of new nonindigenous species into our estuaries (Ruiz & Carlton, 2003).

Ships have been transporting species to new and exotic places since the first sailing ships were built (Carlton, 1987) in approximately 3,000 BC (Åkesson, 2007). Ships can transport species in dry ballast, ballast water, on the hull, or in the sea chest (Carlton, 1987). Ballast water and hull-fouling from ships are two of the leading primary vectors

for nonindigenous species transport around the world (Ruiz & Carlton, 2003; Molnar *et al.*, 2008). Several notorious invaders transported by ballast water include the zebra mussel, *Dreissena polymorpha* to the Great Lakes in 1988, which has caused continuous problems by blocking water filtrations systems (Cohen & Weinstein, 1998) and the comb jelly, *Mnemiopsis leidyi*, to the Black Sea that was blamed for the collapse of the fishing industry (Faasse & Bayha, 2006) but see (Bilio & Niermann, 2004) for an alternate explanation.

Infrastructure development such as canal building and dredging has also been responsible for a large number of species invasions. Over 300 species have migrated between the Red Sea and the Mediterranean Sea since the opening of the Suez Canal in 1869 (Galil, 2000; Goren & Aronov, 2002; Briggs, 2007). The opening of the Panama Canal in 1914 provided a similar opportunity for species to migrate between the Pacific and Caribbean waters (Smith *et al.*, 2004) though the freshwater Gatun Lake reduces the likelihood of direct transport of many marine species (Minchin *et al.*, 2009). Likewise, dredging of the Danube-Main-Rhine Canal between the Black and North Seas in 1836 and the Volga-Don Canal between the Black and Caspian Seas in 1952 has promoted similar opportunities for the exchange of species in those regions (Alexandrov *et al.*, 2007).

Movable structures such as buoys, dry docks and drilling platforms have been documented to provide transport across great distances. For example, the Japanese native grapsid crab, *Plagusia dentipes* was transported from Japan to California on a semisubmersible exploratory drilling platform in 1976 (Benech, 1978).

The aquaculture and fisheries vector includes intentional stocking and release, hitchhikers associated with aqua-culture species, and illegal release. The Pacific oyster was introduced to the Northeast Pacific coast from Japan in 1918. Because the water is colder in the Northeast Pacific, the Japanese oyster was unable to establish and reproduce, even though it was able to grow well, so seed oysters were continually imported from Japan to replenish stocks. During that period of time, other hitchhiking nonnative species came with the oyster seed. Another example is the Chinese mitten crab, *Eriocheir sinensis*, a recent invader to California. It is believed that the Chinese mitten crab was intentionally, illegally released to establish a fishery in California as this species is considered a delicacy in its native Asian countries ([http://www.wsg.washington.edu/mas/ecohealth/invasive\\_crabs/mitten\\_crab.html](http://www.wsg.washington.edu/mas/ecohealth/invasive_crabs/mitten_crab.html)).

The live seafood trade vector includes processing, packing, shipping and release of animals being sold for human consumption. There is a global market for the sale of live seafood that spreads organisms that have the potential to become established in new areas (Chapman *et al.*, 2003). One example is the discovery in Puget Sound, Washington of a live American lobster, *Homarus americanus*, native to the Northwest Atlantic, during a survey in 1999 (<http://www.peter.unmack.net/archive/acn/acnlsep99/0007.html>). It is believed to have been released after it was purchased for consumption. Other than this incidental discovery, *H. americanus* is not known to occur anywhere on the Northwest Pacific coast.

Another vector is research and education. Classroom education for K-12 can include acquiring and raising species that are not native to the area. Frequently, once the class

project is over, the organisms are released into the new environment to avoid killing the specimens (“mercy release”), providing an opportunity for a new nonindigenous species to become established. Outreach and education for K-12 has been developed by Sea Grant for the safe disposal of experimental organisms once a project is complete that includes freezing the organisms or how to preserve them in alcohol, and why this is a necessary step to protect the environment in which we live (Sea Grant, 2006).

The Habitat restoration and mitigation vector includes restoration and biocontrol releases for mitigation of other nonindigenous species. In the 1970’s in San Francisco Bay, California, the East Coast cordgrass, *Spartina alterniflora*, was planted to restore coastal wetlands damaged by human activities (Ayres *et al.*, 2004). Unfortunately, the new invader hybridized with the native cordgrass, *Spartina foliosa*. The new hybrid is rapidly replacing the native cordgrass and significantly altering the wetlands by occupying mudflat habitats that are important nursery grounds for bird and fish species (California Sea Grant, 1999-2000).

Recreational boats and fishing provide another vector for biota of all kinds to migrate into and out of a place beyond their natural dispersal ability. The zebra mussel in the Great Lakes has been well distributed by recreational boats being transported from one Lake to another within the United States. A 100<sup>th</sup> meridian initiative, established to monitor boats crossing it to prevent the mussel from spreading any further west, was successful in slowing down the migration. However, a close relative of the zebra mussel, the quagga mussel, has just been reported in Lake Mead, indicating that they have finally managed to breach the 100<sup>th</sup> meridian and are on the move west. Another invader that hitchhikes on

boats and the soles of fishermen's boots is the New Zealand mud snail, *Potamopyrgus antipodarum*. Although it is a freshwater snail, it has begun to invade estuaries along the west coast that have higher salinities than the NZ mud snail was believed to tolerate (Boese *et al.*, unpublished data).

Vectors in the aquarium trade include aquarium escapees, ornamental plant escapees, and hitchhikers in the transport of live aquarium plants and animals. The Indo-Pacific lionfish, *Pterois volitans*, has become established on the western Atlantic coast after an aquarium release in Florida (Whitfield *et al.*, 2002; Semmens *et al.*, 2004).

Many of these vectors can be either primary or secondary vectors or both. In addition, it is often difficult to determine the specific primary transport mechanism by which a nonindigenous species arrived into a new environment. For example, we have confidence that shipping is the major vector for many species, but it is unclear whether the invader traveled in the ballast water or on the hull of the ship. These species are referred to as polyvectoric species (Carlton & Ruiz, 2005).

### *Survival*

The attributes that help species survive in their native environment are also applicable for survival in a new environment. Key insights on nonindigenous species behavior when colonizing a new area can be gained by studying elements of population biology such as the role of carrying capacity, mortality, and reproduction (McMahon *et al.*, 2006). Generalists, species which can adapt to broad environmental conditions, tend to have a higher probability of survival than specialist species in need of specific environmental conditions (Roman & Palumbi, 2004). High fecundity is also a common characteristic of

many successful invaders (Sakai *et al.*, 2001; Mihalka *et al.*, 2006; Barnes, 2008). Community composition in the new environment also plays role. A nonindigenous species that has no natural predators and no parasites or diseases in the new community has a competitive advantage over native species; this is the “enemy release” hypothesis (Elton, 1958; Colautti *et al.*, 2004). An example of an invader that has expanded at least in part due to the release from predation/parasitism is the zebra mussel, *Dreissena polymorpha*, in the Great Lakes.

### *Population Establishment*

Elton’s hypothesis is that disturbed communities with their low biodiversity are more easily invaded by nonindigenous species than intact, more diverse communities (Elton, 1958). Higher diversity means more kinds of every type of species, which means a higher probability that one or more species could be predators of the potential invader or have high enough numbers to win in a competition for food or space. Existence of predators and low food availability due to competition could contribute to biotic resistance in diverse communities. However, diversity per se is not a measure of biotic resistance. Elton’s hypothesis does not hold true for communities with naturally low biodiversity such as intertidal mud flats that contain the ghost shrimp, *Neotrypaea californiensis*. The disturbance from high bioturbation by the ghost shrimp makes it difficult for most other species, native or nonindigenous to survive. Here, a community with very low diversity is just as resistant to invasion as a community with high diversity.

According to the intermediate-disturbance hypothesis, some disturbance increases diversity because disturbance creates open niches in a community (Grime, 1973). In other

words, if a nonindigenous species arrived at an opportune time, a community might be susceptible to the establishment of the invader if food or space availability was sufficient and there was no or reduced competition and/or predation from native species. Disturbance has been shown to provide new opportunities for shifts in community structure, by either native or nonindigenous species (Cardinale & Palmer, 2002; Sanders *et al.*, 2003; Altman & Whitlatch, 2007).

### *Range Expansion – Patterns of Invasion*

Both natural dispersal and secondary vectors play a role in range expansion, creating the patterns of distribution for a nonindigenous species in a new environment. One important secondary vector is intracoastal shipping. Ballast water standards are not as stringent for ships moving between ports within a country. Once a nonindigenous species arrives anywhere along the Northeast Pacific coast, transport by ships north and south along the coast from one port to the next is a growing concern (Cordell *et al.*, 2009).

Natural larval dispersal, estuarine circulation, and coastal currents also facilitate the secondary distribution of several known invaders in the Northeast Pacific. As an example, planktotrophic larvae of many species use vertical migration in the water column to promote dispersal into or out of an estuary. The mud crab, *Rhithropanopeus harrisi*, an invader in San Francisco Bay, uses vertical migration on the flood tide to promote estuarine retention (Cronin & Forward, 1979). Conversely, another invader in San Francisco Bay, the green crab, *Carcinus maenas*, uses vertical migration on nocturnal ebb tides to flush its planktotrophic larvae out to sea (Zeng & Naylor, 1996; Queiroga *et al.*, 1997). *C. maenas* larvae attempt to remain in the nearshore environment

feeding on plankton for up to eight weeks. Then, the developing megalopae depend on currents to recruit back into the estuary (Zeng *et al.*, 1997; Hedvall *et al.*, 1998; Queiroga, 1998). This estuarine-ocean-estuarine migration cycle during the larval phase of development is not unique: several species of brachyuran crabs on the southern African coast exhibit this same behavior cycle (Papadopoulos *et al.*, 2002). The native ghost shrimp, *Neotrypaea californiensis*, also exhibits this behavior in Oregon estuaries with great success (Johnson and Gonora 1981).

Invaders with this estuarine-ocean-estuarine migration cycle can be dispersed north and south along a coast through natural nearshore oceanographic processes (Grosholz, 1996). During normal years, the California Current in the Northeast Pacific is characterized by southward and offshore currents during the summer from about mid-April – October and northward currents during the winter (Peterson 2006). These conditions would theoretically prevent larval dispersal of *R. harrisi* and *C. maenas* northward. However, El Niño events typically bring more rain to California than under normal climate conditions. Coupled with strong northward and shoreward currents that are indicative of El Niño events in the Northeast Pacific, these events provided the transport mechanism for the northward expansion of *C. maenas* during the El Niño event of 1997-1998 (Behrens Yamada, 2001). This phenomenon is not restricted to nonindigenous species. The northward range expansions of many different kinds of native marine species during this and other El Niño events have also been documented (Percy & Schoener, 1987; Peterson *et al.*, 2002).

The nonindigenous amphipod, *Grandidierella japonica* also has a very broad distribution on the Northeast Pacific Coast. From the first record in the San Francisco Estuary in 1966 (Chapman & Dorman, 1975), *G. japonica* has since spread to 41 estuaries from Puget Sound, Washington to Tijuana, California (Lee & Reusser, 2006). However the extensive spread of this invader cannot be explained by larval dispersal, as *G. japonica* broods its young. One potential secondary dispersal mechanism, in lieu of a larval phase, is rafting on eelgrass or other material that provides transport for this species out to sea and into neighboring estuaries along the coast (Thiel, 2003).

As discussed here, anthropogenic and natural mechanisms both contribute to the secondary dispersal and establishment success of nonindigenous species. Together they can form a “hub and spoke” model where by the anthropogenic mechanisms introduce new invaders into regional hubs and then both secondary vectors natural dispersal processes create spokes of distribution in a dispersal wheel (Carlton, 1996b). The mechanisms both individually and together form different patterns of invasion that threaten the biodiversity of estuarine ecosystems in the Northeast Pacific.

### ***Ecological Problem Formulation – Research Objectives***

Changes in biodiversity alter ecosystem functions and services (e.g., (Mack, 2000), which in turn result in economic losses from aquatic invaders measuring in the hundreds of millions of dollars (Pimentel, 2000, 2001). In a review of coastal invasions, Carlton, (2001) suggested that aquatic introductions have had fundamental impacts on fisheries resources, industrial development and infrastructure, human welfare, and ecosystem resources. In San Francisco Bay for example, the invasion of a single species, the Asian

clam *Corbula amurensis* (= *Potamocorbula amurensis*), (Carlton, 1990) has reduced recruitment of the previously dominant “dry-period” benthic community (Nichols, 1990), suppressed seasonal phytoplankton blooms (Alpine, 1992), and reduced copepod populations presumably because of reduced phytoplankton (Kimmerer, 1994). As already discussed, primary causes for the continued global distribution of these nonindigenous species are human activities such as shipping and aquaculture (Carlton, 1987).

### *Ecosystem Informatics for Marine Biogeography and Natural History*

While the best management option is to prevent introductions altogether, even with the best-case scenarios, a certain number of invaders will “slip through”. The best strategy then becomes early detection and rapid response to eradicate the invaders (see the [National Invasive Species Management Plan](#)). An inherent component of this strategy is baseline knowledge of the nonindigenous species present within a system, their primary distribution vectors (Ruiz *et al.*, 2000), and a baseline of the native species to assess what resources are at risk. Additionally, these baselines could provide information necessary to formulate ecological performance standards related to the rate of new introductions and/or to the extent of impacts on native biodiversity. As pointed out by the GAO (2003), the lack of such ecological performance standards is a major limitation of the [National Invasive Species Management Plan](#).

Unfortunately information on the pathways, distributions, and habitat requirements of native and nonindigenous species in marine/estuarine ecosystems is widely scattered in the literature and across various databases (Lee *et al.*, 2008b). Examples of local and regional nonindigenous species databases include the Guide to the Exotic Species in San

Francisco Bay database, (<http://www.exoticguide.org/>); the Delivering Alien Invasive Species Inventories for Europe database, (DAISIE: <http://www.europe-alien.org/index.jsp>); and the Nonindigenous Aquatic Species database (NAS: <http://nas3.er.usgs.gov/>) to name a few. While these databases provide information on nonindigenous species, they do not provide information on the current distribution and biodiversity of native species that may be displaced by continued invasions. Additionally, most of the existing nonindigenous species databases do not allow queries to extract information on multiple species.

An ecosystem informatics solution that integrates the distribution, dispersal, and natural history information for both native and nonindigenous marine and estuarine species was needed. Therefore, the first goal of this research was to design and describe a spatially explicit integrated framework for the storage and retrieval of biologic information for native and nonindigenous marine and estuarine species as described in chapter 2, submitted to the Journal of Biogeography.

Data collected in the information system described in Chapter 2 was utilized in Chapter 3 to determine invasion rates for Northeast Pacific estuaries. In addition, the hierarchical structure in the biogeographic framework described in Chapter 2 provided the spatial organization of the data for ecological modeling at two spatial scales within the hierarchy as described in Chapter 4.

### *Management Strategy for Ballast Water*

Recognizing the potential economic and ecological damage caused by nonindigenous species, government agencies are seeking ways to prevent new invasions, manage current

distributions, and monitor aquatic environments at risk. Because ballast water is recognized as one of the primary vectors for new aquatic invaders (Carlton & Geller, 1993; Ruiz & Carlton, 2003; Molnar *et al.*, 2008), the International Maritime Organization (IMO) is implementing an international treaty on ballast water exchange policies (Fofonoff *et al.*, 2003a; Minton *et al.*, 2005; Cordell *et al.*, 2009). Additionally, the United States Coast Guard (USCG) recently released proposed ballast water discharge limits based on organism concentrations (USCG, 2009) and the U.S. Environmental Protection Agency is in the process of evaluating similar standards under the Clean Water Act. The second goal of this research was to identify a potential approach to set criteria for managing the current influx of nonindigenous species through ballast water exchange described in Chapter 3 to be submitted to Biological Invasions. The per capita invasion probability approach is also being reviewed by the National Academy of Science along with several other approaches for establishment of ballast water discharge criteria.

### *Ecological Niche Modeling for Prediction*

While control and management of transport vectors will slow the influx of nonindigenous species, insights into where existing or new invaders are likely to spread enables best management practices to inhibit migration to new habitats. For example, the 100<sup>th</sup> meridian initiative was established to keep the zebra mussel from invading waters west of the Mississippi River (<http://www.100thmeridian.org/>). An alternative management strategy is rapid response and eradication similar to California's response to the first occurrence of *Caulerpa taxifolia* in Southern California (Jousson *et al.*, 2000). Or, in the

worst case, an adaptive management plan can be developed. All of these strategies can be optimized if the vulnerability of specific locales to invasion can be predicted.

Recent advancements in niche modeling techniques provide tools to utilize knowledge of native and invaded habitats and environments along with detailed environmental data layers to predict the potential range of a nonindigenous species. Niche modeling can also be used to predict new distribution patterns of established nonindigenous species in response to climate change. One niche modeling technique, McCune's HyperNiche (<http://www.hyperniche.com>), using Nonparametric Multiplicative Linear Regression (NMPR) methods, was evaluated in Chapter 4 to determine suitability for predicting potential range extent of already existing nonindigenous and native estuarine species in the Northeast Pacific. This article was published in the ICES Journal of Marine Science 65: 742-745. 2008.

## ***References***

- Åkesson, P. (2007) Boat & ship types.  
URL <http://www.abc.se/~pa/bld/shiptype.htm>.
- Alexandrov, B., Boltachev, A., Kharchenko, T., Lyashenko, A., Son, M., Tsarenko, P. & Zhukinsky, V. (2007) Trends of aquatic alien species invasions in ukraine. *Aquatic Invasions*, **2**, 215-242.
- Alpine, A.E.a.J.E.C. (1992) Trophic interactions and direct physical effects control phytoplankton biomass and production in an estuary. *Limnol. Oceanogr.*, **37**, 946-955.
- Altman, S. & Whitlatch, R.B. (2007) Effects of small-scale disturbance on invasion success in marine communities. *Journal of Experimental Marine Biology and Ecology*, **342**, 15-29.
- Ayres, D.R., Smith, D.L., Zaremba, K., Klohr, S. & Strong, D. (2004) Spread of exotic cordgrasses and hybrids (*spartina sp.*) in the tidal marshes of san francisco bay, california, USA. *Biological Invasions*, **6**, 221-231.

- Barnes, M.A. (2008) Fecundity of the exotic applesnail, *potamocorbula amurensis*. *Journal of the North American Benthological Society*, **27**, 738-745
- Benech, S.V. (1978) Ocean transport of a community of the grapsid crab *plagusia dentipes* (de haan, 1833). *Crustaceana*, **35**, 104-104
- Bilio, M. & Niermann, U. (2004) Is the comb jelly really to blame for it all? *Mnemiopsis leidyi* and the ecological concerns about the caspian sea. *Marine Ecology Progress Series*, **269**, 173-183
- Boudouresque, C.F. (1994) Introduced species in european coastal waters: Les especes introduites dans les eaux cotieres d'eupe et de mediterranee: Etat de la question et consequences. In: *Ecosystems Report Series* (ed. C.F. Boudouresque, Briand, F., Nolan, C.), pp. 8-27. Commission European Communities, Brussel
- Behrens Yamada, S. (2001) *Global invader: The european green crab*. Oregon State University, Corvallis, OR.
- Briggs, J.C. (2007) Marine biogeography and ecology: Invasions and introductions. *Journal of Biogeography*, **34**, 193-198
- California Sea Grant (1999-2000) California sea grant college program annual report 1999-2000. In, pp. 1-63. University of California, La Jolla, California
- Cardinale, B.J. & Palmer, M.A. (2002) Disturbance moderates biodiversity-ecosystem function relationships: Experimental evidence from caddisflies in stream mesocosms. *Ecology*, **83**, 1915-1927
- Carlton, J., J. Thompson, L. Schemel, and F.H. Nichols (1990) Remarkable invasion of san francisco bay (california, USA) by the asian clam *potamocorbula amurensis*. I. Introduction and dispersal. *Marine Ecological Progress Series*, **66**, 81-94
- Carlton, J.T. (1979) History, biogeography, and ecology of the introduced marine and estuarine invertebrates of the pacific coast of north america. *Ph. D. thesis. University of California, Davis*, 904 pages
- Carlton, J.T. (1987) Patterns of transoceanic marine biological invasions in the pacific ocean. *Bulletin of Marine Science*, **41**, 452-465
- Carlton, J.T. (1992) Introduced marine and estuarine mollusks of north america: An end-of-the-20th-century perspective. *Journal of Shellfish Research*, **11**, 489-505
- Carlton, J.T. (1996a) Marine bioinvasions: The alteration of marine ecosystems by nonindigenous species. *Oceanography*, **9**, 36-43

- Carlton, J.T. (1996b) Pattern, process, and prediction in marine invasion ecology. *Biological Conservation*, **78**, 97-106.
- Carlton, J.T. (2009) Deep invasion ecology and the assembly of communities in historical time. *Biological invasions in marine ecosystems. Ecological studies* (ed. by G. Rilov, And Crooks, J.A.), pp. 13-56. Springer-Verlag, Berlin Heidelberg.
- Carlton, J.T. & Geller, J.B. (1993) Ecological roulette: The global transport of nonindigenous marine organisms *Science*, **261**, 78-82.
- Carlton, J.T. & Ruiz, G.M. (2005) Vector science and integrated vector management in bioinvasion ecology: Conceptual frameworks. *Invasive alien species: A new synthesis* (ed. by H.A. Mooney, R.N. Mack, J.A. Mcneely, L.E. Neville, P.J. Schei and J.K. Waage), pp. 36-58. Island Press, Covelo, California.
- Chapman, J.W., Blakeslee, A.M.H., Carlton, J.T. & Bellinger, M.R. (2008) Parsimony dictates a human introduction: On the use of genetic and other data to distinguish between the natural and human-mediated invasion of the european snail *littorina littorea* in north america. *Biological Invasions*, **10**, 131-133.
- Chapman, J.W. & Carlton, J.T. (1991) A test of criteria for introduced species: The global invasion by the isopod *synidotea laevidorsalis* (miers, 1881). *Journal of Crustacean Biology*, **11**, 386-400.
- Chapman, J.W. & Carlton, J.T. (1994) Predicted discoveries of the introduced isopod *synidotea laevidorsalis* (miers, 1881). *Journal of Crustacean Biology* **14**, 700-714.
- Chapman, J.W. & Dorman, J.A. (1975) Diagnosis, systematics, and notes on *grandidierella japonica* (amphipoda: Gammaridea) and its introduction to the pacific coast of the united states. *bulletin of Southern California Academy of Science*, **74**, 104-108.
- Chapman, J.W., Miller, T.W. & Coan, E.V. (2003) Live seafood species as recipes for invasion. *Conservation Biology*, **17**, 1386-1395.
- Cohen, A.N. & Carlton, J.T. (1998) Accelerating invasion rate in a highly invaded estuary. *Science*, **279**, 555-558.
- Cohen, A.N., Harris, L.H., Bingham, B.L., Carlton, J.T., Chapman, J.W., Lambert, C.C., Lambert, G., Ljubenkov, J.C., Murray, S.N., Rao, L.C., Reardon, K. & Schwindt, E. (2005) Rapid assessment survey for exotic organisms in southern california bays and harbors, and abundance in port and non-port areas. *Biological Invasions*, **7**, 995-1002.

- Cohen, A.N. & Weinstein, A. (1998) The potential distribution and abundance of zebra mussels in California. In, p. 16. San Francisco Estuary Institute, San Francisco.
- Cohen, A.N. & Carlton, J.T. (1998) Accelerating invasion rate in a highly invaded estuary. *Science*, **279**, 555-558.
- Colautti, R.I., Ricciardi, A., Grigorovich, I.A. & Macisaac, H.J. (2004) Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, **7**, 721-733.
- Coleman, A.W. (1996) DNA analysis methods for recognizing species invasion: The example of *Codium*, and generally applicable methods for algae. *Hydrobiologia*, **326-327**, 29-34.
- Collin, R. (2003) Phylogenetic relationships among calyptraeid gastropods and their implications for the biogeography of marine speciation. *Systematic Biology*, **52**, 618-640.
- Cordell, J.R., Lawrence, D.J., Ferm, N.C., Tear, L.M., Smith, S.S. & Herwig, R.P. (2009) Factors influencing densities of non-indigenous species in the ballast water of ships arriving at ports in Puget Sound, Washington, United States. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **19**, 322-343.
- Cox, C.B. & Moore, P.D. (2005) *Biogeography: An ecological and evolutionary approach*, 7 edn. Blackwell Publishing, Malden, MA.
- Cronin, T.W. & Forward, R.B.J. (1979) Tidal vertical migration: An endogenous rhythm in estuarine crab larvae. *Science*, **205**, 1020-1022.
- Elton, C.S. (1958) *The ecology of invasions by animals and plants*, Methuen, London.
- Faasse, M.A. & Bayha, K.M. (2006) The ctenophore *Mnemiopsis leidyi* A. Agassiz 1865 in coastal waters of the Netherlands: An unrecognized invasion? *Aquatic Invasions*, **1**, 270-277.
- Fofonoff, P.W., Ruiz, G.M., Steves, B. & Carlton, J.T. (2003) In ships or on ships? Mechanisms of transfer and invasion for non-native species to the coasts of North America. *Invasive species: Vectors and management strategies* (ed. by G.M. Ruiz and J.T. Carlton), pp. 152-182. Island Press, Washington DC.
- Galil, B.S. (2000) A sea under siege - alien species in the Mediterranean. *Biological Invasions*, **2**, 177-186.
- Goren, M. & Aronov, A. (2002) First record of the Indo-Pacific parrot fish *Scarus ghobban* in the Eastern Mediterranean. *Cybium*, **26**, 239-240.

- Grantham, B., A., Eckert, G.L. & Shanks, A.L. (2003) Dispersal potential of marine invertebrates in diverse habitats. *Ecological Applications*, **13**, S108-S116.
- Grime, J.P. (1973) Competitive exclusion in herbaceous vegetation. *Nature*, **242**, 344-347.
- Grosholz, E.D. (1996) Contrasting rates of spread for introduced species in terrestrial and marine systems. *Ecology*, **77**, 1680-1686.
- Grosholz, E.D., Ruiz, G.M., Dean, C.A., Shirley, K.A., Maron, J.L. & Connors, P.G. (2000) The impacts of a nonindigenous marine predator in a california bay. *Ecology*, **81**, 1206-1224.
- Hedvall, O., Moksnes, P.-O. & Pihl, L. (1998) Active habitat selection by megalopae and juvenile shore crabs *carcinus maenas*: A laboratory study in an annular flume. *Hydrobiologia*, **375/376**, 89-100.
- Jeschke, J.M.a.S., David L. (2005) Invasion success of vertebrates in europe and north america. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 7198-7202.
- Johnson, L.E., Bossenbroek, J.M. & Kraft, C.E. (2006) Patterns and pathways in the post-establishment spread of non-indigenous aquatic species: The slowing invasion of north american inland lakes by the zebra mussel. *Biological Invasions*, **8**, 475-489.
- Jousson, O., Pawlowski, J., Zaninetti, L., Zechman, F.W., Dini, F., Di Guiseppe, G., Woodfield, R., Millar, A. & Meinesz, A. (2000) Invasive alga reaches california. *Nature*, **408**, 157-158.
- Kimmerer, W.J., E. Gartside, and J.J. Orsi (1994) Predation by an introduced clam as the probable cause of substantial declines in zooplankton in san francisco bay. *Marine Ecological Progress Series*, **113**, 81-93.
- Klautau, M., Russo, C.a.M., Lazoski, C., Boury-Esnault, N., Thorpe, J.P. & Sole-Cava, A.M. (1999) Does cosmopolitanism result from overconservative systematics? A case study using the marine sponge *chondrilla nucula*. *Evolution*, **53**, 1414-1422.
- Kolar, C.S.a.L., David M. (2001) Progress in invasion biology: Predicting invaders. *TRENDS in Ecology & Evolution*, **16**, 199-204.
- Lassuy, D.R. (1995) Introduced species as a factor in extinction and endangerment of native fish species. *American Fisheries Soc. Symp.*, **15**, 391-396.
- Lee, C.E. (1999) Rapid and repeated invasions of fresh water by the copepod *eurytemora affinis*. *Evolution*, **53**, 1423-1434.

- Lee, H., Reusser, D., Ranelletti, M. & Nehmer, R. (2008) *User guide for the pices nonindigenous species information system templates, vers. Gamma. Newport, or, USA.*
- Lee, H. & Reusser, D.A. (2006) Pacific coast ecosystem information system (*pceis*). Beta version of geospatial database of native and nonindigenous estuarine/coastal species of the northeast pacific and associated landscape data (internal epa and usgs release). In. Office of Research and Development, National Health and Environmental Effects Research Laboratory.
- Mack, R., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. Bazzaz (2000) Biotic invasions: Causes, epidemiology, global consequences and control. *Ecological Applications*, **10**, 689-710.
- Mihulka, S., Pysek, P., Martinkova, J. & Jarosik, V. (2006) Invasiveness of oenothera congener alien to europe: Jack of all trades, master of invasion? *Perspectives in Plant Ecology, Evolution and systematics*, **8**, 83-96.
- Minchin, D., Gollasch, S., Cohen, A.N., Hewitt, C.L. & Olenin, S. (2009) Chapter 5 characterizing vectors of marine invasion. *Biological invasions in marine ecosystems* (ed. by G. Rilov, And Crooks, J.A.), pp. 109-116. Springer-Verlag, Berlin, Heidelberg.
- Minton, M.S., Verling, E., Miller, A.W. & Ruiz, G.M. (2005) Reducing propagule supply and coastal invasions via ships: Effects of emerging strategies. *Frontiers in Ecology and the Environment*, **3**, 304-308.
- Molnar, J.L., Gamboa, R.L., Revenga, C. & Spalding, M.D. (2008) Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, **6**, 485-492.
- Nehring, S. (2006) Four arguments why so many alien species settle into estuaries, with special reference to the german river elbe. *Helgol. Marine Res.*, **60**, 127-134.
- Nichols, F.H., J.K. Thompson and L.E. Schemel (1990) Remarkable invasion of san francisco bay (san francisco, ca) by the asian clam potamocorbula amurensis. II. Displacement of a former community. *Marine Ecological Progress Series*, **66**, 95-101.
- Paavola, M., Olenin, S. & Leppakoski, E. (2005) Are invasive species most successful in habitats of low native species richness across european brackish water seas? *Estuarine, Coastal and Shelf Science*, **64**, 738-750.
- Papadopoulos, I., Wooldridge, T.H. & Newman, B.K. (2002) Larval life history strategies of sub-tropical southern african estuarine brachyuran crab and the implications for tidal inlet management. *Wetlands Ecology and Management*, **10**, 249-256.

