Title
Targeting Conservation Activities: Cost-effective Wetlands Restoration in the Central Valley of California

Author
Newbold, Stephen C.

Publication Date
2002

Peer reviewed

Targeting Conservation Activities:
Cost-effective Wetlands Restoration in the Central Valley of California

by

Stephen Carlisle Newbold

B.S. (University of Florida) 1995
M.S. (University of California, Davis) 1999

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Ecology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA
DAVIS

Approved:

Committee in charge

2002

UMI Number: 3062252

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Those areas that are inundated or saturated by surface or ground water at a frequency and duration sufficient to support, and that under normal circumstances do support, a prevalence of vegetation typically adapted for life in saturated soil conditions. Wetlands generally include swamps, marshes, bogs, and similar areas.

Jurisdictional definition of wetlands
40 CFR 232.2(r)

Wetlands are places where heavy equipment is likely to churn up mud even after the weather has been dry for some time.

William Lewis, Jr., wetland scientist
In response to Senator Lauch Faircloth (R-NC),
who asked for a practical definition of wetlands
that could be given to a farmer (Williams 2001).

All models are wrong. Some models are useful.

George Box

Errors using inadequate data are much less than those using no data at all.

Charles Babbage

If you want to convince me of something, Mrs. Landingham, give me numbers.

Jed Bartlet
NBC's *West Wing*
Final episode, season two

Table of Contents

Preface v
Chapter 1 – Introduction 1
1.1 Public and private benefits and costs of wetlands conservation 1
1.2 Outline of the dissertation 5
1.3 Consider a square landscape 7
1.4 Conservation biology and reserve design 11
1.5 Wetlands assessment methods 18
Chapter 2 – Figuring the effects of configuration 23
2.1 Optimization and land use change 23
2.2 Methods 26
2.2.1 Generating a hypothetical watershed 26
2.2.2 Specifying production functions for ecosystem services 27
2.2.3 Integrating benefits and costs into an optimization framework 31
2.2.4 Comparing site selection algorithms 32
2.2.5 Comparing management goals 34
2.3 Results 38
2.3.1 Selection algorithms 39
2.3.2 Different management goals 42
2.3.3 Monte Carlo and sensitivity analyses 45
2.4 Conclusions 51
Chapter 3 – Wetlands as habitat 54
3.1 Wetlands and species protection 54
3.2 Focusing on mallards 57
3.2.1 The data 57
3.2.2 Regression models of bird abundances 63
3.2.3 Management significance of the spatial effects 87
3.2.4 Do mallards look up from their ponds? 91
3.3 Regression models for other bird species 94
3.4 An ideal free regression model 99
Parameter 109
3.5 Conclusions 110
Chapter 4 – Wetlands and water quality 112
4.1 Pollution abatement in constructed and natural wetlands 112
4.2 The Central Valley water quality model 114
4.2.1 Model overview 115
4.2.2 The data 118
4.2.3 Water and mass balance calculations 124
4.2.4 Routing runoff through the landscape 133
4.2.5 Results 138
Chapter 5 – The Costs of Wetlands Restoration 150
5.1 Opportunity costs 150

5.2 The costs of constructing restored wetlands 157
Chapter 6 – An Integrated Model for Prioritizing Wetlands Restoration Activities 160
6.1 The reserve site selection problem revisited 160
6.2 The integrated model 163
6.3 Solving the wetlands restoration problem 165
6.4 Two case studies 168
6.4.1 The watersheds case study 169
6.4.2 The WRP case study 183
Chapter 7 – Conclusions 188
7.1 Summary of the main results 188
7.2 Potential improvements and directions for future research 194
References 197
Appendix – Benefits and costs, optimization, and the Wetlands Reserve Program 210

Preface

This dissertation is the culmination of six years of graduate study, two and a half of which were devoted specifically to the research described in these pages. The research was supported by an EPA grant, under the Science to Achieve Results (STAR) program, and a grant from the University of California Water Resources Center.

I have written the dissertation as if I was solely responsible for all of the work, but of course that is not true. I had a superb major advisor and dissertation committee to help me throughout, and I benefited from many conversations with fellow students and friends as well. First and foremost, Marca Weinberg was the best advisor a graduate student could hope for. I must have spent more time talking with Marca in my first year – about classes, research ideas, and environmental economics in general – than most spend with their advisors in their entire five- (or six- or seven-) year tenure as PhD students. And she kept it up, even after she left U.C. Davis for greener pastures. It was also a great pleasure to have Jim Wilen on my dissertation committee. If some small part of his ability to think clearly and produce useful insights about virtually any question he is presented with shows up anywhere in my future work, then he will have done me a great service. I only wish that I could have spent even more time with Jim. John Eadie is an expert on all things related to waterfowl, and he provided an invaluable check on my statistical models of mallard abundances. John is involved in real-world wetlands management in the Central Valley of California, and I benefited greatly from tagging along to a few meetings where decision-makers in state, federal, and non-governmental organizations were actually trying to figure out how best to manage wetlands in the region. These were enlightening experiences, and I hope that similar opportunities will arise in the future. I also always appreciated John's boundless enthusiasm for research. Finally, Greg Pasternack never failed to provide very insightful advice on the parts of my dissertation research related to hydrology, which is his area of expertise. Greg was also always able to draw my attention to the more practical aspects of my models and results. But of course, the invaluable mentorship and advice from my dissertation committee notwithstanding, all remaining mistakes in this dissertation are mine and mine alone.

My dissertation committee was not my only source of support. It was during time spent with friends and classmates, even more so than with my professors and faculty advisors, that I experienced most acutely the hints of discovery, the feeling of progress, and the great satisfaction of learning new things. And these things, I came to realize, are what can make a career doing research FUN. The friends and classmates who made my time in graduate school such a rewarding experience include: James Eaves, Catherine Hickey, Sandeep Mohapatra, Brian Paciotti, Neil Pelkey, Dan Pollock, Deborah Salon, Marty Smith, and Matt Zafonte. Thanks so much to all of you.

And finally, without the unconditional love and support from my family, nothing I have accomplished to date could have been possible, and nothing I hope to accomplish in the future is even conceivable. Thanks Mom, Dad, John, and Anje. I love you all.

Chapter 1 – Introduction

This dissertation is about the public benefits that wetlands can provide and how those benefits depend on where wetlands are located in the landscape. The two broad objectives of the research were (1) to further our understanding of the role that landscape configuration plays in the provision of ecosystem services from wetlands, and (2) to enhance our practical ability to account for spatial effects and tradeoffs between competing environmental goals when evaluating, designing, and implementing wetlands policies. I have addressed only a few key aspects of these broad objectives, but the methods, and – with some qualifications – many of the conclusions, could be applied to other aspects of wetlands policy as well.

1.1 Public and private benefits and costs of wetlands conservation

If one considers public policy related to wetlands as a collective effort to maximize the net benefits they provide to society, then measuring the benefits and costs of wetlands conservation becomes a central task for evaluating, designing, and implementing wetlands policies. I have not attempted a comprehensive benefit-cost analysis of wetlands conservation, but such an ideal provides a useful context for the research described here. So to start, I offer the following categorization of the benefits and costs of wetlands conservation.

Public benefits are benefits that accrue to the public at large. These have all of the attributes of "public goods" – many individuals in society can benefit from them simultaneously without affecting the benefits other individuals receive (unlike private goods, which can only be owned, used, or consumed by one person at a time). Public benefits from wetlands conservation include ecosystem services and environmental amenities. Ecosystem services are those benefits provided to society through the normal workings of wetland ecosystem functions. The three classes of ecosystem services that wetlands are most often said to provide are: habitat benefits, water quality benefits, and flood control benefits. Habitat benefits arise from the role that wetlands play in supporting populations of particular wildlife species of concern, or in maintaining biodiversity in general. Water quality benefits arise from the physical, chemical, and biological processes in wetlands that serve to reduce the concentration of pollutants entering downstream water bodies. Flood control benefits arise from the effects wetlands have on hydrologic processes in a watershed. Wetlands can function as temporary detention basins for floodwaters during large storm events, thereby reducing flood damages to downstream areas. Environmental amenities are those benefits provided to society from the structural characteristics (as opposed to functions) of wetlands.[1] Environmental amenities depend more on wetlands "just being there" than actually "doing something." Environmental amenities from wetlands include use values, such as opportunities for water sports, hiking, camping, bird watching, etc., as well as non-use values, such as the existence value of open space and a "pristine" or "natural" environment.

Private benefits are benefits that accrue to the private owners of land on which wetlands are found. These could include profits from operating duck clubs, the sale of crops, or timber harvested in wetlands.

[1] You will not find this definition of environmental amenities in the environmental economics literature. There the term is generally used as a catch-all for virtually any type of non-market environmental benefit. However, for the present purposes the definition is useful, to distinguish between ecosystem services, which are often best measured using a production function approach that integrates methods from economics and ecology, and other types of public environmental benefits, many of which can be measured using hedonic, travel cost, or contingent valuation methods alone.

Public costs of wetlands preservation would include any disamenities (negative externalities) from wetlands, e.g. nuisance and increased disease transmission from mosquitoes, and increased damages to neighboring landowners' crops from wildlife associated with wetlands. In the case of wetlands restoration, public costs would also include any public funds used to finance their construction and maintenance.

Private costs of wetlands preservation include the opportunity costs of the land (the present value of the entire stream of expected future benefits from the land in its highest economic use), and disamenities from wetlands borne by the landowners themselves, such as damage to a landowner's own crops from wildlife associated with wetlands. In the case of wetlands restoration, private costs would also include the private landowner's share (if any) of construction and maintenance costs. The net benefits wetlands provide to society are the sum of public and private benefits less the sum of public and private costs.

On the benefit side, this research focused on two of the major classes of wetland ecosystem services: habitat benefits and water quality benefits. In the environmental economics literature ecosystem services have been largely ignored (beyond the admission of their existence), and in the ecology literature they have not often been treated in ways useful for policy evaluation. On the cost side, this research focused on the private costs of wetlands restoration, which can be approximated by the market value of the land that is to be restored to wetlands, plus the costs of converting the land from its present use back to a fully functioning wetland ecosystem.

The research described here proceeded in two fairly distinct stages. First, I developed production functions for habitat and water quality ecosystem services from wetlands. Second, I integrated the production functions, along with estimates of restoration costs, into an optimization-based decision framework that can determine the configuration of restoration activities that maximizes the expected levels of wetland ecosystem services. The methods, and to some degree the models themselves, could be transferred to different areas and policy contexts, but they were developed and applied specifically to wetlands restoration in the Central Valley of California, in hopes of gaining insights that might be immediately relevant to decision makers in the region.

In its fully realized incarnation the modeling framework described here would include production functions for all public benefits that wetlands could conceivably provide in a particular study area. This might include models of habitat relationships for all species of concern in the region, models of transport and kinetics for all water constituents that could have an impact on human health or ecosystem integrity, hydrologic models of the potential for flooding at all locations in the study area, economic models of the amenity value of wetlands, and so on. Because my focus was on particular aspects of wetlands policy evaluation (namely, spatial effects and multiple objectives), I considered only a small subset of the ecosystem services that make up the complete package of public benefits from wetlands. This research focused on two specific examples of wetland ecosystem services – the provision of habitat for mallards in the breeding season and the attenuation of nutrients from non-point source runoff. These ecosystem services provide substantial contrast between required modeling techniques and management implications, so this strategy should produce an informative case study of multi-objective land use decision-making in the context of wetlands conservation.

1.2 Outline of the dissertation

The dissertation is organized as follows. The next section presents a stylized wetlands restoration scenario to show how spatial effects can be important for decision-making. In

Sections 1.4 and 1.5, I briefly review the reserve site selection literature and the wetland assessment literature, to provide context for the models and applications in Chapters 2 through 6.

Chapter 2 presents a numerical model of multi-objective land use decision-making that can account for spatial effects. The simple story is that when spatial effects are important one should consider them explicitly when making decisions regarding land use; otherwise such decisions will likely be sub-optimal. The model's usefulness comes from its ability to provide a means for investigating in a rigorous way some of the most important features of land use decisions. The model described in Chapter 2 can be used to measure the importance of spatial effects, determine the distribution of wetlands restoration activities that maximizes the provision of ecosystem services, and compare outcomes under different management strategies.[2]

In Chapter 2, I illustrate the framework and demonstrate its utility using stylized representations of ecosystem services. Chapters 3 and 4 describe more realistic models of ecosystem services from wetlands in the Central Valley of California. Chapter 3 presents regression models that relate bird abundances to the distribution of land use in the study area, with special attention paid to mallards. The models are best understood as a description of how birds in a population of fixed size would distribute themselves across the landscape. Some important assumptions are required to use the model for predicting the total population size that the landscape can support, which will often be the policy endpoint of most interest. I maintained these assumptions for applications described in later sections of Chapter 3 and in later chapters, but in section 3.4 I discuss the assumptions and present a

[2] Chapter 2 is based largely on Newbold (2002).

more general model that can be used to predict the equilibrium population size as a function of landscape configuration based on a static analysis of species abundances.

Chapter 4 describes a spatially distributed hydrologic simulation model that I developed to estimate nitrogen and phosphorus loads to rivers and streams from non-point source runoff, and to predict reductions in nutrient loads that would result from restoring wetlands at different locations in the landscape. I applied the model to the entire Central Valley, and Section 4.2.5 presents baseline results from the model.

Chapter 5 addresses the cost side of wetlands conservation. The cost of wetlands protection will consist solely of opportunity costs associated with alternative uses of the land, which can be measured to a reasonable first approximation by the market value of nearby agricultural or urban parcels. This research focused on wetlands restoration decisions, the costs of which will include opportunity costs (the market value of the parcel itself) plus construction and maintenance costs associated with converting the parcel from its present use to wetlands. Chapter 5 presents estimates of both opportunity costs and construction costs for wetlands restoration in the Central Valley.

Chapter 6 describes an integrated optimization model that can target wetlands restoration activities and investigate the importance of spatial effects and the magnitude of tradeoffs between wetland ecosystem services in the Central Valley. To develop the integrated model, I combined the production functions for ecosystem services from Chapters 3 and 4 and the estimates of restoration costs from Chapter 5 into a numerical optimization model, similar to the one presented in Chapter 2. The optimization model can determine the configuration of wetlands restoration activities that maximizes some weighted combination of mallard abundance in the breeding season and nutrient attenuation. The model can be applied to either small watersheds within the Central Valley, or the entire

Central Valley but with substantial constraints imposed. The scope of application is limited because of the difficulties involved in the optimization modeling itself – I will have something to say about this at various points in the dissertation. The model was applied to several case studies in the Central Valley, and results from the case studies are presented in Section 6.4.

In Chapter 7, I conclude with a recap of some of the main conclusions from the research and a brief discussion of the limitations of the models to suggest possible directions for future research.

1.3 Consider a square landscape

One of the main themes of this dissertation is the importance of spatial effects. By "spatial effects," I mean the effects of land use change on some policy endpoint of interest (this research focused on environmental endpoints) above and beyond what can be attributed to changes in the total area of each land use type alone. A controlled experiment that could be used to investigate such effects might involve holding constant the amount of each land use type in a region, moving parcels of land around, and then measuring the endpoints of interest in the re-configured landscapes. In such a controlled experiment the differences in the endpoints would solely be due to spatial effects, because the total amount of each land use type was held constant across the trials. Researchers generally do not have the luxury of performing such experiments on the large scales that are of most interest, so they must rely on other techniques for investigating spatial effects, including statistical modeling and simulation modeling, both of which were used extensively in this research.

But even if spatial effects can be measured... So what? Are the substantively important? Spatial effects could be important in two ways. First, if policy outcomes are (at least in part) a function of spatial effects, then only by fully accounting for them can policy be as effective as possible. Predicting the outcomes of different policy options is easier when all of the important causal factors are understood. The second way that spatial effects could be important, beyond this standard motivation for policy-relevant research, is that spatial effects can complicate land use decisions by making the outcomes of otherwise independent choices interdependent. In this section I present a stylized example of a wetlands restoration decision scenario to show how this can happen.