- Pearcy, W.G. & Schoener, A. (1987) Changes in the marine biota coincident with the 1982-1983 el niño in the northeastern subarctic pacific ocean. *Journal of Geophysical Research*, **92**, 14,417-14,428.
- Peterson, W.T., Keister, J.E. & Feinberg, L.R. (2002) The effects of the 1997-99 el niño/ la niña events on hydrography and zooplankton off the central oregon coast. *Progress in Oceanography*, **54**, 381-398.
- Pimentel, D., Lach, L., Zuniga, R., Morrison, D. (2000) Environmental and economic costs associated with non-indigenous species in the united states. *BioScience*, **50**, 53-65.
- Pimentel, D., S. McNair, J. Janecka, J. Wightman, C. Simmonds, C. O'connell, E. Wong, L. Russel, J. Zern, T. Aquino, and T. Tsomondo (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems and Environment*, **84**, 1-20.
- Queiroga, H. (1998) Vertical migration and selective tidal stream transport in the megalopa of the crab *carcinus maenas*. *Hydrobiologia*, **375/376**, 137-149.
- Queiroga, H., Costlow, J.D. & Moreira, M.H. (1997) Vertical migration of the crab *carcinus maenas* first zoea in an estuary: Implications for tidal stream transport. *Marine Ecology Progress Series*, **149**, 121-132.
- Ribera, M.A. & Boudouresque, C.F. (1995) Introduced marine plants, with special reference to macroalgae: Mechanisms and impact. In: *Progress in phycological Research* (ed. D.J. Chapman), Bristol.
- Roman, J. & Palumbi, S.R. (2004) A global invader at home: Population structure of the green crab, *carcinus maenas*, in europe. *Molecular Ecology*, **13**, 2891-2898.
- Ruiz, G.M. & Carlton, J.T. (2003) Invasion vectors: A conceptual framework for management. *Invasive species: Vectors and management strategies* (ed. by G.M. Ruiz and J.T. Carlton), p. 518. Island Press, Washington.
- Ruiz, G.M., Carlton, J.T., Grosholz, E.D. & Hines, A.H. (1997) Global invasion of marine and estuarine habitats by nonindigenous species: Mechanisms, extent, and consequences. *American Zoologist*, **37**, 621-632.
- Ruiz, G.M., Fofonoff, P., Hines, A.H. & Grosholz, E.D. (1999) Non-indigenous species as stressors in estuarine and marine communities: Assessing invasion impacts and interactions. *Limnology and Oceanography*, **44**, 950-972.

- Ruiz, G.M., Fofonoff, P.W., Carlton, J.T., Wonham, M.J. & Hines, A.H. (2000) Invasion of coastal marine communities in north america: Apparent patterns, processes, and biases. *Annual Review of Ecology and Systematics*, **31**, 481-531.
- Sakai, A.K., Allendorf, F. W., Holt, J.S., Lodge, D. M., Molofsky, J., with, K.A., Baughman, S., Cabin, R.J., Cohen, J. E., Ellstrand, N.C., Mccauley, D.E., O'neal, P., Parker, I.M., Thompson, J.N., Weller, S.G. (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305-322.
- Sala, O.E., Chapin Iii, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M. & Wall, D.H. (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770-1774.
- Sanders, N.J., Gotelli, N.J., Heller, N.E. & Gordon, D.M. (2003) Community disassembly by an invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 2474-2477.
- Scheltema, R.S. (1988) Initial evidence for the transport of teleplanic larvae of benthic invertebrates across the east pacific barrier. *Biological Bulletin*, **174**, 145-172.
- Scheltema, R.S. & Williams, I.P. (1983) Long-distance dispersal of planktonic larvae and the biogeography and evolution of some polynesian and western pacific mollusks. *Bulletin of Marine Science*, **33**, 545-565.
- Sea Grant (2006) Don't release a pest: Freezing is best! Public Service Announcement and animated video for k-12 education. University of California and University of Central Florida.
- Semmens, B.X., Buhle, E.R., Salomon, A.K. & Pattengill-Semmens, C.V. (2004) A hotspot of non-native marine fishes: Evidence for the aquarium trade as an invasion pathway. *Marine Ecological Progress Series*, **266**, 239-244.
- Smith, S.A., Bell, G. & Bermingham, E. (2004) Cross-cordillera exchange mediated by the panama canal increased species richness of local freshwater fish assemblages. *The Royal Society Biological Science*, **272**, 1889-1896.
- Sytsma, M.D., Cordell, J.R., Chapman, J.W. & Draheim, R.C. (2004) Lower columbia river aquatic nonindigenous species survey 2001-2004. Final technical report. In, p. 69. Prepared For The United States Coast Guard And The United States Fish And Wildlife Service.
- Thiel, M. (2003) Rafting of benthic macrofauna: Important factors determining the temporal succession of the assemblage on detached macroalgae. *Hydrobiologia*, **503**, 49-57.

- Toft, J.D., Cordell, J.R. & Fields, W.C. (2002) New records of crustaceans (amphipoda, isopoda) in the sacramento/san joaquin delta, california, and application of criteria for introduced species. *Journal of Crustacean Biology*, **22**, 190-200.
- USCG. 2009. Proposed Ballast Water Discharge Standard Rulemaking. Fact sheet produced for U.S. Coast Guard Office of Operating and Environmental Standards. ([http://www.safinah.co.uk/uploads/assets/PDFs/Ballast\\_Water\\_Discharge\\_Standard.pdf](http://www.safinah.co.uk/uploads/assets/PDFs/Ballast_Water_Discharge_Standard.pdf)). The proposed rulemaking can be found at: <http://www.regulations.gov>. Enter docket number USCG-2001-10486 in the Search box.
- Vitousek, P.M., Mooney, H.A., Lubchenco, J. & Melillo, J.M. (1997) Human domination of earth's ecosystems. *Science*, **277**, 494-499.
- Waples, R.S. (2004) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438-450.
- Watts, P.C., Thorpe, J.P. & Taylor, P.D. (1998) Natural and anthropogenic dispersal mechanisms in the marine environment: A study using cheilostome bryozoa. *Philosophical Transactions of The Royal Society Biological Sciences*, **353**, 453-464.
- Westheide, W. & Schmidt, H. (2003) Cosmopolitan versus cryptic meiofaunal polychaete species: An approach to a molecular taxonomy. *Helgoland Marine Research*, **57**, 1-6.
- Whitfield, P.E., Gardner, T., Vives, S., Gilligan, M.R., Courtenay Jr., W.R., Ray, G.C. & Hare, J.A. (2002) Biological invasion of the indo-pacific lionfish *pterois volitans* along the atlantic coast of north america. *Marine Ecology Progress Series*, **235**, 289-297.
- Wilcove, D.S., D. Rothstein, J. Dubow; A. Phillips and E. Losos. (1998) Quantifying threats to imperiled species in the united states. Assessing the relative importance of habitat destruction, alien species, pollution, overexploitation, and disease. *BioScience*, **48**, 607-615.
- Wolff, W. (1999) Exotic invaders of the meso-oligohaline zone of estuaries in the netherlands: Why are there so many? *Helgolander Meeresunters*, **52**, 393-400.
- Wolff, W. (2005) Non-indigenous marine and estuarine species in the netherlands. *Zool. Med. Leiden*, **79**, 1-116.
- Zeng, C. & Naylor, E. (1996) Endogenous tidal rhythms of vertical migration in field collected zoea-1 larvae of the shore crab *carcinus maenas*: Implications for ebb tide offshore dispersal. *Marine Ecology Progress Series*, **132**, 71-82.

Zeng, C., Naylor, E. & Abello, P. (1997) Endogenous control of timing of metamorphosis in megalopae of the shore crab *Carcinus maenas*. *Marine Biology*, **128**, 299-3005.

**Chapter 2: Evolution of natural history information in the 21<sup>st</sup> Century  
– Developing an integrated framework for biological and geographic  
data**

Authors: Deborah A. Reusser and Henry Lee II

## ***Abstract***

**Aim:** Threats to marine and estuarine species operate over many spatial scales, from nutrient enrichment at the watershed/estuarine scale to invasive species and climate change at regional and global scales. To help address issues across these scales, we developed a conceptual framework for a spatially explicit biological information system containing queryable biological data that allows extraction of information on multiple species, across a variety of spatial scales based on species' distributions, natural history attributes, and habitat requirements.

**Location:** Global with examples for near-coastal species in the Northeast Pacific.

**Methods:** A suite of classification schemas were developed or integrated to capture species' biogeographic distributions, natural history, and habitat attributes. When possible, a hierarchical classification typology was developed to capture information at multiple levels of detail. Reproduction, development, feeding, life style, salinity, and habitat association attributes all fit into hierarchical schemas. In cases where the species' attribute was not hierarchical, a multi-dimensional schema was designed. These biotic attributes were integrated with a modified version of the Marine Ecoregion of the World (MEOW) biogeographic schema for near-coastal waters. To connect near-shore environments to watershed characteristics, a land-sea framework was developed.

**Results:** An integrated framework was designed to capture life history characteristics, environmental preferences and geographic distribution information for marine and

estuarine flora and fauna. Key aspects of this framework include: 1) consistent terminology; 2) translation of numerical habitat values and physiological requirements into classes; 3) classification schemas for natural history, environmental attributes, and geographic distributions; and 4) integration of biotic attributes to allow database queries on single or multiple species across different spatial scales. Examples of how biotic attributes and biogeographic distributions can be integrated are illustrated with chitons in the Northeast Pacific.

**Main Conclusions:** As scientists shift from research on localized impacts on individual species to regional and global scale threats, macroecological approaches of studying multiple species over broad geographical areas are becoming increasingly important. A structured framework as described here, for biological and geographical data storage and retrieval, is a critical first step towards addressing these macroecological questions.

**Keywords:** Hierarchical schemas, Classification schemas, Natural history, Biological information systems, Biogeography, Marine Ecosystems, Chitons

### ***Introduction***

The extensive biological and ecological knowledge collected on flora and fauna of the world over the last two to three hundred years primarily resides in a wealth of encyclopedias, books, and journal articles. Two recent examples of attempts to synthesize this diffuse information are *Fishes of Alaska* (Mecklenburg *et al.*, 2002) and *The Light and Smith Manual – Intertidal Invertebrates from Central California to Oregon* (Carlton, 2007). Each of these tomes contain over a thousand pages of biological information

including habitat, distributions, and life history characteristics. Today, advanced search engines available on the World Wide Web (WWW) also make it possible to discover a vast amount of text-based literature on life histories and habitat requirements for many species one at a time. In addition, some of these data have been captured in electronic database format and can be accessed via the WWW (e.g., FishBase, <http://www.fishbase.org/search.php> and GBIF, <http://data.gbif.org>). However, with very few exceptions (e.g., Gulf of California Invertebrate Database; <http://www.desertmuseum.org/center/seaofcortez/searchdb.php>), the existing formats do not provide a simple way to relate biotic attributes of marine and estuarine species to their biogeographic distributions. Furthermore, although these text-based reports and web searches make it possible to extract considerable information about a single species, they do not provide a straightforward approach to extracting groups of species with similar attributes, environmental preferences, and/or distributions. Nonetheless, there is a growing need to synthesize information on multiple species from multiple locations to address regional and global issues, such as climate change and invasive species (Lee *et al.*, 2008b). Unfortunately, because of the time investment required to extract information for multiple species and/or locations from the plethora of text-based sources, existing data are often underutilized (Kerr *et al.*, 2007). Another challenge in synthesizing these data is the idiosyncratic nature of much of the natural history data. For example, depth ranges of marine species are given in feet, fathoms, and meters, while the term “littoral” can mean either intertidal or the intertidal and shallow subtidal (Open University, 2000).

To utilize the existing knowledge better, integrated biological information systems are needed that allow the organization and extraction of both quantitative and qualitative natural history and distributional data. One approach to help organizing such diverse types of information is to develop classification systems. There have been a few attempts to develop classification schemas for the natural history of marine/estuarine organisms. For example, Bush *et al.* (2007) and Bambach *et al.* (2007) developed a schema for “ecospace utilization” in marine fossil assemblages based on three attributes – life position relative to the sediment-water interface (tiering), motility, and feeding mechanism. Fauchald & Jumars (1979) developed feeding guilds of polychaete worms based on three criteria: what the species feeds on, motility of the species, and its feeding morphology. Both of these systems are multi-dimensional in the sense that they combine several attributes to define each class or guild. Although these classification systems are useful in capturing specific types of natural history attributes, integrating multiple classification schemas within a single framework can increase the utility of the information and provide insights not possible when evaluating a single type of biotic attribute.

To address this need, we designed a biological information framework for the synthesis of natural history, environmental, and geographic information for marine and estuarine species. Key aspects of this framework include: 1) consistent terminology for natural history requirements; 2) translation of specific numerical values for habitat/physiological requirements (e.g., salinity tolerances) into classes; 3) either multi-dimensional or hierarchical classification schemas for natural history, environmental attributes, and

geographic distributions depending on which schema best fits the data; and 4) an integration of the various biotic attributes to allow database queries on single or multiple species across different spatial scales. This integrated biological information framework was developed for data storage and retrieval of marine/estuarine organisms in the “Pacific Coast Ecosystem Information System” (*PCEIS*). However, the concepts presented here are applicable to biological information systems in general, including freshwater and terrestrial biota. To illustrate how some of the concepts can be applied, we present examples based on the species richness of soft-bottom assemblages in the Pacific Northwest and on the biogeographic and bathymetric distributions of chitons (Mollusca: Polyplacophora) in the Northeast Pacific.

### ***Classification Schemas to Capture Environmental and Natural History Attributes***

Quantifying patterns of similarity/dissimilarity across multiple species is difficult because much of the existing natural history data are qualitative and often anecdotal. Additionally, there is often a lack of accepted terminology across different taxa and/or habitat types. The use of classes is a practical approach to capturing and organizing such natural history information and simplifies data queries to analyze patterns for multiple species and/or locations. Additionally, translating natural history attributes into standardized classes imposes a rigorous terminology as well as allowing an estimation of a class value when quantitative values are not available. For example, exposure classes are often used in assessing intertidal habitats because of the difficulty in quantifying actual wave energy (Lindgarth & Gamfeldt, 2005). Another use of environmental or habitat classes is to fill

in gaps when there are missing quantitative values. For example, it is possible to estimate temperature classes for a species from an analysis of the spatial locations where the species is found. Probably the greatest limitation of classes is that they require simplification at the potential loss of some ecological nuances. The importance of the loss of species-specific or quantitative environmental information will depend upon the nature of the questions being addressed as well as the inherent resolution of the classes.

The simplest type of classification structure occurs when an attribute can be divided into two or more unique classes without the need for subclasses. Examples of characteristics and habitats of biota that can be captured in a non-hierarchical class structure include habitat regime, temperature regimes, wave energy, and population status of nonindigenous species (e.g., established, not established, and stocked). The classification schemas adopted for temperature and wave/current energy are described below.

### *Temperature Classification Schema*

Temperature classes used in *PCEIS* are based on annual and seasonal temperature ranges for geographical regions adapted from (Hall, 1964). These classes (Table 2-1) are based on mean monthly temperatures in the ecoregions occupied by the species and do not take into account local or habitat differences in temperature (e.g., upper intertidal versus subtidal).

A more up-to-date analysis of coastal temperature regimes may change the criteria for the classes as well as how the classes are calculated (e.g., using site-specific temperature

<b>Temperature Classes</b>	<b>Definition</b>
Cold water	No months $>10^{\circ}$ C with minimum approaching $0^{\circ}$ C.
Cool temperate	Less than 4 months $>10^{\circ}$ C
Mild temperate	Six months at $10^{\circ}$ C and $\leq 4$ months at $15^{\circ}$ C
Warm temperate	No months cooler than $10^{\circ}$ C and $\geq 4$ months $\geq 15^{\circ}$ C
Outer tropical	No months cooler than $10^{\circ}$ C and approx. 4 months at or near $20^{\circ}$ C
Inner tropical	No months cooler than $18^{\circ}$ C and $\geq 6$ months at or near $20^{\circ}$ C

Table 2-1: Temperature classes for marine/estuarine species based on annual temperature ranges for geographical regions, adapted from (Hall, 1964). These classes are based on mean monthly temperatures in the ecoregions occupied by the species and do not take into account local differences in temperature.

ranges instead of those based on regions). However, even a coarse classification such as this can be useful for some ecological questions. For example, such coarse temperature classes have been used in an environmental matching risk assessment to identify marine/estuarine species in one area that could be transported via shipping and potentially invade another area with similar environmental conditions (Gollasch, 2006). With freshwater fishes, simply assigning them to cold-water or warm-water thermal guilds was used to identify habitat availability under climate induced changes in lakes and streams (Eaton & Scheller, 1996; Stefan *et al.*, 2001).

### *Wave Energy Classification Schema*

Wave energy is another important habitat attribute for open coast intertidal communities, as their biotic assemblages are generally defined by the intensity of wave exposure (Howes *et al.*, 1999). However, actual wave energy is difficult to quantify so classes have been substituted in many past analyses. Lindegarth and Gamfeldt (2005) compared categorical wave exposure classes to continuous values and concluded that categorical values are meaningful substitutions. They point out however, that the mechanism used to define wave classes has important consequences, identifying wave height as a good tool

<b>Wave Energy Classes</b>	<b>Definition</b>
Exposed	“High ambient wave conditions usually prevail within this exposure category, which is typical of open-Pacific type conditions.” Max. fetch distances >500 km.
Semi-exposed	“Swells, generated in areas distant from the shore unit create relatively high wave conditions. During storms, extremely large waves create high wave exposures.” Max. fetch distance between 50 and 500 km.
Semi-protected	“Waves are low most of the time except during high winds.” Fetch in range of 10 – 50 km.
Protected	“Usually areas of provisional anchorages and low wave exposure except in extreme winds.” Fetch <10 km.
Very Protected	“Usually the location of all-weather anchorages, marinas and harbors.” Max. fetch <1 km.

Table 2-2: Wave energy classes for marine/estuarine species based on engineering methods for estimating wave heights using wind speed and direction from (Howes *et al.*, 1999).

for defining wave classes. Accordingly, we adopted wave energy classes for marine/estuarine species (Table 2-2) that are derived from maximum fetch and modified effective fetch as defined by (Howes *et al.*, 1999), who provide specific procedures for calculating wave energy based on engineering methods used to estimate wave heights from wind speed and direction.

### ***Hierarchical Schemas for Environmental and Natural History Attributes***

Darwin's observation that "All organic beings are found to resemble each other in descending degrees" (Darwin, 1859) can be applied to most life history characteristics. Specifically, a hierarchical classification structure allows capturing biotic attributes at descending levels of detail. This philosophy was used in developing life history and habitat classification schemas to the extent practical. An advantage of a hierarchical structure is that information with different levels of resolution can be integrated into a single system. For example FishBase only provides three broad classes for the salinity tolerances of fish (freshwater, brackish, or marine). Other sources however, provide salinity information at a higher level of detail, such as the National Exotic Marine and Estuarine Species Information System (NEMESIS), that provides quantitative salinity tolerances for nonindigenous species in Chesapeake Bay (Fofonoff *et al.*, 2003b). The hierarchical classification schema for salinity discussed below provides the ability to capture both levels of information. When analyzing data, an advantage of the hierarchical structure is that it provides the ability to extract information synthesized at an appropriate level of detail for a specific question. For example, using niche models to predict the distribution of near-coastal benthos Reusser and Lee (2008) showed that habitat and

physiological classes are nearly as predictive of distributions as quantitative values, indicating that the small loss in model precision was more than offset by the efficiency of using class values for multiple species. Hierarchical schemas also provide flexibility, allowing the addition of new subclass levels if they prove necessary.

To evaluate the practicality of natural history hierarchies, we developed an integrated, multi-level classification schema for a range of life history characteristics and physiological tolerances. Some biotic characteristics and habitat requirements were readily adapted to a hierarchical schema, such as taxonomy, salinity, and depth. Other attributes were not as obviously hierarchical, such as reproductive mode, trophic level, feeding mode, substrate type, invasion vectors for nonindigenous species, and ecosystem type. We propose schemas for these attributes, recognizing that alternative structures are possible. To the extent possible, this framework either utilizes existing classifications as one of the levels or a cross-walk is provided between the two. For example, Fauchald and Jumars (1979) had two classes of deposit feeders: surface deposit feeding species and burrowers (= subsurface deposit feeders). In our system, we added a higher level (deposit feeders) to capture the general feeding mode when the specifics are not known. With other attributes, we added finer resolution classes to capture more detailed information. For the suspension feeding mode, we added lower level classes that define whether species are obligate or facultative suspension feeders and whether they are active or passive suspension feeders. A portion of these hierarchical classification schemas are highlighted below as examples.

### *Taxonomic Hierarchical Schema*

The grandfather of all biological hierarchical classification schemas is the Linnaean Taxonomic Tree of Life. In 1758, one hundred years before Darwin's theory of evolution, Carolus Linnaeus published his hierarchical classification schema for flora and fauna (Linnaeus, 1758). The current version of Linnaeus's scientific classification schema, shown in Figure 2-1, has undergone modifications through the centuries, but is still the recognized standard. Web based versions of this standard taxonomic structure are available from several sources, such as the Integrated Taxonomic Information System (ITIS; <http://www.itis.gov>) and the World Register of Marine Species (WoRMS; <http://www.marinespecies.org>).

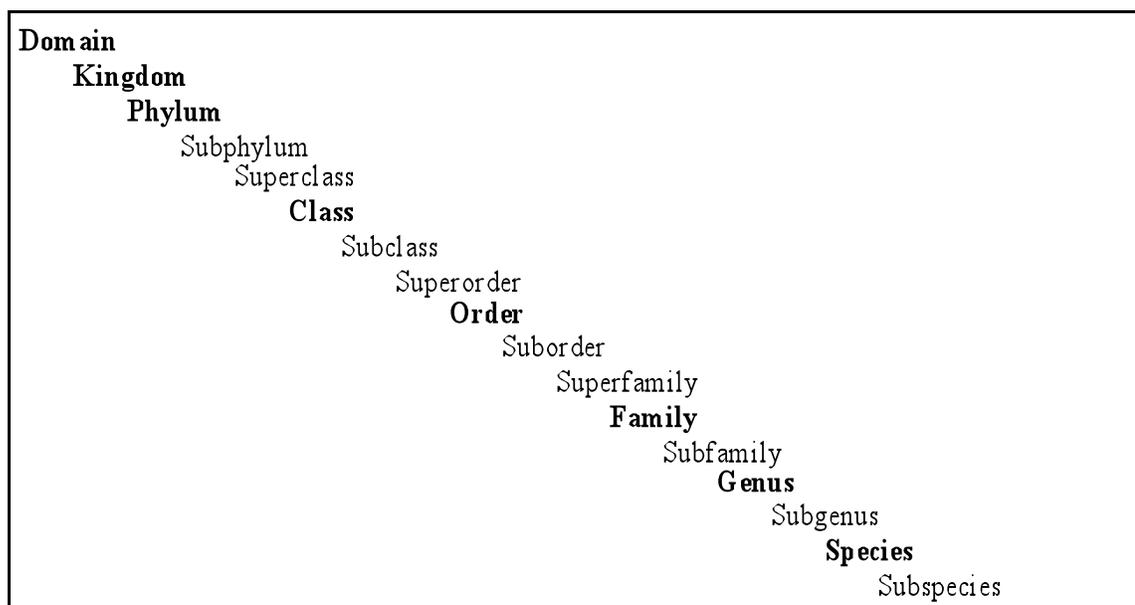


Figure 2-1: Taxonomic Hierarchy often referred to as the Tree of Life. Bold text indicates a level from the original Linnaean Taxonomic Hierarchy developed by Carolus Linnaeus in 1758 (Linnaeus, 1758).

Biological information systems should capture the full taxonomic tree for biota to allow detailed analyses of community assemblages using metrics such as taxonomic distinctness (Clarke & Warwick, 1999), and analyses within and among taxonomic groups (Koleff *et al.*, 2003). However, the constant revision of species names and their taxonomies is a major challenge for researchers. To assist the non-expert with these changes, an integrated biological information framework should capture and search for synonyms and common misspellings as well as the most current name. This provides a mechanism for conversion of historical names to current naming conventions for the standardization of species occurrences through time and in different places, which is critical in assessing historical changes in species distributions in response to climate change or species invasions. For example, the literature and information available on chitons on the Pacific Coast spans more than a century. During that time, genus and species names have been modified, changed, merged or split into a compendium of different names that confound the integration of natural history and distribution data for this taxon. Almost half of the chiton names for the California species listed in Hemphill (1890) have changed, and Eernisse (1986) gives a detailed example of the tortuous changes in the names of three species of *Lepidochitona* from 1892 to 1983 (Eernisse, 1986). Capturing these invalid names provides a cross walk through time of name changes for a single species.

### *Salinity Hierarchical Schema*

Salinity is considered a “master variable” in controlling the distribution of marine/estuarine species (Hodgkin, 1987) and provides another good example of a

hierarchical classification scheme. A standard hierarchical classification system for salinity was defined in 1958 at the “Symposium on the Classification of Brackish Waters” (IUBS, 1958). The resulting “Venice System” consists of six classes (freshwater, oligohaline, mesohaline, polyhaline, euhaline, and hypersaline), and this non-hierarchical system has been widely adopted. This symposium also defined a higher resolution hierarchical schema for the Baltic Sea where the oligohaline, mesohaline, and polyhaline classes were each divided into two classes (Figure 2-2). This higher resolution classification apparently has not been used outside of the Baltic region even though it allows a better characterization of estuarine habitats. Because the higher resolution subclasses were based on the lower salinity Baltic, the euhaline class was not subdivided. Accordingly, we divided the Venice euhaline class into two subclasses, using 36 practicalsalinity units (psu) as the breakpoint. To allow more flexibility, we then created a hierarchical system modeled after the Venice system by incorporating the coarser salinity levels used in FishBase (Figure 2-2). This schema generates a three-level hierarchical classification, from the coarse scale (freshwater, brackish, and marine) to the six-level Venice system, and finally to the ten-level “modified Venice system” if higher resolution data are available.

To assess the extent of information lost by going to a classification system versus using numerical salinity values, we evaluated the relationship between salinity and species richness in the soft-bottom benthos. While we recognize that salinity is only one environmental characteristic of an estuarine environment affecting the number of species,

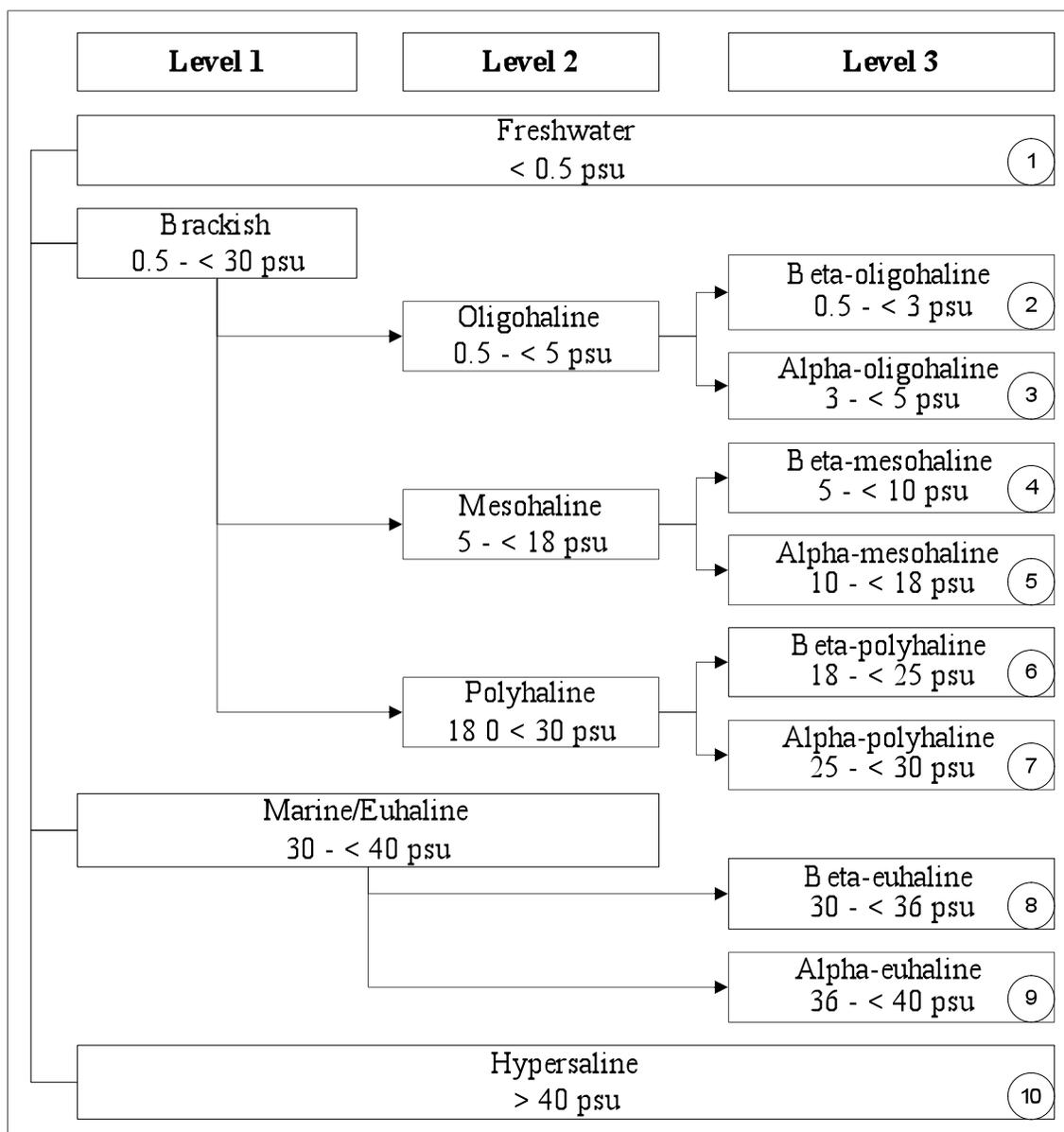


Figure 2-2: Three level hierarchical schema for salinity. The first 3 classes (Freshwater, Brackish and Marine) of Level 1 are used in FishBase while Level 2 is the Venice System developed by the International Union for Biological Sciences (IUBS, 1958). Level 3 is the modified Venice System based on the Baltic system (IUBS, 1958) enhanced in the hierarchical schema to include beta- and alpha-euhaline levels for marine waters. Salinity is given as practical salinity units (psu). Numbers in the small circles in the lower right hand side of the class boxes refer to the salinity classes used in the analysis shown in Figure 2-4.

salinity-species richness relationships have been used to assess the ecological condition of the benthos (Josefson & Hansen, 2004). Using non-impacted sites from a database of benthic samples compiled from multiple ecological monitoring studies of the near-coastal waters of California, Oregon, and Washington (N=531) (see (Nelson *et al.*, 2005) for discussion of methods), we regressed the number of species per sample on continuous salinity (psu) measurements taken at the time of sampling (Figure 2-3). This analysis was repeated after converting the numerical salinity values to classes based on the modified Venice System defined in Figure 2-2. An ANOVA analysis was used because the intervals between salinity classes are not equally spaced, (Figure 2-4). The regression analysis with salinity modeled as a continuous variable yielded an  $R^2 = 0.21$ , whereas, the analysis with salinity as a categorical variable (i.e., modified Venice) had an  $R^2 = 0.22$ . The similar  $R^2$  values of the two models suggest there is little or no loss of precision by using salinity classes compared to the site-specific numerical values. In addition, more samples could be evaluated using the modified Venice System classes because samples that did not have a site-specific salinity measurement could be plotted in a Geographic Information System (GIS) and a salinity class assigned based on its proximity to other samples where the salinity was known. Depending upon the density and spatial location of previous salinity measurements, the sites without salinity measurements can be classified at any of the three levels of the salinity hierarchy.

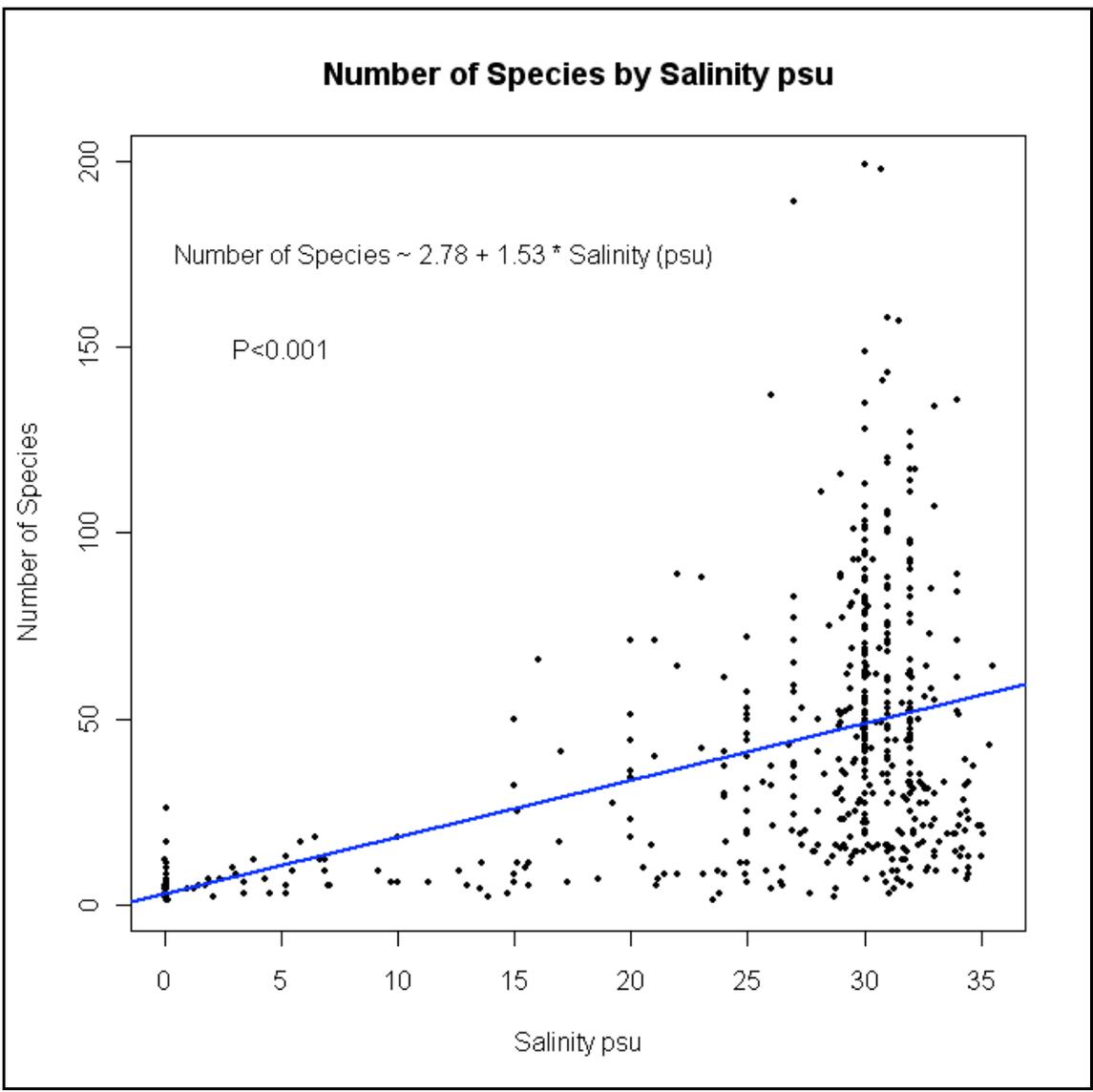


Figure 2-3: Number of species per benthic sample relative to the salinity value at the time of sampling. Data are from near-coastal monitoring studies in California, Oregon, and Washington. Only sites classified as undisturbed were used in order to reduce the potential for anthropogenic influences on species richness. Predictive equation based on least squares regression model ( $R^2 = 0.21$ ;  $N=526$ ;  $df=1, 524$ ;  $t\text{-value}=11.78$ ;  $P < 0.001$ ).

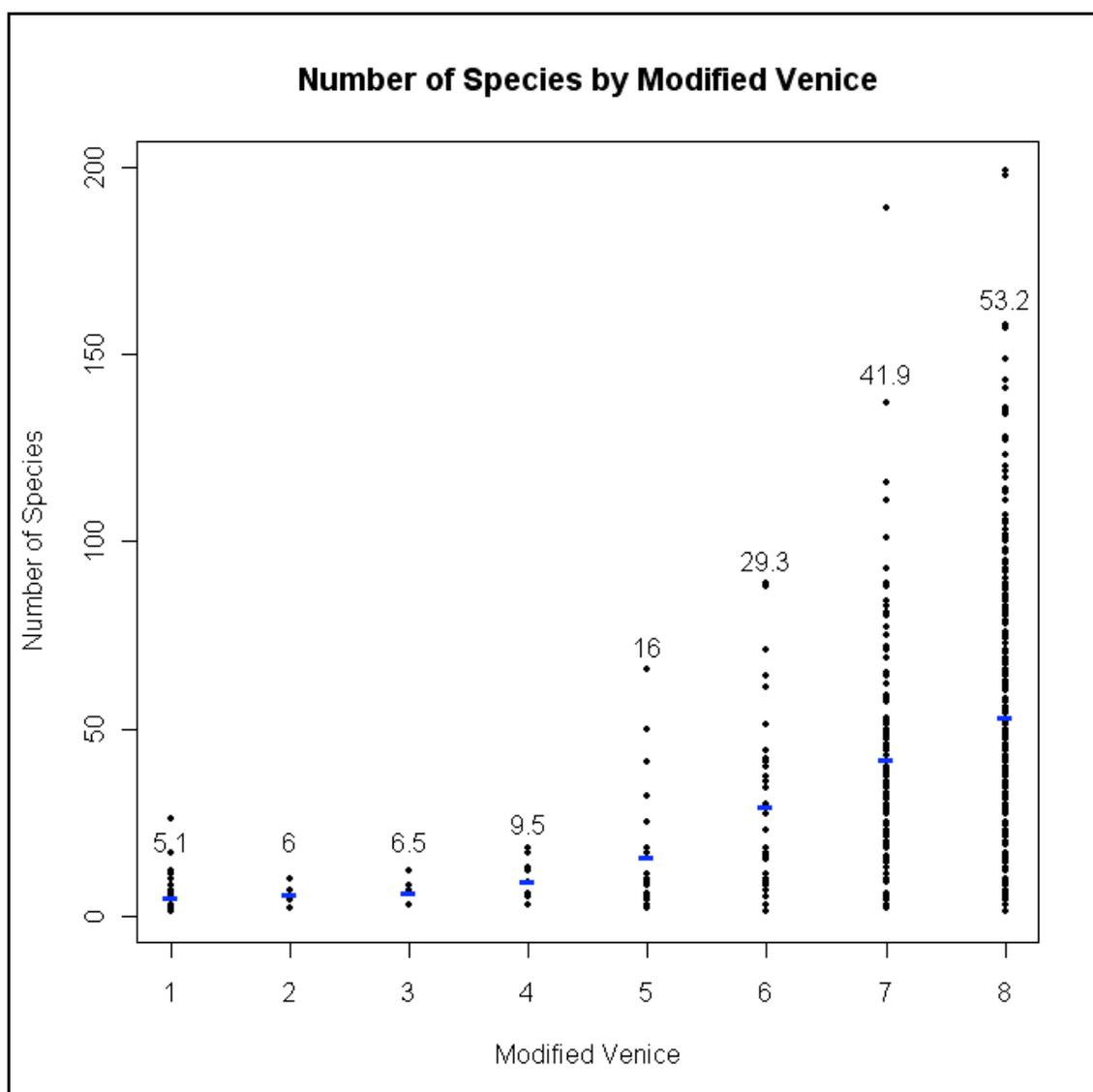


Figure 2-4: Salinity values from benthic samples in Figure 2-3 converted to modified Venice salinity classes. Modified Venice class numbers relate to the salinity classes shown in Figure 2-2 where 1=0-0.5 psu; 2=0.5-<3 psu; 3=3-<5 psu; 4=5-<10 psu; 5=10-<18 psu; 6=18-<25 psu; 7=25-<30 psu; 8=30-<36 psu. Values and blue horizontal bars indicate mean number of species within each salinity class (ANOVA;  $R^2 = .22$ ;  $N=531$ ;  $df = 7, 523$ ;  $F=23.71$ ;  $P < 0.001$ ). As indicated in figure 2-3, actual salinity values of the samples ranged from 0 - 35 psu. Since no samples existed for salinity classes 9 & 10, these classes were excluded from this graph.

### *Depth Hierarchical Schema*

Depth is another key factor affecting the distribution of marine and estuarine species (Madden *et al.*, 2005). Historically, different bathymetric classification systems have been used for benthic versus pelagic species, and we maintain this dichotomy as the first level of the depth hierarchy (Figure 2-5). The pelagic zone is divided into five generally recognized classes (epipelagic to hadopelagic). To better capture the distribution of near-surface pelagic flora and fauna, we adopted a three level split of the epipelagic class into surface, shallow, and deep zones. The oceanic environment is commonly divided into five main benthic classes, from the intertidal to the hadal. To expand the range to include semi-terrestrial environments, we incorporated the supralittoral zone and coastal fringe (e.g., dunes). Other than splitting the intertidal into upper, middle, and lower zones, there is no generally agreed upon subdivision of the neritic, bathyal, and abyssal zones. The 30 meter depth threshold for shallow subtidal habitats in the neritic zone is based on Madden *et al.* (2005), though some other classification systems use 20 meter as the threshold (e.g., Biological Traits Information Catalogue, <http://www.marlin.ac.uk/biotic>).

With the potential effects of ocean acidification on deeper water calcareous species (Kleypas *et al.*, 2006), higher resolution depth subclasses for deep oceanic waters have become important to help identify species at greatest risk. To capture this higher resolution, the subdivision of the bathyal and abyssal zones proposed by Berggren and Miller (1989) for benthic foraminifera were incorporated with one modification. Berggren and Miller divided the bathyal zone into three subclasses but the abyssal into only two.

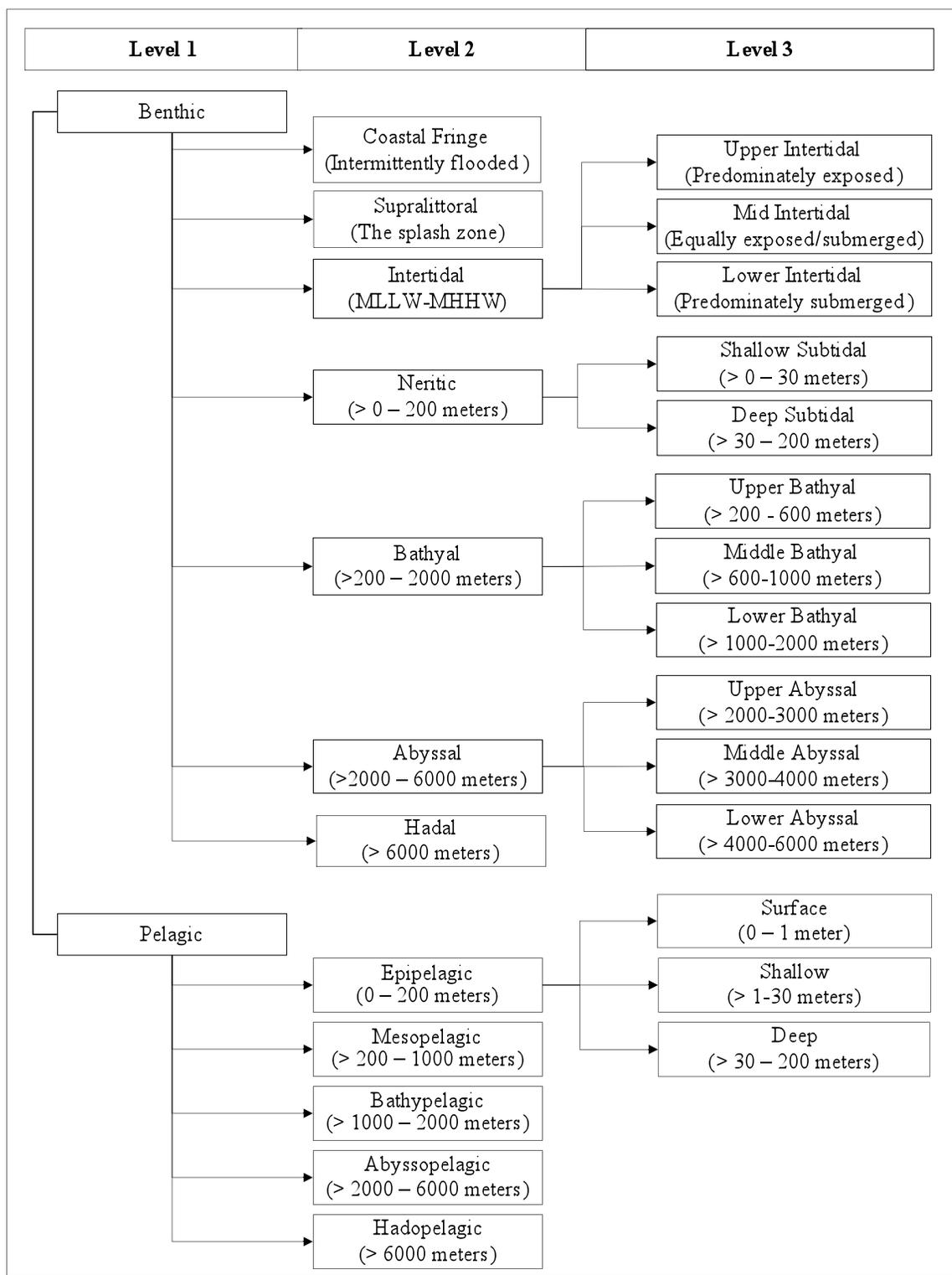


Figure 2-5: Hierarchical schema for depth for benthic and pelagic marine and estuarine species.

The limitation with this division of the abyssal zone is that the split does not adequately resolve the depth of the calcite compensation depth (CCD, the depth below which the rate of dissolution of calcium carbonate is equal to the rate of its formation). The depth of the CCD varies by location but ranges from about 3600 m to 4800 m in the North Pacific (Berger *et al.*, 1976); thus, we created an additional subdivision at 4000 m (Figure 2-5). The 4000 m break is considered preliminary, and will be modified as additional information becomes available on the distribution of calcareous and non-calcareous deep-sea organisms as well as the predicted effects of climate change on the depth of the CCD. When data are available, quantitative bathymetric depth ranges are stored as well as depth classes for a species, simplifying such a reclassification if the class threshold changes based on new information.

The value of capturing information in a hierarchical classification schema can be seen in Figure 2-6 where the numbers of chitons in the Northeast Pacific are shown by depth class. Species that occur only in the intertidal or only extend into the shallow subtidal are presumably at the greatest risk to both global warming and sea level rise. Species that extend into the deep subtidal to abyssal depths are at decreasing risk to temperature increases and sea level rise. Conversely, because chitons have calcareous shells, those species that occur in the deeper depth zones are at greater potential risk to the shallowing of the CCD in response to ocean acidification.



### *Reproductive Mode Hierarchical Schema*

Some natural history characteristics of biota are inherently categorical such as whether a species has an asexual or sexual reproductive strategy. While this initial split is straightforward, reproductive strategies can be further subdivided to capture more detailed aspects of a species' reproductive mode such as whether a sexual species is monoecious (= hermaphrodite) or dioecious (= separate sexes) or whether an asexual species is one of five different classes such as parthenogenic (= animals that reproduce from unfertilized eggs). These kinds of data are “messy”, and different hierarchical structures are possible. Nonetheless, we suggest that the hierarchical schema presented in Figure 2-7 imposes a useful structure in organizing this important biotic trait. One example of how this information can be used is in the identification of species that have characteristics of successful invaders, such as parthenogenic reproduction like the New Zealand mud snail *Potamopyrgus antipodarum* (Kerans *et al.*, 2005).

### *Geographic Hierarchical schemas*

Location matters – spatial relationships are fundamental to the evolving organization of life on Earth (Gaile & Willmott, 2003), with ecological processes operating across many different spatial scales. To address this phenomenon, we developed hierarchical spatial groupings to allow us to link macroecological patterns and processes at a range of scales to address a variety of ecological questions and management issues.

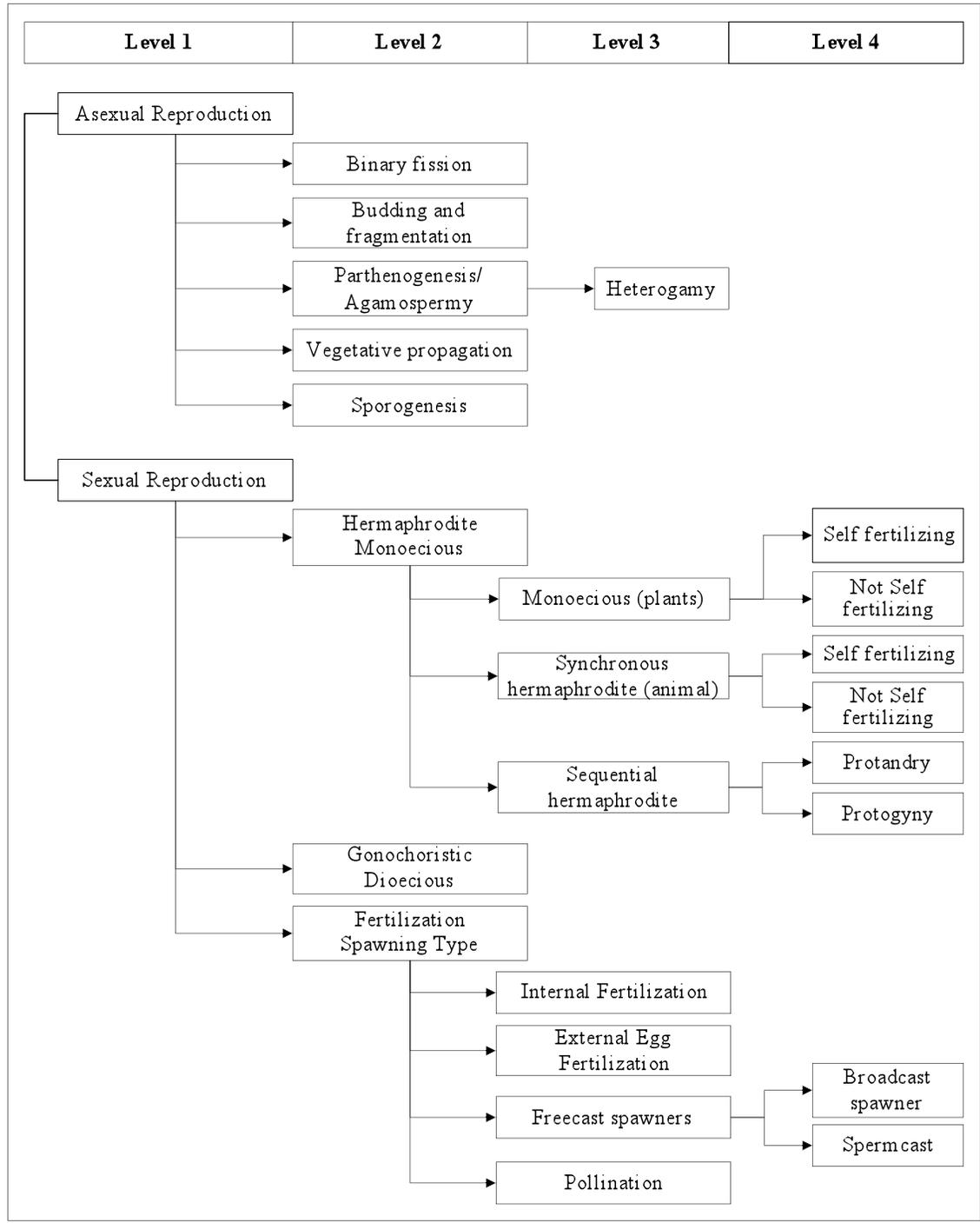


Figure 2-7: Hierarchical schema for reproduction as defined by (Lee *et al.*, 2008a). Spermcast refers to the dispersal of male gametes into the water column that fertilize eggs retained by the female (Pemberton *et al.*, 2003) while a broadcast spawner refers to species that discharge both sperm and eggs into the water column.

### *Biogeographic Hierarchical Schema*

Biogeographic frameworks attempt to define regions with similar fauna and flora compared to other geographical regions (Briggs, 1995). Although boundaries between marine biogeographic regions can be defined conceptually and visualized on a map, they are human constructs that have no distinct physiological demarcations between biogeographic regions except in cases where there are fundamental changes in biome type (e.g., marine to terrestrial shorelines). In particular, boundaries in the ocean tend to be characterized by gradual transitions. Due to the lack of discrete boundaries, a number of schemas have been developed for the delineation of the world's oceans. The boundaries defined by some of these schemas are very large ocean areas designed for fisheries management and conservation purposes. For example, the Large Marine Ecosystems (LMEs) schema developed by the National Oceanic and Atmospheric Administration (NOAA; <http://www.lme.noaa.gov/>) consists of 64 marine areas with delineations based on a combination of political and biogeographical boundaries. In comparison, the International Union for Conservation of Nature (IUCN) schema consists of more than four times as many delineations, with 264 marine ecoregions (Kelleher *et al.*, 1995).

These different biogeographical systems are usually presented as static alternatives. However, the recently developed Marine Ecoregions Of World (MEOW) (Spalding *et al.*, 2007) uses a three-tiered schema (realm, province, and ecoregion). These three levels are closely related to the boundaries identified in other marine ecoregion schemas, with the province level approximating the LMEs and the ecoregions approximating the IUCN

ecoregions so it is not necessary to treat the different systems as distinct alternatives. Thus, species' distributions can be entered at the ecoregion scale if available and aggregated to higher levels or entered at the higher level if detailed location information is not known. An example of an analysis using the MEOW ecoregions is the distribution of the total number of chiton species and number of endemic chitons in the Northeast Pacific by ecoregion (Figure 2-8). While this distributional pattern raises a number of interesting biogeographical questions, the analysis was conducted to help identify species at risk to climate change (Lee et al., in progress). Because of their limited range, the thirty-eight endemic species are assumed to be more vulnerable to climate changes than species with broader ranges (Rabinowitz, 1981), with the Southern California and Cortezian (= Gulf of California) ecoregions having the greatest number of endemics at risk.

The species in the Chukchi and Bering seas are also assumed to be at greater risk because of the greater projected temperature increases in the Arctic (Corell, 2006), the limited habitat availability for Arctic species to migrate to new areas as climate changes, and because CCD is higher in colder water (Andersson *et al.*, 2008).

One limitation of the MEOW schema is that it does not split the Pacific and Atlantic into east/west components. These wide stretches of ocean act as effective barriers for near-coastal species (e.g., East Pacific barrier) (Cox & Moore, 2005), separating species on opposite sides of oceans. Additionally, most invasions in near-coastal waters in the

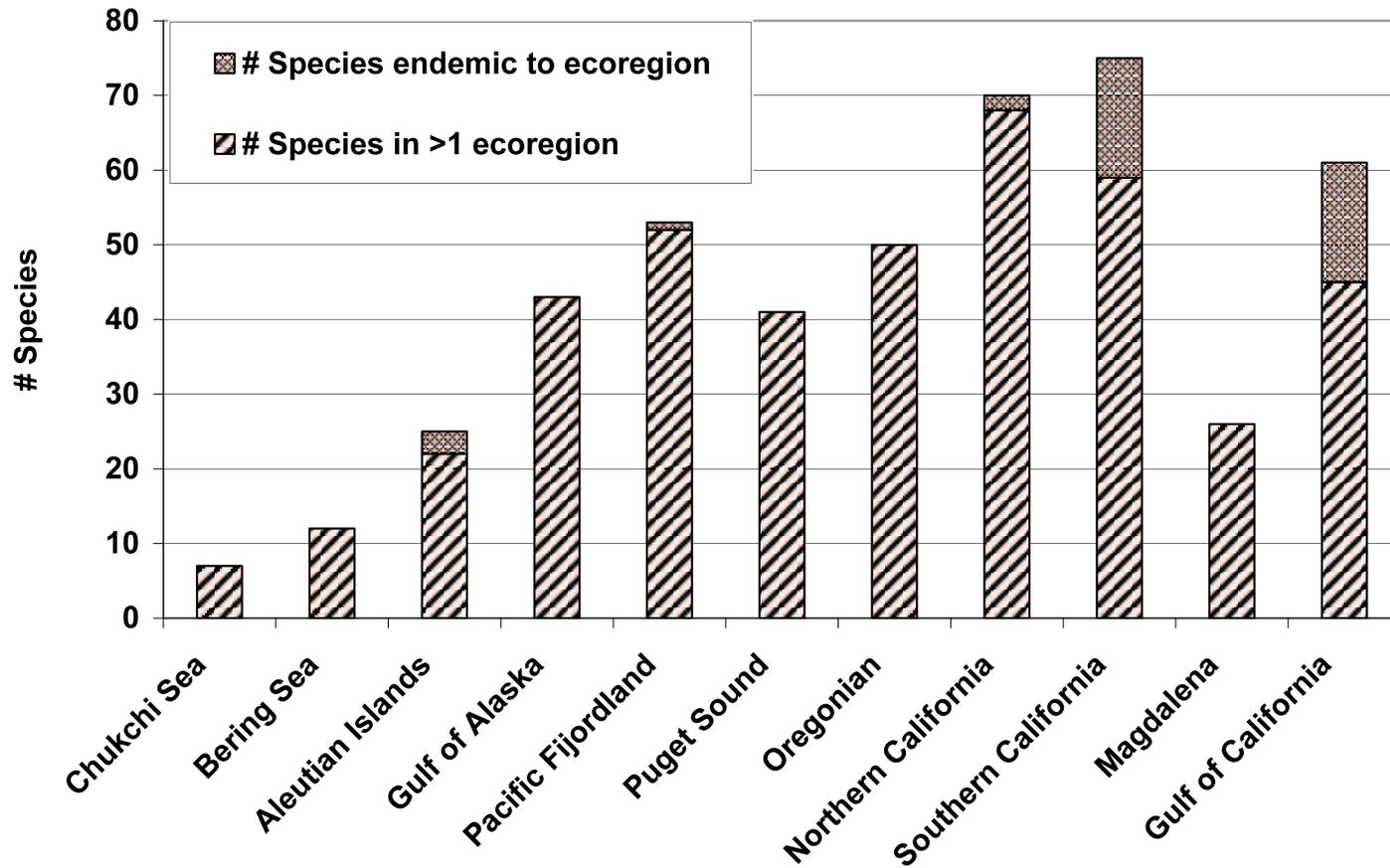


Figure 2-8: The total number of chiton species and the number of endemic species in the Northeast Pacific (Gulf of California to Chukchi) by MEOW ecoregion.

Northern Hemisphere have occurred across oceans rather than from lower or higher latitudes (Ruiz *et al.*, 2000) and the current version of MEOW ecoregions does not readily facilitate capturing such across-ocean invasions. Therefore, as shown in Figure 2-9, a new “region” level was added to the MEOW schema between the realm and province level to provide east-west breaks in the Pacific and Atlantic Ocean realms. Adding this region level simplifies addressing questions such as the relative importance of eastern versus western sides of oceans as donor regions for nonindigenous species (Chapman, 2000). It is interesting to note that, there are no known established nonindigenous chiton species in the Northeast Pacific region.

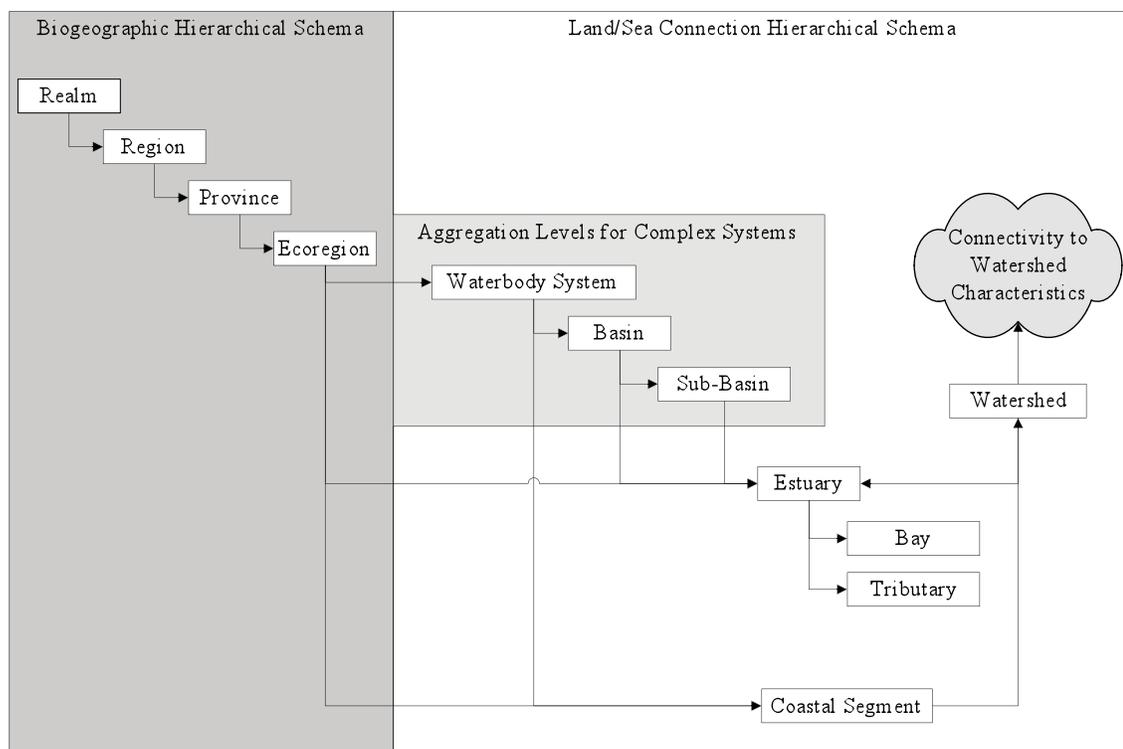


Figure 2-9: Hierarchical schemas for marine biogeographic areas and the land/sea connection.