Consider a square landscape with seven parcels of land, as in Figure 1.1. The five

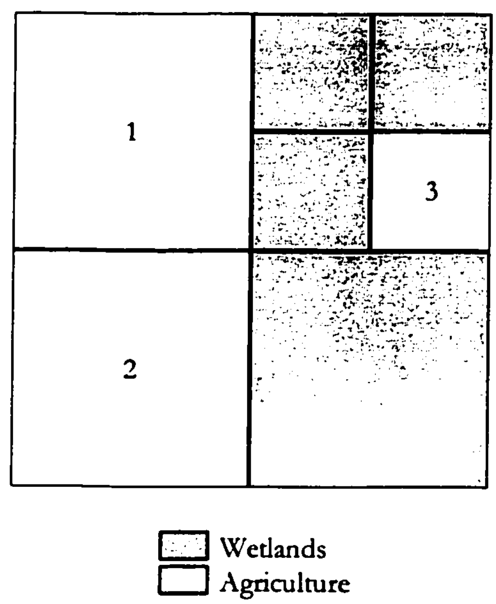


Figure 1.1 – A square landscape.

gray parcels are wetlands and the three white parcels are agriculture. Now consider a manager whose task is to increase the population of a particular wildlife species of concern. The species relies on wetlands and is negatively affected by agriculture. Specifically, the sustainable population size of the species is given by the following equation:

$$P = A - 2E \quad (1.1)$$

In equation 1.1, P is the population size of the species, A is the total area of wetlands in the landscape, and E is the total length of edge shared between wetlands and agriculture. Next, assume for simplicity that all three parcels of agriculture cost the same and that the manager has sufficient funds to purchase and restore any two of the three parcels to wetlands. If one unit of area is defined as the size of the smallest parcel, and one unit of edge is defined as the length of the smallest parcel, then initially $A = 7$ and $E = 7$, which means $P = 7 - 2 \times 7 = -7$. The fact that the sustainable population size is negative means that the landscape is currently a "sink" for the species; any individuals that migrate into the area will not survive.

One strategy the manager could use to choose which two parcels to restore to wetlands is: (1) calculate the sustainable population size if each agriculture parcel is restored individually, (2) choose from the three agriculture parcels the one that yields the greatest increase in population size, (3) calculate the sustainable population size if each of the two remaining parcels is restored individually, and (4) choose from the remaining two parcels the one that yields the greatest increase in population size. By using this decision strategy, the manager would make the best decision at each iteration, choosing first the parcel that yields the greatest benefit to the species, and second the parcel that yields the greatest benefit given that the first was already chosen. However, this iterative strategy will not yield the largest increase in population size possible. The parcel that yields the largest increase in population size initially is parcel 3. If parcel 3 were restored to wetlands, then $A = 8$ and $E = 4$, which means $P = 8 - 2 \times 4 = 0$, an increase of 7. If either parcel 1 or 2 were restored to wetlands first, then $A = 11$ and $E = 7$, which means $P = 11 - 2 \times 7 = -3$, an increase of only 4. If parcel 3 were restored first the manager would be indifferent between parcels 1 and 2 in the second round. If either parcel 1 or 2 were restored, in addition to parcel 3, then $A = 12$ and

$E = 4$, which means $P = 12 - 2 \times 4 = 4$, an overall increase of 11. However, if instead of the iterative decision strategy described above all possible combinations of parcels were considered, the manager would see that choosing parcels 1 and 2, not 1 and 3, would yield the greatest overall increase. If parcels 1 and 2 were restored, then $A = 15$ and $E = 3$, which means $P = 15 - 2 \times 3 = 9$, an overall increase of 16. The two site selection strategies are summarized in Table 1.1.

Table 1.1 – The performance of two sites selection strategies for the stylized wetlands restoration scenario.

	A	E	P
Baseline	7	7	-7
Alternative 1: (iterative strategy)			
Choose parcel 3...	8	4	0
then choose parcel 1 or 2	12	4	4
Alternative 2: (optimizing strategy)			
Choose parcels 1 and 2	15	3	9

Compare the above scenario to one in which there are no spatial effects. If $P = A$ (or any increasing function of A alone), then the iterative strategy would yield the same outcome as one where all possible combinations were considered. In the above case, parcels 1 and 2 would be chosen using either strategy. It is the spatial effects in the scenario described above that make the iterative strategy sub-optimal, and this occurred because the benefits of restoring parcel 1 depended on whether or not parcel 2 was restored, and vice versa. The benefits of choosing either of these parcels are interdependent (or "endogenous") in the presence of spatial effects; they would be independent (or "exogenous") in the absence of spatial effects. In this way, spatial effects can make an otherwise simple decision situation complicated. This example had only three options from which to choose, but many real-world situations may have dozens, hundreds, even thousands of options. In such cases it would not be possible to consider all combinations – the computational demands would exceed the capacity of even the fastest computers – so the design of site selection strategies that perform well in the presence of spatial effects is an important endeavor.

The example presented above was meant to provide motivation for the chapters to come. To provide context, in the next section I briefly review some of the related research that has appeared in the scientific literature. Two strands of the literature that are especially relevant to this dissertation are (1) the reserve site selection literature, which deals with methods for choosing sites for a network of nature reserves to protect biodiversity in general, and (2) the wetlands assessment literature, which deals with methods for measuring the functions and values of wetlands in particular.

1.4 Conservation biology and reserve design

The design of nature reserves is one of the central themes of conservation biology. In *The Theory of Island Biogeography*, MacArthur and Wilson (1967) provided a conceptual foundation for describing relationships between landscape configuration and the diversity of species. MacArthur and Wilson focused on explaining differences in diversity and turnover rates of species on islands of varying size and isolation, but the theory found immediate application to the design of nature reserves on mainlands as well (Diamond 1975), and soon the so-called "SLOSS debate" ensued (e.g. Simberloff and Abele 1982; Soulé and Simberloff 1986). SLOSS refers to "single large or several small" nature reserves, and the debate was over which was better for protecting biodiversity. The debate produced no definitive answer, except that the preferred design will depend largely on particular circumstances (Prendergast et al. 1999; Simberloff 1988; Saunders et al. 1991). Nevertheless, island biogeography theory and the large amount of research it inspired, including the SLOSS debate and more recent contributions (e.g. Hubbell 2001), continue to provide some of the key foundations for research on the spatial aspects of reserve design (Noss et al. 1997; Margules and Pressey 2000; Veech 2000). However, much of the current research on reserve design is not really spatial at all.

Current research on reserve design uses mathematical programming techniques to select sites for inclusion in a network of protected areas. Using these methods reserves are "designed" by choosing sites from a candidate set of sites, where each site in the candidate set is of predetermined size, shape, and location. Thus, the body of literature reporting applications of these techniques is aptly named "the reserve site selection" literature. Any particular reserve site will protect only some of the species in need of protection; it is the network as a whole that is intended to protect as many species as possible, and reserve site selection strives to choose the smallest set of sites required to meet that goal. The selection of sites is based primarily on the species that are known to occur on the sites, not on any theoretically determined equilibrium level of diversity that the network of reserves would support (which is what reserve site selection based on island biogeography theory might do). In the standard reserve site selection problem, the only information necessary for designing an optimal network of nature reserves is a complete description of the current distribution of species in the landscape. All considerations of species' interactions with their environment and with each other are ignored. Furthermore, no spatial information is used in the standard formulation. Therefore, standard reserve site selection techniques represent more of a natural history approach than an ecological approach to designing nature reserves.

The standard reserve site selection problem can be stated mathematically as follows:

$$\underset{x_j}{Max}\left[\sum_{i=1}^{T} y_i\right] \tag{1.2a}$$

$$y_i = \begin{cases} 1, & \sum_{j=1}^{N} x_j z_{ij} \geq k_i \\ 0, & \sum_{j=1}^{N} x_j z_{ij} < k_i \end{cases} \tag{1.2b}$$

$$\sum_{j=1}^{N} x_j \leq C \tag{1.2c}$$

In expressions 1.2a-c, T is the total number of species that occur on all sites under consideration for the reserve network, $z_{ij} = 1$ if species i occurs on site j and 0 otherwise, x_j = 1 if site j is included in the reserve network and 0 otherwise (the x_j's are the choice variables), C is the maximum number of sites that could be included in the network, and k_i is the minimum number of sites in the reserve network on which species i must occur for it to be adequately protected. Therefore, $y_i = 1$ if species i is adequately protected (if it occurs on at least k_i sites in the reserve network) and 0 otherwise. In words, expressions 1.2a-c say: Maximize the number of species protected by choosing up to C sites for inclusion in a network of nature reserves.

Despite the low level of ecological detail in the standard formulation, the reserve site selection literature has made an exceedingly useful contribution to conservation biology simply through the structure it imposes on conservation problems. The objective must be defined clearly, the relevant ecology must be summarized in numerical form, and the constraints must be accounted for explicitly. The value of the formality of this approach cannot be overstated, especially for conservation biology, a discipline with positive and normative components that are sometimes hard to separate. However, just because a problem can be written down in a form similar to expressions 1.2a-c does not mean that the solution immediately follows. This formulation is merely a problem statement and a description of the characteristics that the solution must have (the solution will maximize a specific objective, and satisfy certain constraints). Actually finding the solution (or good candidate solutions) requires mathematical programming techniques.

The note by Underhill (1994) provides a convenient entry into the reserve site selection literature. Underhill pointed out that a class of heuristics conservation biologists often used to solve reserve site selection problems, so-called "greedy algorithms," could not guarantee optimal solutions in general. Greedy algorithms work by breaking the problem down into a sequence of smaller problems and finding optimal solutions to each in turn. The iterative strategy used to select parcels for wetlands restoration in the previous section was a greedy algorithm, and Underhill's point was similar to the point of that section: greedy algorithms sometimes fail when benefits are endogenous. However, the reason that benefits are endogenous in the standard reserve site selection problem is different than the reason they were endogenous in the stylized wetlands restoration scenario in the previous section.

Benefits are endogenous in the standard reserve site selection problem because of complementarities between the sets of species that occur on each site. The example given by Underhill (1994) is worth reproducing here to make this point clear. Consider five sites on which species occur according to Table 1.2 (the values in the table are the z_{ij}'s in expressions 1.2a-c). To protect all species, a greedy algorithm – one that chooses sites in decreasing order of the number of new species added to the protected set at each step – would choose site 1 first, then 2, and then 3. However, it is easy to see that only sites 2 and 3 are required to protect all eight species. Notice that the endogeneity of benefits, and therefore the failure (or underperformance) of the greedy algorithm, in this example does not stem from spatial effects. Again, spatial relationships are not accounted for in the standard reserve site selection problem.[3]

Table 1.2 – Hypothetical reserve site selection problem from Underhill (1994).

		Species							
		1	2	3	4	5	6	7	8
Sites	1	0	0	1	1	1	1	1	0
	2	1	1	1	1	0	0	0	0
	3	0	0	0	0	1	1	1	1
	4	1	1	1	0	0	0	0	0
	5	0	0	0	0	0	1	1	1

In spite of Underhill's note, greedy algorithms and other heuristics are still used for reserve site selection problems. This is because optimizing algorithms (those that guarantee an optimal solution), such as branch and bound, are not feasible for many real-world problems, and because heuristics often perform quite well for the standard reserve site selection problem even when they cannot guarantee optimality (Pressey et al. 1996).[4] Less is known about the performance of greedy algorithms when spatial effects are important. Some preliminary results relevant to this issue are presented in Chapter 2, where a greedy algorithm is compared to an optimizing algorithm on a site selection problem with spatial effects.

Many reserve site selection applications have appeared in the conservation biology literature in recent years. Some have focused on the general performance of alternative heuristics and maximization criteria (e.g. Lomolino 1994; Camm et al. 1996; Pressey et al.

[3] Though see Williams and RaVelle (1997) for a discussion of spatial optimizing models from the operations research literature that could be applied to the reserve site selection problem.
[4] However, the literature has not investigated in much depth the conditions that will cause a greedy algorithm to perform poorly for the standard reserve site selection problem. In what kinds of environments, or in the presence of what kinds of ecological relationships, will species be distributed across sites in a way similar to Table 1.1?

1996; Csuti et al. 1997; Pressey et al. 1997; Polasky et al. 2001); some have been more concerned with applications to specific conservation problems (e.g. Davis et al. 1996; Williams et al. 1996); some have investigated the importance of incorporating economic costs into the standard formulation (e.g. Walpole and Sinden 1997; Ando et al. 1998; van Langevelde et al. 2000; Polasky et al. 2001a); and some have presented general models and discussed their variants (e.g. Wright et al. 1983; Williams and ReVelle 1997; Margules and Pressey 2000). In a recent contribution, Polasky and Solow (2001) addressed the value of information in reserve site selection. They applied a Bayesian framework to a stylized site selection problem and demonstrated that the value of alternative survey strategies depends on, among other things, the number of sites that can be selected (*C* in expression 1.2c).[5]

Though not usually associated with the reserve site selection literature, Hof and others (Hof and Flather 1996; Hof and Bevers 1998; Hof et al. 1999), and Nevo and Garcia (1996) have also used optimization techniques to address conservation problems. These researchers generally focus on issues of ecological management on a smaller scale than most applications in the reserve site selection literature. They also generally use more explicitly specified relationships between species population dynamics and management decisions, often including spatial effects.

And it is only by including spatial effects in optimization algorithms that the reserve site selection literature can begin to incorporate some of the ideas from island biogeography theory and the SLOSS debate. It is not immediately clear how these two approaches can be completely reconciled, but advances in reserve site selection methods may be possible by incorporating some of the general results from island biogeography theory and its intellectual

---

[5] See the biodiversity bibliography compiled by Steve Polasky for a comprehensive listing of the reserve site selection literature and more: http://www.apec.umn.edu/faculty/spolasky/Biobib.html.

offspring (such as metapopulation ecology (Hanski 1999)). It was not my main goal, but the research in this dissertation makes some steps in that direction. The model presented in Chapter 2 can incorporate various types of spatial effects, which means that it could be used to "derive" rules for reserve design – for individuals with particular movement or dispersal characteristics, for populations with particular migratory patterns, for communities with particular types of interactions, etc. I will have more to say about how this dissertation fills some of the gaps in the reserve site selection literature in Chapter 6, after the details of the numerical models are presented in Chapters 2 through 5.

The reserve site selection literature provides the background for the general class of optimization models used in this dissertation, but as yet no applications to wetlands conservation have appeared in the literature. However, there have been a number of methods developed for assessing wetland functions and values, usually for the purpose of informing permitting decisions for wetlands conversions.[6] These methods could, in theory, provide the foundation of an integrated framework for prioritizing wetlands conservation activities, so a brief review is warranted.

1.5 Wetlands assessment methods

Most methods for assessing wetland functions and values can be put into one of two categories: field-based methods and GIS-based methods. Field-based methods rely on extensive field visits for all existing wetlands or potential restoration sites under consideration. Data are collected on hydrologic conditions, soils, and flora and fauna at the sites, and are used to make inferences about the functions the wetlands might perform.

---

[6] Much has been written about the distinction between functions and values in the wetlands literature (Mitsch and Gosselink 1993, ch. 15; Brinson and Rheinhardt 1998; Lewis 2001). Wetland functions are all of the physical, chemical, and biological processes that occur in wetlands. Wetland values are the benefits provided to society by those functions.

GIS-based methods rely on remotely sensed data, and therefore can apply an assessment model more consistently across large spatial extents. The range of wetland functions that can be inferred from remotely sensed data is limited, however. In general, field-based methods use many types of data on a few sites, while GIS-based methods use a few types of data on many sites.