### *Land/Sea Connection Hierarchical Schema*

Biogeographic schemas are useful for addressing questions related to the evolutionary history of taxa, patterns of invasion, and potential effects of climate change. However, a finer resolution schema is needed to address questions related to the interactions between watersheds and near-coastal water bodies, such as nutrient runoff (Bricker *et al.*, 2007), or local distributions of species. To address these issues, we developed a five tier hierarchical classification schema incorporating local to regional areas where the land meets the sea, which is linked to the biogeographic hierarchical schema (Figure 2-9). The land/sea connection hierarchical schema includes three aggregation levels (water body system, basin, and sub-basin) for compound systems, which are defined as water bodies that receive runoff from multiple watersheds (e.g., Puget Sound and Monterey Bay). For example, the entire Puget Sound is considered a single water body system which is composed of basins such as Whidbey Basin that in turn are composed of sub-basins such as Possession Sound and Skagit Bay. One-to-many estuaries and/or coastal segments may be embedded within any of these higher-order aggregated systems. Pragmatically, we have used the presence of National Wetland Inventory (NWI; <http://www.fws.gov/wetlands>) estuarine polygons to define estuaries (Lee & Brown, 2009). Coastal segments are stretches of the coast that drain directly into the ocean or water body system rather than an estuary, and are equivalent to NOAA's coastal drainage areas (CDA; <http://coastalgeospatial.noaa.gov/>). The defining characteristic of both estuaries and coastal segments is that a single continuous watershed can be defined for them. For small coastal creeks, the watershed may be very small ( $< 1 \text{ km}^2$ ) while the entire watersheds draining into the Columbia River and San Francisco Estuary are greater

than 120,000 km<sup>2</sup>. Regardless of size, the key aspect is that there is a one-to-one relationship between the estuary or coastal segment and the watershed draining into it. At the finest level are the bays and tributaries that make up a single estuary. Depending upon the size and complexity of an estuary it may or may not have identified tributaries or may be composed of multiple embayments and tributaries (e.g., San Francisco Estuary).

This connectivity to watershed areas provides the ability to analyze differences in estuarine and near-shore communities based on watershed characteristics and land use practices within the watershed (e.g., Lee and Brown, 2009). The aggregation levels for complex systems provide the tools necessary to conduct analyses within different segments of a complex system. This schema also allows researchers to summarize species distributions and biodiversity data at the spatial scale(s) appropriate to the specific scientific or managerial question as well as to match the scale of biological data to the scale of available environmental data.

### ***Populating the Information System***

Developing an integrated framework is only part of the task of synthesizing biotic data at regional and global scales – the other is populating the information system. Graphical interfaces, compared to a spreadsheet format, increases the efficiency of extracting multiple types of information from text-based sources for a single species. For example, data entry of species habitat regimes based on the classification structure of habitat regimes seen in Table 2-3 can be implemented in a graphical user interface shown in Figure 2-10. However, regardless of how efficiently information can be entered for a single species, a major challenge still remains for entering data for the large number of

existing species - the Census of Marine Life estimates about 230,000 described marine species to date (<http://www.coml.org/about>). One approach to capturing certain types of natural history data is to fill in the information for multiple species at higher taxonomic levels. While it is not always possible to capture the specifics of natural history attributes at taxonomic levels above the species level, use of hierarchical schemas provides a framework for information to be captured at a general level. For example, one attribute that can often be captured at the genus or family level is feeding type. As was previously mentioned, Fauchald and Jumars (1979) assigned polychaete feeding types by family and genus, and the biogeographic pattern for the relative frequency of carnivorous marine snails was determined by assigning feeding type by family (Valentine *et al.*, 2002). Other natural history attributes that can often be captured at higher taxonomic levels include general habitat type (e.g., pelagic versus benthic), reproductive mode (e.g., sexual versus asexual), and certain aspects of development type (e.g., pelagic larvae versus direct development).

Drawing on the precept that “geography is biology”, another powerful approach of capturing natural history attributes for multiple species is to use automated routines to extract habitat ranges or preferences from coupled biotic and environmental data from field surveys. From the benthic database used for the salinity analysis mentioned above, we were able to automate the extraction of environmental ranges for a suite of key habitat

<b>Regime Class</b>	<b>Definition</b>
Terrestrial	Land areas not directly impinging upon aquatic ecosystems.
Lakes and Ponds (Lentic)	Body of standing fresh water, including wetlands.
Rivers, Streams, and Creeks (Lotic)	Flowing bodies of freshwater, including riparian zones.
Estuaries and Lagoons	Estuary: A semi-enclosed coastal water body with one or more rivers or streams flowing into it and with a connection to the ocean. Salinities in estuaries are normally below that of the bordering ocean water. Lagoons are shallow coastal water bodies separated from the ocean by a barrier island or by shallow or exposed sandbanks or coral reefs. Depending upon freshwater inputs and connection to the ocean, salinity in lagoons can range from essentially fresh to hypersaline.
Coastal Fringe	Area between terrestrial and nearshore or estuarine ecosystems with primarily terrestrial characteristics but strongly affected by bordering aquatic ecosystem (e.g., sand dunes, estuarine shrub/scrub wetland, estuarine forest wetland).
Coastal Bay	An area of water mostly surrounded by land on the open coast, creating calmer waters than the open sea.
Nearshore	0 – 30m. The outer coast; from the intertidal to 30m bathymetric isopleth.
Shelf	30 – 200m bathymetric isopleths. The benthos and water above that borders the continent and extends out to where there is an increased slope of the seafloor, approximately 200m depth.
Oceanic	> 200m bathymetric isopleths. Includes the benthos and water above the continental slope and ocean floor.

Table 2-3: Habitat Regimes. Regimes are the broad physical and environmental divisions based on a combination of salinity, geomorphology and depth (Madden *et al.*, 2005).

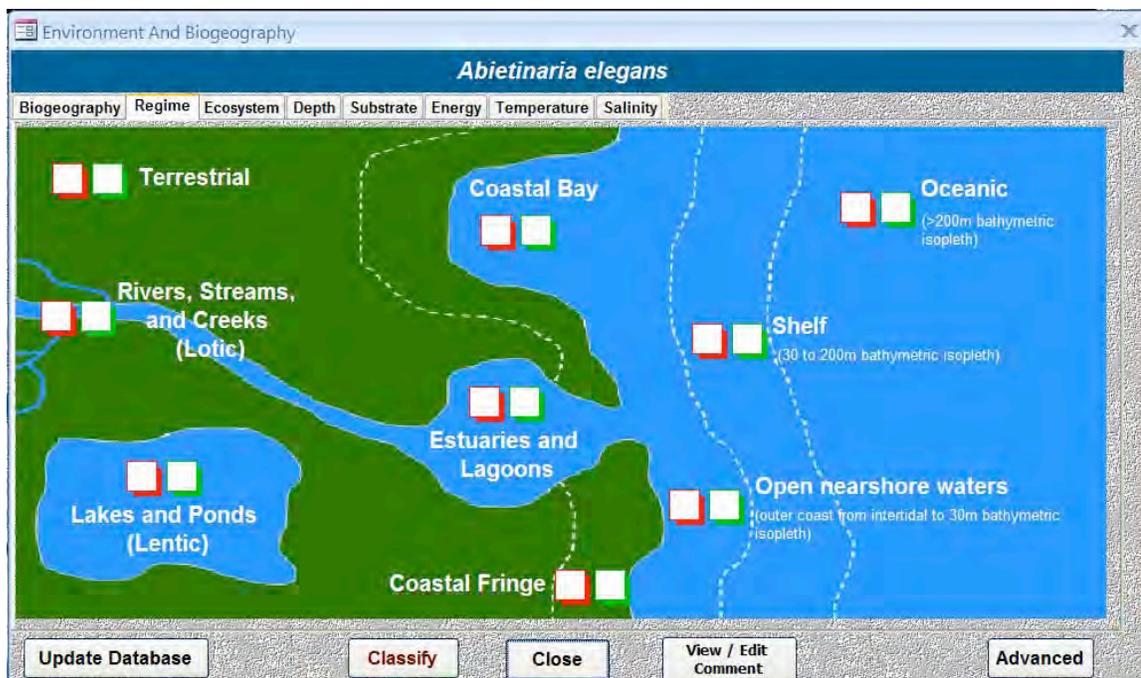


Figure 2-10: Graphical interface for data entry and display of regime data for a single species. Clicking on a box will cause a check to be displayed in the box. A check in a box with a red border indicates a regime the species has been observed in, while a check in a box with a green border indicates the preferred habitat for the species. Use of observed and preferred options for several attributes allows a separation of the “normal” range for a species versus the extreme range, which may include individuals displaced out of its normal habitat.

attributes including salinity, depth, and percent silt/clays for over 3000 benthic taxa from more than 4000 samples. Development of flexible tools to allow such bulk data collection and entry from different types of monitoring studies and downloads from biotic web sites will allow the generation of environmental ranges for thousands of species within specific regions or globally.

### ***Conclusions***

As the primary insults to the environment shift from localized pollution impacts to regional and global scale threats, macroecological approaches of studying multiple species over broad geographical areas are becoming an increasingly important research

thrust (Kerr et al., 2007). In response to this need for macroecological tools, we developed the *PCEIS* biological information system. The multi-dimensional and hierarchical topologies in *PCEIS* provide a standardized, integrated framework to capture complex environmental, biological, and geographical data on marine/estuarine species and near-coastal water bodies. Perhaps its major contribution is that it facilitates matrix overlays between and among the natural history and environmental ranges for individual or multiple species across different geographical scales. An example of a multifactor analysis is the crosswalk between the biogeographic ranges of chiton species in the Northeast Pacific by their bathymetric distribution (Table 2-4). As mentioned above, species that occupy fewer ecoregions are assumed, in general, to be at greater risk to climate change and most other perturbations compared to species with wider ranges. Additionally, species limited to the intertidal and shallow subtidal (1-30 m) are directly exposed to the impacts of both temperature increases and sea level rises. Because of their calcareous shells, the deepest chiton species (>2000 m) are at potential risk to a shallowing of the CCD in response to ocean acidification. While any analysis based on broad scale patterns is not definitive, we suggest that analyses like that presented in Table 2-4 can provide important insights into the identity and number of near-coastal species at greatest risk to climate change. Thus, with integrated biological information systems like *PCEIS*, it becomes practical to utilize a suite of different distributional and natural history traits to address a range of research and management questions. Though the present framework was designed for marine/estuarine species, many of the classes and hierarchies are directly transferable to freshwater species. Greater modifications would be

required for terrestrial species, though we suggest that the concepts of classifying natural history and the use of hierarchical topologies apply to terrestrial species as well.

Number of Ecoregions/ Depth Distribution	One Ecoregion	Two Ecoregions	Three Ecoregions	Four Ecoregions	> Four Ecoregions
Intertidal Only	9	3	5	2	2
Intertidal & Shallow Subtidal	10	7	7	11	13
Intertidal to Deep Subtidal	8	4	7	8	21
Intertidal to Lower Bathyal	4	2	0	3	14
Deep Subtidal to Abyssal	0	1	0	1	2

Table 2-4: Crosswalk of the biogeographic and bathymetric ranges of chitons in the Northeast Pacific (Gulf of California to Chukchi Sea; N = 144 species). Biogeographic ranges are defined by the number of MEOW ecoregions occupied in the world. The bathymetric ranges are based on consolidated classes from those given in Figure 2-6. The solid black boxes highlight the number of species assumed to be at greatest risk due to limited distributional and bathymetric range combinations from temperature increases and sea level rise (shallow species) or to ocean acidification (deep species). Dashed boxes indicate high risk combinations.

### ***Acknowledgements***

DAR was partially funded through AMI/GEOSS IAG # DW-14-92231501-0 from the U.S. Environmental Protection Agency (EPA) and by the North Pacific Marine Science Organization (PICES) through funding from the Japanese Government. Special thanks to Dr. Melanie Frazier, U.S. EPA postdoc, for assistance with statistical analyses, Jody Stecher, Katie Marko and Karen Ebert for their continued efforts populating the database,

and Rachel Nehmer and Tad Larsen for implementation of the graphical user interface. This publication was subjected to review by the Geosciences and Zoology Departments at Oregon State University, the National Health and Environmental Effects Research Laboratory's Western Ecology Division of the EPA, and the Western Fisheries Research Center of U.S. Geological Survey, and is approved for publication. However, approval does not signify that the contents reflect the views of the U.S. EPA. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader; such use does not constitute official endorsement or approval by the U. S. Department of Interior, the U. S. Geological Survey, or the U.S. Environmental Protection Agency of any product or service to the exclusion of others that may be suitable.

### ***References***

- Andersson, A.J., Mackenzie, F.T. & Bates, N.R. (2008) Life on the margin: Implications of ocean acidification on mg-calcite, high latitude and cold-water marine calcifiers. *Marine Ecology Progress Series*, **373**, 265-273.
- Bambach, R.K., Bush, A.M. & Erwin, D.H. (2007) Autecology and the filling of ecospace: Key metazoan radiations. *Palaeontology*, **50**, 1-22.
- Berger, W.H., Adelseck, J.C.G. & Mayer, L.A. (1976) Distribution of carbonate in surface sediments of the Pacific Ocean. *Journal of Geophysical Research*, **81**.
- Bricker, S., Longstaff, B., Dennison, W., Jones, A., Boicourt, K., Wicks, C. & Woerner, J. (2007) Effects of nutrient enrichment in the nation's estuaries: A decade of change. In: *NOAA Coastal Ocean Program Decision Analysis Series No. 26*, p. 328. NOAA, National Centers for Coastal Ocean Science, Silver Springs, MD.
- Briggs, J.C. (1995) *Global Biogeography*. Elsevier Science B. V., Amsterdam, The Netherlands.

- Bush, A.M., Bambach, R.K. & Daley, G.M. (2007) Changes in theoretical ecospace utilization in marine fossil assemblages between the mid-paleozoic and late cenozoic. *Paleobiology*, **33**, 76-97.
- Carlton, J.T. (2007) *The Light and Smith Manual: Intertidal invertebrates from central California to Oregon*, 4 edn. University of California Press, Berkeley and Los Angeles, California.
- Chapman, J.W. (2000) Climate effects on the geography of nonindigenous peracaridan crustacean introductions in estuaries. In: *First National Conference on Marine Bioinvasions* (ed. J. Pederson), pp. 66-80. Massachusetts Institute of Technology.
- Clarke, K.R. & Warwick, R.M. (1999) The taxonomic distinctness measure of biodiversity: Weighting of step lengths between hierarchical levels. *Marine Ecology Progress Series*, **184**, 21-29.
- Corell, R.W. (2006) Challenges of climate change: An arctic perspective. *AMBIO: Journal of the Human Environment*, **35**, 148-152.
- Cox, C.B. & Moore, P.D. (2005) *Biogeography: An ecological and evolutionary approach*, 7 edn. Blackwell Publishing, Malden, MA.
- Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*, 1st edn. John Murray, London.
- Eaton, J.G. & Scheller, R.M. (1996) Effects of climate warming on fish thermal habitat in streams of the United States. *Limnology and Oceanography*, **41**, 1109-1115.
- Eernisse, D., J. (1986) The genus *lepidochitona* Gray, 1921 (mollusca: Polyplacophora) in the Northeastern Pacific Ocean (Oregonian and Californian provinces). *Zoologische Verhandelingen*, **228**, 3-52.
- Fauchald, K. & Jumars, P.A. (1979) The diet of worms: A study of polychaete feeding guilds. *Oceanography and Marine Biology: An Annual Review*, **17**, 193-284.
- Fofonoff, P.W., Ruiz, G.M., Steves, B., Hines, A.H. & Carlton, J.T. (2003) National exotic marine and estuarine species information system (nemesis). URL <http://invasions.si.edu/nemesis/>.
- Gaile, G.L. & Willmott, C.J. (2003) *Geography in America at the dawn of the 21st century*. Oxford University Press, New York.

- Gollasch, S. (2006) Assessment of the introduction potential of aquatic alien species in new environments. In: *Assessment and control of biological invasion risks* (ed. by F. Koike, M.N. Clout, M. Kawamichi, M. De Poorter and K. Iwatsuki), pp. 88-91. Shoukadoh Book Sellers, Kyoto, Japan and IUCN, Gland, Switzerland.
- Hall, C.A. (1964) Shallow-water marine climates and molluscan provinces. *Ecology* **45**, 226-234.
- Hemphill, H. (1890) *Catalogue of North American shells*. Privately Published, San Diego, California.
- Hodgkin, E.P. (1987) The hydrology of the Swan River estuary: Salinity the ecological master factor. In: *The Swan River estuary - ecology and management* (ed. by J. John), pp. 34-44. Curtin University Environmental Studies Group, Perth, Western Australia.
- Howes, D.E., Morris, M. & Zacharias, M. (1999) *British Columbia estuary mapping system*. Resource Inventory Committee, Province of British Columbia, Victoria, British Columbia.
- IUBS (1958) The venice system for the classification of marine waters according to salinity. *International Symposium for the Classification of Brackish Waters*, p. 2. The International Union of Biological Sciences, Venice, Italy.
- Josefson, A.B. & Hansen, J.L.S. (2004) Species richness of benthic macrofauna in Danish estuaries and coastal areas. *Global Ecology and Biogeography*, **13**, 273-288.
- Kelleher, G., Bleakley, C. & Wells, S.M. (1995) A global representative system of marine protected areas. Washington, DC.
- Kerans, B.L., Dybdahl, M.F., Gangloff, M.M. & Jannot, J.E. (2005) *Potamopyrgus antipodarum*: Distribution, density, and effects on native macroinvertebrate assemblages in the Greater Yellowstone ecosystem. *Journal of the North American Benthological Society*, **24**, 123-138.
- Kerr, J.T., Kharouba, H.M. & Currie, D.J. (2007) The macroecological contribution of global change solutions. *Science*, **316**, 1581-1584.
- Kleypas, J.A., Feely, R.A., Fabry, V.J., Langdon, C., Sabine, C.L. & Robbins, L.L. (2006) Impacts of ocean acidification on coral reefs and other marine calcifiers: A

guide for future research, report of a workshop held 19-20 april 2005. Sponsored by NSF, NOAA and the U. S. Geological Survey, St. Petersburg, FL.

- Koleff, P., Gaston, K.J. & Lennon, J.J. (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367-382.
- Lee II, H. & Brown, C.A. (2009) Classification of regional patterns of environmental drivers and benthic habitats in Pacific Northwest estuaries. (eds. H. Lee and C.A. Brown). U.S. EPA, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Newport, Oregon.
- Lee II, H., Reusser, D., Ranelletti, M. & Nehmer, R. (2008b) *User guide for the PICES nonindigenous species information system templates, vers. Gamma. Newport, or, USA.*
- Lee II, H., Reusser, D.A., Olden, J.D., Smith, S.S., Graham, J., Burkett, V., Dukes, J.S., Piorkowski, R.J. & Mcphedran, J. (2008a) Integrated monitoring and information systems for managing aquatic invasive species in a changing climate. *Conservation Biology*, **22**, 575-584.
- Lindegarh, M. & Gamfeldt, L. (2005) Comparing categorical and continuous ecological analyses: Effects of "wave exposure" on rocky shores. *Ecology*, **86**, 1346-1357.
- Linnaeus, C. (1758) *Systema naturae per regna tria naturae :Secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. "System of nature through the three kingdoms of nature, according to classes, orders, genera and species, with [generic] characters, [specific] differences, synonyms, places".* 10TH edn.
- Madden, C.J., Grossman, D.H. & Goodin, K.L. (2005) Coastal and marine systems of North America: Framework for an ecological classification standard: Version II. NatureServe, Arlington, Virginia.
- Mecklenburg, C.W., Mecklenburg, A.T. & Thorsteinson, L.K. (2002) *Fishes of Alaska*, 1 edn. American Fisheries Society, Bethesda, Maryland.
- Nelson, W.G., Lee II, H., Lamberson, J.O., Engle, V., Harwell, L. & Smith, M.L. (2005) Condition of estuaries of the Western United States for 1999: A statistical summary. Office of Research and Development, National Health and Environmental Effects Research Laboratory. EPA/620/R-04/200.
- Open University (2000) *Waves, tides and shallow water processes*, 2 edn. Elsevier Science & Technology Books, Boston.

- Pemberton, A.J., Hughes, R.N., Manriquez, P.H. & Bishop, J.D.D. (2003) Efficient utilization of very dilute aquatic sperm: Sperm competition may be more likely than sperm limitation when eggs are retained. *The Royal Society biology letters*, **(Supplement) 270**, S223-S226.
- Rabinowitz, D. (1981) Seven forms of rarity. In: *The biological aspects of rare plant conservation* (ed. by H. Synge), pp. 205-217. Wiley.
- Reusser, D.A. & Lee II, H. (2008) Predictions for an invaded world: A strategy to predict the distribution of native and non-indigenous species at multiple scales. *ICES Journal of Marine Science*, **65**, 742-745.
- Ruiz, G.M., Fofonoff, P.W., Carlton, J.T., Wonham, M.J. & Hines, A.H. (2000) Invasion of coastal marine communities in North America: Apparent patterns, processes, and biases. *Annual Review of Ecology and Systematics*, **31**, 481-531.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdana, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lonbana, A., Lourie, S.A., Martin, K.D., Mcmanus, E., Molnar, J., Recchia, C.A. & Robertson, J. (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience*, **57**, 573-583.
- Stefan, H.G., Fang, X. & Eaton, J.G. (2001) Simulated fish habitat changes in North American lakes in response to projected climate warming. *Transactions of the American Fisheries Society*, **130**, 459-477.
- Valentine, J.W., Kaustuv, R. & Jablonski, D. (2002) Carnivore/noncarnivore ratios in Northeastern Pacific marine gastropods. *Marine Ecological Progress Series*, **228**, 153-163.

### **Chapter 3: Per capita invasion probabilities: A linear model to predict rates of invasion via ballast water**

Authors: Deborah A. Reusser, Henry Lee II, Melanie R. Frazier Gregory M. Ruiz, Paul Fofonoff, Mark S. Minton, and Whitman Miller

***Abstract***

Ballast water discharges are a major source of species introductions into marine, estuarine, and freshwater ecosystems. To predict the potential rate of invasion from ballast water, we developed a linear invasion model based on historic invasion rates, estimated organism concentrations, and the recorded foreign ballast water discharges in the United States. Propagule pressure from each ship was estimated from previous measurements of organism densities in untreated ballast water multiplied by the ballast water discharged. The per capita invasion probability was calculated by dividing the total number of ballast water invaders per year by the total propagule pressure discharged via ballast water into a waterbody or coast. The resulting per capita invasion probability represents the likelihood that a single discharged organism will become established as a new nonindigenous species within the waterbody or coast. Analysis was done at the individual estuary scale as well as a coast-wide scale for the East, Gulf and Pacific coasts. There is less uncertainty with the coast-wide scale analysis as it removes the unknown influence from secondary invasion vectors between estuaries on a single coast. In addition, because the Pacific Coast has been very well studied, there is less uncertainty in the total number of ballast water invaders on the Pacific Coast. Based on data from the Smithsonian Institute, an average of 14,788,369 metric tons of foreign ballast water per year was discharged into ports on the Pacific Coast from 2005-2007. Between 1981-2006, a total of 67 invertebrate and macroalgal species >50 microns are estimated to have invaded the Pacific Coast through ballast water. Assuming a linear dose-response relationship, the median per capita invasion probability for the Pacific Coast is  $3.61 \times 10^{-11}$ . Using per capita invasion probabilities, a range of discharge volumes, a range of

organism concentrations and a safety factor, it is possible to calculate a range of ballast water organism-based performance standards for different risk levels, defined as the probability that a new species will invade per year. Depending upon the assumptions used in the risk analysis, this approach predicts that approximately one new species will invade every 10 to 100 years with the current discharges rates for the Pacific Coast of the United States and the International Maritime Organization (IMO) standard of <10 organisms >50 microns per m<sup>3</sup> of ballast. We suggest that this approach is a viable method to quantitatively evaluate the risk of different organism-based ballast water performance standards based on reasonably well known input parameters.

**Key words:** Ballast Water Discharge; Invasion Probabilities; Propagule Pressure; Aquatic Invaders; IMO standards.

## ***Introduction***

*“Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.” (Box & Draper, 1987)*

Aquatic invasions are a key factor causing environmental stress on freshwater, estuarine and marine ecosystems (Ruiz *et al.*, 1999; Occhipinti-Ambrogi & Savini, 2003). The primary source for these biological invasions is shipping (Ruiz & Carlton, 2003; Molnar *et al.*, 2008). In the past century, both the increase in shipping traffic as well as the reduced time for transoceanic voyages has increased the number of aquatic organisms (propagule supply) surviving transport to new environments around the world (Carlton & Geller, 1993; Ruiz *et al.*, 1997). Increasing propagule pressure has increased the rates of invasion in a number of aquatic ecosystems (Carlton & Geller, 1993; Cohen & Carlton, 1998; Ruiz & Carlton, 2003; Cordell *et al.*, 2009). Of the potential shipping vectors, ballast water is one of, if not, the most important (Carlton, 1996b; Fofonoff *et al.*, 2003a). For example, since the opening of the St. Lawrence Seaway in 1959, ballast water is the suspected source for over 70% of the nonindigenous species found in the Great Lakes (Holeck *et al.*, 2004). The importance of ballast water discharge is likely to increase in the future as other vectors, such as purposeful introductions and oyster-associated introduction are better managed.

The first approach to managing this vector was to implement mid-ocean ballast water exchange, where ballast was exchanged using either a flow-through or empty-and-refill method (Cordell *et al.*, 2009). While this management strategy reduced organism concentrations in ballast water, it was not sufficiently effective (Locke *et al.*, 1991;

Minton *et al.*, 2005). In response, national and international efforts began, under the auspices of the International Maritime Organization (IMO), to evaluate other options for managing ballast water discharges and to develop an international ballast water treaty. An important early IMO decision was to establish ballast water performance standards based on organism concentrations in the discharged ballast water. The proposed IMO D-2 standards are given in Table 3-1. The fundamental assumption behind establishing organism-based ballast water performance standards is that invasion risk decreases with decreasing propagule supply. This assumption is supported by a wide body of evidence showing that the establishment probability for nonindigenous species either increases with propagule pressure due to a higher concentration of organisms in an inoculation, and/or an increase in the frequency of inoculations (Ruiz *et al.* 2000, Kolar and Lodge 2001, Colautti *et al.* 2006, Simberloff 2009).

While the IMO standards were recognized as a major step forward, there was concern by some coastal states in the United States that the proposed IMO standards were not sufficiently protective. The United States Coast Guard (USCG) responded to this concern by proposing a two-phase implementation of performance standards (Table 3-1). The first phase is equivalent to the IMO standards while phase two is 1000-fold more stringent. The State of California has also proposed alternative standards with the ultimate goal of *no detectable* discharge of organisms in ballast water (Table 3-1).

<b>Organism Class</b>	<b>IMO D-2 Standard</b>	<b>U.S. Coast Guard Phase I</b>	<b>U.S. Coast Guard Phase II</b>	<b>CA &amp; NY Interim Standards</b>	<b>CA &amp; NY Long-term Standard</b>
Organisms >50 microns	<10 per m <sup>3</sup>	<10 per m <sup>3</sup>	<1 per 100 m <sup>3</sup> (< 0.01 per m <sup>3</sup> )	No detectable living organisms	No detectable living or culturable organisms
Organisms 10-50 microns	<10 per ml	<10 per ml	<1 per 100 ml	≤1 per 100 ml	No detectable living or culturable organisms
Organisms <10 microns	None	None	<1,000 bacteria & 10,000 viruses per 100 ml	≤10 <sup>3</sup> cfu of bacteria per 100 ml	No detectable living or culturable organisms
<i>Escherichia coli</i>	<250 cfu per 100 ml	<250 cfu per 100 ml	<126 cfu per 100 ml	≤126 cfu per 100 ml	No detectable living or culturable organisms
Intestinal enterococci	<100 cfu per 100 ml	<100 cfu per 100 ml	<33 cfu per 100 ml	≤33 cfu per 100 ml	No detectable living or culturable organisms
Toxicogenic <i>Vibrio cholerae</i> (serotypes O1 and O139)	<1 cfu per 100 ml Or 1 cfu per g wet weight zooplankton	<1 cfu per 100 ml	<1 cfu per 100 ml	≤1 cfu per 100 ml Or ≤1 cfu per g wet zoological sample	No detectable living or culturable organisms

Table 3-1: Summary of existing or proposed ballast water performance standards applicable to United States waters. All organism dimensions are for the “minimum dimension”. Standards for the >50 micron and 10-50 micron classes are for “viable” or “living” organisms. Note that Phase II of the Coast Guard standard can be implemented incrementally. cfu = “colony forming units”.

This broad range in proposed performance standards highlights the complexity of establishing organism concentrations that are protective of the environment while being technologically feasible and economically viable.

One uncertainty in predicting new ballast-associated invasions is the nature of the propagule supply dose-response relationship. As shown in Figure 3-1, the dose-response relationship could be an exponential (*a*), multiphasic (*b*), linear (as assumed in the models described here) (*c*), or hyperbolic curve (*d*). As pointed out by Ruiz and Carlton (2003), the exact shape of the dose-response curve is unknown. However, at the low organism densities associated with the proposed ballast discharge standards (Table 3-1), it is possible that Allee effects reduce the probability of a successful invasion for many species (Drake, 2004), as illustrated in lines *a* and *b* in Figure 3-1. Allee effects are reductions in the per capita population growth rate in sparse populations. Such depressions in individual growth rates in rarefied populations may occur due to several, potentially interacting, mechanisms (i.e., mate limitation, increased predation, genetic inbreeding, and/or increased dispersal) (Drake, 2004; Gascoigne & Lipcius, 2004; Kramer *et al.*, 2009). In addition, for low concentrations of species (stippled area of Figure 3-1), a linear estimate is likely to be adequately protective if Allee effects reduce the probability of invasion success relative to a linear relationship (*c*) for a given propagule supply.

Our goal was to define a linear relationship of type *c* based on propagule pressure for the development of ballast water performance standards that would be protective.