The Wetland Evaluation Technique (WET) was one of the earliest field-based methods developed for assessing wetland functions. It was created by researchers under contract with the Federal Highway Administration in the early 1980's, and has been updated several times since (Adamus et al. 1987, Adamus et al. 1991; National Research Council 1995). To implement WET, an analyst answers a series of questions on the hydrology, vegetation, soils, location, and other characteristics of the wetland, and then applies a rule-based protocol that assigns rank-ordered values (low, medium, or high) to the probability that the wetland will perform a given function. The functions addressed by WET are: groundwater exchange, flood flow alteration, sediment stabilization, sediment/toxicant retention, nutrient removal/transformation, production export, aquatic diversity/abundance, wildlife diversity/abundance, and recreation and uniqueness/heritage.

The hydrogeomorphic (HGM) approach is another field-based wetlands assessment method. The HGM approach was developed by researchers under contract with the U.S. Army Corps of Engineers, and was designed for determining mitigation ratios required to offset wetlands conversions (Brinson 1993; Brinson and Rheinhardt 1996). The first step in applying the HGM approach is to identify unaltered reference sites for a particular class of wetlands in a particular region (e.g. wet pine flats in southeastern North Carolina, as in Rheinhardt et al. 1997). Next, biotic and abiotic characteristics are measured in the reference wetlands and are designated as benchmarks against which the functions of impacted wetlands are then compared. Like WET, the HGM approach is based on a set of comprehensive conceptual models of the relationships between the structure and functions of wetlands. The conceptual models, however, generally do not include clear links between structural attributes and levels of functions. For example, Rheinhardt et al. (1997) developed indices for "maintenance of the characteristic hydrologic regime," "maintenance of characteristic nutrient and elemental cycling processes," "maintenance of characteristic plant community," and "maintenance of characteristic physiognomic structure." Each of these factors was measured by an index of the form:

$$Index = \{V_1 + [V_2 + (V_3 + V_4)/2]/2\}/2 \tag{1.3}$$

where V_1, V_2, V_3, and V_4 are normalized indices of individual structural attributes. In Rheinhardt et al. (1997), V_1 in the index for "maintaining characteristic plant communities" was the ratio of the percent cover of herbaceous canopy in the assessed site relative to that in the reference sites. In general, the functional form of the index – the manner in which the V's are combined – depends on the perceived relative importance of each of the structural attributes included in the model for each factor. The end result is intended as an indicator of the degree to which functions have been affected by alterations, and is usually expressed on a scale from 0 to 1. Field-based methods such as WET and the HGM approach are useful when rapid assessments of a few sites are required, but because they only deal in indices and relative values they are less useful for looking explicitly at the importance of spatial effects and tradeoffs between competing environmental goals, which were the two main objectives of this research.

Apart from approaches based on field observations, other researchers have developed GIS-based methods for assessing wetland functions (or identifying areas suitable for wetlands restoration). Russell et al. (1997) developed a GIS-based model that ranks sites for wetlands restoration based on their predicted average wetness (a function of the topography of the watershed based on a digital elevation model), their size, and their proximity to existing riparian vegetation. The model was applied to a watershed in southern California. O'Neill et al. (1997) applied a similar model to the Upper Arkansas River basin in Colorado. Their model added a measure of floodplain disturbance, which was thought to affect restoration potential. Like many field methods, GIS methods generally use a rule-based approach to comparing sites. For example, Figure 1.2 shows the rule-based model used to rank sites in Russel et al. (1997), which gives highest priority to sites that meet certain thresholds for wetness values and proximity to existing riparian areas. The GIS-based model developed by Cedfeldt et al. (2000) incorporates all three of the major classes of ecosystem services that wetlands are often said to provide: habitat, water quality, and flood

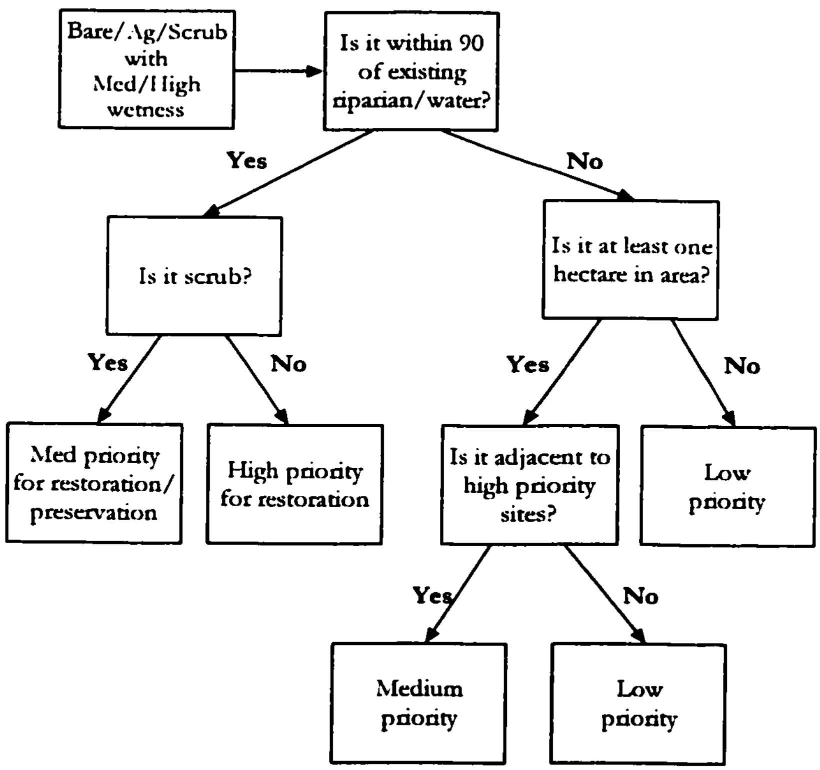


__Figure 1.2__ – Rule-based model for ranking sites in a GIS wetlands assessment method from Russell et al. (2000).

control benefits. Their model also incorporates more variables related to wetland functions than most GIS-based models, including some that address spatial interactions between wetlands. For example, the model recognizes that wetlands downstream of other wetlands would be less effective in attenuating floods because the upstream wetlands would have already stored much of the floodwater during any given storm. Under these conditions, two wetlands in a row would not do twice the work of a single wetland.[7] In lieu of empirical models of wetland functions, the Cedfeldt model relies on a rule-based approach similar to Russel et al. (1997). The GIS-based model presented by McAllister et al. (2000) has the most in common with the models used for the present research. The McAllister model prioritizes sites for restoration based on their expected contribution to flood reduction and the expected cost of restoring them. However, unlike the models in this dissertation, the McAllister model is suitable only for fairly large-scale applications. It can identify watersheds within which wetlands restoration for flood control should be the most cost-effective, but it cannot prioritize parcels within watersheds.

The main objective of this research was not to develop a new wetlands assessment method, nor was it to integrate island biogeography theory with the reserve site selection literature. The purpose of reviewing these two strands of the literature was to provide context for the approach and methods described in the following chapters. The objective of this research was to develop methods for investigating the importance of spatial effects and tradeoffs between objectives, for land use decisions in general and for wetlands conservation in particular. In the next chapter I present a general numerical model of land use decision-making, and I demonstrate the utility of the model through several simulation exercises. The

[7] We will see another example of this kind of spatial effect in Chapter 4, but in relation to water quality benefits.

general model in Chapter 2 provides the foundation for the empirical research described in Chapters 3 through 6.

Chapter 2 – Figuring the effects of configuration

2.1 Optimization and land use change

This chapter describes a framework for analyzing the potential environmental and economic impacts of land use changes. Land use changes can have profound effects on the quality of the environment. The conversion of natural lands to agriculture and urban uses can increase species extinction rates (Boulinier et al. 2001), affect landscape hydrological processes (Knox 2001), and even exacerbate climate change (Dale 1997). This chapter focuses on wetlands restoration, but most of the ideas will apply to other types of land that deliver public benefits as well.

The framework described in this chapter is designed to help decision makers and analysts (1) explicitly consider multiple environmental objectives when contemplating or analyzing management decisions on a watershed or landscape scale, and (2) investigate the importance of spatial effects for the outcomes of those decisions. Implementation of the framework requires two distinct modeling phases. Phase 1 involves describing relationships between the extent and configuration of wetlands and other land use types to the provision of different classes of valued ecosystem services. Phase 2 involves incorporating the functions estimated in Phase 1 into a spatial optimization model that can compare the expected environmental impacts and economic costs of alternative management strategies.

In this chapter, I demonstrate the utility of the framework with stylized functions for two classes of ecosystem services: water quality and habitat quality. I use the spatial optimization model to compare the expected environmental impacts and economic costs of various management goals and strategies in a hypothetical watershed. "Management goals" refers to the intentions of the manager with respect to the ecosystem services considered – whether the manager wants to minimize nutrient loads, maximize species abundance, or some combination of the two. "Management strategies" refers to the means by which the manager attempts to achieve the specified goal – the algorithm used to select the sites for wetlands restoration. These strategies could range from the simplest of heuristics (e.g. maximize the area of wetlands in the watershed), to more sophisticated site selection rules that incorporate, to greater or lesser degrees, information on the specific processes that affect ecosystem functions.

An optimization framework can be used to analyze the impacts of the goals and strategies on management outcomes by quantifying: (1) the effects of maximizing one ecosystem service without considering the others, and (2) the differences between those strategies that account for spatial effects and those that do not. Given a fixed restoration budget, a set of sites chosen to maximize water quality will generally result in lower habitat quality than a set of sites chosen specifically to maximize habitat quality, and vice versa. Also, the benefits of restoring any particular site to wetlands may depend on which other sites are also restored. If this is the case, then choosing sites iteratively will generally provide lower levels of ecosystem services than if all sites were chosen simultaneously (Underhill 1994; and Section 1.3).

To motivate the site selection problem, consider a hypothetical watershed that contains ten parcels of agricultural land offered for inclusion in an easement program (I will call these the "sites"). The watershed manager has resources sufficient to purchase and restore only five of the ten sites and must decide which five to choose. One way the manager could approach the problem would be to estimate for each site the expected environmental benefits were it to be purchased and restored, estimate for each site the costs of restoration, and then undertake restoration on those sites with the highest benefit-to-cost ratios. If the benefits and costs could be determined a priori, then this strategy would guarantee the greatest level of environmental benefits possible given the budget constraint (Martello and Toth 1990; Hyman and Leibowitz 2000).[8] However, if the benefits of restoring a given site depend not only upon the characteristics of the site itself, but also upon the nature of the surrounding landscape, then the benefits could be affected by the decisions regarding restoration of the other sites. In this case the benefits would be endogenous with respect to the other restoration decisions, and, as shown in Section 1.3, if benefits are endogenous then the simple approach described above will not be sufficient to guarantee the optimal solution.[9]

In the remaining sections of this chapter, I describe the framework that provides the foundation for the entire dissertation and discuss its important features using a stylized example. To begin, think of the hypothetical manager mentioned earlier and imagine that the manager's goal is to increase as much as possible the provision of ecosystem services that wetlands provide. As in the real world, the hypothetical manager will be faced with

---

[8] There is a common misunderstanding regarding the arithmetic of benefit-cost analysis that is worth mentioning here. It may seem that choosing sites in decreasing order of net benefits would yield the greatest total benefit, but this is not the case. To see this, consider the situation where the benefits and costs of restoring each of the ten sites in the hypothetical watershed are: $2,000 (benefits) and $1,000 (costs) for site one, and $750 (benefits) and $250 (costs) for the other nine sites. If sites were chosen in decreasing order of net benefits, site one would be chosen first, since its net benefits are $1,000 and the net benefits of each of the other sites are $500 each. If sites were chosen in decreasing order of their benefits-cost ratio, sites would be chosen from sites two through ten first, since each of their benefit-cost ratios are three and the benefit-cost ratio for site one is two. Now assume that the manager has exactly $1,000 to spend. By the net benefits criterion, site one alone would be chosen for a total net benefit of $1,000. By the benefit-cost criterion, any four of sites two through ten would be chosen, for a total net benefit of $2,000. This example shows that the selection algorithm based on the benefit-cost ratio performs better. In addition, an algorithm based on net benefits would require benefits and costs to be measured in the same units; the algorithm based on the benefit-cost ratio does not.

[9] Furthermore, when the endogeneity is due to spatial effects, the strength of the spatial effects relative to the site-specific effects will influence the degree to which a greedy algorithm will under-perform relative to an optimizing algorithm. Since it will not always be possible to apply optimizing algorithms to these kinds of problems, it will be important to have an idea of how strong the spatial effects are since this may provide the only indication of how sub-optimal the site selection heuristics are likely to be.

unavoidable tradeoffs. Not only are there usually more potential projects to fund than money to fund them, but the manager also must decide how to prioritize multiple environmental objectives.

2.2 Methods

2.2.1 Generating a hypothetical watershed

The watershed used to illustrate the following example contains three land use types: urban, agriculture, and wetlands. The watershed is a 25×25 grid of square cells, indexed by $i = 1, 2, \ldots 625$. A river runs through the middle of the watershed, and twenty of the agriculture cells have been offered for enrollment in an easement program. The manager's job is to choose a set of sites from the twenty offered and restore them to wetlands to maximize some combination of ecosystem services. The watershed, which is pictured in Figure 2.1, was created by: (1) randomly assigning seven square urban areas of 36, 16, 9, 4, 4, 4, and 4 cells, (2) randomly assigning 20 potential restoration sites – cells offered for enrollment in the easement program, (3) randomly assigning 50 wetland cells, (4) assigning the remaining 478 cells to agriculture with no potential for restoration, and (5) putting a river in the middle of the landscape, between cell columns 12 and 13. Most of the results presented in this chapter are based on Figure 2.1, but I also present results from a Monte Carlo analysis based on 100 different simulated landscapes, each created according to the five steps listed above.

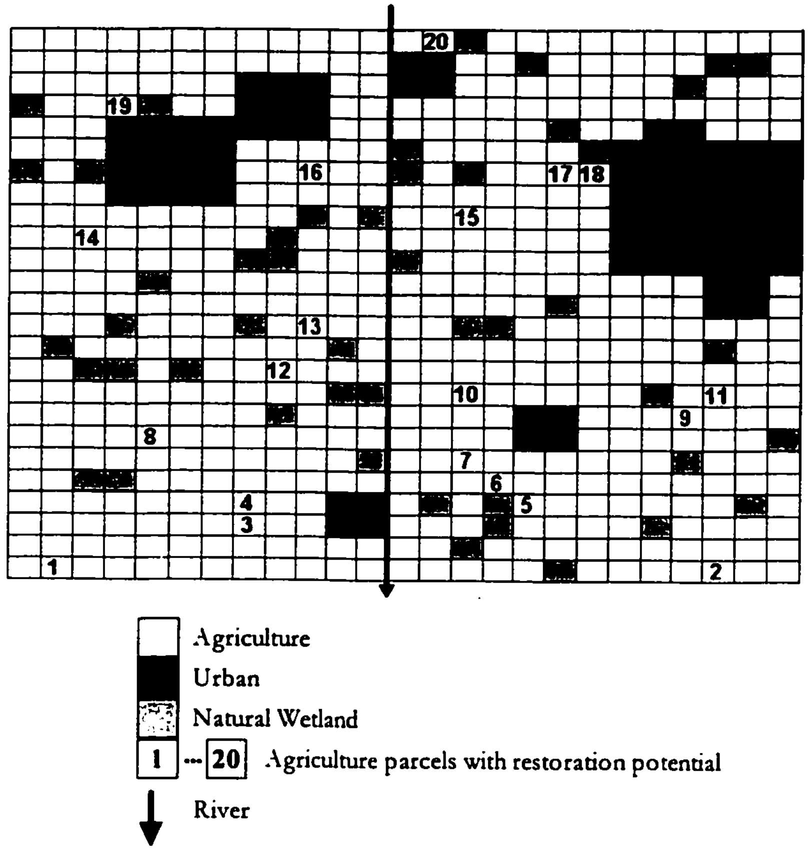


Figure 2.1 – The hypothetical watershed, baseline conditions.

2.2.2 Specifying production functions for ecosystem services

To implement an optimization approach to selecting sites, the manager must solicit from the relevant experts mathematical descriptions of the ecosystem services of interest in the watershed.[10] In this example the manager is concerned with water quality and habitat quality only. According to hydrologists the manager has consulted, the best current understanding

[10] In this chapter I am glossing over what may be one of the most difficult steps in this process in the real world. The relevant experts may not have all of the information that the manager wants, or they may not be able (or willing) to distill the best available scientific understanding of the key hydrologic and ecological processes into a form that the manager can use for decision making. I am assuming here that the manager has completed this difficult task and can now get to the business of prioritizing sites for restoration. In Chapters 2 and 3, I address the problem of specifying production functions for ecosystem services directly, for an application of the framework to the Central Valley of California.

of the determinants of water quality in the watershed can be described by the following equation for N, the nutrient load to the river:

$$N = \sum_{i=1}^{625} n_i (1-\theta)^{h_i} \tag{2.1}$$

In equation 2.1, n_i is the annual nutrient load in direct runoff from cell i, which depends on the cell's land use type; h_i is the distance of cell i from the river's edge; and θ is the fraction of the nutrient load in surface or subsurface flow that is dissipated per unit distance (the edge length of a cell) as it makes its way towards the river. Soranno et al. (1996) used this functional form to investigate phosphorus loads to lakes from non-point source runoff in Wisconsin. In a real application one could parameterize and use a simple model like equation 2.1, or one could use a more sophisticated simulation model, such as AGNPS (Bosch et al. 1998; Grunwald and Norton 2000) or SWAT (Brown and Hollis 1996; Krysanova et al. 1998). Chapter 4 presents an intermediate alternative: a spatially distributed hydrologic simulation model developed using a mass balance approach. In the meantime I will use equation 2.1 to illustrate the framework.

According to ecologists the manager has consulted, the best current understanding of the determinants of habitat quality in the watershed can be described by the following equation for S, the total expected abundance for an umbrella species in the watershed:

$$S = \sum_{i=1}^{625} \rho_i \left(1 - \prod_{q=1}^{625} \left(1 - \frac{\rho_q / \rho^{\max}}{d_{iq}^{\eta}} \right) \right) \tag{2.2}$$

Hof and Bevers (1998) used a function similar to equation 2.2, which is based on a model of habitat connectivity for species that exhibit random radial dispersal. The model assumes that the probability of a cell being connected to other cells (i.e., the probability that the species can successfully disperse from one cell to another) is the joint probability that it is connected to any (not all) of the other cells in the landscape, and that the individual connectivity probabilities are independent (Hof and Bevers 1998; pp 18-20). In equation 2.2, ρ_i is the suitability of cell i, the abundance on the cell if it were completely connected to other cells; q is an alternative index for the cells; $\rho^{\max}$ is the abundance on a completely connected cell of the most suitable land use type; and η determines the rate at which connectivity decreases with distance. As in the water quality model, the unit of distance is the edge length of a cell. The probability that cell i is connected to another cell q is $\frac{\rho_q / \rho^{\max}}{d_{iq}^{\eta}}$, therefore connectivity is a function of the distance between the cells and the suitability of the cells (normalized by the maximum possible suitability). The joint probability that cell i is not connected to any other cell is $\prod_{q=1}^{625}\left(1-\frac{\rho_q / \rho^{\max}}{d_{iq}^{\eta}}\right)$. Therefore, the probability that cell i is connected (the cell's "connectivity") is $1-\prod_{q=1}^{625}\left(1-\frac{\rho_q / \rho^{\max}}{d_{iq}^{\eta}}\right)$. The abundance on cell i is the product of its suitability and connectivity, and the total abundance in the watershed is the sum of the abundances on the individual cells. Again, the model used in a real application would depend on the species of concern and the data and technical resources available. Chapter 3 presents regression models that relate mallard abundances to the distribution of land use in the Central Valley of California. In the meantime I will use equation 2.2 to illustrate the framework.

To provide the maximum possible level of ecosystem services with a limited budget, the manager must also consider the costs of restoration. In the hypothetical watershed, the cost of purchasing and restoring a site is a function of its distance from urban areas, according to the following equation:

$$c_i = \alpha \sum_{q=1}^{625} \frac{U_q}{d_{iq}^{\varphi}} \tag{2.3}$$

In equation 2.3, $U_q = 1$ for urban cells and 0 for all other cells, φ determines the rate at which cost drops off with distance, and α is a scaling parameter. Figure 2.2 shows the cost surface for the watershed in Figure 2.1, as determined by equation 2.3. Chapter 5 describes

Figure 2.2 – The cost surface for the hypothetical watershed.

estimates of wetlands restoration costs based on county assessor data, which provide information on parcel values in most counties in the study area. However, the data were not adequate for estimating a spatially explicit model of land values. In this chapter, I use the cost structure in equation 2.3 partly in recognition of the fact that costs as well as benefits could be influenced by spatial effects, which could also be important for land use decision-making.

2.2.3 Integrating benefits and costs into an optimization framework

Equations 2.1 through 2.3 are sufficient to calculate the benefits and costs of restoration in the hypothetical watershed. Next, they must be related to the restoration decisions and combined in an optimization framework. The manager can affect the levels of ecosystem services through her (limited) control over the configuration of the landscape by purchasing and restoring some of the agriculture parcels offered for enrollment in the easement program. The optimization problem the manager must solve is:

$$\underset{m_1,\ldots,m_{625}}{Max}\left[W_N N + W_S S\right] \tag{2.4a}$$

Subject to:

$$N = N_0 - \sum_{i=1}^{625} n_i (1-\theta)^{b_i} \tag{2.4b}$$

$$S = \sum_{i=1}^{625} \rho_i \left(1 - \prod_{q=1}^{625}\left(1 - \frac{\rho_q / \rho^{\max}}{d_{iq}^{\eta}}\right)\right) - S_0 \tag{2.4c}$$

$$\rho_i = \rho_i^0 + m_i\left(\rho^W - \rho_i^0\right) \tag{2.4d}$$

$$n_i = n_i^0 + m_i\left(n^W - n_i^0\right) \tag{2.4e}$$

$$c_i = \alpha \sum_{q=1}^{625} \frac{U_q}{d_{iq}^{\varphi}} \tag{2.4f}$$

$$\sum_{i=1}^{625} m_i c_i \leq Budget \tag{2.4g}$$

In the objective function (expression 2.4a), W_N and W_S are weighting factors that depend on the relative value the manager places on improvements in water quality and habitat quality, and m_1,... m_{625} are binary [0,1] choice variables that indicate whether or not each site is restored (m_i must equal 0 for all cells not offered for enrollment in the easement program).

In equations 2.4b and 2.4c, N and S are the improvements in water quality and habitat quality in the watershed, and N_0 and S_0 are the baseline levels of nitrogen load and species abundance.[11] In equations 2.4d and 2.4e, which link the manager's restoration decisions to the final levels of ecosystem services, n_i^0 and ρ_i^0 are the initial nutrient loading and habitat suitability values for site i, n^W and ρ^W are the values for wetland cells, and n_i and ρ_i are the final values for site i. If $m_i = 0$, then nothing changes, but if the site is chosen for restoration, then $m_i = 1$, $n_i = n^W$, and $\rho_i = \rho^W$. Equation 2.4f is the cost function, and expression 2.4g is the budget constraint. Notice that while the benefits of management are endogenous, the costs are not. This is by assumption only, and will not generally be true in the real world. If land values are affected by surrounding land use types (Doss and Taff 1996; Geoghegan et al. 1997), then restoring a wetland in a particular location may change the costs of restoring nearby sites to wetlands later. I did not address this complication, but many of the insights regarding the importance of spatial effects on the benefits side that come out of this exercise will apply to the cost side as well.

2.2.4 Comparing site selection algorithms

In the real world it will not always be feasible to implement a systematic optimization strategy for selecting sites. This would be the case if policy or economic constraints were such that there were very little flexibility in determining which sites could be selected, or if the manager had no or little information on the factors that influence the provision of ecosystem services in the watershed. In these situations, opportunistic management is the

[11] Notice that I have changed slightly the definition of N and S from equations 2.1 and 2.2, which first introduced the production functions. There they were the total values for the ecosystem services; now they are the increases above the baseline values.

best that could be done.[12] However, in cases where there is sufficient flexibility and information to support a systematic optimization approach, it still could be difficult to find the optimal solution simply because of the combinatorial nature of the problem. Consider trying to find an optimal set of restoration sites by checking all affordable combinations of sites. If 100 sites were up for consideration and any 50 were affordable, then there would be more than 10^{29} different possible combinations. At one million checks per second it would take over three million, billion years to enumerate all of the combinations. Little wonder there is significant interest in admittedly sub-optimal site selection algorithms (Csuti et al. 1997; Pressey et al. 1997). This is perhaps an unfair yardstick against which to measure real-world management, but one still might want to know just how sub-optimal the alternative heuristics would be. What are the environmental costs of choosing sites merely to maximize total wetland area, or using some other rule of thumb, as opposed to choosing sites optimally? Put another way, what are the potential environmental benefits of incorporating more realistic ecological and hydrological information into the decision-making process and using more effective strategies for selecting sites?

To answer these questions for the hypothetical watershed, I compared three site selection algorithms: (1) a simple heuristic that maximizes wetland area, (2) an iterative site selection algorithm where at each step the site with the largest benefit-cost ratio is chosen, and (3) an optimizing algorithm that enumerates all affordable sets of sites. None of these algorithms is intended to perfectly mimic the decision process used by managers in the real world, but as a group they are intended to span the range of possible decision strategies. It is clear that managers in the real world take more into consideration than merely the area of

[12] Nevertheless, there are some easy improvements that managers could incorporate into their site selection methods that follow directly from this research. See the Appendix for an application of some simple benefit-cost and optimization concepts to the California Wetlands Reserve Program.

wetlands preserved or restored. However, insofar as "no-net-loss" goals drive policy making – and at least at the national level the no-net-loss rhetoric is pervasive – one could think of the first algorithm as a first approximation to a "naïve" manager's decision rule.[13] The second algorithm is a much more generous characterization of a watershed manager's decision strategy. The manager is assumed to know exactly the form and parameter values of the production functions for ecosystem services, and to choose, one at a time, sites that yield the greatest increase in those services. The third algorithm is an ideal case. The manager has full information about the production functions *and* the foresight and technical know-how to choose the truly optimal set of sites. It is clear that the former – full information about the production functions – is a very generous assumption, but in the sections that follow I will show that the latter – choosing the best set of sites – could be difficult as well, even if the production functions were known with certainty.

2.2.5 Comparing management goals

I have greatly simplified the multi-objective nature of real world environmental problems in this example – only two objectives are built in – but it still retains the basic feature. The manager might *like* to maximize both of the environmental benefits, but of course she cannot maximize more than one at a time. However, with both functions specified explicitly one can investigate the tradeoffs that result under a variety of management goals.

The utility of an optimization framework for analyzing tradeoffs comes from its ability to answer the following kinds of questions: What is the maximum reduction in the nutrient load to the river achievable given the budget (N^{max}), and what level of species

[13] In other words, "a busy manager," or "a manager with limited information," or "a manager who is tightly constrained by bureaucratic protocol," etc.

abundance results from that solution ($S \mid N^{max}$)? What is the maximum species abundance achievable given the budget (S^{max}), and what level of nutrient load results from that solution ($N \mid S^{max}$)? The differences between N^{max} and $N \mid S^{max}$, and S^{max} and $S \mid N^{max}$ are the maximum possible tradeoffs associated with the different management goals.

In a continuous world there would be an infinite number of solutions between these two extremes, and the family of these solutions would define a production possibility frontier, or "PPF," for ecosystem services in the watershed. In the present example, with only two ecosystem services under consideration, the PPF is the curve on a graph of $-N$ vs. S that indicates the maximum water quality attainable with a fixed budget for any given level of habitat quality; see Figure 2.3. At each point on the PPF the only way to get more water

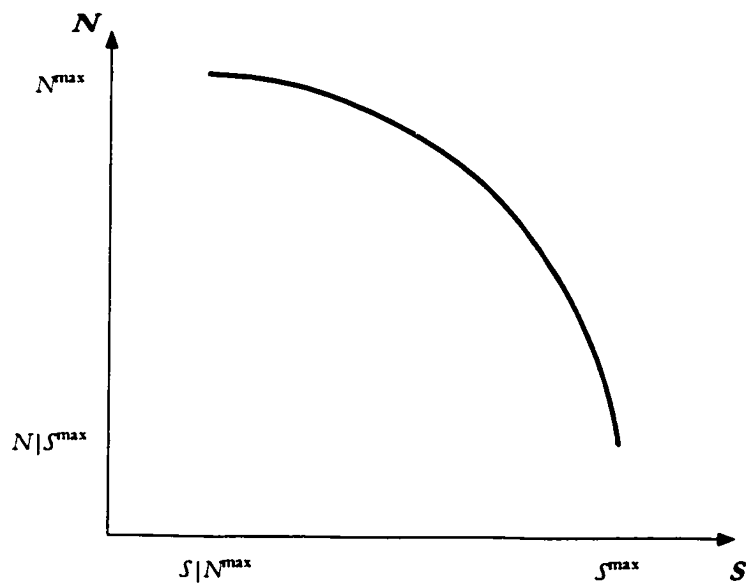


Figure 2.3 - Production possibility frontier for water quality and habitat quality.

quality is to decrease habitat quality, and vice versa. The PPF is sometimes called the "non-dominated front" for this reason.[14] Each point on the PPF could be thought of as the result of some optimization problem: $W_N = 0$ and $W_S = 1$ to determine $[N \mid S^{max}, S^{max}]$, $W_N = 1$ and $W_S = 0$ to determine $[N^{max}, S \mid N^{max}]$, and different levels of W_N and W_S that sum to one to determine intermediate points on the PPF. I used just such a strategy to trace out the PPF for the hypothetical watershed. A set of optimization problems was solved for the full range of weights on N and S. In general, the shape and position of the PPF will depend on the form and parameters of the production functions, the nature of the landscape, the costs of restoring each potential site, and the restoration budget. Therefore, the PPF summarizes much information about the potential environmental benefits of wetlands restoration and

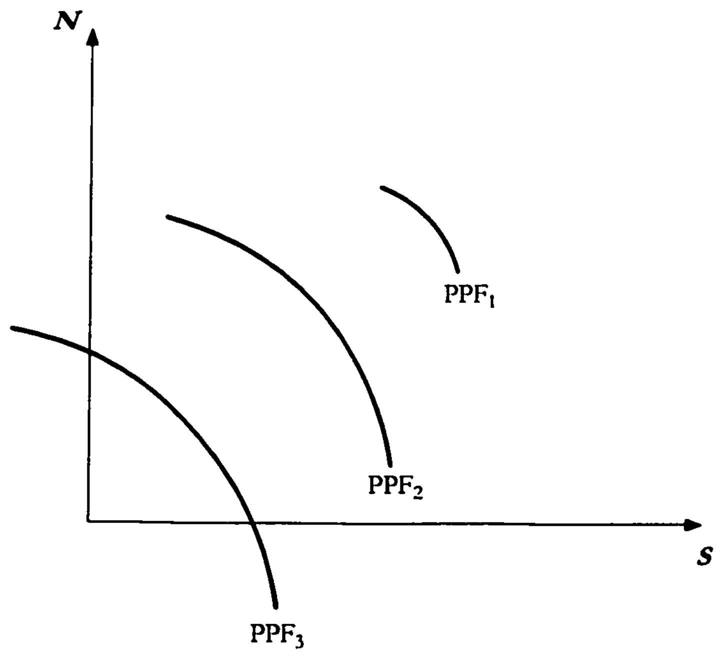


Figure 2.4 – Different possible shapes of the production possibility frontier.