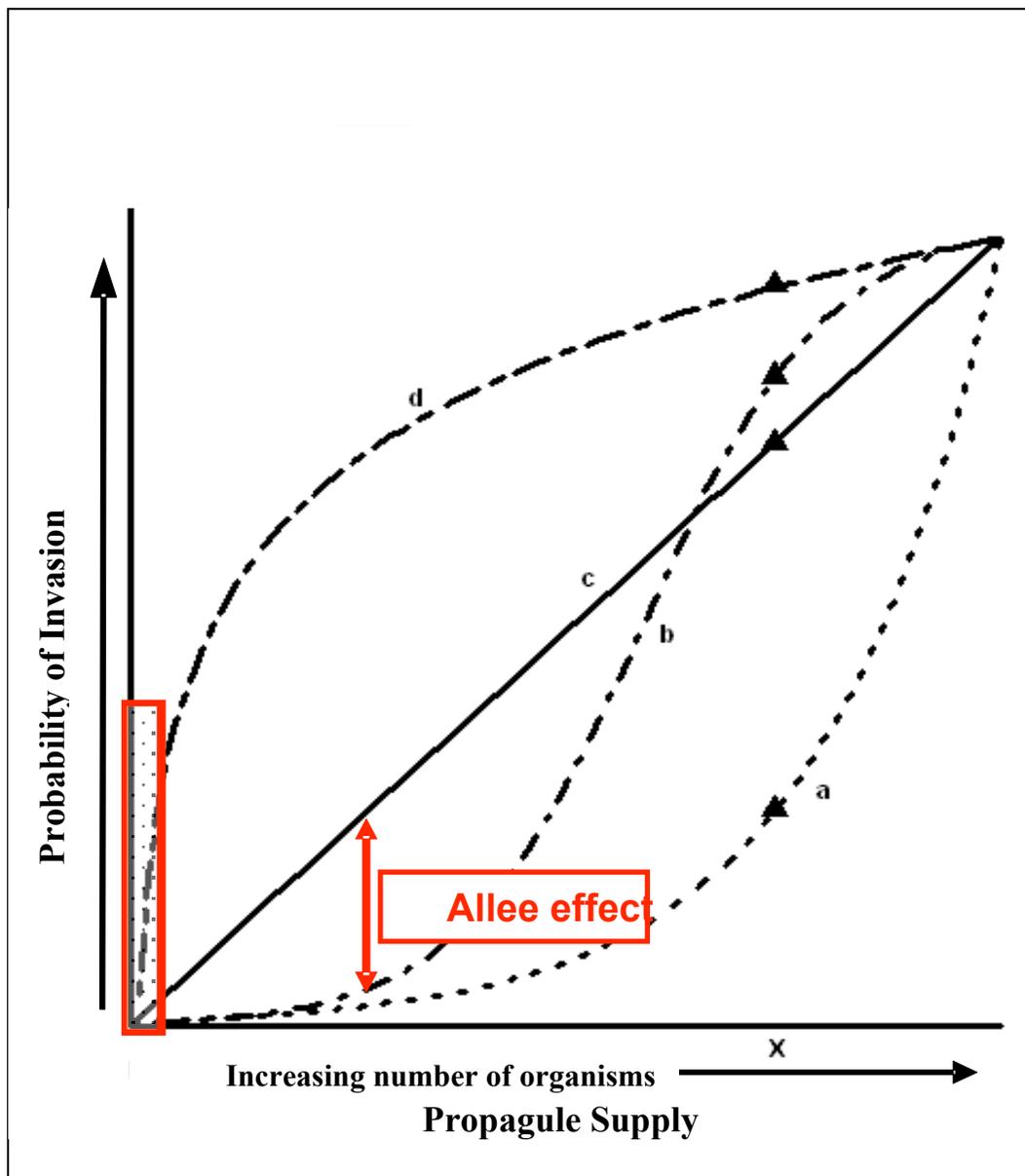


Figure 3-1: Hypothetical propagule supply dose-response curves. Potential responses include; a) exponential; b) multiphasic; c) linear (as assumed in the models described here); and d) logarithmic or hyperbolic. The triangles denote the range in invasion probabilities that can result from different relationships for one propagule dose ( $X$ ). The arrow indicates the reduction in the probability of invasion between the linear response and the exponential and sigmoid responses due to Allee effects. The stippled area in the image illustrates that the propagule doses associated with the proposed performance standards (Table 3-1) are likely to be at the low end of the dose-response curve. (Modified from Ruiz and Carlton, 2003, reprinted here with permission from Island Press).

Based on the premise of a linear dose-response, we developed a “per capita invasion probability” (PCIP) approach to estimate the likelihood of invasion based on historical invasion rates and calculated ballast-associated propagule pressures. The PCIP is the likelihood that a unique, non-native propagule (individual) discharged from ballast water will become established in a specified waterbody per year. Using a linear dose response model, the PCIP is calculated from the historical number of ballast-mediated invasions in a specified waterbody during a time period, the average annual total ballast discharged at a location, and the estimated organism concentrations in the discharged ballast water. We focus on the >50 micron size class of organisms because the data are available, though in theory, the approach could be applied to the 10-50 micron size class.

An advantage of this approach is that it directly relates the risk of invasion to ballast water organism concentrations, so it can be used to generate performance standards. It is important to note, however, that because of all the complexities involved with the invasion process, our objective was not to find a highly predictive relationship between the calculated propagule supply (or ballast discharge rate) and the site-specific rates of invasion. Rather, our objectives were to “cut through” the complexities while documenting our assumptions and identifying whether the assumptions were protective of the environment (i.e., resulted in a lower performance standard). We believe that this approach allows policy makers to use PCIP to establish performance standards based on different risk levels.

## ***Methods***

The linear invasion model to predict the potential rate of invasion from ballast water is defined by:

$$\text{Equation 3-1: } \text{PCIP} = N_h / (D_h * C_h)$$

where PCIP denotes the per capita probability an organism will invade a water body (new invading species \* organism<sup>-1</sup>),  $N_h$  is the historical annual invasion rate of potential ballast-associated invaders for a water body (new invading species \* year<sup>-1</sup>),  $D_h$  is the historic annual foreign ballast discharged into a waterbody (m<sup>3</sup> year<sup>-1</sup>) and  $C_h$  is the historic concentration of organisms in ballast water discharged into a waterbody (organisms \* m<sup>-3</sup>).

As mentioned, the per capita invasion probability (PCIP) is the probability that an individual propagule, or organism, discharged from ballast water will become established as a new nonindigenous species within the waterbody. For example, if one new nonindigenous species became established within a waterbody in which a total of a million individual organisms were discharged in a year, the per capita invasion probability would equal 10<sup>-6</sup>. Because the PCIP only accounts for new invaders, it does not address the issue of multiple invasions of currently existing nonindigenous species into the waterbody.

This model assumes a linear dose-response, with the number of invaders increasing proportionally with larger ballast water organism concentrations and/or greater volumes of ballast water discharged. Accordingly, after calculating a PCIP from a historical

invasion rate, it is possible to predict the number of new, unique invaders per year given ballast water organism concentration and ballast water volume where:

$$\text{Equation 3-2: } N_p = \text{PCIP} * D_p * C_p$$

Where  $N_p$  denotes the predicted annual invasion rate of potential ballast-associated invaders for a waterbody (new invading species \* year<sup>-1</sup>),  $D_p$  is the predicted annual foreign ballast discharge rate into a waterbody (m<sup>3</sup> year<sup>-1</sup>), and  $C_p$  is the predicted concentration of organisms in ballast water discharged into a waterbody (organisms m<sup>-3</sup>).

### *Foreign Ballast Water Discharge Rates for Coastal Waterbodies and the Great Lakes*

Historic average annual foreign ballast discharge rates ( $D_h$  in Equation 3-1) were used to calculate the total propagule supply. Discharge rates for coastal waterbodies were obtained from the Smithsonian Institution ballast water database (see the National Ballast Information Clearinghouse, <http://invasions.si.edu/nbic/search.html>). Average yearly discharge values (Table 3-2) were calculated for the contiguous East, Gulf, and Pacific coasts from discharge records for all ships discharging foreign ballast into coastal ports on the respective coasts from 2005 to 2007. Only ballast identified as coming from a foreign source was included. These dates were chosen because they occur after the implementation of mandatory ballast water reporting and represent the most complete discharge records available. Average annual foreign discharge rates were also calculated for 17 coastal ports, representing a cross section of small to large ports based on discharge records from the same time period, 2005 to 2007. Because the foreign ballast was recorded on a per tank basis, the movement of undischarged foreign ballast among ports could be estimated. That is, by following foreign ballast by tank it was possible to

account for foreign ships that initially entered one port but did not discharge their ballast until they visited another port. Foreign discharge values for multiple ports within a waterbody were summed for a total discharge volume for a waterbody, including freshwater ports in larger systems (e.g., Columbia River). For the Great Lakes, the National Biological Invasion Shipping Study (Reid and Carlton, 1997) reported a total annual foreign ballast water discharge into the Great Lakes of 1,395,461 metric tons in 1991. This was before mandatory ballast water exchange, which was initiated in the Great Lakes in 1993.

### *Estimates of Organism Concentrations in Ballast Water*

Organism concentrations in ballast water discharged in coastal waters ( $C_h$  in Equation 3-1) were estimated from Minton et al. (2005), who reported zooplankton ( $> 80$  microns) concentrations in unmanaged ballast water in 354 ships of various types. Similar organism values were reported in a survey of 429 ships of multiple vessel types that had no ballast water exchange or treatment (MEPC, 2003b). Both of these studies showed that organism concentrations in untreated ballast water can vary by orders of magnitude among ships. For example, about 3.8% of the ships reported by Minton et al. (2005) had organism concentrations less than  $10 \text{ m}^{-3}$  while about 1.1% of the ships had concentrations greater than  $50,000 \text{ m}^{-3}$ . Thus, the actual propagule dose a waterbody receives is unknown and will depend on the distribution of organism concentrations among the ships discharging within a system.

Table 3-2: Historical number of invaders ( $N_h$ ), foreign ballast discharge volumes ( $D_h$ ), and per capita invasion probabilities (PCIP) for the East, Gulf, and Pacific coasts of the United States, 17 coastal ports, and the Great Lakes. The number of coastal invasions is the number of non-native invertebrates and macroalgae >50 microns first reported from 1981 to 2006 that were possibly introduced via ballast water and considered established. The total number of invaders in the coastal ports includes marine, brackish, and freshwater species, while the total without freshwater excludes the freshwater invaders. The foreign ballast discharges for the coastal waterbodies are the annual averages of 2005 to 2007 and include marine, brackish, and freshwater ports within the waterbody. Per capita invasion probabilities for the coastal waterbodies are given for a range of possible values, including the lower quantile (0.025), median, and upper quantile (0.975), based on the randomization algorithm that estimated organism concentrations for each of the ships discharging into a waterbody. The number of invaders for the Great Lakes is given for both macrofauna and phytoplankton for the period 1960 to 1988, while the ballast water discharge volume is for 1991. The sum of the discharge volumes and number of ships from the 17 ports is less than the coastal averages because all ports were included in the coastal values. FW = freshwater.

Waterbody	Total # Invaders / Total # w/o FW species	Average Annual Foreign Ballast Water Discharge Vol. ( $m^3 \text{ year}^{-1}$ )	# Ships with Foreign Ballast Water 2005-2007	Per Capita Invasion Probability (lower 0.025 quantile)	Per Capita Invasion Probability (median)	Per Capita Invasion Probability (upper 0.975 quantile)
<b>East Coast</b>	<b>40</b>	<b>7,407,832</b>	<b>12,860</b>	<b>4.00E-11</b>	<b>4.31E-11</b>	<b>4.64E-11</b>
Charleston	13/12	281,160	563	3.05E-10	3.70E-10	4.46E-10
Chesapeake	17/14	3,011,982	1315	3.85E-11	4.51E-11	5.28E-11
Jacksonville	14/13	130,296	791	7.48E-10	8.58E-10	9.83E-10
Miami	4/4	578,482	2515	5.04E-11	5.51E-11	6.02E-11
Narragansett Bay	13/13	21,030	19	2.38E-09	5.41E-09	1.35E-08
Portsmouth	9/9	6,377	10	3.26E-09	1.54E-08	6.16E-08

Waterbody	Total # Invaders / Total # w/o FW species	Average Annual Foreign Ballast Water Discharge Vol. (m <sup>3</sup> year <sup>-1</sup> )	# Ships with Foreign Ballast Water 2005-2007	Per Capita Invasion Probability (lower 0.025 quantile)	Per Capita Invasion Probability (median)	Per Capita Invasion Probability (upper 0.975 quantile)
<b>Gulf Coast</b>	<b>18</b>	<b>19,605,340</b>	<b>11,821</b>	<b>6.98E-12</b>	<b>7.31E-12</b>	<b>7.67E-12</b>
Corpus Christi	5/5	1,254,845	621	2.65E-11	3.18E-11	3.84E-11
Galveston	4/4	748,136	778	3.53E-11	4.28E-11	5.22E-11
Pensacola	3/3	1,121	8	8.72E-09	2.45E-08	7.88E-08
Tampa Bay	7/1	734,718	923	5.37E-11	6.54E-11	7.88E-11
<b>Pacific Coast</b>	<b>67</b>	<b>14,788,369</b>	<b>5998</b>	<b>3.41E-11</b>	<b>3.61E-11</b>	<b>3.83E-11</b>
Columbia River	22/12	5,533,618	1759	2.89E-11	3.17E-11	3.47E-11
Coos Bay	22/22	583,517	87	2.18E-10	3.04E-10	4.40E-10
Humboldt Bay	29/29	5,539	10	1.42E-08	5.24E-08	1.85E-07
Los Angeles / Long Beach	31/31	2,676,874	1693	8.20E-11	9.23E-11	1.05E-10
Puget Sound	23/21	3,960,438	1167	4.12E-11	4.64E-11	5.23E-11
San Diego Bay	23/21	31,271	112	4.20E-09	5.92E-09	8.52E-09
San Francisco Estuary	53/45	1,548,116	1015	2.33E-10	2.74E-10	3.22E-10
<b>Great Lakes – Macrofauna</b>	17	1,395,461	Unknown	NA	9.10E-11	NA
<b>Great Lakes – Phytoplankton</b>	14	1,395,461	Unknown	NA	NA	NA

Because the distribution of organism concentrations in ballast water among ships is highly skewed, the number of organisms being released into coastal waterbodies may be quite variable, and the mean concentration may over or underestimate the true propagule pressure. Consequently, rather than estimating PCIP values using the mean concentration of organisms we developed a randomization algorithm to calculate a range of PCIP values for a water body (Appendix A). The algorithm randomly assigns each ship discharging foreign ballast in a waterbody a concentration of organisms, selected from the distribution of values reported by Minton et al. (2005; their Figure 3a). The randomly selected concentration was then multiplied by the volume of foreign ballast discharged by that particular ship (see Table 3-2 for number of ships in each waterbody). These total organism values for each ship within a waterbody were then summed, generating a total propagule dose from which the PCIP value was calculated. This process was repeated 10,000 times to create a distribution of PCIPs for each waterbody from which the lower (0.025), median, and upper (0.975) quantile values were determined and are shown in Table 3-2.

Figure 3-2 shows the range of PCIPs for the Pacific Coast generated with this method. Using a range of possible PCIP values allows predictions to be made that do not underestimate the risk of invasion, which might occur if only the mean concentration of organisms is used. (Note that with a fixed historical invasion rate, higher PCIP values result from lower discharge values since the same number of invaders occurred with a lower propagule pressure.) Because we did not have individual ship records for the Great Lakes during 1991, the mean ballast water organism concentration from the IMO

baseline study ( $4640 \text{ m}^{-3}$ , MEPC, 2003b) was used to calculate the PCIP for the Great Lakes.

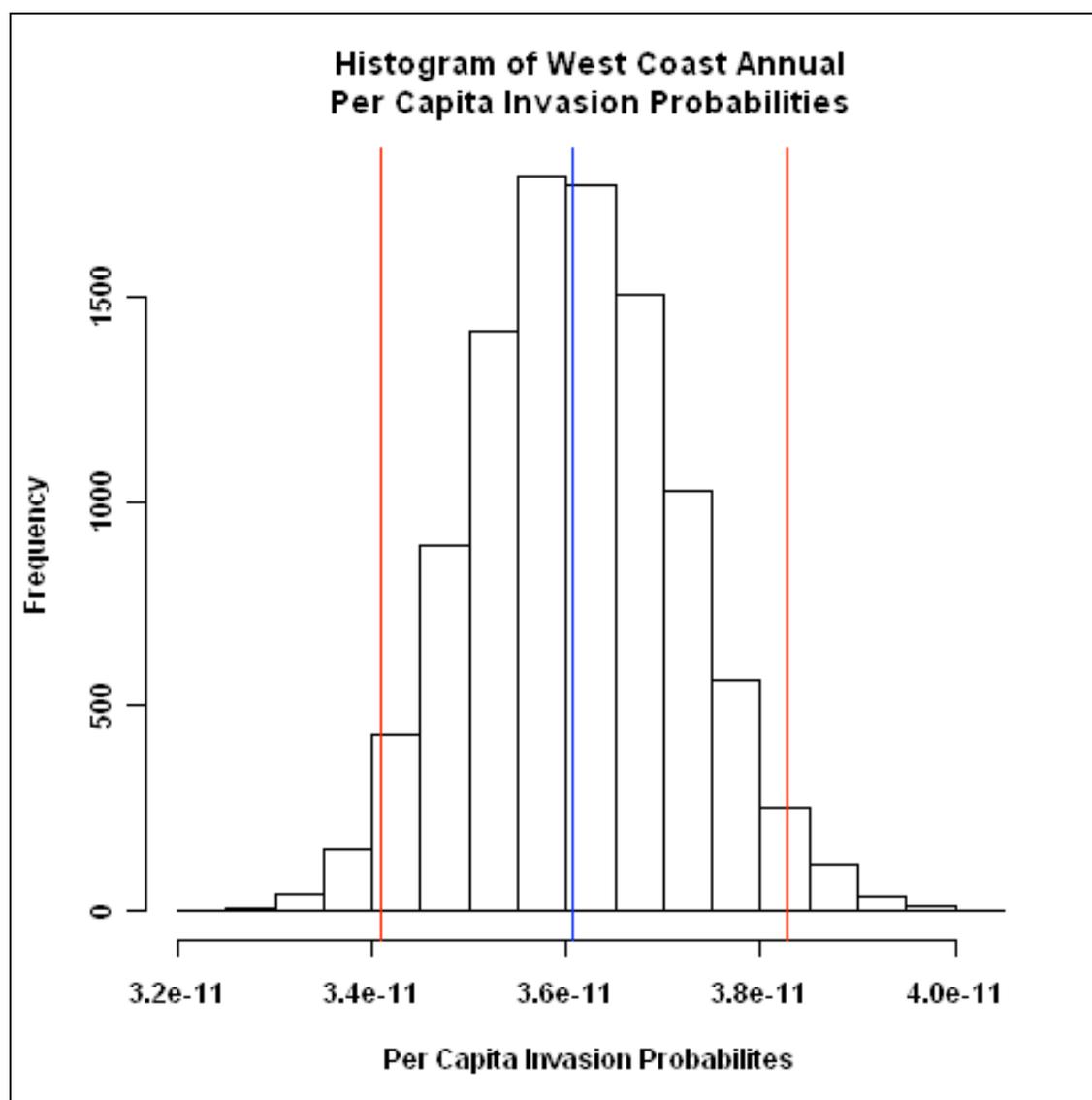


Figure 3-2: Distribution of per capita invasion probabilities (PCIPs) for the Pacific Coast based on 10,000 iterations of the randomization algorithm (Appendix A) for organism concentrations among the 5998 ships discharging foreign ballast. The red lines indicate the lower 0.025 quantile and the upper 0.975 quantile while the blue line indicates the median. Approximately 95% of the values fall between the red lines.

### *Estimates of Historical Invasion Rates*

The total numbers of invaders reported from 1981 to 2006 were synthesized for the contiguous United States Pacific, East, and Gulf Coasts as well as for 17 individual coastal waterbodies (Table 3-2). The 1981 to 2006 time period was before the implementation of mandatory mid-ocean ballast water exchange for coastal waterbodies, allowing the use of the estimates of organism concentrations in unexchanged ballast. The 25 year time period was chosen to smooth out short term variations in invasion rates as well as variations in monitoring efforts. A longer time period also helps to mitigate effects of the lag between an actual invasion event and when the species is first discovered (e.g., Costello and Solow, 2003).

The number of invaders is based on non-native invertebrates and macroalgae >50 microns; fishes and vascular plants were not included. Besides being reported in each coast or waterbody within the 25 year window, the species included in the analyses had to be considered established and potentially introduced via ballast water. The coastal invaders were classified into three salinity tolerance regimes: marine/estuarine (>20 psu), brackish (0.5-20 psu), and freshwater (<0.5 psu). This broad classification allows an evaluation of the importance of freshwater invaders in river-dominated estuaries such as the Columbia River. Because of the poor resolution between native versus nonindigenous phytoplankton species in coastal waters (Carlton, 2009), no attempt was made to estimate the number of invaders in the 10-50 micron size class. The numbers of invaders were generated from the Smithsonian Institution invasive species database (Fofonoff et al. 2003b). The majority of the East, Gulf, and Pacific invaders and their vectors are listed in Appendix A of Ruiz et al. (2000).

The 1960 to 1988 time period was chosen for the Great Lakes because it occurred before the implementation of mandatory ballast water exchange in 1993. During this interval, 17 macrofaunal ballast-associated invaders were reported (<http://www.glerl.noaa.gov/res/Programs/ncrais/docs/great-lakes-list.xls>, accessed September 26, 2009), resulting in an invasion rate of 0.58 invaders per year. This rate is based on all shipping-related invaders as well as three macrofaunal invaders with unknown vectors. The invasion rate for phytoplankton was similar (Table 3-2), resulting in a total rate of slightly more than 1 invader per year which is similar to that reported by Ricciardi (2006).

### ***Uncertainties in Historical Invasion Rates and Safety Factors***

Of the three parameters going into the calculation of a PCIP, the historical invasion rate has the greatest uncertainty. One source of this uncertainty is that many coastal nonindigenous species can potentially invade through multiple vectors, such as both ballast water and hull fouling (e.g., Fofonoff et al., 2003a). Inclusion of these “polyvetric” invaders (Ruiz and Carlton, 2003) in the historic invasion numbers in Table 3-2 potentially inflates the ballast-associated invasion rate, resulting in an artificially high PCIP. Because of differences in the relative importance of different vectors among estuaries, uncertainty related to multiple vectors is probably greater when comparing among estuaries than for the coast-wide estimates. For example, San Diego Bay, which has a high invasion rate relative to the ballast discharge volume, is the home to the largest naval base on the Pacific Coast consisting of approximately 54 naval ships. Ballast discharges from military ships are not included in the volumes in Table 3-2, but most naval ships tend to discharge relatively small amounts of ballast (see Table 3-2 in

Appendix A of U.S. EPA, 1999), which suggests a higher propagule pressure from hull fouling in San Diego Bay. Hull fouling may also be relatively more important in smaller ports that have low ballast discharge rates but a relatively large number of commercial fishing and recreational boats with no foreign ballast.

Secondary invasions could also inflate estimates of historical ballast-associated invasion rates in individual waterbodies. After the primary invasion and establishment of a new NIS into a biogeographic region, the invader may spread via secondary invasions from the initially established population. Likely mechanisms for secondary invasions include ballast water discharges and hull fouling via intracoastal commercial traffic emanating from the infected waterbody (e.g., Simkanin et al., 2009; Cordell et al., 2009) as well as hull fouling on recreational boats. Secondary invasions may also occur via natural dispersal mechanisms, such as currents and rafting, as suggested by occurrence of soft-bottom NIS in Pacific Northwest estuaries with no ballast discharges or oyster aquaculture (Lee et al., 2006; Lee, unpublished data).

An important source of uncertainty that could result in underestimating PCIP values is the underestimation of historical invasion rates. Carlton (2009) identified 12 sources of error leading to invader underestimation including unknown, unreported, misclassified, and rare invaders. In some parts of the world, such as Denmark, South Africa, and Chile where no invasions prior to mid-nineteenth century are recognized, the number of known invaders could be underestimated by as much as 5 to 10 times (Carlton 2009). For California, Cohen (in Falkner et al., 2006) suggested that unrecognized invaders could increase the invasion rate by 50% to 100%. A recent analysis of California invaders lists

457 cryptogenic species versus 358 nonindigenous species (California Dept. of Fish and Game, 2009); the California invasion rate would more than double if all these cryptogenic species were actually nonindigenous. While some of these cryptogenic species are likely unrecognized native sibling species (e.g., Knowlton, 1993), the high number of cryptogenic species suggests that the reported number of invaders may underestimate actual numbers by 50% to 100% within the United States.

Other sources of uncertainty could also cause us to underestimate the risk of introducing new invaders through ballast discharges: the relationship between propagule pressure and the probability of invasion could be steeper than the proportional relationship we assume in this model, in particular at very low concentrations (curve *d* in Figure 3-1); survival in ballast tanks could improve if voyage durations decrease due to faster ships; and waterbodies may become increasingly susceptible to invasion due to climate change or other environmental changes. While it is not possible to quantify the total uncertainty from these various sources, safety factors on the order of 5 to 20-fold have been proposed when calculating the potential risk to endangered and threatened species from exposure to pesticides (U.S. EPA, 2004b), and similar ranges could be used in the generation of discharge standards. We strongly suggest using a single safety factor rather than multiplying a string of individual safety factors for each potential source of uncertainty, which quickly results in unrealistic values (see Chapman et al., 1998).

### ***Among Port Patterns of Invasion Risk***

There is considerable range in the PCIP values among the 17 individual ports both along a single coast and across coasts (Table 3-2). The largest difference, more than 1600-fold,

is between the Humboldt Estuary and Columbia River. We suspect these among-estuary differences are due to a suite of non-exclusive factors. Part of this broad range may reflect differences in the invasibility among waterbodies, whether due to differences in biotic resistance or local environmental drivers. For example, the lower invasion probability in the Columbia River compared to other large Pacific Coast ports may be partially explained by wide seasonal and tidal salinity fluctuations (e.g., Hickey et al., 1998) that limit estuarine invaders to euryhaline species.

One pattern observed on all three coasts is that the smaller ports had more invaders than expected from the amount of foreign ballast water (i.e., higher PCIP values). Humboldt Bay, a small port in northern California, had only ten ships discharging foreign ballast from 2005 to 2007 (Table 3-2). Even with this small ballast input, Humboldt had the third largest number of invaders of the 17 estuaries, only exceeded by the San Francisco Estuary and the Los Angeles/Long Beach port. It is possible that these smaller ports have a greater invasibility than larger systems, but we suggest secondary invasions and invasions via mechanisms other than foreign ballast water discharges are relatively more important in these systems, which inflate the PCIP values. In particular, Humboldt Bay's proximity to the San Francisco Estuary and the prevailing northward oceanographic currents along the coast from San Francisco Estuary (particularly in El Niño years) may provide one mechanism of secondary invasion (Grosholz, 1996; Behrens Yamada *et al.*, 2005) in addition to intracoastal shipping.

We evaluated the potential effect of these factors on Humboldt by removing NIS from the Humboldt list if they: 1) had been observed in Pacific Coast estuaries that do not receive

ballast water discharges; 2) were found on the outer coast; and/or 3) had a potential vector other than ballast water. Of the 29 potential ballast-water invaders reported from Humboldt between 1980 and 2005, the introduction of only two could not be explained by mechanisms other than foreign ballast water discharges in Humboldt. The corresponding PCIP value (median = 3.58E-09) with the reduced invader list is only about 5% of the value when all potential invaders are included. We suspect that secondary invaders and polyvetric invaders also inflate the PCIP values in the other small ports. Another issue for estimating invasion probabilities in small estuaries is the large statistical variability in estimates based on small sample sizes. Consequently, ports with small amounts of ballast discharge will have high PCIP values even with the occurrence of a single ballast associated invader.

Because of these factors, we believe the PCIP values for the moderate to large ports are more reliable, with moderate/large ports defined as those having an average annual foreign discharge volume of  $\geq 100,000 \text{ m}^3$ . This threshold was chosen because of a distinct break in ballast discharge volumes that occurs between  $31,271 \text{ m}^3$  (San Diego) and  $130,296 \text{ m}^3$  (Jacksonville). The 12 moderate/large ports contribute 99.67% of the total ballast from the 17 estuaries. The range in PCIP values among these moderate to large ports is about 28-fold compared to the more than 1000-fold range when the small ports are included.

Discharge standards can be generated for individual ports by rearranging Equation 3-2 to calculate the organism concentration in ballast water ( $C_p$ ) associated with a projected ballast discharge volume ( $D_p$ ), acceptable risk as represented by the number of new

invaders per year ( $N_p$ ), PCIP value from Table 3-2 or otherwise calculated, and a safety factor of the form:

$$\text{Equation 3-3: } C_p = N_p / (D_p * \text{PCIP} * \text{Safety Factor})$$

$$\text{Safety factor} = \text{number} \geq 1 \text{ (unitless)}$$

In Equation 3-3, PCIP values can be the 0.5 (median) or the 0.975 quantile estimates calculated in which organism concentrations were randomly assigned to ships based on the distribution of concentrations measured by Minton et al. (2005). The 0.975 quantile represents the upper possible probability that a propagule discharged from ballast water will become established as a new invader. The median represents the “average” probability of establishment. Because it is in the denominator, the safety factor is set to 1 if no adjustment is made for uncertainties.

Because of the uncertainties surrounding invasion rates for single estuaries, we believe a better alternative is to base the standard on a specified confidence interval (e.g., upper 95% CI) around the PCIP values for the 12 moderate/large ports. An advantage of this approach is that it incorporates the among estuary variation in PCIP values in the calculation of the discharge standard. Using this approach, the formula to calculate the discharge standard is of the form:

$$\text{Equation 3-4: } C_p = N_p / (D_p * \text{PCIP}_{\text{CI}} * \text{Safety Factor})$$

Where  $\text{PCIP}_{\text{CI}}$  = an upper estimate of the probability that a single propagule from ballast discharge will become established as a new invasive species; calculated from the upper confidence interval estimates of PCIP for the 12 moderate to large ports. PCIP values for

the 12 individual ports are based on the 0.5 (median) or 0.975 quantile estimates from the randomization algorithm for organism concentrations for each ship.

	Upper 90% CI	Upper 95% CI	Upper 99% CI	Upper 99.9% CI
Median	3.48E-10	3.77E-10	4.41E-10	5.34E-10
0.975 quantile	3.71E-10	4.05E-10	4.80E-10	5.90E-10

Table 3-3: PCIP<sub>CI</sub> based on upper 90%, 95%, 99%, and 99.9% confidence intervals around the median and 0.975 quantile PCIP values for the 12 moderate to large estuaries in Table 3-2.

Table 3-3 gives the 90%, 95%, 99%, and 99.9% upper confidence intervals generated for the 12 moderate and large ports around the median and 0.975 quantile values. These are two-tailed confidence intervals so, for example, 5% of the values are larger than the 90% confidence interval values. Assuming an annual ballast water discharge rate of 30 million m<sup>3</sup> for the Pacific Coast, an acceptable risk as represented by an invasion rate of one new invader per thousand years, the upper 99.9% confidence interval value for the 0.975 quantile PCIP for the Pacific Coast, and a 10-fold safety factor, the discharge standard equation is of the form:

$$\text{Equation 3-5: } C_p = (1 \times 10^{-3} \text{ invader/yr}) / (30 \times 10^6 \text{ m}^3 \text{ ballast water/yr} * 5.90 \times 10^{-10} \text{ invader/organism} * 10) = 0.006 \text{ organisms m}^{-3}$$

The resulting discharge standard of 0.006 organisms m<sup>-3</sup> is similar to the USCG Phase II standard for >50 micron organisms (0.01 organisms m<sup>-3</sup>). The value derived from Equation 3-5 is based on a number of protective assumptions, including doubling the current Pacific Coast ballast discharge volume, using the 0.975 quantile for the estimated PCIP values, using the upper 99.9% CI value, and including a 10-fold safety factor.

Modifying the safety factor changes the discharge standard by varying degrees, and one way to visualize the “regulatory landscape” is to plot the invasion probabilities as a contour plot, or “risk diagram”, as a function of ballast water discharge volumes and organism concentrations. Figure 3-3 shows the risk diagrams based on three different safety factors (1, 10, and 20), using the PCIP value for the 99.9% confidence interval of the 0.975 quantile value from the 12 moderate/large estuaries. We consider these risk diagrams as complements to Equation 3-4, and the R code (R Development Core Team, 2008.) to generate these diagrams based on different input values is given in Appendix A.

### ***Across Coast Patterns of Invasion Risk***

Due to the significant potential for secondary invasions, we believe the best way to develop discharge standards is to use Equation 3-3 with PCIP values derived from the aggregated data for a particular coast. The aggregated data eliminate the uncertainty of secondary invaders as the historical invasion rate is based on the unique invaders to a coast so no invader is counted more than once. This approach is also supported by the fact that the overall variance in PCIP values among the coastal regions is small. In particular, there is only a 19% difference between the East and Pacific coasts (Table 3-2). The Gulf Coast PCIP is less than 6-fold smaller than the East or Pacific coasts, while the PCIP value for macrofauna for the Great Lakes is about 2-fold larger than those for the East and Pacific Coasts. Thus, even when comparing across three different coasts and the Great Lakes, there is only slightly more than a 12-fold range in the PCIP values. This relatively small range across diverse environments with different ballast discharge volumes and donor regions indicates that the analysis at this spatial scale captures many of the sources of variation.

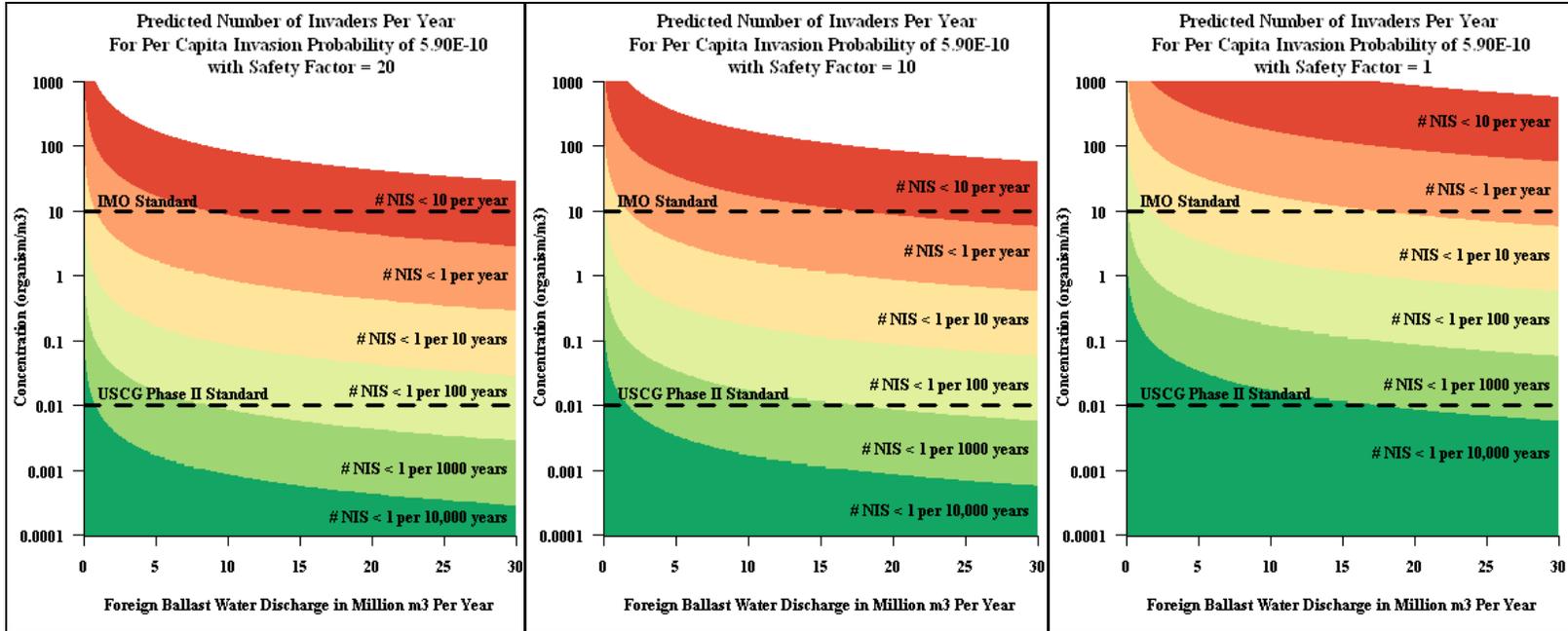


Figure 3-3: Risk diagrams illustrating the effect of three different safety factors (1, 10, and 20). Calculations are based on the 99.9% confidence interval of the 0.975 quantile value of PCIP from the 12 moderate to large estuaries.

We focus our analysis on the Pacific Coast because the extensive research on the distribution of NIS in this region (e.g., Cohen and Carlton, 1995; Cohen et al., 2001; Lee et al., 2003; deRiveria et al., 2005; California Dept. Fish Game, 2009) produces the most complete historical invasion rate. Using the same inputs for an acceptable invasion rate, ballast water discharge volume, and safety factor as for the estuary calculation (Equation 3-5), and the upper 0.975 quantile PCIP value specific to the Pacific Coast, the discharge standard becomes:

$$\text{Equation 3-6: } C_p = (1 \times 10^{-3} \text{ invaders/yr}) / (30 \times 10^6 \text{ m}^3 \text{ ballast water/yr} * 3.83 \times 10^{-11} \text{ invaders/organism} * 10) = 0.087 \text{ organisms m}^{-3}$$

Based on this set of assumptions, the discharge standard for >50 micron organisms would be approximately 100-fold lower than the proposed IMO standard, about 9-fold higher than the Phase II USCG standard, and about 10-fold higher than the standard derived from the multiple estuaries (Equation 3-5). As another example, we set the acceptable risk at one new invader per 100 years, the safety factor to 2, and use the median PCIP value instead of the upper quantile. With these less protective assumptions, the standard is 4.6 organisms m<sup>-3</sup>, about 2-fold lower than the IMO standard. Both of these predictions are illustrated as risk diagrams in Figure 3-4.

### ***Assumptions and Limitations***

As with any approach used to establish ballast water discharge standards, the per capita invasion probabilities make a number of assumptions. We list the major assumptions in Table 3-4 along with an assessment of how they affect the calculation of the PCIPs and the discharge standard derived from these probabilities.

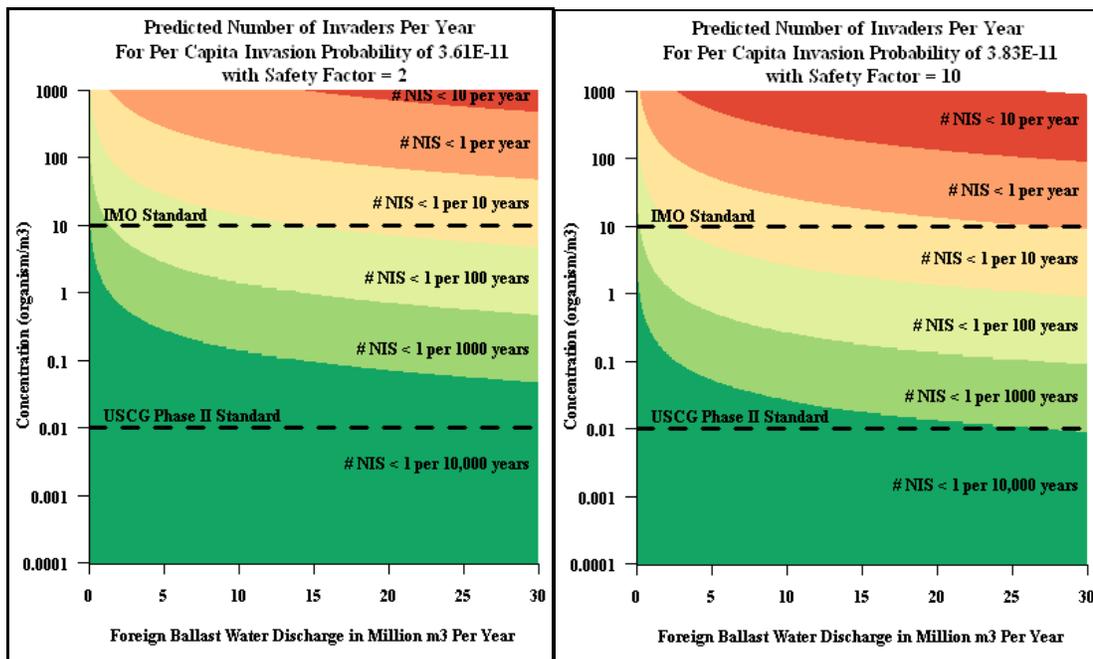


Figure 3-4: Risk diagrams based on less protective (left diagram) and more protective (right diagram) assumptions. The risk diagram on the left is based on the median PCIP for the Pacific Coast and a safety factor of 2. The diagram on the right is based on the upper 0.975 quantile PCIP value and a safety factor of 10.

The PCIP values for the smaller ports are substantially higher than those for systems with moderate to large ballast discharge volumes. As discussed, we believe this is largely a result of secondary invasions inflating the presumed ballast-associated invasion rate. However, if the higher invasion rates are actually a result of the smaller ports having a greater invasibility, the standards generated from the coast values or the moderate/large ports would not be protective of these systems. Another way that the present analysis could underestimate risk is by failing to account for the introduction of species that can become established with a single or very small number of individuals, such as a

parthenogenic species. As discussed in Section II, the only absolute protection against such invaders is a true zero discharge standard.

Our analysis is limited to organisms >50 microns, though the PCIP approach is theoretically applicable to smaller size classes. The practical limitations, however, are the difficulty in distinguishing native from nonindigenous protozoa, phytoplankton, and microbes and the corresponding lack of data on historical invasion rates. As pointed out by Carlton (2009), “no introduced diatoms, dinoflagellates, or other phytoprotists are recognized in San Francisco Bay, at either the morphospecies or genospecies level” despite the abundance of phytoplankton in ballast water. However, it would be possible to conduct an analysis for the Great Lakes given the reported historical invasion rate for phytoplankton (Table 3-2) if an estimate for the historical ballast water phytoplankton concentrations could be obtained.

By using past invasion rates to predict future rates, fundamental assumptions of the per capita probability approach are that neither the invasion potential of any new invaders or the invasibility of the waterbody itself will change in the future. If the best colonizers tended to invade first, then the PCIPs derived from these historical data would over predict the number of new invaders for a given propagule pressure. However, the apparent increase in the rate of invasions in a number of aquatic ecosystems (e.g., Cohen and Carlton, 1998; Holeck et al., 2004) is the opposite of what would be expected if there had been a general decrease in the virility of new invaders. Changes in the invasibility of aquatic ecosystems are more difficult to assess. In particular, environmental change associated with climate change is a “wild card” for any approach to set discharge

standards. Development due to port expansion could also change the invasibility of a system. Probably the only practical near-term solution is to incorporate a safety factor in anticipation of such changes. Over a longer-term, it is possible to periodically re-evaluate PCIP values for a coast to determine if there have been any substantial changes.

### ***Recommendations/Conclusions***

The per capita invasion probability approach attempts to cut through the “Gordian Knot” of uncertainties associated with predicting ballast water invasions, and Equations 3-3 and 3-4 and the risk diagrams (Figures 3-3 and 3-4) can be used to set organism-based discharge standards. As with all approaches, there are a number of assumptions (see Table 3-4). Accordingly, our strategy was to develop an approach that allows risk managers the option to develop discharge standards with different risk levels based on different sets of assumptions. Specifically, the following inputs can be set: 1) acceptable invasion risk as measured by an invasion rate; 2) ballast water discharge volume; 3) use of PCIPs based on median ballast water organism concentration or upper quantile values; 4) median or an upper confidence interval around the PCIP with the among-port analysis; and 5) magnitude of the safety factor.

The uncertainty around the parameters going into the per capita invasion probability model is relatively small. Even with the historical invasion rate, the uncertainty is only on the order of 2-fold for the Pacific Coast. In comparison, there are much greater levels of uncertainty in predicting the population vital rates that are needed for reaction-diffusion or PVA models. Additionally, the PCIP model does not have to be parameterized for each species or type of species as with population modeling

approaches. Finally, the data going into the per capita probability approach are readily understandable by managers and the public, which is beneficial in gaining acceptance for any ballast water discharge standard.

Of the three approaches to setting discharge standards discussed here (PCIP from individual estuaries; values based on upper confidence intervals from distributions of PCIPs about individual estuaries; PCIP values based on aggregated coastal values), we suggest that the coastal approach has the lowest inherent uncertainty. Furthermore, since most invaders spread along the coast, analysis at this scale is ecologically appropriate. Because of the extensive effort in documenting invaders on the Pacific Coast, the PCIP values for the Pacific Coast are the most reliable and we recommend using this coast to generate discharge standards for marine and estuarine ports.

The PCIP value for macrofauna for the Great Lakes is about 2-fold larger than those for the East and Pacific coasts, suggesting that there may be a greater likelihood of any individual propagule becoming established as a new invader in the Great Lakes. However, less complete data were available for ballast discharge volume and organism concentrations, and we consider the calculations for the Great Lakes a preliminary analysis. While there is the complicating factor of mandatory ballast water exchange after 1993, it may be possible to generate more up-to-date data for an analysis using the PCIP approach with a detailed study of the Great Lakes. As mentioned above, a study focused on the Great Lakes may also allow an analysis on phytoplankton invasion rates.

<b>Assumption</b>	<b>Effect on Estimate of Per Capita Invasion Probability</b>	<b>Effect on Discharge standard</b>	<b>Mitigation Approaches</b>
Linear dose-response	Likely over estimates invasion probability for many sexual species due to Allee effects; potentially under estimates for asexual and parthenogenic species.	Protective against most sexual invaders; possibly under protective for asexual and parthenogenic species.	Use upper bound estimates for input values and/or safety factor.
Secondary invasions did not contribute to historical invasion rate.	Inflates PCIP to the extent that invaders did not invade via foreign ballast water discharged into the waterbody.	Erroneously decreases discharge standard.	Exclude small ports from analysis and/or conduct analysis on a coastal scale.
Exclusion of small ports from across-port calculations.	Generates more accurate PCIPs if invasions in small ports from secondary vectors. Artificially decreases PCIP if actual primary invasions into the small ports.	Depends whether secondary or primary ballast water invaders in small ports.	Conduct analysis on a coastal scale.
All ship-associated invaders actually invaded via foreign ballast water.	Increases PCIP to extent that species invaded via a mechanism other than foreign ballast water.	Erroneously decreases discharge standard.	Account for possible overestimate of invasion rate when choosing other inputs or the safety factor. Coastal scale analysis may reduce effect of polyvectic invaders.
No change in invasibility of waterbody over time or change in the invasion potential of new invaders.	Either increases or decreases PCIP depending upon type & magnitude of environmental changes in waterbody. Decrease in viability of new invaders results in PCIPs based on historical rates over predicting new invasions.	Protective or under protective depending upon the type & magnitude of changes.	Use upper bound estimates for input values and/or safety factor to account for changes in environment. No adjustment for historical invasion rates.

Table 3-4: Major assumptions of the per capita invasion probability approach to setting ballast water discharge standards.

Secondary invasions appear to be an important source of uncertainty. To understand the role of secondary invasions better, future surveys for nonindigenous species should not only focus on the larger ports but should also include smaller ports and estuaries with no foreign ballast input. Additionally, further studies of the role of intracoastal shipping and ballast discharges are needed to help elucidate their role in spreading invaders into smaller ports with minimal foreign ballast water discharges.

### *Acknowledgements*

The invasion data (collection and compilation) were supported by past funding: Department of Defense Legacy Program, National Sea Grant Program, Smithsonian Institution. Shipping data, was extracted from the National Ballast Water Information Clearinghouse. This publication was subjected to review by the Geosciences Department of Oregon State University, the National Health and Environmental Effects Research Laboratory's Western Ecology Division of the EPA, and the Western Fisheries Research Center Biological Resources Discipline of USGS, and is approved for publication. However, approval does not signify that the contents reflect the views of the U.S. EPA. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader; such use does not constitute official endorsement or approval by the U. S. Department of Interior, the U. S. Geological Survey, or the U. S. Environmental Protection Agency of any product or service to the exclusion of others that may be suitable.

## References

- Behrens Yamada, S., B. R. Dumbauld, A. Kalin, C. E. Hunt, R. Figlar-Barnes, and A. Randall. 2005. Growth and persistence of a recent invader *Carcinus maenas* in estuaries of the northeastern Pacific. *Biological Invasions* **7**:309-321.
- Blake, A. C., D. B. Chadwick, A. Zirino, and I. Rivera-Duarte. 2004. Spatial and Temporal Variations in Copper Speciation in San Diego Bay. *Estuaries* **27**:437-447.
- Box, G. E. P., and N. R. Draper. 1987. *Empirical Model-Building and Response Surfaces*. Wiley.
- Boyd, M. J., T. J. Mulligan, and F. J. Shaughnessy. 2002. Non-indigenous marine species of Humboldt Bay, California. California Department of Fish and Game.
- Carlton, J. T. 1996. Pattern, process, and prediction in marine invasion ecology. *Biological Conservation* **78**:97-106.
- Carlton, J. T. 2009. Deep Invasion Ecology and the Assembly of Communities in Historical Time. Pages 13-56 *in* G. Rilov, and Crooks, J.A., editor. *Biological Invasions in Marine Ecosystems*. Ecological Studies. Springer-Verlag, Berlin Heidelberg.
- Carlton, J. T., and J. B. Geller. 1993. Ecological Roulette: The Global Transport of Nonindigenous Marine Organisms *Science* **261**:78-82.
- Cohen, A. N., and J. T. Carlton. 1998. Accelerating invasion rate in a highly invaded estuary. *Science* **279**:555-558.
- Colautti, R. I., I. A. Grigorovich, and H. J. MacIsaac. 2006. Propagule pressure: a null model for biological invasions. *Biological Invasions* **8**:10.
- Cordell, J. R., D. J. Lawrence, N. C. Ferm, L. M. Tear, S. S. Smith, and R. P. Herwig. 2009. Factors influencing densities of non-indigenous species in the ballast water of ships arriving at ports in Puget Sound, Washington, United States. *Aquatic Conservation: Marine and Freshwater Ecosystems* **19**:322-343.
- Costello, C. J., and A. R. Solow. 2003. On the Pattern of Discovery of Introduced Species. *Proceedings of the National Academy of Sciences of the United States of America* **100**:3321-3323.
- Drake, J. M. 2004. Allee effects and the risk of biological invasion. *Risk Analysis* **24**:795-802.
- Fofonoff, P. W., G. M. Ruiz, B. Steves, and J. T. Carlton. 2003a. In ships or on ships? Mechanisms of transfer and invasion for non-native species to the coasts of North

- America. Pages 152-182 in G. M. Ruiz and J. T. Carlton, editors. *Invasive species: vectors and management strategies*. Island Press, Washington DC.
- Fofonoff, P. W., G. M. Ruiz, B. Steves, A. H. Hines, and J. T. Carlton. 2003b. National Exotic Marine and Estuarine Species Information System (NEMESIS).
- Gascoigne, J., and R. N. Lipcius. 2004. Allee effects in marine systems. *Marine Ecological Progress Series* **269**:49-59.
- Grosholz, E. D. 1996. Contrasting rates of spread for introduced species in terrestrial and marine systems. *Ecology* **77**:1680-1686.
- Hickey, B. M., L. J. Pietrafesa, D. A. Jay, and W. C. Boicourt. 1998. The Columbia River Plume Study: Subtidal variability in the velocity and salinity fields. *Journal of Geophysical Research*.
- Holeck, K. T., E. L. Mills, H. J. MacIsaac, M. R. Dochoda, R. I. Colautti, and A. Ricciardi. 2004. Bridging Troubled Waters: Biological Invasions, Transoceanic Shipping, and the Laurentian Great Lakes. *BioScience* **54**:919-929.
- Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* **16**:199-204.
- Kramer, A. M., B. Dennis, A. Liebhold, M., and J. M. Drake. 2009. The evidence for Allee effects. *Population Ecology* **51**:341-354.
- Lee, H., II, and D. A. Reusser. 2006. Pacific Coast Ecosystem Information System (*PCEIS*). Beta version of geospatial database of native and nonindigenous estuarine/coastal species of the Northeast Pacific and associated landscape data (internal EPA and USGS release). Office of Research and Development, National Health and Environmental Effects Research Laboratory.
- Locke, A., D. M. Reid, W. G. Sprules, J. T. Carlton, and H. C. Van Leeuwen. 1991. Effectiveness of mid-ocean exchange in controlling freshwater and coastal zooplankton in ballast water. *Canadian Technical Reports in Fisheries and Aquatic Sciences*.
- MEPC. 2003. Harmful Aquatic Organisms in Ballast Water: Comments on draft Regulation E-2 Concentrations of organisms delivered in ships' ballast water in the absence of any treatment: Establishing a baseline for consideration of treatment efficacy. Submitted by the International Council for the Exploration of the Sea (ICES). Marine Environment Protection Committee, International Maritime Organization. MEPC 49/2/21, London (May 23, 2003).
- Minton, M. S., E. Verling, A. W. Miller, and G. M. Ruiz. 2005. Reducing propagule supply and coastal invasions via ships: effects of emerging strategies. *Frontiers in Ecology and the Environment* **3**:304-308.

- Molnar, J. L., R. L. Gamboa, C. Revenga, and M. D. Spalding. 2008. Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment* **6**:485-492.
- Occhipinti-Ambrogi, A., and D. Savini. 2003. Biological invasions as a component of global change in stressed marine ecosystems. *Marine Pollution Bulletin* **46**:542-551.
- Reid, D. M., and J. T. Carlton. 1997. Shipping Study 1 - A: A Study of the Introduction of Aquatic Nuisance Species by Vessels Entering the Great Lakes and Canadian Waters Adjacent to the United States. CG-D-17-97, Maritime Studies Program, Williams College - Mystic Seaport, Springfield, Virginia.
- Ricciardi, A. 2006. Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distributions* **12**:425-433.
- Ruiz, G. M., and J. T. Carlton. 2003. Invasion Vectors: A Conceptual Framework for Management. Page 518 *in* G. M. Ruiz and J. T. Carlton, editors. *Invasive Species: Vectors and Management Strategies*. Island Press, Washington.
- Ruiz, G. M., J. T. Carlton, E. D. Grosholz, and A. H. Hines. 1997. Global invasion of marine and estuarine habitats by nonindigenous species: Mechanisms, extent, and consequences. *American Zoologist* **37**:621-632.
- Ruiz, G. M., P. Fofonoff, A. H. Hines, and E. D. Grosholz. 1999. Non-indigenous species as stressors in estuarine and marine communities: Assessing invasion impacts and interactions. *Limnology and Oceanography* **44**:950-972.
- Ruiz, G. M., P. W. Fofonoff, J. T. Carlton, M. J. Wonham, and A. H. Hines. 2000. Invasion of coastal marine communities in North America: apparent patterns, processes, and biases. *Annual Review of Ecology and Systematics* **31**:481-531.
- Simberloff, D. 2009. The Role of Propagule Pressure in Biological Invasions. *Annual Review of Ecology, Evolution, and Systematics* **40**.
- Simkanin, C., I. Davidson, M. Falkner, M. Sysma, and G. M. Ruiz. 2009. Intra-coastal ballast water flux and the potential for secondary spread of non-native species on the US West Coast. *Marine Pollution Bulletin* **58**:366-374.

**Chapter 4: Predictions for an invaded world: a strategy to predict the distribution of native and non-indigenous species at multiple scales**

Authors: Deborah A. Reusser and Henry Lee II

***Abstract***

Habitat models can be used to predict the distributions of marine and estuarine non-indigenous species (NIS) over several spatial scales. At an estuary scale, our goal is to predict the estuaries most likely to be invaded, but at a habitat scale, the goal is to predict the specific locations within an estuary that are most vulnerable to invasion. As an initial step in evaluating several habitat models, model performance for a suite of benthic species with reasonably well-known distributions on the Pacific coast of the United States needs to be compared. We discuss the utility of non-parametric multiplicative regression (NPMR) for predicting habitat- and estuary-scale distributions of native and NIS. NPMR incorporates interactions among variables, allows qualitative and categorical variables, and utilizes data on absence as well as presence. Preliminary results indicate that NPMR generally performs well at both spatial scales and that distributions of NIS are predicted as well as those of native species. For most species, latitude was the single best predictor, although similar model performance could be obtained at both spatial scales with combinations of other habitat variables. Errors of commission were more frequent at a habitat scale, with omission and commission errors approximately equal at an estuary scale.

**Keywords:** ecological niche modeling, geographic scale, habitat modeling, non-indigenous species, non-parametric multiplicative regression, Northeast Pacific.

## ***Introduction***

Many new habitat-modeling techniques are emerging in the environmental sciences for predicting distributions of plants and animals. These new techniques have been developed by conservation biologists to identify critical habitats for threatened and endangered species (Peterson, 2001), and in invasion biology to identify areas at risk of invasion (Peterson and Vieglais, 2001; Herborg et al., 2007). One limitation has been that the data available for these types of modeling exercises have been sparse for many species over large geographic areas. Recently, however, museums, universities, and government agencies have been collating and distributing large biological and environmental datasets on the Internet, making it possible to apply the new modeling techniques to many different species and environments. The difficulty comes in knowing how to apply the techniques with different datasets. As pointed out by McNyset and Blackburn (2006), it is crucial to understand the specific habitat model being used, including how errors and uncertainties associated with the model affect model performance (Barry and Elith, 2006).

Drawing on the lessons from Elith et al. (2006) and Barry and Elith (2006), we are currently developing a strategy to evaluate a suite of habitat models for predicting distributions of native and non-indigenous estuarine species. Estuarine environments present additional challenges compared with terrestrial and marine environments. For example, estuaries functionally represent habitat islands, and the larvae of many species are episodically dispersed great distances by currents. In addition, estuarine science does not have the luxury of continuous distributions for key environmental data layers, such as

sediment composition and water temperature, that are available in terrestrial and, to a lesser extent, marine environments. Most Pacific coast estuaries are not detectable at a one-degree cell size, as used by Wiley et al. (2003) to predict the distributions of fish in the Atlantic Ocean and Caribbean Sea.

The strategy we are developing to evaluate the performance of habitat-modeling techniques for estuarine species is: (i) initially to model native and non-indigenous species (NIS) with reasonably well-known distributions; (ii) to include species with a variety of different spatial extents; (iii) to challenge the approaches by modeling large spatial extents, including areas outside the known range of the target species, when possible; (iv) to evaluate outcomes over different ecologically relevant spatial scales; and (v) to validate model outputs using independent data from the Northeast Pacific while attempting to minimize the effects of autocorrelation. After evaluating the various habitat models with known species, we will attempt to predict the distributions of new invaders.

Here, we present a preliminary evaluation of a relatively new modeling technique, non-parametric multiplicative regression (NPMR). To that end, we evaluate NPMR using marine/estuarine benthic species with reasonably well-documented distributions at two spatial scales, habitat and estuary. The reason for modeling a suite of species at this stage, rather than conducting a detailed analysis on one or two species, is to evaluate model behavior across a range of taxonomic and functional groups, and to evaluate the utility of different habitat variables at two spatial scales.

### ***Methods***

We modeled species distributions using NPMR as implemented in HyperNiche version

1.20 (McCune and Mefford, 2004). NPMR represents a species' response surface in multidimensional environment space by smoothing its response in a local area of environmental space through the combination of information from neighboring observations in environmental space (McCune, 2006). The reported advantages of NPMR are that it: (i) incorporates interactions among multiple ecological variables; (ii) can represent complex species response surfaces; and (iii) controls for overfitting (McCune, 2006). Three additional advantages for estuarine species modeling are that it: (iv) does not require continuous environmental data layers; (v) can incorporate categorical habitat variables; and (vi) incorporates absence as well as presence data. The Gaussian weighting function with a local mean estimator was used in all NPMR modeling. Several different approaches are available to evaluate model performance, and for this analysis, we used the area under the curve (AUC) of the receiver operating characteristic curve (Elith et al., 2006). An AUC value greater than 0.5 indicates that the model is performing better than random in predicting a species' presence/absence. Elith et al. (2006) used a cut-off of AUC greater than 0.75 for models that had "a useful amount of discrimination", and we add the criterion that an AUC greater than 0.90 indicates that the model has high discrimination. Additionally, NPMR uses a "leave-one-out cross validation", so it is possible to estimate omission and commission errors. These errors were calculated using a threshold value of greater than 0 probability of occurrence equaling presence.

### *Habitat (point) Scale*

The objective of the habitat-scale modeling was to predict the presence/absence of a species at specific points as defined by individual benthic samples. Benthic samples for the modeling were obtained from US Environmental Protection Agency's (EPA) Western

Coastal Environmental Monitoring and Assessment Program (EMAP; see Nelson et al., 2005). We used samples from estuarine surveys in California, Oregon, and Washington in 1999, 2000, 2002, and 2003, as well as samples from the 2002 survey of near-coastal and estuarine sites in south central Alaska and the shelf survey (30–120 m) of Washington in 2003. Therefore, habitat scale includes both estuarine and nearshore sites, whereas the estuary-scale analysis included only estuarine studies as defined below. Most samples were taken with a 0.1-m<sup>2</sup> grab and sieved through a 1.0-mm mesh screen. However, because different sample sizes or meshes were used in the San Francisco and 2002 intertidal surveys, the current analysis utilizes presence/absence data rather than abundance data. In all, there were 664 benthic samples and greater than 2500 taxa across all stations.

The 23 most frequently occurring species at greater than 50 stations were chosen for modeling. An additional four species with occurrence at between 38 and 49 sites were also modeled to include additional species whose ranges extended into Alaska or the Washington continental shelf. In all, 13 native species, 11 NIS, and 4 cryptogenic species were modeled at a habitat scale. The most frequently occurring species (greater than 100 stations) included native species of amphipod (*Americorophium salmonis*) and polychaete (*Glycinde polygnatha*), NIS of amphipod (*Grandidierella japonica* and *Monocorophium acherusicum*), a bivalve (*Mya arenaria*), polychaetes (*Polydora cornuta*, *Pseudopolydora kempfi*, and *Streblospio benedicti*), and a cryptogenic species of tanaid (*Leptochelia dubia*).

Quantitative habitat variables at each station included percentage silt and clay, total organic carbon (TOC) of the sediment, and sample depth. Because the values for overlying salinity were not available for the intertidal sites, we used two categorical salinity classifications. One was the Venice system, consisting of five classes: fresh water, oligohaline, mesohaline, polyhaline, and euhaline. The second system consisted of subdividing the oligohaline, mesohaline, and polyhaline into two classes each, for a total of eight classes. The salinity class for each site was determined from the overlying water sample when available; otherwise, the location of each sampling point was plotted in a GIS system, and the salinity class was estimated by the location's proximity to existing salinity records. Another suite of categorical variables was the presence/absence of four ecological engineering guilds: burrowing shrimp (*Neotrypaea californiensis* or *Upogebia pugettensis*), submerged aquatic vegetation (*Zostera marina* or *Zostera japonica*), marsh plants, and/or macroalgae. The presence of burrowing shrimp, rooted aquatic plants, or macroalgae can alter the structure of intertidal and shallow-water benthic assemblages through a variety of mechanisms, including the effects on dissolved oxygen, sedimentation, and bioturbation.

### *Estuary Scale*

The objective at an estuary scale was to predict the presence/absence of a species within a specific estuary. Data for modeling at an estuary scale were obtained from the Pacific Coast Ecosystem Information System (*PCEIS*). *PCEIS* is a regional database of native and non-indigenous invertebrates, plants, and fish found in estuaries on the Pacific coast of the US, with associated landscape and watershed characteristics for each estuary (Lee and Reusser, 2006). The information contained in *PCEIS* comes from a variety of

sources, including historical sampling efforts, museum records, published literature, and ongoing monitoring efforts such as the US EPA's EMAP program. From a set of 180 estuaries with biological information in *PCEIS*, a subset of 28 was selected, where at least 100 species had been reported, to reduce errors of omission from false negative occurrences. These estuaries were well distributed along the Pacific coast from Grays Harbor, Washington, to the Tijuana River in California, and varied in size from 4 to 14 518 km<sup>2</sup>.