[14] The PPF in Figure 2.3 is drawn concave, but this need not be true in the real world. Its shape will depend on the particulars of the production functions for the ecosystem services of interest. However, if the manager's (or society's) preferences are linear, as assumed in the objective function in expression 2.4a, then only the concave portions (in the extreme case only the endpoints) of the PPF will be relevant.

the tradeoffs involved. The farther away the PPF is from the origin, which represents current conditions, the greater the environmental improvements achievable, and the longer the arc of the PPF the greater the tradeoffs between environmental benefits facing the manager. For example, PPF_1 in Figure 2.4 describes a watershed where relatively large improvements in water quality and habitat quality would be possible given the budget, and there would be very little compromise between the two to consider. At the other extreme, PPF_3 describes a watershed where there would be less improvement possible given the budget, and there would be a large tradeoff between the objectives to consider. Because the origin represents current conditions, the fact that PPF_3 crosses the N and S axes implies an on-going degradation process that the restoration budget would be inadequate to completely stem. This would be the case if there were a steady conversion of wetlands to agriculture over time, so even though restoration might be taking place in some parts of the watershed, wetlands conversions would still be taking place elsewhere. In this case, if the manager allocated funds to maximize species abundance, the nutrient load to the river would still increase, and vice versa. PPF_2 describes an intermediate case, where wetlands conversion pressures would not be great enough to diminish current levels of ecosystem services, but the manager would still face substantial tradeoffs between the gains in the two objectives.

These concepts will figure prominently in later sections of this chapter, as well as in Chapter 6, which presents several empirical case studies in the Central Valley. This research is largely concerned with developing methods for delineating PPFs for real watersheds. However, even if a PPF cannot be delineated for a real watershed, the concept is still useful as it forces one to think about multiple objectives, constraints, and tradeoffs – features that virtually all environmental policy problems will share.

2.3 Results

I applied the optimization model to the hypothetical watershed to answer questions posed in the preceding sections regarding the importance of the selection algorithms and the management objectives. The parameter values used for the baseline example are given in Table 2.1. The superscripts U, A, and W indicate values for urban, agriculture, and wetland cells. The parameters for the water quality function were set so that urban and agriculture

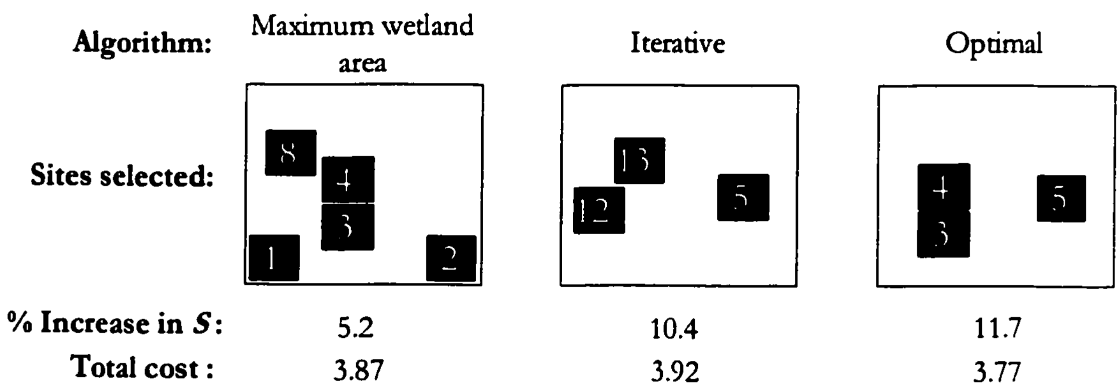
Table 2.1 – Parameter values used for the simulations.

Parameter	Value	Description
n^U	1	Parameters for the water quality production function, equation 2.4b
n^A	1	
n^W	-5	
θ	0.5	
ρ^U	0	Parameters for the habitat quality production function, equation2.4c
ρ^A	0	
ρ^W	1	
η	3	
α	2	Parameters for the cost function and constraint, equations 2.4f and 2.4g
φ	2	
Budget	4	

cells contributed equally to nutrient loading; a wetland cell could attenuate the equivalent nutrient loads of five upland cells; and the nutrient load, or attenuation in the case of wetlands, diminished by half for each unit of distance away from the river. The parameters for the habitat quality function were set so only wetland cells, either natural or restored, were suitable for the species, and connectivity decreased sharply with distance. The parameters for the cost function were set so there was substantial variation in costs across the landscape (refer back to Figure 2.2), and the budget was set so that at most five sites could be purchased.

2.3.1 Selection algorithms

For the purposes of comparing selection algorithms I focused on just one of the ecosystem services: habitat quality. The results from the different site selection algorithms for maximizing S are presented in Figure 2.5, which shows the sites selected and the expected increases in species abundance for each algorithm.



<u>Figure 2.5</u> - Results from the site selection algorithms for increasing S in the hypothetical watershed.

For comparison to the baseline watershed, the selected sites are loosely arranged in Figure 2.5 according to their relative positions in Figure 2.1. The first panel in Figure 2.5 shows the results of the heuristic that maximizes wetland area. All cells in the watershed were the same size, so this algorithm merely selected the largest set of affordable sites, those that were farthest from urban areas. Restoring these five sites yielded a 10% increase in the total area of wetlands in the watershed, and an approximately 5% increase in species abundance.

The second panel of Figure 2.5 shows the results of the algorithm that selected sites iteratively. This algorithm uses information regarding the relationship between wetland configuration and species abundance, so it should generally perform better than a heuristic based merely on maximizing wetland area (because total wetland area is only part of what determines species abundance). Figure 2.5 shows that in this example it did perform better. Only three sites were selected, which is a 6% increase in wetland area, but restoring this set yielded a greater than 10% increase in species abundance. The third panel of Figure 2.5 shows the results of the optimizing algorithm. This set of sites was selected by considering all possible combinations of affordable sets (of which there happened to be 3,322) and directly comparing the resulting increases in S. Again only three sites were chosen, but the optimal solution resulted in a more clumped configuration of wetlands than the iterative solution (restored sites 3 and 4 are adjacent to each other). The optimal solution yielded a larger increase in species abundance, nearly 12% compared to just over 10% for the iterative solution.

The reason that the iterative algorithm did not select the optimal set of sites is that only one site was considered at a time. The effect of choosing each site on the benefits from sites chosen in later rounds was not considered; i.e., the algorithm is not "forward-looking." Walking through the iterative selection process will make this point clear. With no sites selected, site 5 had the highest benefit-cost ratio at 0.744 (increase in $S = 1.83$ and cost = 2.46), so it was selected first. With site 5 selected, site 13 had the highest benefit-cost ratio at 0.709, so it was selected second. With both sites 5 and 13 selected, site 12 had the highest benefit-cost ratio at 0.704, so it was selected third. At this point no more sites were affordable, so the selection process was complete. This algorithm did consider the effects that sites selected in earlier rounds had on the benefits of sites selected in later rounds. It could do this because benefits were calculated anew at each stage, assuming the previously selected sites would be restored to wetlands. But it failed to consider the effects that selected sites would have on the benefits of those selected in later rounds. There was no way for the algorithm to see that by selecting site 3 in an early round extra benefits could be had by selecting site 4 in a later round because of their adjacency. The algorithm resulted in a sub-optimal set of sites because it could not fully account for the endogeneity of management benefits.

This failing is of more than just theoretical interest. The iterative strategy could be considered analogous to a manager who used a year-to-year site selection strategy. If a manager failed to consider the probable nature of the landscape as it evolved over time, due to the manager's own actions or due to forces outside of the manager's control, then opportunities for greater environmental improvements would be missed. The iterative strategy could also be considered analogous to uncoordinated decision-making on the part of different government agencies or other organizations involved in wetlands conservation. For example, the Army Corps of Engineers, the Environmental Protection Agency, the Fish and Wildlife Service, the Nature Conservancy, and more are involved in wetlands conservation activities on some level, but they do not always coordinate their efforts. Uncoordinated decision-making could yield less than optimal results because the spatial interactions between wetlands (and other land use types) could not be fully accounted for if different actors were making restoration decisions without knowing where others were restoring, or planning to restore, wetlands as well.

The iterative solution performed only slightly worse than the optimal solution in this example. The difference between the iterative solution and the max-wetland-area solution was much greater than the difference between the optimal and iterative solutions. So at least in the present example using an optimizing algorithm appears to be of less importance than accounting for the spatial interactions that affect ecosystem services in the first place, even if by way of a sub-optimal heuristic. However, in the real world one may not know how well different algorithms will perform beforehand, so it still would be desirable to apply the best available methods to the problem. Operations researchers have developed a number of heuristics that can achieve optimal or near-optimal results for many types of problems where benefits are endogenous with respect to the selection process, as was the case here (Reeves 1993). Comparing these algorithms was beyond the scope of this research, but see Pressey et al. (1997) and Csuti et al. (1997) for more discussion on this topic in the context of conservation decision-making.[15]

2.3.2 Different management goals

Results from varying the management goal are presented in Figures 2.6 and 2.7. These results would help the manager select a set of sites for restoration when both water quality and habitat quality were considered important. Because the manager could not maximize both ecosystem services simultaneously, she would have to choose a relative weighting for the objectives; i.e., the manager would have to choose a point on the production possibilities frontier. In a real application the choice between the options represented by the PPF might be made based on economic information on how the effected public values water quality relative to species abundance, if that information were available.

To trace out the PPF, I solved the optimization problem for the full range of combinations of W_N and W_S, starting with $W_N = 0$ and $W_S = 1$ and iterating in 0.05 increments – $(W_N, W_S) = (0.05,0.95); (0.1,0.9); \ldots (1.0,0.0)$. Each solution yielded a particular

[15] As a result of this research, as well as my reading of similar research, I have come to the tentative conclusion (call it a working hypothesis) that the degree to which heuristics will depart from the truly optimal solution will likely be much less than the degree to which a thoroughly naïve algorithm – one that ignores spatial effects altogether – will diverge from a sub-optimal heuristic solution that at least accounts for the spatial character of the production function at each iteration.

level of N and S, and a particular set of selected sites. Figure 2.6 shows the five unique sets of selected sites for all combinations of W_N and W_S, and Figure 2.7 shows the PPF for the watershed in terms of the percent improvement in N and S. There were only five unique sets of sites, so there were only five points on the PPF.

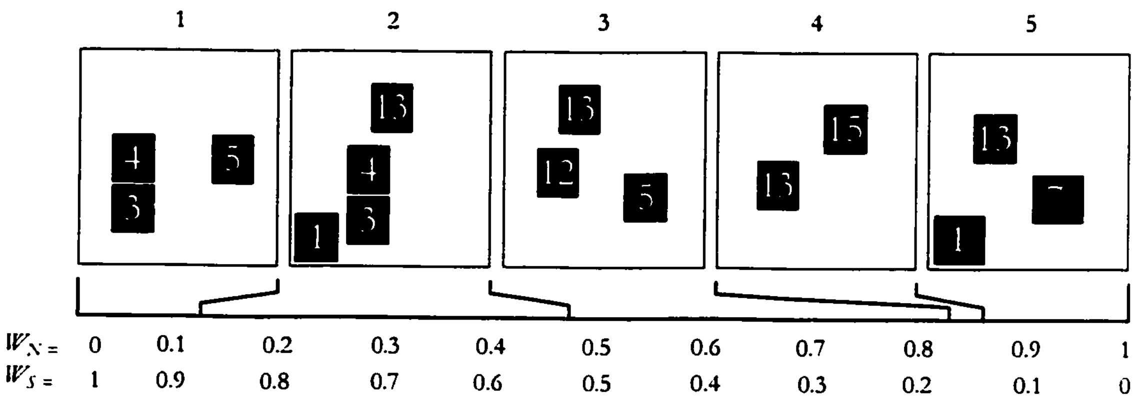


Figure 2.6 – Optimal sets of sits for all possible combinations of W_N and W_S.

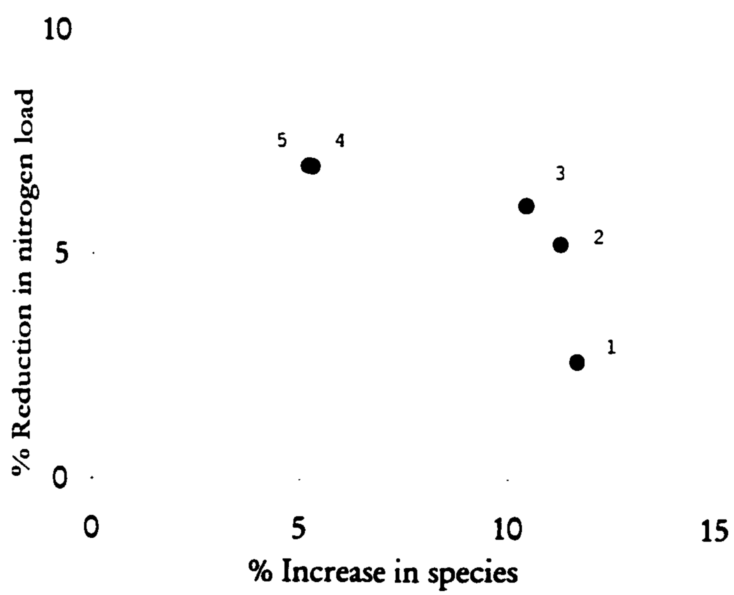


Figure 2.7 – The PPF for the hypothetical watershed.

Set 1 in Figure 2.6 and point 1 in Figure 2.7 correspond to the S^{max} solution, and set 5 and point 5 correspond to the N^{max} solution. The S^{max} solution was the "clumpiest" of the affordable sets of sites, and the N^{max} solution was made up of sites closer to the river. Excluding site 20, which was adjacent to an urban area and therefore very expensive, no sites were closer to the river than sites 7 and 13, and these two sites were the most "rural" of the sites near the river, and therefore the least expensive. Site 1 was included in the N^{max} solution merely because it was so cheap and though far from the river still provided some benefits. The intermediate sets of sites, sets 2, 3, and 4 in Figure 2.6, struck compromises between proximity to the river (good for water quality) and proximity to other wetlands (good for habitat quality).

Figure 2.6 also shows the ranges of W_N and W_S for which each set of sites was optimal. The S^{max} solution was the optimal set unless W_N/W_S exceeded 0.1/0.9. For W_N/W_S between 0.15/0.85 and 0.45/0.55 set 2 was optimal, and so on. In the real world these ratios would be scale dependent. The main point here is that one ccould identify meaningful thresholds analogous to these in real watersheds. For example, suppose that in a real application the units of measurement for nutrient load and species abundance were [kg/year] and [individuals]. If the above results were obtained and if the manager assigned less value to a 10 kg/year decrease in average nutrient loads than a 90 individual increase in species abundance, i.e. $W_N/W_K < 0.1/0.9$, then set 1 in Figure 2.6 should be chosen.

The practical utility of this approach comes in part from its ability to greatly reduce the number of options a manager needs to consider. Instead of having to consider 3,322 sets of affordable sites, the manager would only need to consider the five sets in Figure 2.6. One of these five provides more of one or both N and S than all of the other 3,317 affordable sets. Also, if the manager wanted to choose between these five sets based on economic valuation criteria, she would only need to know the relative value of N and S with sufficient accuracy to place it in one of the intervals shown in Figure 2.6. In fact, Figure 2.7 shows that the decision situation would be even less demanding than this. Sets 4 and 5 yielded essentially identical increases in both ecosystem services, so the manager would likely be indifferent between these two choices. In this way the framework could also allow a manager to determine how much flexibility there is in the choices she faces.

2.3.3 Monte Carlo and sensitivity analyses

The results presented so far came about because of the nature of the production functions for N and S and the spatial arrangement of the potential restoration sites and other parcels in the watershed. If S were merely an increasing function of wetland area (i.e., if spatial relationships did not matter), then the solution that maximized wetland area would also maximize S. And if the spatial effects were weak enough, or if management sites did not vary significantly across the dimensions that determined S (if they were of approximately equal distance from each other by being evenly distributed across the landscape), then an iterative selection algorithm would have been more likely to achieve the S^{max} solution. Under these conditions the management choices in early rounds would not much affect the benefits of choices in later rounds. It is both the form of the production functions for ecosystem services and the heterogeneity of the landscape that complicate the site selection problem.

To investigate the importance of the configuration of the landscape, I repeated the optimization exercises 100 times using a different randomly generated landscape each time. Table 2.2 shows the results of this Monte Carlo analysis. The first column in Table 2.2 is the performance measure for the different scenarios. The "max" superscripts indicate that the

Table 2.2 – Monte Carlo results show that the magnitude of the tradeoffs depend on the initial configuration of the landscape.

Performance measure	Average	Standard deviation	Frequency of ones
$\frac{N^{\max}}{N \mid S^{\max}}$	41	80	0
$\frac{S^{\max}}{S \mid N^{\max}}$	3.3	3.5	0
$\frac{N^{\max}}{N \mid Area^{\max}}$	97	106	0
$\frac{S^{\max}}{S \mid Area^{\max}}$	2.0	1.5	0.04
$\frac{S^{\max}}{S^{literature}}$	1.18	0.3	0.38

model was solved to optimize that variable. For example, $\frac{N^{\max}}{N \mid S^{\max}}$ in the first row of the table is the water quality improvement from selecting sites to minimize nutrient loads divided by the water quality improvement from selecting sites to maximize species abundance. Table 2.2 demonstrates several important points. First, rows 1 and 2 of the table show that the magnitude of the tradeoffs between objectives was generally quite large for the simulated watersheds, but they did vary with the configuration of the landscape. Sites selected to maximize water quality delivered, on average, more than 40 times the water quality improvement of sites selected to maximize habitat quality. Sites selected to maximize habitat quality delivered, on average, more than three times the habitat improvement of sites selected to maximize water quality. The reason the water quality benefits were more sensitive to the configuration of the selected sites was that the potential restoration sites varied with respect to the distance from the river (of which there was only one, in the middle of the watershed) more than they did with respect to the distance from other wetlands or restoration sites (of which there were many, scattered throughout the watershed). In other words, there were more opportunities to increase wetland connectivity than there were to restore wetlands close to the river.

Second, rows 3 and 4 show that the optimizing algorithm nearly always outperformed the strategy of maximizing wetland area, but this also depended on the configuration of the landscape. Four of the 100 randomly generated watersheds happened to be configured such that the solution that maximized wetland area also delivered the maximum habitat benefits. Finally, row 5 shows that the optimizing algorithm increased species abundance 18% more than the iterative algorithm on average, and in this case there were many more instances in which the two solutions coincided (38 of the 100 cases).

The Monte Carlo analysis allows one to generalize to the "family" of watersheds randomly generated by the five rules described earlier, which produced clumpy urban areas and remnant natural wetlands in a matrix of agriculture. This is more satisfying than the very specific results for the baseline watershed in Figure 2.1, but they were still all simulated watersheds and therefore only suggestive of the tradeoffs a manager might face in the real world. The main points to take from the Monte Carlo analysis are that (1) the configuration of the landscape matters, (2) when spatial effects are important, optimizing algorithms will frequently outperform a strategy of merely maximizing wetland area, and (3) an iterative strategy that accounts for spatial effects at each iteration will generally perform worse than an optimizing strategy, but usually much better than a strategy of merely maximizing wetland area.

As a final exercise to demonstrate the utility of the framework, I performed a sensitivity analysis on the key parameters in the production functions for ecosystem services. Tables 2.3 and 2.4 show how the results for the baseline watershed changed when the parameters of the production functions changed. Each parameter was varied around its baseline value while the other parameters were held constant. Only the values on either side of a transition from one optimal set of sites to another are included in the tables. For example, Table 2.3 indicates that for $1.0 \leq \eta \leq 2.6$, the sites 1, 2, 3, 4, and 8 were chosen if the objective were to maximize S. Note that this was the solution that maximized wetland area. For $2.8 \leq \eta \leq 6$, sites 3, 4, and 5 were chosen, the clumpiest set of sites. Also note that the sites chosen were insensitive to the values of ρ^W and n^W. Given the functional forms of the production functions (equations 2.1 and 2.2), these parameters do not influence the relative benefits of different spatial arrangements of wetland parcels. They do affect the expected levels of ecosystem services, but not which sites maximize the levels of services.

A sensitivity analysis such as this would allow a manager to determine the importance of the uncertainty associated with each of the hydrological and ecological parameters in the production functions for ecosystem services. In the same way that the framework could delineate the ranges of relative values of N and S for which the set of sites selected would be the same, it could also delineate analogous ranges for the parameters of the production functions. For the purposes of maximizing S, the manager in this example would only need to know if η were less than or greater than about 2.7. This could facilitate an important link between basic research and management. Using this type of framework, managers could provide feedback to hydrologists and ecologists regarding which components of the production functions were most important to better pin down.

Table 2.3 – Sites selected to maximize *S* were sensitive to η only.

ρ^W	η	**Sites selected**
0.1	3	3, 4, 5
1.0	3	3, 4, 5
1	1	1, 2, 3, 4, 8
1	2.6	1, 2, 3, 4, 8
1	2.8	3, 4, 5
1	6	3, 4, 5

Table 2.4 – Sites selected to minimize *N* are sensitive to θ only.

n^W	θ	Sites selected
-1	0.5	1, 7, 13
-8	0.5	1, 7, 13
-5	0.025	1, 2, 3, 8, 12
-5	0.125	1, 3, 12, 13
-5	0.375	1, 3, 12, 13
-5	0.400	1, 7, 13
-5	0.900	1, 7, 13

2.4 Conclusions

This chapter presented an optimization framework for prioritizing sites for wetlands restoration on a watershed scale. Adequately addressing problems of this type requires a truly multidisciplinary approach. At the outset, one must develop a good understanding of the hydrologic and ecological processes by which the configuration of the landscape affects ecosystem services. Once the production functions are specified, an economic perspective – structuring the problem in cost-effectiveness terms and using optimization techniques – provides a powerful framework for illuminating the inevitable tradeoffs between the many options that a manager may face. Finally, operations research tools, in the form of effective site selection heuristics, will often be required to tackle the large problems that will come with most real-world decision situations.

Developing adequate representations of all of the important factors that affect water quality and habitat quality in a watershed is a monumental task in itself. It should be understood clearly up front that the information that comes from the application of this type of framework to a real watershed would be only as good as the basic hydrology and ecology that went into it. However, this does not mean that these functions must be known with certainty before an optimization approach could be used to prioritize restoration sites, just that any temptation to take the outputs of the model more seriously than the functions describing ecosystem services used as its foundation should be resisted. On the other hand, if managers demur from using uncertain scientific information, merely on the grounds that it is uncertain, then no improvements in effectiveness would be possible. Developing the tools necessary for incorporating the relevant scientific information into the decision making process – be it already available or still in the process of discovery – is an important and necessary step in the direction of making more effective environmental management decisions.

The use of numerical optimization techniques could allow a manager to apply information in a rational and replicable manner for making decisions, or for analyzing past decisions. Because the framework is structured as an optimization model, not only could two scenarios be compared to each other, they could be compared to the "best possible" scenarios (subject to the limitations involved in describing real-world processes in numerical forms amenable to optimization modeling). Furthermore, through sensitivity analyses like the one presented above, managers could determine the required resolution of ecological and economic information that would be crucial for decision-making, which could yield focused recommendations for future research.

Real-world wetlands management decisions will often turn on more than just two objectives, and usually more than 5 out of 20 potential restoration sites will be up for consideration. However, by finding only the non-dominated sets of sites these methods could still allow a manager to reduce significantly the number of alternatives she has to select from, just as for the hypothetical watershed the manager was left with only five out of 3,322 affordable sets from which to choose. Guaranteeing optimal solutions to these types of problems is difficult if the number of sites available for restoration is large, but through a combination of appropriate simplifications and the application of modern optimization heuristics, this approach could be applied to real watersheds to address real management questions. Spatial effects and tradeoffs between water quality, habitat quality, and other types of environmental benefits are inherent in real-world management decisions, though they often are left unexplored. The framework presented in this chapter provides a means for exploring them. The following chapters describe an empirical version of this framework that I applied to several case studies in the Central Valley of California.

Chapter 3 – Wetlands as habitat

3.1 Wetlands and species protection

Habitat support for wildlife figures prominently on the list of valuable public benefits that wetlands can provide. Species that rely on wetlands include many that are of conservation concern as well as many that are hunted or harvested. At least one third of the species listed as threatened and endangered by under the Federal Endangered Species Act live in wetlands, and nearly half rely on wetlands for at least part of their life cycle (U.S. EPA 1995). Murdock (1994) found that wetlands provide habitat for nearly 90% of rare, threatened, or endangered species in the southern Appalachian region, and according to Edwards and Weakley (2001), wetlands harbor a large proportion of the endangered plant species in the southeastern United States. Kirkland and Ostfeld (1999) found that the number of federally endangered mammals was positively related to the loss of wetlands in each state. Eighty percent of America's breeding bird populations, more than fifty percent of the 800 species of protected migratory birds, and many commercially or recreationally harvested species – such as alligators, muskrats, nutria, beaver, mink, and otter – also rely on wetlands (Mitsch and Gosselink 1993).

Wetland losses have been substantial throughout the continental United States, with approximately 50% converted to other uses since European settlers arrived, but losses have been especially severe in California, where more than 90% of the historic wetland acres have been converted (Dahl 1990). The study area for this research was the Central Valley of California, which consists of more than 5 million hectares (approximately 23,000 square miles) between the Sierra Nevada mountain range to the east and the Coast Ranges to the west (see Figure 3.1). The Central Valley is a vast, flat region, much of which was historically inundated by spring snowmelt runoff from the Sierra Nevada in most years. Comprehensive government water works projects, including at least one dam on virtually every major river, tributary, and stream in the state, have drastically altered the hydrologic regime of the region and allowed large areas of historically flooded or saturated lands to be converted to agriculture and urban uses (Mount 1995). Before European settlers arrived in California, more than 40% of the Central Valley would have qualified as wetlands (CVPIA 2000), but today only approximately seven percent of the valley exists as wetlands (National Wetlands Inventory data: see Section 3.2.1 and Table 3.1).

Figure 3.1 - The Central Valley of California is bounded on the east by the Sierra Nevada mountain range, and on the west by the Coastal Range. It stretches over 677 kilometers (420 miles) from northwest to southeast, and covers more than 5,863,000 hectares (22,640 square miles).

Despite substantial losses, wetlands in the Central Valley remain important for many species of concern. The Central Valley is one of the most important regions in western North America for migrating shorebirds. In winter and spring, wetlands in the valley support more shorebirds than any other inland region, and in fall the Great Salt Lake is the only inland site in western North America consistently surpassing the Central Valley in shorebird numbers (Shuford et al. 1998). According to the California Wildlife Habitat Relations model, 59 of 103 species with special management status in California (e.g. threatened or endangered) rely to some degree on fresh or saline emergent wetlands (California Department of Fish and Game 1999).

Because wetlands are important for so many species in the Central Valley, many wetlands restoration programs in the region have been motivated in large part by the habitat benefits they are expected to provide. Models of species-habitat relationships could help managers design wetlands restoration programs more effectively. The development of habitat relationships models for all species of concern in the Central Valley was well beyond the scope of this project. This chapter presents results from statistical models designed to describe habitat preferences for a number of bird species that breed in the Central Valley. This research focused on mallards, but I will also present preliminary results for several other birds species, and bird richness and diversity. The models use data from throughout the Central Valley that are reasonably spatially explicit, so the results could be used to inform valley-wide rankings of potential restoration sites.

3.2 Focusing on mallards

The bulk of this chapter describes statistical models of the relationships between mallard (*Anas platyrhynchos*) abundances in the breeding season (May through July) and landscape characteristics in the Central Valley of California. Mallards are one of the most common waterbird species that breed in the Central Valley, and management often centers on them as a general indicator species for wetlands-waterfowl relationships (Central Valley Habitat Joint Venture 1990, McLandress et al. 1996). The Central Valley is a crucial area for waterfowl migrating on the Pacific Flyway. Nowhere else in North America do so many waterfowl spend the winter on such a small wetland base (Heitmeyer et al. 1989). The general question that the models described here were designed to address is: How do the amount and arrangement of land use types in the Central Valley affect the distribution of abundances of mallards in the breeding season? The models described in this chapter cannot account for the small-scale habitat preferences of mallards (those related to within-patch heterogeneity), but they should add to our currently less-developed understanding of medium-scale (several patches, 5-50 hectares) and large-scale (landscape-level, thousands of hectares) factors that affect mallard distributions.

3.2.1 The data

The models in this chapter are based on bird abundance data collected in the Central Valley of California in the years 1997-2000 by the North American Breeding Bird Survey (BBS). These data served as dependent variables in regression models relating bird counts to a set of variables that describe the nature of the landscape surrounding each of the BBS survey locations. Measures of the landscape variables came from a GIS land use dataset developed by the California Department of Water Resources (DWR), and from the U.S. Fish and Wildlife Service's National Wetlands Inventory (NWI).

The North American Breeding Bird Survey is a large scale monitoring effort run by the U.S. Geological Survey.[16] Each year during the breeding season, skilled volunteer surveyors drive hundreds of 25-mile routes across North America, and at every half-mile point they stop and count all birds they can identify by sight or sound within 400 meters. More than 800 route-stops have been surveyed at least once between 1997 and 2000 in the Central Valley. BBS surveyors usually count more than one hundred species in the Central Valley in a given year, but only a handful of species are sighted frequently. In 1997, 1998, and 1999 only 26 out of 117, 29 out of 121, and 28 out of 113 species were seen at more than five percent of the route-stops.

The BBS data is not collected for the purposes of analyzing species-habitat relationships. Its main purpose is to assess long-term trends of the relative abundance of breeding birds, generally over large spatial scales. Since the survey protocol was designed with this goal in mind, the data are not ideally suited for landscape-level spatial analyses of habitat preferences. The data collection protocol does not call for the measurement of any habitat or land use variables.[17] Also, surveyors are instructed to count birds at the same locations from year to year, but they will not always return to the exact same locations every year because considerations of roadside safety or convenience also come into play when they decide where to stop and conduct the counts. This may not introduce significant error into large-scale trend summaries, but the same may not be true for detailed analyses of habitat relationships. Another drawback of using BBS data for modeling species-habitat relationships, especially with management objectives in mind, is that those species that occur most frequently are likely not those of the highest conservation concern. Rare species may

[16] See the BBS website: http://www.mp2-pwrc.usgs.gov/bbs/.
[17] Though this may change in the future (O'Connor et al. 2000).

not be observed frequently enough use statistical models to investigate their habitat preferences. These qualifications notwithstanding, the BBS dataset is still the best set of location-specific species abundance data available that is collected over a significant period of time, over a large area, and at locations that are randomly distributed across the landscape.[18]

Even though the BBS data is collected mainly for analyzing overall trends in abundance, a number of researchers have used the data for other purposes, including investigating species-habitat relationships. For example, Flather and Sauer (1996) used BBS data to analyze relationships between abundances of neotropical migrants and landscape characteristics in the Eastern United States. Herkert (1998) used BBS data to compare the attractiveness of Conservation Reserve Program lands relative to other land use types for grasshopper sparrows in the mid-west. Koenig (1998) took advantage of the spatial nature of the dataset to measure the degree of synchrony in recruitment of land birds in California by analyzing spatial autocorrelation in the trends in abundances across the state.