Species occurrence data for the 28 estuaries were extracted from *PCEIS* for the 28 species used in the habitat-scale model evaluation. Based on model data predictor guidelines, this set of 28 species was reduced to a subset of 13 species (seven native, five non-indigenous, and one cryptogenic) that had a minimum of 10 species present and 10 species absent in the 28 selected estuaries. A suite of 13 landscape-scale characteristics for each estuary were extracted from *PCEIS* across four broad categories: geography (biogeographic province and latitude), climate (mean annual air temperature and mean annual precipitation), watershed (land, water, intertidal, subtidal, and riverine area), and geomorphology (ratio of water to land, ratio of subtidal to intertidal, ratio of riverine to estuarine, and ratio of intertidal to estuarine area).

### *Geographic Variables and Datasets*

Two geographic variables, latitude and biogeographic province, were evaluated in both the habitat and estuary analyses. The biogeographic provinces were based on the study by Croom et al. (1995) that divides the outer coasts of California, Oregon, and Washington into three provinces, and classifies Puget Sound as a fourth. For the habitat analysis,

south central Alaska and the Washington shelf samples were considered separate biogeographic provinces, and a categorical variable was added to indicate whether the site was coastal or estuarine. In the habitat analysis, model runs including all sites (n = 664) are referred to as “All sites with geography” or “All sites without geography”, depending on whether or not latitude and biogeographic province were included. We could not include quantitative measures of overlying water salinity or temperature in the “all site” scenarios because our overall data included intertidal sites, precluding the measurement of overlying water variables. To evaluate their importance, the subtidal sites (n = 454) were modeled using the site-specific quantitative values for temperature and salinity, as well as the categorical salinity classes mentioned above.

### ***Results and Discussion***

A prime objective of this preliminary analysis was to determine whether NPMR provided sufficient power in predicting the distributions of native and non-indigenous benthic species to warrant more detailed analyses. For the “All sites with geography” models, the average AUC for all 28 species at a habitat scale was 0.87. The average AUC for the 13 species used in the estuarine-scale analysis was 0.79. Based on the relatively high AUC values for both scales, we conclude that NPMR was sufficiently predictive to continue its evaluation with more detailed analyses.

A related question is whether errors of omission (false negatives) or commission (false positives) were more prevalent. Are the models more likely to predict incorrectly that a species is absent, or are they more likely to predict incorrectly that a species is present? At the habitat scale, the commission error was higher than the omission error for 23 of the 28 species, based on the “All Sites with geography” models. One possible reason for

the higher frequency of commission errors was that we were unable to include temperature and salinity in this model, because our overall data included intertidal sites. However, the inclusion of a quantitative measure of salinity and overlying water temperature in the subset of subtidal samples did not change this pattern, and commission errors were still larger in 21 of the 28 species. Another possible cause is the inherent small-scale variability of benthic organisms. In this case, the model would correctly predict that the habitat type is suitable, but the species could be absent in an individual sample because of processes operating locally, such as small-scale variation in recruitment or predation. In future habitat-scale analyses, we will attempt to address this issue by aggregating samples over a larger spatial area.

At an estuary scale, the frequency of the two types of errors was more similar, with the commission error more prevalent than the omission error for 7 of the 13 species, with one species containing no omission or commission errors. These omission errors are real errors—the model predicted that a species would not occur, even though it had been found within the estuary. Conversely, the commission errors could reflect reporting biases, and future sampling of these estuaries may reveal that the species is indeed present.

Another question was whether our model would detect a difference in ability to predict the presence/absence of native species vs. NIS. One possibility was that model performance could be degraded if the NIS had not yet approached an “equilibrium” distribution. Using the models based on the “All sites with geography” scenario at a habitat level, we found no difference in predictive power between native and NIS. The average

AUC for the 13 native species was 0.87, and for the 11 NIS, it was 0.86. Similarly, at the estuary scale, the average AUC of both the native and NIS was 0.79. These results indicate that it is possible to use NPMR to evaluate the relationship of well-established NIS with environmental factors at habitat and estuary scales. This is not to be construed as concluding that the distribution of some of these NIS might not expand in the future or that the habitat models will work as well with rarer or recently introduced species.

Our last line of inquiry focused on evaluating the key environmental predictors. One specific question relates to the inclusion of geographic variables. This raises a conundrum. If the goal is to predict a species' distribution within a proscribed geographic extent, then inclusion of geographic variables may be appropriate. However, if the goal is to predict the "equilibrium" range of a species, then inclusion of geographic variables may overestimate the range for a species undergoing contraction or, conversely, underestimate the range for a species undergoing expansion. For example, predicting the distribution of a recently introduced NIS based on models using geographic variables may substantially underestimate its final range.

To assess the influence of the geographic variables, we compared model performance with and without their inclusion. At an estuary scale, removing the geographic variables had no effect on the average AUC of the 13 species, though there was a decrease of one species in the greater than 0.90 category and an increase of one species in the minimally predictive range (greater than 0.5 and less than 0.75; Figure 4-1). It appears that the combination of mean air temperature and mean precipitation of the watershed captures much of the environmental information contained in latitude and/or biogeographic

province. Future modeling efforts at an estuary scale will explore whether the inclusion of high-resolution coastal sea surface temperature improves the predictions at this scale.

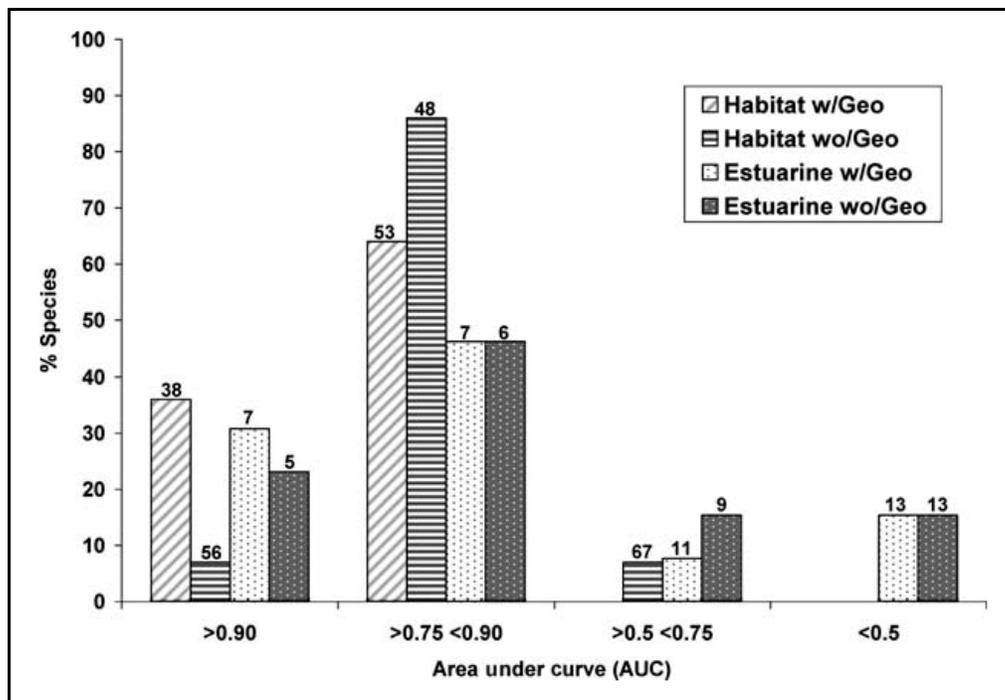


Figure 4-1: Percentage of species falling into different classes of the AUC in the habitat- and estuary-scale modeling with and without the geographic (w/Geo and wo/Geo, respectively) variables, latitude, and biogeographic provinces. “Habitat w/Geo” are the habitat-scale models with  $n = 664$  samples and the geographic variables included. “Habitat wo/Geo” are the results from the same dataset without the geographic variables. “Estuarine w/Geo” and “Estuarine wo/Geo” are the estuary-scale models ( $n = 28$  estuaries) with and without the geographic variables, respectively. The habitat results were based on 28 species and the estuarine results on 13 species. Numbers above each bar indicate the average neighborhood size used in the model.

At the habitat scale, latitude was the single best predictor for 21 species, and either latitude or biogeographic province was incorporated into the final model for all 28 species. This is not surprising, because sampling station locations ranged from Tijuana to Alaska. Nonetheless, removal of the geographic variables resulted in just a relatively small decline in overall model performance, with average AUC decreasing from 0.87 to 0.83. The more apparent effect of removing geographic variables was that the number of

species models that displayed high discriminatory power (AUC .0.90) declined from ten to two, with a concomitant increase in the number of species with an AUC of 0.75–0.90 (Figure 4-1).

Quantitative values of overlying water salinity and temperature were considered as potentially key variables at a habitat scale. To evaluate their effects, we took the subset of 454 subtidal samples that included these measurements. The results of these subtidal samples mirrored those of the “All Sites with geography” models, with an average AUC of 0.87 when the geographic variables were included. Therefore, inclusion of these variables did not have a discernible effect on average model performance when the geographic variables were included, and latitude was still the single most predictive variable for 14 of the 28 species in the sub-tidal subset. Excluding the geographic variables reduced the average AUC of the subtidal subset slightly (0.85). When the geographic variables were excluded, water temperature was the single most important variable for 16 species, and the quantitative measure of salinity was the most important for seven species. Although never the most important single variable, the categorical measurements of salinity were included in the final models approximately as often as the quantitative salinity measurements. The results suggest that, although latitude remained the single best predictor, combinations of variables including water temperature and salinity can generate comparable models. The results also suggest that, although quantitative salinity measurements are preferred, the use of salinity classes is a reasonable substitute when quantitative measurements are not available.

### ***Conclusions***

This preliminary analysis has suggested that NPMR can predict the distributions of many native and non-indigenous benthic species with a reasonable degree of accuracy at both habitat and estuary scales. However, more detailed analysis with NPMR is required along with comparisons with other modeling approaches to conclude whether it is the “best” approach to predicting the potential distributions of newly introduced NIS. The preliminary analysis also generated insights into the types of habitat variables that can be used in such predictions. Geographic variables were the strongest single predictors at both scales, although combinations of watershed-and estuarine-landscape characteristics (estuary scale) or site-specific quantitative or categorical habitat variables (habitat scale) could be used to generate models of similar predictive power.

### ***Acknowledgements***

DAR was partially funded through AMI/GEOSS IAG # DW-14-92231501-0 from the US Environmental Protection Agency. The publication was subjected to review by the National Health and Environmental Effects Research Laboratory’s Western Ecology Division and the USGS, and is approved for publication. However, approval does not signify that the contents reflect the views of the US EPA or the USGS. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader; such use does not constitute official endorsement or approval by the US Department of Interior, the US Geological Survey, or the US Environmental Protection Agency of any product or service to the exclusion of others that may be suitable.

## *References*

- Barry, S., and Elith, J. 2006. Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43: 413–423.
- Croom, M., Wolotira, R., and Henwood, W. 1995. Marine Region 15: Northeast Pacific. In *A Global Representative System of Marine Protected Areas, Vol. IV, The Great Barrier Reef Marine Park Authority*, pp. 55–106. Ed. by G. Kelleher, C. Bleakley, and S. Wells. The World Bank and IUCN, Washington DC.
- Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R. J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29: 129–151.
- Herborg, L., Jerde, C. L., Lodge, D. M., Ruiz, G. M., and MacIsaac, H. 2007. Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications*, 17: 663–674.
- Lee, H., and Reusser, D. A. 2006. The Pacific Coast Ecosystem Information System (PCEIS). US EPA and US Geological Survey Access Database version 1.0.
- McCune, B. 2006. Non-parametric habitat models with automatic interactions. *Journal of Vegetation Science*, 17: 819–830.
- McCune, B., and Mefford, M. J. 2004. HyperNiche. In *Nonparametric Multiplicative Habitat Modeling. Version 1.19*. MjM Software, Gleneden Beach, OR.
- McNyset, K. M., and Blackburn, J. K. 2006. Does GARP really fail miserably? A response to Stockman et al. (2006). *Diversity and Distributions*, 12: 782–786.
- Nelson, W. G., Lee, H., II, Lamberson, J. O., Engle, V., Harwell, L., and Smith, L. M. 2005. Condition of estuaries of the western United States for 1999: a Statistical Summary. EPA Report 620/R-04/200. US EPA, Office of Research and Development.
- Peterson, A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. *The Condor*, 103: 599–605.
- Peterson, A. T., and Vieglais, D. A. 2001. Predicting species invasion using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *BioScience*, 51: 363–371.
- Wiley, E. O., McNyset, K. M., Peterson, A. T., Robins, C. R., and Stewart, A. M. 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography*, 16: 120–127.

## **Chapter 5: Summary**

### ***Uncertainty***

There are different types of uncertainty depending on the research objectives. In chapter 2, I presented an integrated system of schemas to capture biologic and biogeographic information on marine and estuarine species. One source of uncertainty is the design. If the design of the system has not incorporated critical relationships across tables then retrieving information that requires those relationships to be intact will not be possible. For example, there are many different ways to store critical information such as a publication, a species and a location. If the species information is stored in two different tables, (one table for species: publications links and one table for species: location links) it would be possible to identify where a species was located and what publications were associated with the species, but it would not be possible to extract which publication reported that a particular species had been found in a place. Quantifying uncertainty in the design of a system can be accomplished by engaging a broad user community to test the system and identify data storage and retrieval needs, shortcomings and required enhancements. The PCEIS database has been undergoing design enhancements based on user feedback from international, federal, state and academic users over the last several years.

Another form of uncertainty associated with information systems is the correctness of the data. This uncertainty can be minimized by adopting standard operating procedures for data entry, quality assurance checking, data correction, and data analysis. These quality assurance procedures were designed and implemented for PCEIS and are on file at the U. S. Environmental Protection Agency, Western Ecology Branch, Newport, OR.

The uncertainties associated with the Per Capita Invasion Probability model have been detailed in chapter 3 along with the known assumptions and how those assumptions might affect the PCIP calculations (see Table 3-4).

Uncertainty in niche modeling is quantified using several statistics including Akaike's information criterion (AIC) and the AUC statistic. The AIC statistic is used as a measure of goodness of fit and includes a penalty for more complex models. Given two models, the model with the smallest AIC value is usually the better model. The AUC statistic is the measure of sensitivity versus 1-specificity. Given a confusion matrix as shown in Table 5-1, the AUC measures how well the model predicts compared to the random model. The best model fit will optimize between sensitivity and specificity so that errors of commission and omission are minimized.

		<b>Predicted</b>	
		<b>Present</b>	<b>Absent</b>
<b>Observed</b>	<b>Present</b>	Sensitivity	Error of Omission
	<b>Absent</b>	Errors of Commission	Specificity

Table 5-1: Confusion matrix identifying performance measures for niche models.

The AUC statistic is discussed in chapter 4. It is used in chapter 4 to quantifying the accuracy of the best models. It is also used in chapter 4 to compare the models for predicting species occurrences at the habitat scale versus the estuary scale.

### ***Niche Modeling***

Over the course of my studies, my original research plans were modified as new information became available. I had originally planned to use the Genetic Algorithms for

Ruleset Production (GARP) technique to model the distribution of a set of native and nonindigenous marine and estuarine species found in the Northeast Pacific. While GARP had been used in terrestrial environments with some success, it had not been applied to estuarine environments. Adapting GARP to a very broad coastal area with very small estuaries created very large environmental grid layers which caused the software to hang. In addition, a study done by Elith *et al.* in 2006 indicated that GARP did not perform as well as other niche modeling techniques. I decided at that point to look for an alternative niche modeling technique. HyperNiche had many advantages as discussed in chapter 4 so I decided to try it.

Since that time, several other niche modeling techniques have gained in popularity including Generalized Linear Models (GLM), Generalized Additive Models (GAM), Classification Tree Analysis (CTA), Artificial Neural Networks (ANN), Multivariate Adaptive Regression Splines (MARS), Boosting Regression Trees (BRT), and Maximum Entropy to name just a few. This is an active area of research and several recent publications, a few identified in chapter 4, have compared the predictive ability of these techniques.

In addition, a new software package has been developed for the R statistical programming environment called BIOMOD. This package does many different niche modeling techniques and builds an ensemble model that uses the best performing models. Information on it can be found here: <http://www.biochange-lab.eu/wordpress/wp-content/uploads/thuiller-et-al-ecography2009.pdf>. The BIOMOD manual (in the R package download) contains a description of each niche modeling technique available in

BIOMOD with references and valuable information on the advantages/ disadvantages of each of the models.

My future research goals include applying some of these different niche modeling techniques to identify current and future estuarine species distributions relative to changing climate and sea level rise.

### ***Conclusions***

As the primary anthropogenic insults to the environment shift from localized impacts to regional and global scale threats, macroecological approaches of studying multiple species over broad geographical areas are becoming increasingly important research thrusts (Kerr et al., 2007). In response to this need for macroecological tools, a conceptual framework was developed that integrates natural history and distributional data across spatial scales ranging from ocean basins to estuarine tributaries. This framework was then implemented in the Pacific Coast Environmental Information System (*PCEIS*). The multi-dimensional and hierarchical topologies in *PCEIS* provide a standardized, integrated framework to capture complex environmental, biological, and geographical data on marine/estuarine species and near-coastal water bodies. Perhaps its major contribution is that it facilitates matrix overlays between and among the natural history and environmental ranges for individual or multiple species across different geographical scales. Thus, with integrated biological information systems like *PCEIS*, it becomes practical to utilize a suite of different distributional and natural history traits to address a range of research and management questions such as current nonindigenous

species distributions, modeling historical invasion rates to identify potential performance criteria or niche modeling to predict future nonindigenous species distributions.

The per capita invasion model predicts the probability of the introduction of a new invader based on historical rates of invasion and ballast water discharges. This approach provides several advantages for the establishment of performance criteria for ballast water discharges. First, many of the complexities of the invasion processes are captured by using historical invasion rates to derive the per capita probability. Second, it is a relatively simple linear model with identifiable assumptions, which can be modified as new data become available. Third, the model is not restricted to a particular guild or taxa. Finally, new information can be used to refine the model to address some uncertainties. Some analyses indicate that there is not a simple relationship between ballast water discharge and rate of invasion, but there is agreement that without ballast water discharge there would be fewer invaders. Therefore, while this model doesn't capture all the complexities of the invasion process, it does provide an approach that is protective, adaptable, and transparent.

When a nonindigenous species arrives however, niche models are useful tools to predict where that species has the potential to survive. The analysis of NPMR suggests it can predict the distributions of many native and nonindigenous benthic species with a reasonable degree of accuracy at both habitat and estuary scales. This analysis also provided insights into the types of habitat variables needed for predicting future species distributions, including that use of classes for variables such as salinity are often as predictive as quantitative values. However, more detailed analysis with NPMR is

required along with comparisons with other modeling approaches to conclude whether it is the “best” approach to predicting the potential distributions of newly introduced nonindigenous species.

While the thrust of this research was centered around nonindigenous species, the hierarchical framework designed and implemented for the storage and retrieval of biological information at a variety of spatial scales has already been shown to be useful in addressing other macroecological scale issues such as climate change and species at risk from sea-level rise. In addition, habitat variables identified using NPMR will be useful in using niche modeling techniques to model of potential distributions of native species under alternative future environmental conditions.

## Chapter 6: Bibliography

- Åkesson, P. (2007) Boat & ship types. URL <http://www.abc.se/~pa/bld/shiptype.htm>.
- Alexandrov, B., Boltachev, A., Kharchenko, T., Lyashenko, A., Son, M., Tsarenko, P. & Zhukinsky, V. (2007) Trends of aquatic alien species invasions in ukraine. *Aquatic Invasions*, **2**, 215-242.
- Alpine, A.E.a.J.E.C. (1992) Trophic interactions and direct physical effects control phytoplankton biomass and production in an estuary. *Limnol. Oceanogr.*, **37**, 946-955.
- Altman, S. & Whitlatch, R.B. (2007) Effects of small-scale disturbance on invasion success in marine communities. *Journal of Experimental Marine Biology and Ecology*, **342**, 15-29.
- Andersson, A.J., Mackenzie, F.T. & Bates, N.R. (2008) Life on the margin: Implications of ocean acidification on mg-calcite, high latitude and cold-water marine calcifiers. *Marine Ecology Progress Series*, **373**, 265-273.
- Ayres, D.R., Smith, D.L., Zaremba, K., Klohr, S. & Strong, D. (2004) Spread of exotic cordgrasses and hybrids (*spartina sp.*) in the tidal marshes of san francisco bay, california, USA. *Biological Invasions*, **6**, 221-231.
- Bambach, R.K., Bush, A.M. & Erwin, D.H. (2007) Autecology and the filling of ecospace: Key metazoan radiations. *Palaeontology*, **50**, 1-22.
- Barnes, M.A. (2008) Fecundity of the exotic applesnail, *pomacea insularum*. *Journal of the North American Benthological Society*, **27**, 738-745.
- Barry, S., and Elith, J. (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, **43**: 413–423.
- Behrens Yamada, S., Dumbauld, B.R., Kalin, A., Hunt, C.E., Figlar-Barnes, R. & Randall, A. (2005) Growth and persistence of a recent invader *carcinus maenas* in estuaries of the northeastern pacific. *Biological Invasions*, **7**, 309-321.
- Behrens Yamada, S. (2001) *Global invader: The european green crab*. Oregon State University, Corvallis, OR.
- Benech, S.V. (1978) Ocean transport of a community of the grapsid crab *plagusia dentipes* (de haan, 1833). *Crustaceana*, **35**, 104-104.

- Berger, W.H., Adelseck, J.C.G. & Mayer, L.A. (1976) Distribution of carbonate in surface sediments of the pacific ocean. *Journal of Geophysical Research*, **81**.
- Bilio, M. & Niermann, U. (2004) Is the comb jelly really to blame for it all? *Mnemiopsis leidyi* and the ecological concerns about the caspian sea. *Marine Ecology Progress Series*, **269**, 173-183.
- Blake, A.C., Chadwick, D.B., Zirino, A., and Rivera-Duarte, I. (2004) Spatial and temporal variations in copper speciation in San Diego Bay. *Estuaries*, **27**, 437-447.
- Boudouresque, C.F. (1994) Introduced species in european coastal waters: Les especes introduites dans les eaux cotieres d'eupe et de mediterranee: Etat de la question et consequences. In: *Ecosystems Report Series* (ed. C.F. Boudouresque, Briand, F., Nolan, C.), pp. 8-27. Commission European Communities, Brussel.
- Box, G.E.P. & Draper, N.R. (1987) *Empirical model-building and response surfaces*. Wiley.
- Boyd, M.J., Mulligan, T.J., and Shaughnessy, F.J. (2002) Non-indigenous marine species of Humboldt Bay, California. In, p. 118. California Department of Fish and Game.
- Bricker, S., Longstaff, B., Dennison, W., Jones, A., Boicourt, K., Wicks, C. & Woerner, J. (2007) Effects of nutrient enrichment in the nation's estuaries: A decade of change. In: *NOAA Coastal Ocean Program Decision Analysis Series No. 26*, p. 328. NOAA, National Centers for Coastal Ocean Science, Silver Springs, MD.
- Briggs, J.C. (1995) *Global biogeography*. Elsevier Science B. V., Amsterdam, The Netherlands.
- Briggs, J.C. (2007) Marine biogeography and ecology: Invasions and introductions. *Journal of Biogeography*, **34**, 193-198.
- Bush, A.M., Bambach, R.K. & Daley, G.M. (2007) Changes in theoretical ecospace utilization in marine fossil assemblages between the mid-paleozoic and late cenozoic. *Paleobiology*, **33**, 76-97.
- California Sea Grant (1999-2000) California sea grant college program annual report 1999-2000. In, pp. 1-63. University of California, La Jolla, California.
- Cardinale, B.J. & Palmer, M.A. (2002) Disturbance moderates biodiversity-ecosystem function relationships: Experimental evidence from caddisflies in stream mesocosms. *Ecology*, **83**, 1915-1927.

- Carlton, J., J. Thompson, L. Schemel, and F.H. Nichols (1990) Remarkable invasion of san francisco bay (california, USA) by the asian clam *potamocorbula amurensis*. I. Introduction and dispersal. *Marine Ecological Progress Series*, **66**, 81-94.
- Carlton, J.T. (1979) History, biogeography, and ecology of the introduced marine and estuarine invertebrates of the pacific coast of north america. *Ph. D. thesis. University of California, Davis*, 904 pages.
- Carlton, J.T. (1987) Patterns of transoceanic marine biological invasions in the pacific ocean. *Bulletin of Marine Science*, **41**, 452-465.
- Carlton, J.T. (1992) Introduced marine and estuarine mollusks of north america: An end-of-the-20th-century perspective. *Journal of Shellfish Research*, **11**, 489-505.
- Carlton, J.T. (1996a) Marine bioinvasions: The alteration of marine ecosystems by nonindigenous species. *Oceanography*, **9**, 36-43.
- Carlton, J.T. (1996b) Pattern, process, and prediction in marine invasion ecology. *Biological Conservation*, **78**, 97-106.
- Carlton, J.T. (2007) *The light and smith manual: Intertidal invertebrates from central california to oregon*, 4 edn. University of California Press, Berkeley and Los Angeles, California.
- Carlton, J.T. (2009) Deep invasion ecology and the assembly of communities in historical time. *Biological invasions in marine ecosystems. Ecological studies* (ed. by G. Rilov, And Crooks, J.A.), pp. 13-56. Springer-Verlag, Berlin Heidelberg.
- Carlton, J.T. & Geller, J.B. (1993) Ecological roulette: The global transport of nonindigenous marine organisms *Science*, **261**, 78-82.
- Carlton, J.T. & Ruiz, G.M. (2005) Vector science and integrated vector management in bioinvasion ecology: Conceptual frameworks. *Invasive alien species: A new synthesis* (ed. by H.A. Mooney, R.N. Mack, J.A. Mcneely, L.E. Neville, P.J. Schei and J.K. Waage), pp. 36-58. Island Press, Covelo, California.
- Chapman, J.W. (2000) Climate effects on the geography of nonindigenous peracaridan crustacean introductions in estuaries. In: *First National Conference on Marine Bioinvasions* (ed. J. Pederson), pp. 66-80. Massachusetts Institute of Technology.
- Chapman, J.W., Blakeslee, A.M.H., Carlton, J.T. & Bellinger, M.R. (2008) Parsimony dictates a human introduction: On the use of genetic and other data to distinguish between the natural and human-mediated invasion of the european snail *littorina littorea* in north america. *Biological Invasions*, **10**, 131-133.

- Chapman, J.W. & Carlton, J.T. (1991) A test of criteria for introduced species: The global invasion by the isopod *synidotea laevidorsalis* (miers, 1881). *Journal of Crustacean Biology*, **11**, 386-400.
- Chapman, J.W. & Carlton, J.T. (1994) Predicted discoveries of the introduced isopod *synidotea laevidorsalis* (miers, 1881). *Journal of Crustacean Biology* **14**, 700-714.
- Chapman, J.W. & Dorman, J.A. (1975) Diagnosis, systematics, and notes on *grandidierella japonica* (amphipoda: Gammaridea) and its introduction to the pacific coast of the united states. *bulletin of Southern California Academy of Science*, **74**, 104-108.
- Chapman, J.W., Miller, T.W. & Coan, E.V. (2003) Live seafood species as recipes for invasion. *Conservation Biology*, **17**, 1386-1395.
- Clarke, K.R. & Warwick, R.M. (1999) The taxonomic distinctness measure of biodiversity: Weighting of step lengths between hierarchical levels. *Marine Ecology Progress Series*, **184**, 21-29.
- Cohen, A.N. & Carlton, J.T. (1998) Accelerating invasion rate in a highly invaded estuary. *Science*, **279**, 555-558.
- Cohen, A.N., Harris, L.H., Bingham, B.L., Carlton, J.T., Chapman, J.W., Lambert, C.C., Lambert, G., Ljubenkov, J.C., Murray, S.N., Rao, L.C., Reardon, K. & Schwindt, E. (2005) Rapid assessment survey for exotic organisms in southern california bays and harbors, and abundance in port and non-port areas. *Biological Invasions*, **7**, 995-1002.
- Cohen, A.N. & Weinstein, A. (1998) The potential distribution and abundance of zebra mussels in california. In, p. 16. San Francisco Estuary Institute, San Francisco.
- Colautti, R.I., Ricciardi, A., Grigorovich, I.A. & Macisaac, H.J. (2004) Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, **7**, 721-733.
- Colautti, R.I., Grigorovich, I.A., & Macisaac, H.J. (2006) Propagule pressure: A null model for biological invasions. *Biological Invasions*, **8**, 10.
- Coleman, A.W. (1996) DNA analysis methods for recognizing species invasion: The example of codium, and generally applicable methods for algae. *Hydrobiologia*, **326-327**, 29-34.
- Collin, R. (2003) Phylogenetic relationships among calyptraeid gastropods and their implications for the for biogeography of marine speciation. *Systematic Biology*, **52**, 618-640.

- Cordell, J.R., Lawrence, D.J., Ferm, N.C., Tear, L.M., Smith, S.S. & Herwig, R.P. (2009) Factors influencing densities of non-indigenous species in the ballast water of ships arriving at ports in Puget Sound, Washington, United States. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **19**, 322-343.
- Corell, R.W. (2006) Challenges of climate change: An arctic perspective. *AMBIO: Journal of the Human Environment*, **35**, 148-152.
- Cox, C.B. & Moore, P.D. (2005) *Biogeography: An ecological and evolutionary approach*, 7 edn. Blackwell Publishing, Malden, MA.
- Cronin, T.W. & Forward, R.B.J. (1979) Tidal vertical migration: An endogenous rhythm in estuarine crab larvae. *Science*, **205**, 1020-1022.
- Croom, M., Wolotira, R., and Henwood, W. (1995) Marine Region 15: Northeast Pacific. In *A Global Representative System of Marine Protected Areas*, Vol. IV, The Great Barrier Reef Marine Park Authority, pp. 55-106. Ed. by G. Kelleher, C. Bleakley, and S. Wells. The World Bank and IUCN, Washington DC.
- Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*, 1st edn. John Murray, London.
- Drake, J.M. (2004) Allee effects and the risk of biological invasion. *Risk Analysis*, **24**, 795-802.
- Eaton, J.G. & Scheller, R.M. (1996) Effects of climate warming on fish thermal habitat in streams of the United States. *Limnology and Oceanography*, **41**, 1109-1115.
- Eernisse, D., J. (1986) The genus *lepidochitona* Gray, 1921 (Mollusca: Polyplacophora) in the northeastern Pacific Ocean (Oregonian and Californian provinces). *Zoologische Verhandelingen*, **228**, 3-52.
- Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R. J., et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**: 129-151.
- Elton, C.S. (1958) *The ecology of invasions by animals and plants*, Methuen, London.
- Faasse, M.A. & Bayha, K.M. (2006) The ctenophore *mnemiopsis leidyi* A. Agassiz 1865 in coastal waters of the Netherlands: An unrecognized invasion? *Aquatic Invasions*, **1**, 270-277.
- Fauchald, K. & Jumars, P.A. (1979) The diet of worms: A study of polychaete feeding guilds. *Oceanography and Marine Biology: An Annual Review*, **17**, 193-284.