However, all of these studies used species abundance data aggregated to the route level. Counts at all 50 stops on each route were summed for each year and associated with either the route start point or center point. This results in a substantial loss of information because the models cannot take advantage of differences between landscape characteristics across stops. Also, at this level of aggregation the sum of the counts for the route are associated with landscape features within a circle large enough to encompass the entire linear route, which includes a great deal of land over which counts were not conducted. The models presented in this chapter are among the first to be based on stop-level BBS data.

[18] The routes were randomly distributed along roads, but the stops were distributed systematically along the routes, at every half-mile point. The fact that only roadside locations were surveyed is another factor makes the data less than ideal. However, Austin et al. (2000) found no bias in estimates of mallard abundance per area of wetlands as a function of distance from roads.

The DWR land use dataset contains a detailed classification of land use types, including ten types of agricultural land, six types of urban land, and four types of native vegetation. The data were collected between 1986 and 1998 using aerial photography and extensive field visits. The NWI dataset, which is maintained by the U.S. Fish and Wildlife Service, contains wetlands boundaries and attributes for most states in the continental U.S. The NWI is based on a detailed classification of wetland types, including several general system and sub-system classes with varying levels of detail under each (Cowardin et al. 1979). Data for the Central Valley were collected between 1972 and 1987, primarily using aerial photography.

I combined the DWR and NWI datasets to create a single, consistent dataset from which I could extract land use characteristics from the neighborhood of all BBS route-stops. NWI wetland boundaries and types were given precedence over any DWR land use polygons that coincided with them, and all uplands were classified according to the DWR data. Figure 3.2 shows the combined land use dataset (with land uses aggregated to six general types), and Table 3.1 lists the total area of each land use type in the Central Valley. I used the Patch Analyst extension (Schumaker 1998) in ArcView to extract landscape attributes from the combined dataset within 400 meters of each BBS route-stop. Figure 3.3 shows the locations of the routes in the Central Valley and a close up view of several stops along one route.

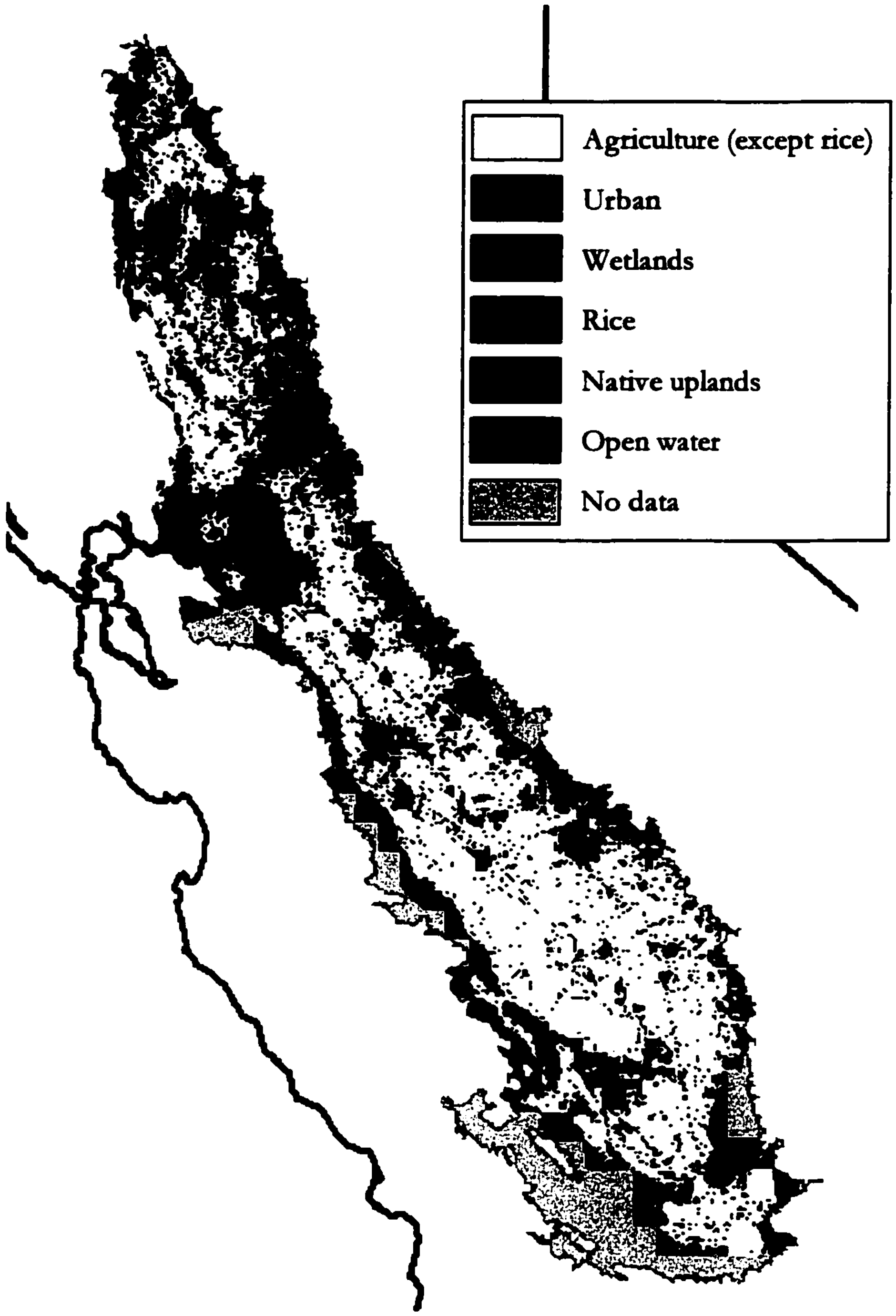


Figure 3.2 – Land use in the Central Valley of California, from the combined DWR and NWI dataset.

Table 3.1 – Land use types in the Central Valley, from the combined DWR and NWI dataset.

Land use type	Area [ha]	Percent	Aggregate percent	
Field crops	882,488	15.94	Agriculture	54.7
Deciduous fruits nuts	475,977	8.60		
Pasture	389,250	7.03		
Grain and hay crops	269,013	4.86		
Vineyards	244,004	4.41		
Truck nursery berry	235,691	4.26		
Rice	232,154	4.19		
Idle ag	125,766	2.27		
Citrus subtropical	110,737	2.00		
Semi-ag & incidental	63,112	1.14		
Native vegetation	1,398,952	25.27	Native uplands	25.4
Native barren	6,982	0.13		
Native misc	82	0.00		
Palustrine	265,160	4.79	Wetlands	6.7
Estuarine intertidal	23,935	0.43		
Native riparian	21,242	0.38		
Riverine lower perennial	18,728	0.34		
Lacustrine littoral	17,580	0.32		
Riverine tidal	17,331	0.31		
Riverine intermittent	4,243	0.08		
Riverine upper perennial	564	0.01		
Urban	228,405	4.13	Urban	6.1
Vacant urban	40,093	0.72		
Residential	34,865	0.63		
Industrial	18,205	0.33		
Urban landscape	9,515	0.17		
Commercial	7,417	0.13		
Lacustrine limnetic	18,897	0.34	Deep water	0.9
Native open water	18,229	0.33		
Estuarine subtidal	10,575	0.19		
Total:	5,536,937			

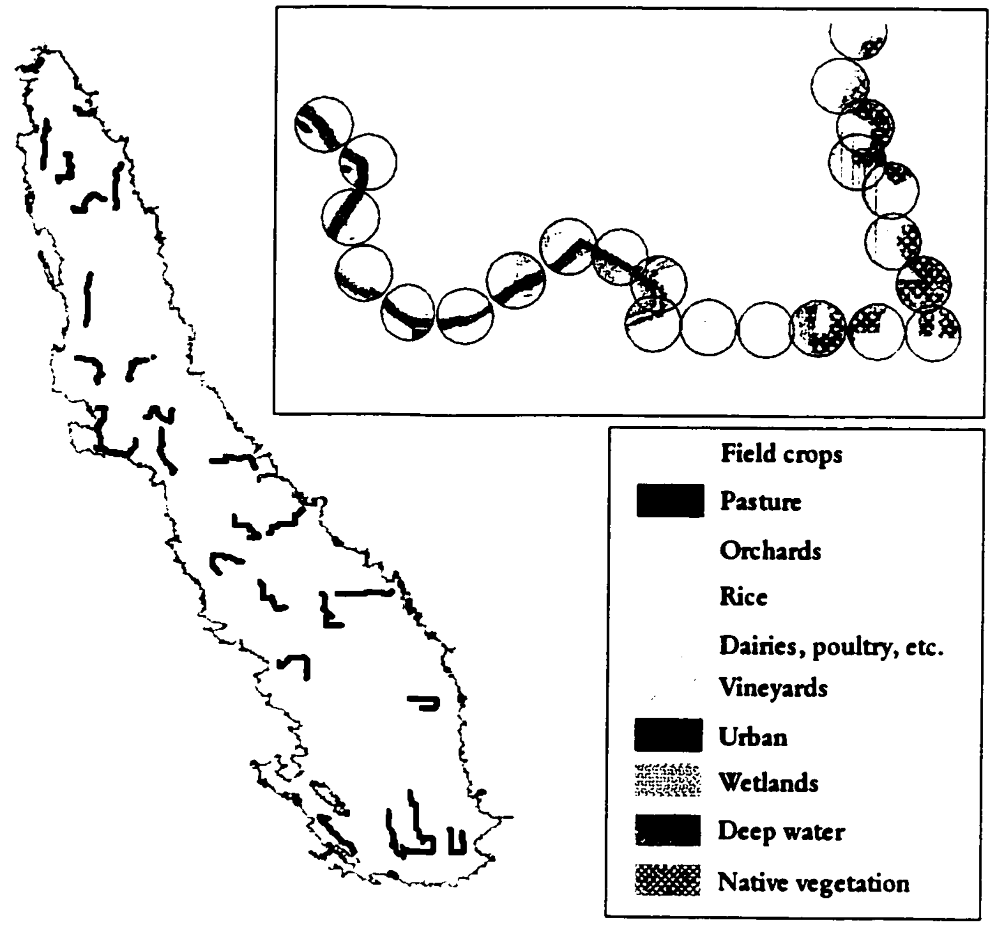


Figure 3.3 – BBS route-stops in the Central Valley, and a close-up view of stops along one route.

3.2.2 Regression models of bird abundances

One of the central tasks of ecology is explaining the distribution and abundance of species, and when management goals include maintaining or enhancing the populations of one or more species of concern, ecological models of species abundances can be very useful. In this section I present regression models of mallard abundances that were estimated using the data described above. Some of the regression results were used in optimization models described in Section 3.2.3 and in Chapter 6 to analyze wetlands restoration site selection strategies in the Central Valley.

3.2.2.1 Model specification

I used maximum likelihood regression to relate mallard abundances from the BBS dataset to a set of land use variables and other covariates. Three types of regression models were estimated: a logistic model, a Poisson model, and a negative binomial model. The logistic model is:

$$\Pr[y_i > 0] = \frac{\exp(\boldsymbol{x}_i \boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}_i \boldsymbol{\beta})} \tag{3.1}$$

In equation 3.1, y_i is the abundance of mallards counted for observation i, x_i is a row vector of independent variables for observation i, and β is a column vector of parameters to be estimated. Logistic regression is a standard method for analyzing species-habitat relationships when only presence-absence data are available (Manel et al. 1999, Pearce and Ferrier 2000). In this case the logistic model was not ideal because it ignores the information in the counts greater than one. However, it is nonetheless instructive to compare the logistic results to the count regression results. The Poisson regression model with exponential mean function is:

$$\Pr[y_i = y] = \frac{\mu_i^{\,y} \exp(-\mu_i)}{y!} \tag{3.2}$$

In equation 3.1, $\mu_i = \exp(\boldsymbol{x}_i \boldsymbol{\beta})$ is the mean. The Poisson model is the standard model for analyzing count data (Cameron and Trivedi 1998). Exponentiation of $\boldsymbol{x}_i \boldsymbol{\beta}$ ensures non-negativity, which is required for the mean, and unlike linear regression the realizations of the Poisson distribution take on only discrete non-negative values, which is consistent with the raw data. One limitation of the Poisson model is its inherent assumption that the mean and variance are equal. In many applications, including this one, the variance is greater than the mean. When this happens the data are "overdispersed." The most severe consequence of overdispersion is that in its presence the usual formulas will underestimate standard errors. However, if the variance is modeled as a multiple of the mean the correction is straightforward.[19] All Poisson regression results presented in this chapter have been corrected for overdispersion.

A more general way to handle overdispersion is to incorporate a random component into the specification of the mean itself: $E[y_i] = \exp(\boldsymbol{x}_i\beta + \varepsilon_i)$. If ε_i is assumed to follow a gamma distribution the negative binomial model results. The mallard abundance data exhibited a fair degree of overdispersion, so it was important to consider the negative binomial along side the standard Poisson model. In practice the negative binomial model usually gives results that are qualitatively, and very often quantitatively, similar to Poisson regression results.

3.2.2.2 Model selection

One of the most difficult aspects of any statistical modeling exercise is model selection – choosing the functional form and set of variables to include in the model. The Poisson and negative binomial models provided the most appropriate functional forms, and data were available for a number of land use types, several weather-related covariates, and dummy variables for each of the nineteen routes. Table 3.2 describes the land use types used in the regression models. To avoid perfect colinearity among the independent variables, one land

---

[19] The t-statistics can be corrected by dividing them by $\sqrt{\varphi}$, where $\varphi = \frac{1}{N-K}\sum_{i=1}^{N}\frac{(\hat{\mu}_i - y_i)^2}{\hat{\mu}_i}$, $\hat{\mu}_i$ is the predicted mean for observation i, y_i is the observed abundance for observation i, N is the number of observations, and K is the number of parameters in the model (Cameron and Trivedi 1998, p 64).

Table 3.2 – Descriptions of the land use types used in the regression models.

Land use type	Description
Field crops…………….	Includes row crops (e.g. beans, corn, and cotton) and grain crops (e.g. barley, wheat, and oats)
Pasture………………..	Includes clover, alfalfa, native, and mixed pastures
Orchards……………..	Includes deciduous fruits and nuts (e.g. apples, cherries, and almonds) and citrus and subtropical fruit orchards (e.g. oranges, avocadoes, and olives)
Rice……………………	Medium and short grain, flooded in the summer
Vineyards……………..	Includes table, wine, and raisin grapes
Dairy and feedstock operations……………	Includes dairies, livestock feed lots, farmsteads, and poultry farms
Urban…………………	Includes residential, commercial, industrial, and vacant lots
Wetlands………………	Includes palustrine - shallow non-tidal wetlands usually dominated by emergent vegetation; littoral - wetlands with less than 30% persistent vegetation and usually associated with deep water habitats such as lakes or reservoirs; and riverine - wetlands associated with natural or artificial channels
Deep-water habitats….	Includes areas generally inundated with water at depths of 2 meters or more

use type had to be excluded from the models. The excluded land use type was native uplands, so all parameters were estimated relative to the native uplands baseline.

The simplest regression model would include only those variables that would certainly affect mallard abundances – wetlands, rice, and urban lands perhaps. However, using a model that was too parsimonious would run the risk of omitted variable bias. A very general model would include all variables that *might* affect mallard abundances – all the land use variables plus second-order and interaction terms between them. However, with too many variables the model would quickly run into multicolinearity and degrees of freedom problems. The strategy I used for this analysis was to compare a number of specifications, starting with a fairly parsimonious model and ending with a fairly general one.

I considered several sets of variables, based on the land use classification in Table 3.2, for six alternative versions of the logistic, Poisson, and negative binomial regression models. First, the percent of each land use type within a 400-meter circle surrounding each route-stop should capture differences in breeding habitat quality across the land use types. Reproductive success for mallards should be highest in and near wet habitats (wetlands and rice primarily), and lower elsewhere, so the land use percent variables should be important predictors of mallard abundances in the breeding season. Next, second-order terms for each land use type (the percent of each land use squared) would allow for a more flexible fit. If the log of the abundance were merely a linear function of the percent of each land use type in the vicinity, then these variables would have no significant explanatory power. If the relationship were more complex, then the second-order terms would allow a more realistic picture to emerge. Figure 3.4 shows the flexibility of the exponential mean function when second-order terms are included.