- Fofonoff, P.W., Ruiz, G.M., Steves, B. & Carlton, J.T. (2003a) In ships or on ships? Mechanisms of transfer and invasion for non-native species to the coasts of north america. *Invasive species: Vectors and management strategies* (ed. by G.M. Ruiz and J.T. Carlton), pp. 152-182. Island Press, Washington DC.
- Fofonoff, P.W., Ruiz, G.M., Steves, B., Hines, A.H. & Carlton, J.T. (2003b) National exotic marine and estuarine species information system (nemesis). URL <http://invasions.si.edu/nemesis/>.
- Gaile, G.L. & Willmott, C.J. (2003) Geography in america at the dawn of the 21st century. In, p. 820. Oxford University Press, New York.
- Galil, B.S. (2000) A sea under siege - alien species in the mediterranean. *Biological Invasions*, **2**, 177-186.
- Gascoigne, J. & Lipcius, R.N. (2004) Allee effects in marine systems. *Marine Ecological Progress Series*, **269**, 49-59.
- Gollasch, S. (2006) Assessment of the introduction potential of aquatic alien species in new environments. *Assessment and control of biological invasion risks* (ed. by F. Koike, M.N. Clout, M. Kawamichi, M. De Poorter and K. Iwatsuki), pp. 88-91. Shoukadoh Book Sellers, Kyoto, Japan and IUCN, Gland, Switzerland.
- Goren, M. & Aronov, A. (2002) First record of the indo-pacific parrot fish *scarus ghobban* in the eastern mediterranean. *Cybiurn*, **26**, 239-240.
- Grantham, B., A., Eckert, G.L. & Shanks, A.L. (2003) Dispersal potential of marine invertebrates in diverse habitats. *Ecological Applications*, **13**, S108-S116.
- Grime, J.P. (1973) Competitive exclusion in herbaceous vegetation. *Nature*, **242**, 344-347.
- Grosholz, E.D. (1996) Contrasting rates of spread for introduced species in terrestrial and marine systems. *Ecology*, **77**, 1680-1686.
- Grosholz, E.D., Ruiz, G.M., Dean, C.A., Shirley, K.A., Maron, J.L. & Connors, P.G. (2000) The impacts of a nonindigenous marine predator in a california bay. *Ecology*, **81**, 1206-1224.
- Hall, C.A. (1964) Shallow-water marine climates and molluscan provinces. *Ecology* **45**, 226-234.
- Hedvall, O., Moksnes, P.-O. & Pihl, L. (1998) Active habitat selection by megalopae and juvenile shore crabs *carcinus maenas*: A laboratory study in an annular flume. *Hydrobiologia*, **375/376**, 89-100.

- Hemphill, H. (1890) *Catalogue of North American shells*. Privately Published, San Diego, California.
- Herborg, L., Jerde, C. L., Lodge, D. M., Ruiz, G. M., and MacIsaac, H. (2007) Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications*, 17: 663–674.
- Hodgkin, E.P. (1987) The hydrology of the swan river estuary: Salinity the ecological master factor *The swan river estuary - ecology and management* (ed. by J. John), pp. 34-44. Curtin University Environmental Studies Group, Perth, Western Australia.
- Holeck, K.T., Mills, E.L., Macisaac, H.J., Dochoda, M.R., Colautti, R.I. & Ricciardi, A. (2004) Bridging troubled waters: Biological invasions, transoceanic shipping, and the laurentian great lakes. *BioScience*, **54**, 919-929.
- Howes, D.E., Morris, M. & Zacharias, M. (1999) British columbia estuary mapping system. In, p. 62. Resource Inventory Committee, Province of British Columbia, Victoria, British Columbia.
- IUBS (1958) The venice system for the classification of marine waters according to salinity. In: *International Symposium for the Classification of Brackish Waters*, p. 2. The International Union of Biological Sciences, Venice, Italy.
- Jeschke, J.M.a.S., David L. (2005) Invasion success of vertebrates in europe and north america. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 7198-7202.
- Johnson, L.E., Bossenbroek, J.M. & Kraft, C.E. (2006) Patterns and pathways in the post-establishment spread of non-indigenous aquatic species: The slowing invasion of north american inland lakes by the zebra mussel. *Biological Invasions*, **8**, 475-489.
- Josefson, A.B. & Hansen, J.L.S. (2004) Species richness of benthic macrofauna in danish estuaries and coastal areas. *Global Ecology and Biogeography*, **13**, 273-288.
- Jousson, O., Pawlowski, J., Zaninetti, L., Zechman, F.W., Dini, F., Di Guisepe, G., Woodfield, R., Millar, A. & Meinesz, A. (2000) Invasive alga reaches california. *Nature*, **408**, 157-158.
- Kelleher, G., Bleakley, C. & Wells, S.M. (1995) A global representative system of marine protected areas. In, Washington, DC.
- Kerans, B.L., Dybdahl, M.F., Gangloff, M.M. & Jannot, J.E. (2005) *Potamopyrgus antipodarum*: Distribution, density, and effects on native macroinvertebrate assemblages in the greater yellowstone ecosystem. *Journal of the North American Benthological Society*, **24**, 123-138.

- Kerr, J.T., Kharouba, H.M. & Currie, D.J. (2007) The macroecological contribution of global change solutions. *Science*, **316**, 1581-1584.
- Kimmerer, W.J., E. Gartside, and J.J. Orsi (1994) Predation by an introduced clam as the probable cause of substantial declines in zooplankton in san francisco bay. *Marine Ecological Progress Series*, **113**, 81-93.
- Klautau, M., Russo, C.a.M., Lazoski, C., Boury-Esnault, N., Thorpe, J.P. & Sole-Cava, A.M. (1999) Does cosmopolitanism result from overconservative systematics? A case study usning the marine sponge *chondrilla nucula*. *Evolution*, **53**, 1414-1422.
- Kleypas, J.A., Feely, R.A., Fabry, V.J., Langdon, C., Sabine, C.L. & Robbins, L.L. (2006) Impacts of ocean acidification on coral reefs and other marine calcifiers: A guide for future research, report of a workshop held 19-20 april 2005. In, p. 88. sponsored by NSF, NOAA and the U. S. Geological Survey, St. Petersburg, FL.
- Kolar, C.S. and Lodge, D.M. (2001) Progress in invasion biology: Predicting invaders. *TRENDS in Ecology & Evolution*, **16**, 199-204.
- Koleff, P., Gaston, K.J. & Lennon, J.J. (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367-382.
- Kramer, A.M., Dennis, B., Liebhold, A., M. & Drake, J.M. (2009) The evidence for allee effects. *Population Ecology*, **51**, 341-354.
- Lassuy, D.R. (1995) Introduced species as a factor in extinction and endangerment of native fish species. *American Fisheries Soc. Symp.*, **15**, 391-396.
- Lee, C.E. (1999) Rapid and repeated invasions of fresh water by the copepod *eurytemora affinis*. *Evolution*, **53**, 1423-1434.
- Lee, H. & Brown, C.A. (2009) Classification of regional patterns of environmental drivers and benthic habitats in pacific northwest estuaries. In: (eds. H. Lee and C.A. Brown), p. 298. U.S. EPA, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Newport, Oregon.
- Lee, H., Reusser, D., Ranelletti, M. & Nehmer, R. (2008a) *User guide for the pices nonindigenous species information system templates, vers. Gamma. Newport, or, USA.*
- Lee, H. & Reusser, D.A. (2006) Pacific coast ecosystem information system (*pceis*). Beta version of geospatial database of native and nonindigenous estuarine/coastal species of the northeast pacific and associated landscape data (internal epa and usgs release). In. Office of Research and Development, National Health and Environmental Effects Research Laboratory.

- Lee, H., Reusser, D.A., Olden, J.D., Smith, S.S., Graham, J., Burkett, V., Dukes, J.S., Piorkowski, R.J. & McPhedran, J. (2008b) Integrated monitoring and information systems for managing aquatic invasive species in a changing climate. *Conservation Biology*, **22**, 575-584.
- Lindegarth, M. & Gamfeldt, L. (2005) Comparing categorical and continuous ecological analyses: Effects of "wave exposure" on rocky shores. *Ecology*, **86**, 1346-1357.
- Linnaeus, C. (1758) *Systema naturae per regna tria naturae :Secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. "System of nature through the three kingdoms of nature, according to classes, orders, genera and species, with [generic] characters, [specific] differences, synonyms, places".* 10TH edn.
- Locke, A., Reid, D.M., Sprules, W.G., Carlton, J.T. & Van Leeuwen, H.C. (1991) Effectiveness of mid-ocean exchange in controlling freshwater and coastal zooplankton in ballast water. *Canadian Technical Reports in Fisheries and Aquatic Sciences*.
- Mack, R., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. Bazzaz (2000) Biotic invasions: Causes, epidemiology, global consequences and control. *Ecological Applications*, **10**, 689-710.
- Madden, C.J., Grossman, D.H. & Goodin, K.L. (2005) Coastal and marine systems of north america: Framework for an ecological classification standard: Version ii. In. NatureServe, Arlington, Virginia.
- McCune, B. (2006) Non-parametric habitat models with automatic interactions. *Journal of Vegetation Science*, **17**: 819–830.
- McCune, B., and Mefford, M. J. (2004) HyperNiche. In Nonparametric Multiplicative Habitat Modeling. Version 1.19. MjM Software, Gleneden Beach, OR.
- McNyset, K. M., and Blackburn, J. K. (2006) Does GARP really fail miserably? A response to Stockman et al. (2006). *Diversity and Distributions*, **12**: 782–786.
- Mecklenburg, C.W., Mecklenburg, A.T. & Thorsteinson, L.K. (2002) *Fishes of alaska*, 1 edn. American Fisheries Society, Bethesda, Maryland.
- Mihulka, S., Pysek, P., Martinkova, J. & Jarosik, V. (2006) Invasiveness of oenothera congeneres alien to europe: Jack of all trades, master of invasion? *Perspectives in Plant Ecology, Evolution and systematics*, **8**, 83-96.
- Minchin, D., Gollasch, S., Cohen, A.N., Hewitt, C.L. & Olenin, S. (2009) Chapter 5 characterizing vectors of marine invasion. *Biological invasions in marine*

*ecosystems* (ed. by G. Rilov, And Crooks, J.A.), pp. 109-116. Springer-Verlag, Berlin, Heidelberg.

- Minton, M.S., Verling, E., Miller, A.W. & Ruiz, G.M. (2005) Reducing propagule supply and coastal invasions via ships: Effects of emerging strategies. *Frontiers in Ecology and the Environment*, **3**, 304-308.
- Molnar, J.L., Gamboa, R.L., Revenga, C. & Spalding, M.D. (2008) Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, **6**, 485-492.
- Nehring, S. (2006) Four arguments why so many alien species settle into estuaries, with special reference to the german river elbe. *Helgol. Marine Res.*, **60**, 127-134.
- Nelson, W.G., Lee, H., Lamberson, J.O., Engle, V., Harwell, L. & Smith, M.L. (2005) Condition of estuaries of the western united states for 1999: A statistical summary. In. Office of Research and Development, National Health and Environmental Effects Research Laboratory. EPA/620/R-04/200.
- Nichols, F.H., J.K. Thompson and L.E. Schemel (1990) Remarkable invasion of san francisco bay (san francisco, ca) by the asian clam *potamocorbula amurensis*. Ii. Displacement of a former community. *Marine Ecological Progress Series*, **66**, 95-101.
- Occhipinti-Ambrogi, A. & Savini, D. (2003) Biological invasions as a component of global change in stressed marine ecosystems. *Marine Pollution Bulletin*, **46**, 542-551.
- Open University (2000) *Waves, tides and shallow water processes*, 2 edn. Elsevier Science & Technology Books, Boston.
- Paavola, M., Olenin, S. & Leppakoski, E. (2005) Are invasive species most successful in habitats of low native species richness across european brackish water seas? *Estuarine, Coastal and Shelf Science*, **64**, 738-750.
- Papadopoulos, I., Wooldridge, T.H. & Newman, B.K. (2002) Larval life history strategies of sub-tropical southern african estuarine brachyuran crab and the implications for tidal inlet management. *Wetlands Ecology and Management*, **10**, 249-256.
- Pearcy, W.G. & Schoener, A. (1987) Changes in the marine biota coincident with the 1982-1983 el niño in the northeastern subarctic pacific ocean. *Journal of Geophysical Research*, **92**, 14,417-14,428.
- Pemberton, A.J., Hughes, R.N., Manriquez, P.H. & Bishop, J.D.D. (2003) Efficient utilization of very dilute aquatic sperm: Sperm competition may be more likely

- than sperm limitation when eggs are retained. *The Royal Society biology letters*, **(Suppliment) 270**, S223-S226.
- Peterson, A. T. (2001) Predicting species' geographic distributions based on ecological niche modeling. *The Condor*, 103: 599–605.
- Peterson, A. T., and Vieglais, D. A. (2001) Predicting species invasion using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *BioScience*, 51: 363–371.
- Peterson, W.T., Keister, J.E. & Feinberg, L.R. (2002) The effects of the 1997-99 el niño/ la niña events on hydrography and zooplankton off the central oregon coast. *Progress in Oceanography*, **54**, 381-398.
- Pimentel, D., Lach, L., Zuniga, R., Morrison, D. (2000) Environmental and economic costs associated with non-indigenous species in the united states. *BioScience*, **50**, 53-65.
- Pimentel, D., S. Mcnair, J. Janecka, J. Wightman, C. Simmonds, C. O'connell, E. Wong, L. Russel, J. Zern, T. Aquino, and T. Tsomondo (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems and Environment*, **84**, 1-20.
- Queiroga, H. (1998) Vertical migration and selective tidal stream transport in the megalopa of the crab *carcinus maenas*. *Hydrobiologia*, **375/376**, 137-149.
- Queiroga, H., Costlow, J.D. & Moreira, M.H. (1997) Vertical migration of the crab *carcinus maenas* first zoea in an estuary: Implications for tidal stream transport. *Marine Ecology Progress Series*, **149**, 121-132.
- Rabinowitz, D. (1981) Seven forms of rarity. *The biological aspects of rare plant conservation* (ed. by H. Synge), pp. 205-217. Wiley.
- Reusser, D.A. & Lee II, H. (2008) Predictions for an invaded world: A strategy to predict the distribution of native and non-indigenous species at multiple scales. *ICES Journal of Marine Science*, **65**, 742-745.
- Ribera, M.A. & Boudouresque, C.F. (1995) Introduced marine plants, with special reference to macroalgae: Mechanisms and impact. In: *Progress in phycological Research* (ed. D.J. Chapman), Bristol.
- Roman, J. & Palumbi, S.R. (2004) A global invader at home: Population structure of the green crab, *carcinus maenas*, in europe. *Molecular Ecology*, **13**, 2891-2898.

- Ruiz, G.M. & Carlton, J.T. (2003) Invasion vectors: A conceptual framework for management. *Invasive species: Vectors and management strategies* (ed. by G.M. Ruiz and J.T. Carlton), p. 518. Island Press, Washington.
- Ruiz, G.M., Carlton, J.T., Grosholz, E.D. & Hines, A.H. (1997) Global invasion of marine and estuarine habitats by nonindigenous species: Mechanisms, extent, and consequences. *American Zoologist*, **37**, 621-632.
- Ruiz, G.M., Fofonoff, P., Hines, A.H. & Grosholz, E.D. (1999) Non-indigenous species as stressors in estuarine and marine communities: Assessing invasion impacts and interactions. *Limnology and Oceanography*, **44**, 950-972.
- Ruiz, G.M., Fofonoff, P. W., Carlton, J. T., Wonham, M. J., Hines, A. H. (2000) Invasion of coastal marine communities in north america: Apparent patterns, processes, and biases. *Annual Review of Ecology and Systematics*, **31**, 481-531.
- Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A., Baughman, S., Cabin, R.J., Cohen, J.E., Ellstrand, N.C., Mccauley, D.E., O'neal, P., Parker, I.M., Thompson, J.N. & Weller, S.G. (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305-322.
- Sala, O.E., Chapin Iii, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M. & Wall, D.H. (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770-1774.
- Sanders, N.J., Gotelli, N.J., Heller, N.E. & Gordon, D.M. (2003) Community disassembly by an invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 2474-2477.
- Scheltema, R.S. (1988) Initial evidence for the transport of teleplanic larvae of benthic invertebrates across the east pacific barrier. *Biological Bulletin*, **174**, 145-172.
- Scheltema, R.S. & Williams, I.P. (1983) Long-distance dispersal of planktonic larvae and the biogeography and evolution of some polynesian and western pacific mollusks. *Bulletin of Marine Science*, **33**, 545-565.
- Semmens, B.X., Buhle, E.R., Salomon, A.K. & Pattengill-Semmens, C.V. (2004) A hotspot of non-native marine fishes: Evidence for the aquarium trade as an invasion pathway. *Marine Ecological Progress Series*, **266**, 239-244.
- Sea Grant (2006) Don't release a pest: Freezing is best! Public Service Announcement and animated video for k-12 education. University of California and University of Florida.

- Semmens, B.X., Buhle, E.R., Salomon, A.K. and Pattengill-Semmens, C.V. (2004) A hotspot of non-native marine fishes: Evidence for the aquarium trade as an invasion pathway. *Marine Ecological Progress Series*, **266**, 239-244.
- Smith, S.A., Bell, G. & Bermingham, E. (2004) Cross-cordillera exchange mediated by the panama canal increased species richness of local freshwater fish assemblages. *The Royal Society Biological Science*, **272**, 1889-1896.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdana, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lonbana, A., Lourie, S.A., Martin, K.D., Mcmanus, E., Molnar, J., Recchia, C.A. & Robertson, J. (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience*, **57**, 573-583.
- Stefan, H.G., Fang, X. & Eaton, J.G. (2001) Simulated fish habitat changes in north american lakes in response to projected climate warming. *Transactions of the American Fisheries Society*, **130**, 459-477.
- Sytsma, M.D., Cordell, J.R., Chapman, J.W. & Draheim, R.C. (2004) Lower columbia river aquatic nonindigenous species survey 2001-2004. Final technical report. In, p. 69. Prepared For The United States Coast Guard And The United States Fish And Wildlife Service.
- Thiel, M. (2003) Rafting of benthic macrofauna: Important factors determining the temporal succession of the assemblage on detached macroalgae. *Hydrobiologia*, **503**, 49-57.
- Toft, J.D., Cordell, J.R. & Fields, W.C. (2002) New records of crustaceans (amphipoda, isopoda) in the sacramento/san joaquin delta, california, and application of criteria for introduced species. *Journal of Crustacean Biology*, **22**, 190-200.
- Valentine, J.W., Kaustuv, R. & Jablonski, D. (2002) Carnivore/noncarnivore ratios in northeastern pacific marine gastropods. *Marine Ecological Progress Series*, **228**, 153-163.
- Vitousek, P.M., Mooney, H.A., Lubchenco, J. & Melillo, J.M. (1997) Human domination of earth's ecosystems. *Science*, **277**, 494-499.
- Waples, R.S. (2004) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438-450.
- Watts, P.C., Thorpe, J.P. & Taylor, P.D. (1998) Natural and anthropogenic dispersal mechanisms in the marine environment: A study using cheilostome bryozoa. *Philosophical Transactions of The Royal Society Biological Sciences*, **353**, 453-464.

- Westheide, W. & Schmidt, H. (2003) Cosmopolitan versus cryptic meiofaunal polychaete species: An approach to a molecular taxonomy. *Helgoland Marine Research*, **57**, 1-6.
- Whitfield, P.E., Gardner, T., Vives, S., Gilligan, M.R., Courtenay Jr., W.R., Ray, G.C. & Hare, J.A. (2002) Biological invasion of the indo-pacific lionfish *pterois volitans* along the atlantic coast of north america. *Marine Ecology Progress Series*, **235**, 289-297.
- Wilcove, D.S., D. Rothstein, J. Dubow; A. Phillips and E. Losos. (1998) Quantifying threats to imperiled species in the united states. Assessing the relative importance of habitat destruction, alien species, pollution, overexploitation, and disease. *BioScience*, **48**, 607-615.
- Wiley, E. O., McNyset, K. M., Peterson, A. T., Robins, C. R., and Stewart, A. M. (2003) Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography*, 16: 120–127.
- Wolff, W. (1999) Exotic invaders of the meso-oligohaline zone of estuaries in the netherlands: Why are there so many? *Helgolander Meeresunters*, **52**, 393-400.
- Wolff, W. (2005) Non-indigenous marine and estuarine species in the netherlands. *Zool. Med. Leiden*, **79**, 1-116.
- Zeng, C. & Naylor, E. (1996) Endogenous tidal rhythms of vertical migration in field collected zoea-1 larvae of the shore crab *carcinus maenas*: Implications for ebb tide offshore dispersal. *Marine Ecology Progress Series*, **132**, 71-82.
- Zeng, C., Naylor, E. & Abello, P. (1997) Endogenous control of timing of metamorphosis in megalopae of the shore crab *carcinus maenas*. *Marine Biology*, **128**, 299-3005.

## **Appendix A: Calculation of Coastal Per Capita Invasion Probabilities**

### **Statistical Analysis Using R:**

Analysis was done using the statistical program R (R Development Core Team 2008) because it is widely available and free. The scripts were developed in the text editor Tinn-R (Faria 2009).

The R script below reads foreign ballast water discharge values for ship discharges for each coast in the United States. The calcPCIP function runs a simulation 10,000 times. For each run, a random organism concentration is selected for each ship based on estimates of Minton et al. (2005) sample data. The function then calculates the high, median and low quantiles of PCIP values for each coast. A histogram of the PCIP values is generated for each coast and written to a png file. Code is also provided that uses the PCIP values to generate contour plots indicating the number of invaders per year given organism concentrations and total amount of ballast water discharged. The code to generate contour plots is given, based on a safety factor of 1. If the safety factor is changed, the text locations will need to be modified to plot correctly on the contour plot. Ballast water discharge data are required, along with an organism density file and historical invasion rate to run this code.

Load the library files needed

```
>library(Hmisc)
>library(MASS)
> library(RColorBrewer)
>library(fields)
```

Identify the column definitions for reading in the file

```
>col.defs<-c(rep("numeric",2))
```

Read in the density Values from the Minton graph

```
>ballastDF <- read.csv("DensityVals.csv", colClasses=col.defs)
```

Create the MeanData table from the density values (N=354)

```
>MeanData <- rep(ballastDF$Density, ballastDF$NumShips)
```

The density values data is a table of the number of ships with organism concentrations of a certain value. The MeanData table contains 354 values with approximate organism concentrations extracted from the table in Minton et al. (2005).

Identify the columns and read the ballast water file

```
>col.defs <- c(rep("character", 5),"numeric", rep("character",2), rep("numeric",2))
>allBallast<-read.csv("coastforiegnballast.csv", colClasses = col.defs)
```

Identify the columns and read the number of invaders per coast in

```
>col.defs <- c("character", "numeric")
>ballastInvaders<-read.csv("Ballast_Invaders.csv", colClasses = col.defs)
```

Create a summary table containing the sums for each coast

```
>ballastSums<-tapply(allBallast$DISCHARGE, allBallast$Coast, sum)
>bwSumsdf <- data.frame(ballastSums)
>bwSumsdf$coast<-row.names(bwSumsdf)
>bwSumsdf$annualForeign <- bwSumsdf$ballastSums/3
```

Function CalcPCIP runs 10,000 simulations randomly assigning an organism concentration to each discharge event, summing the total organism concentrations for the run and calculating the PCIP for each run. After all runs are completed, a histogram of the PCIP values is written to a png file and the 2.5, .5, 97.5 quantile values are calculated for the set of PCIP values generated.

```

>calcPCIP <-function(bInfo, bData) {
#Define a dataframe to contain the calculated values
>RandRun=data.frame(MeanConc=rep(NA,10000), TotalProp=rep(NA, 10000),
PCIP=rep(NA,10000))
#Run calculations 10,000 times to get a normal distribution PCIP
  >for (i in 1:10000) {
    # Get a random array of concentrations for all
    >Conc <- sample(MeanData, size=bInfo$shipCount, replace=TRUE)
    # Calculate the mean concentration for this run and store it
    >RandRun$MeanConc[i] <- mean(Conc)
    # Calculate the number of organisms for each ship for this run
    >Prop<- round(Conc*bData$DISCHARGE,0)
    # Calculate the total organism inoculation from all ships for this run and store it
    >RandRun$TotalProp[i] <- sum(Prop,na.rm=TRUE)
    # Calculate the annual per capita probability
    >RandRun$PCIP[i] <- bInfo$TotBWInvaders/(RandRun$TotalProp[i]/3)
  }
# Create a file name and write out the data generated by the Random Run
> csvFile <- paste(bInfo$Coast, "RanRun", ".csv", sep="")
>write.csv(RandRun, file = csvFile, append =FALSE,na = "NA",row.names = TRUE)
#Calculate the lower, median and upper bound of the annual per capita invasion
probability
>tmp <- quantile(RandRun$PCIP, probs=c(0.025,.5, 0.975))
>bInfo$medianPCIP <- tmp[2]
>bInfo$hbPCIP <- tmp[3]
>bInfo$lbPCIP <- tmp[1]
# Create a histogram of all calculated annual PCIPs, write the graphic to a png file
# Create the name of the file to be written
>pngFile <- paste(bInfo$Coast, ".png",sep="")
# Open the png file for writing
>png(pngFile)
# Create a title for the Histogram based on the name of the coast being processed
>hTitle <- paste("Histogram of", bInfo$Coast, "Coast Annual\nPer Capita Invasion
Probabilities")
>hist(RandRun$PCIP, font=2, font.lab=2,main=hTitle, xlab="Per Capita Invasion
Probabilites")
# add lines for the lower, median and upper quantile PCIP values on the histogram
>abline(v=bInfo$lbPCIP, col="red")
>abline(v=bInfo$hbPCIP, col="red")
>abline(v=bInfo$medianPCIP, col="blue")
# Close the png file
>dev.off()
# Return the dataframe of information for the coast to the calling routine
>return(bInfo)
}
## END FUNCTION

```

Create a unique list of Coasts in allBallast

```
>coastlst<-unique(allBallast$Coast)
>allBallastLst <- unique(allBallast$Coast)
>recCount <- length(coastlst)
```

Create a dataframe to hold the information calculated for each coast

```
>CoastInfo=data.frame(CoastName=rep(NA,recCount), shipCount=rep(NA,
recCount),
TotFB=rep(NA,recCount),TotAnnFB=rep(NA,recCount),TotBWInvaders=rep(NA,recCount),lbPCIP=rep(NA,recCount), medianPCIP=rep(NA,recCount),
hbPCIP=rep(NA,recCount))
```

Loop through all the coasts calling the PCIP function

```
>for(j in 1: length(coastlst)){
  ## Get the name of the current coast ##
  >CoastInfo$CoastName[j] <- coastlst[j]
  ## Get the records for the current coast ##
  >CoastData <- allBallast[allBallast$Coast %in% CoastInfo$CoastName[j],]
  ## Get the count of the number of records for the current coast ##
  >CoastInfo$shipCount[j] <-length(CoastData$Coast)
  ## Get only the records that have foreign ballast discharge ##
  >FBCoastData <- CoastData[CoastData$DISCHARGE > 0,]
  ## Calculate the total foreign ballast
  >CoastInfo$TotFB[j] <- sum(FBCoastData$DISCHARGE)
  ## Calculate the total annual foreign ballast
  >CoastInfo$TotAnnFB[j] <- sum(FBCoastData$DISCHARGE)/3
  ## Store the ballast water invaders per year for a coast
  >CoastInfo$TotBWInvaders[j] <- ballastInvaders$invpyr[ballastInvaders$Coast
  %in% CoastInfo$CoastName[j]]
  ## Calculate the PCIP values for the Coast
  >CoastInfo[j,] <- calcPCIP(CoastInfo[j,], CoastData)
}>
```

Write out the results for each coast to a CSV file

```
>write.csv(CoastInfo, file = "RegionalPCIP.csv", append = FALSE, na = "NA",
row.names = TRUE)
```

Build a vector of values for 3D plot- Organism Concentrations 0 - 1000

```
>conc<-c(seq(0.0001, 0.001, by = 0.00001),
```

```
seq(0.0011, 0.01, by = 0.0001),
seq(0.011, 0.1, by=0.001),
seq(0.11, 1, by = 0.01),
seq(1.1, 100, by = .1),
seq(101, 1000, by=1))
```

Build a vector of discharge values from 0 to 30,000,000

```
>discharge<-seq(0,30000000, length=6001)
```

Get the stored value for the upper quantile for West Coast

```
>probinv<- CoastInfo$hbPCIP[3]
```

Set the safety factor

```
>safetyFactor<- 1
```

Create a matrix to contain the number of invaders given a concentration and discharge

```
>num_invaders=matrix(data=NA, nrow=6001, ncol=2251, byrow="T",
dimnames=NULL)
```

Fill the matrix looping through each concentration and discharge value

```
>for (i in 1:6001) {
  >for (j in 1:2251) {
    >num_invaders[i,j]=probinv*discharge[i]*conc[j]*safetyFactor
  }
}>
```

Make a plot of the probability Matrix

```
# Set the Breaks for the Plot
>brk <- c(0,0.0001, 0.001, 0.01, 0.1,1, 10)

#Create a color palette of Red Yellow Green with six different colors
>myPal<-brewer.pal(6,"RdYlGn")

# Identify the png the plot will be written to
>png("WCRegionalPCIPJan2010.png")
>par(xaxs="i", family="serif")
>iTitle <- paste("Predicted Number of Invaders Per Year \n Given Per Capita
Invasion Probability of", format(probinv,scientific = TRUE, digits=4), " \n West
Coast")
```

```

>image.plot(x=discharge, y=log(conc,10),z=num_invaders, axes=F, breaks=brk,
font.lab=2, col=rev(myPal),lab.breaks=names(brk), xlab="Foreign Ballast Water
Discharge in Million m3 Per Year", ylab="Concentration (organism/m3)",
main=iTitle, add=FALSE, legend.shrink=100)
#label the axes
>axis(1, at=c(0, 1000000, 5000000,10000000,15000000,20000000,25000000,
30000000), labels = c(0, '1', '5','10','15','20','25','30'),font=2, las=1)
>axis (2, at = c(-4,-3, -2, -1, 0, 1, 2, 3), labels = c('0.0001','0.001', '0.01', '0.1', '1', '10',
'100', '1000'), font=2, las=1)

# Label the plot with the number of NIS per year for each color
>text(27900000,-2.3, "# NIS < 1 per 10,000 years", cex=1, col="black", font=2,
adj=c(1,0))
>text(27900000,-.4, "# NIS < 1 per 1000 years", cex=1, col="black", font=2,
adj=c(1,0.5))
>text(27900000,.55, "# NIS < 1 per 100 years", cex=1, col="black" font=2,
adj=c(1,0.5))
>text(27900000, 1.5, "# NIS < 1 per 10 years", cex=1, col="black", font=2,
adj=c(1,0))
>text(27900000,2.50, "# NIS < 1 per year", cex=1, col="black", font=2, adj=c(1,0))
>dev.off()

```