With the squared land use variables included, the models could account for the possibility that the arrangement of the land use types, in addition to the total amount of each present in the landscape, is important for determining the distribution of mallards in the breeding season. If the predicted abundance increased with the area of wetlands over the entire range of the data (i.e., the returns to wetland area was always positive), then consolidated wetland areas of 50 hectares (the total area within a 400-meter radius circle) or more would be better for mallards than the same amount scattered as multiple smaller

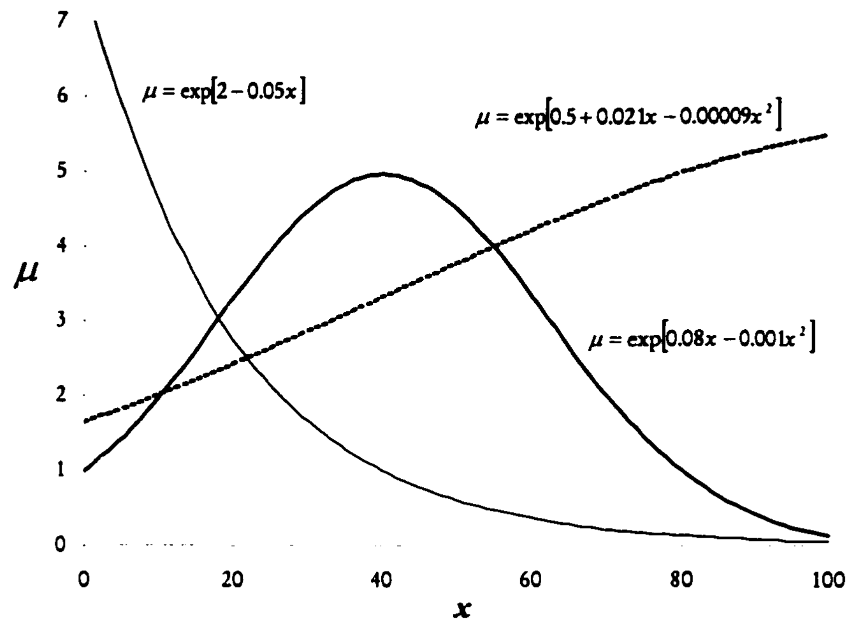


__Figure 3.4__ – The exponential mean function with second-order terms included is flexible enough to represent a wide range of responses.

patches. On the other hand, if the predicted abundance decreased within the range of the data (i.e., before the percent of land in wetlands within 400 meters reached 100%, so the returns to wetland area becomes negative at some point), this would suggest that smaller patches would be better.

Interactive terms between land use types could be included in the models to capture edge effects. Shared edge is not measured directly by interactive terms, but it should be highly correlated with them. Shannon's diversity and evenness indices can capture effects of "generic" landscape heterogeneity. These variables would be important if spatial effects were more complex than could be captured by the land use-specific variables alone. Finally, data were available on several non-land use variables that should control for other potentially important environmental effects not captured by the land use variables: total annual precipitation, to account for some of the spatial and temporal variation in water availability; latitude, to control for any overall north-south differences in climate; average wind conditions and temperature at the time the routes were surveyed; and dummy variables to account for unobserved fixed effects specific to each route and each year.

Of all possible combinations of the variables described above, I chose seven sets (not all of them nested) to be used in different versions of the logistic, Poisson, and negative binomial regression models. The seven versions of the logistic model were estimated for comparison purposes only. The focus here was on explaining abundances, not just occurrences, so some version of the Poisson or negative binomial model was preferred from the outset. To choose the best model from the 14 count models, I used the Akaike Information Criteria (Burnham and Anderson 1998; Anderson et al. 2000), which is based on a relationship between information theory and maximum likelihood and allows one to compare non-nested alternative models. The Akaike Information Criteria is:

$$AIC_i = -2\ln L_i + 2K_i \tag{3.3}$$

In equation 3.3, $\ln L_i$ is the log-likelihood for model *i*, and K_i is the number of parameters in model *i*. A "weight of evidence," w_i, for model *i* being the best approximating model in a set of candidate models, can be calculated as:

$$w_i = \frac{\exp[-\frac{1}{2}(AIC_i - \min AIC)]}{\sum_i \exp[-\frac{1}{2}(AIC_i - \min AIC)]} \tag{3.4}$$

The model with the largest w_i is then taken as the best model.[20]

[20] The idea behind maximum likelihood estimation is that the set of parameters that maximizes the probability of observing the data is the best set of estimates, given the data *and given the functional form of the model*. Model selection using AIC can be though of as an extension of maximum likelihood in that it removes the second conditional in that statement. The AIC can be used to estimate the "likelihood of a model [being the best from a set of a priori candidate models], given the data" (Anderson et al. 2000, p 918). See Burnham and Anderson (1998) for more on model selection using the AIC.

The seven versions shared the following set of variables: rice, urban, deep water, dummy variables for 1998, 1999, and 2000 (1997 served as the baseline), Shannon's diversity and evenness indices, precipitation (total for May through July), day of the year the route was surveyed, latitude, average temperature, and average wind condition. The seven versions were distinguished by the following sets of variables:

Version 1: Wetlands, agriculture
Version 2: Version 1 + second order terms for wetlands, rice, agriculture, urban, deep water
Version 3: Version 2 + wetlands interacted with rice, agriculture, urban, deep water
Version 4: Wetlands, five types of agriculture, second order terms, interactions with wetlands
Version 5: Three types of wetlands, five types of agriculture, second order terms, interactions with wetlands
Version 6: Version 4 + route dummies
Version 7: Version 5 + route dummies

Version 1 was the most parsimonious model, and version 7 was the most general. For versions 1 through 3, I aggregated the five types of agriculture listed in Table 3.2 into a general "agriculture" variable, on the presumption that mallards may not discriminate between types of agriculture (evidence regarding this hypothesis would come from comparing the results of versions 1 through 3 to versions 4 through 7). For versions 5 and 7, I disaggregated wetlands to palustrine, littoral, and riverine types, on the presumption that mallards may discriminate between different types of wetlands (evidence regarding this hypothesis would come from comparing the results of versions 4, 5 and 7).

The disaggregated agriculture types need no explanation, but it may be useful to review the Cowardin wetlands classification system (Cowardin et al. 1979), on which the NWI dataset is based, to describe the disaggregated wetland types used in versions 5 and 7 of the models. The Cowardin system separates wetlands into a hierarchical classification scheme. At the highest level, wetlands are classified into "systems," which include marine, estuarine, riverine, lacustrine, and palustrine. Virtually no marine or estuarine wetlands exist in the Central Valley. The riverine system "includes all wetlands and deep water habitats contained within a channel" (Ibid, p 7). This could confuse the interpretation of the deep water and riverine variables, but there is no obvious sub-category in the Cowardin system that allowed easy separation of the deeper portions of the riverine wetlands (which include the river or stream proper) from the shallower portions (which include riparian wetlands). Littoral wetlands are associated with lacustrine (lake) systems, and have a low proportion of persistent emergent vegetation. The palustrine class consists of small wetlands isolated from larger bodies of water, including ponds and all non-tidal wetlands dominated by trees, shrubs, or persistent emergent vegetation.

3.2.2.3 Regression results

All models were estimated in Limdep, version 7. Table 3.3 shows for each model: (1) the squared correlations between the observed counts and the predicted means, $r^2_{y,\hat{\mu}}$, which serves as a measure of fit analogous to R^2 from the standard linear regression model,[21] (2) the land use variables that were significant at the 5% level plus the direction of their effects, and (3) the Akaike weights (for the count models only).[22] The results shown in Table 3.3 suggest a number of consistent relationships. First, the wetlands and rice variables were significant and positive in all models. The deep-water variable was significant and positive in versions 1 through 5 of the logistic model. The urban variable was significant and negative in about half of the models – mostly in the more parsimonious versions of the logistic and negative

---

[21] $r^2_{y,\hat{\mu}} = \left(\dfrac{\operatorname{cov}(y,\hat{\mu})}{\sqrt{\operatorname{var}(y)\operatorname{var}(\hat{\mu})}} \right)^2$, which equals R^2 in the standard linear regression model.

[22] In the interest of brevity, only the most pertinent results are presented in the text. Data and the Limdep code used to estimate all models are available upon request from the author.

Table 3.3 – Comparison of results from seven versions of the logistic, Poisson, and negative binomial regression models. The w_i's are the Akaike weights for each model, which are a function of the log-likelihood values for all models in the candidate set as per equations 3.3 and 3.4.

	Logistic		Poisson		Negative binomial	
	$r^2_{y,\hat{\mu}}$	Statistically significant variables, 5% level (direction)	$r^2_{y,\hat{\mu}}$ [w_i]	Statistically significant variables, 5% level (direction)	$r^2_{y,\hat{\mu}}$ [w_i]	Statistically significant variables, 5% level (direction)
1	0.1667	deep water (+) rice (+) wet (+)	0.2260 [0.000]	rice (+) wet (+)	0.1659 [0.000]	urban (-) rice (+) wet (+)
2	0.2249	urban (-,+) deep water (+,n.s.) rice (+,n.s.) wet (+,-)	0.2934 [0.000]	rice (+,-) wet (+,-)	0.1442 [0.000]	urban (-,+) rice (+,-) wet (+,-)
3	0.2368	urban (-,+) deep water (+,n.s.) rice (+,n.s.) wet (+,-) wet*rice (-) wet*deep water (-)	0.3367 [0.000]	urban (n.s.,+) rice (+,-) wet (+,-) wet*rice (-) wet*deep water (-)	0.2380 [0.000]	urban (-,+) rice (+,-) wet (+,-) wet*rice (-)
4	0.2281	pasture (+,n.s.) orchards (-,+) urban (-,+) deep water (+,n.s.) rice (+,n.s.) wet (+,-) wet*rice (-) wet*deep water (-)	0.3429 [0.000]	pasture (+,-) orchards (-,n.s.) rice (+,-) wet (+,-) wet*rice (-) wet*deep water (-)	0.2197 [0.000]	orchards (-,+) urban (-,+) rice (+,-) wet (+,-) wet*rice (-) wet*deep water (-)
5	0.2389	orchards (-,+) urban (-,+) deep water (+,n.s.) rice (+,n.s.) palustrine (+,-) littoral (+,-)	0.3632 [0.000]	pasture (+,-) orchards (-,+) rice (+,-) palustrine (+,n.s.) littoral (+,-) riverine (+,n.s.) wet*ag (+) wet*rice (-) wet*urban (+)	0.2133 [0.000]	orchards (-,+) urban (-,+) rice (+,-) littoral (+,-) wet*ag (+)
6	0.2611	orchards (-,+) rice (+,n.s.) wet (+,n.s.) wet*rice (-) wet*deep water (-)	0.3927 [0.000]	pasture (+,-) orchards (-,n.s.) dairy (-,+) rice (+,-) wet (+,-) wet*rice (-) wet*deep water (-)	0.2910 [0.268]	orchards (-,+) rice (+,n.s.) wet (+,n.s.) wet*rice (-)
7	0.2717	pasture (+,n.s.) orchards (-,+) rice (+,n.s.) palustrine (+,-) wet*rice (-)	0.4026 [0.000]	pasture (+,-) orchards (-,+) dairy (-,n.s.) rice (+,-) littoral (+,-) riverine (+,-) wet*ag (+) wet*rice (-) wet*urban (+)	0.2778 [0.732]	orchards (-,+) rice (+,-) littoral (+,-) riverine (+,-) wet*ag (+) wet*rice (-)

binomial models. The aggregate agriculture variable was not significant in any of the models in which it appeared (versions 1 through 3). Of the different types of agriculture in versions 4 through 7: the orchards variable was significant in all models in which it appeared (always negative), the pasture variable was significant in 6 of the 12 models (always positive), and the dairy variable was significant and negative in versions 6 and 7 of the Poisson model only. All of this indicates that the different versions of the models were broadly consistent with each other in terms of the variables that were statistically significant and the direction of their effects. The general results regarding the apparent preferences for "wet lands" in particular – wetlands proper and rice especially – were robust across the different specifications.

Table 3.4 shows the results of the formal model selection process using the AIC statistic. Versions 6 and 7 of the negative binomial were the best approximating models in the candidate set. Mallards apparently did discriminate between types of agriculture, though the evidence for discrimination between types of wetlands was not as strong. The most general model in the candidate set appeared to be the best, which suggested that the data may have been able to support an even more general model. However, the improvement from including the disaggregated wetlands variables was much smaller than the earlier generalizations of the model, which may indicate that the point where the gain in model fit would have been less than the loss in parsimony was imminent. Adding more variables may have soon resulted in an "over-fitted" model.

Table 3.4 – Model selection using the AIC.

Version	AIC	w_i
7N	2783.388	0.7324
6N	2785.402	0.2676
5N	2873.26	0.0000
4N	2887.576	0.0000
3N	2933.132	0.0000
2N	2950.348	0.0000
1N	2990.528	0.0000
7P	3478.986	0.0000
6P	3507.73	0.0000
5P	3760.156	0.0000
4P	3848.192	0.0000
3P	3957.95	0.0000
2P	4072.758	0.0000
1P	4241.906	0.0000

Table 3.5 lists the parameter estimates, standard errors, *t*-ratios and *p*-values for the standard null hypothesis tests (H_0: $\beta = 0$) for the best model, version 7 of the negative binomial (parameter estimates associated with the route dummy variables were not included in the table). Table 3.6 lists the first-order land use type variables, in decreasing order of their estimated effect on mallard abundance. The littoral wetlands variable had the largest estimated coefficient, followed by riverine, rice, deep water, and palustrine wetlands. The *p*-values listed in Table 3.6 suggest that the rice and orchards coefficients were estimated relatively precisely, i.e. their effects on mallard abundances were the most consistent across the observations. The effects of riverine and littoral wetlands were somewhat less consistent. The interpretation that the standard errors are at least partly due to variable effects (and not just sampling error), lends itself to a plausible biological interpretation. Natural wetlands are better for breeding and brood rearing by mallards, partly because they

Table 3.5 – Regression results for version 7 of the negative binomial model.

Variable	Coefficient estimate	Standard error	t - ratio	p - value
Constant	-1.35897	3.34883	-0.40580	0.68489
1998	0.17497	0.56537	0.30949	0.75695
1999	-0.03863	0.26017	-0.14848	0.88197
2000	-0.37168	0.27618	-1.34580	0.17837
Field crops	0.00377	0.01279	0.29440	0.76846
Field crops^2	0.00006	0.00012	0.51379	0.60740
Pasture	0.02496	0.01408	1.77243	0.07632
Pasture^2	-0.00014	0.00016	-0.87594	0.38106
Orchards	-0.06000	0.01836	-3.26726	0.00109
Orchards^2	0.00051	0.00020	2.54676	0.01087
Vineyards	-0.00245	0.08940	-0.02739	0.97815
Vineyards^2	-0.00085	0.00293	-0.28938	0.77229
Dairy	-0.07887	0.04875	-1.61794	0.10568
Dairy^2	0.00156	0.00206	0.75929	0.44768
Urban	-0.03428	0.02622	-1.30736	0.19109
Urban^2	0.00042	0.00030	1.41126	0.15817
Deep water	0.05023	0.14670	0.34241	0.73204
Deep water^2	-0.00033	0.00804	-0.04091	0.96737
Rice	0.05722	0.01562	3.66238	0.00025
Rice^2	-0.00038	0.00017	-2.21440	0.02680
Palustrine	0.03318	0.02923	1.13526	0.25627
Palustrine^2	-0.00011	0.00030	-0.36395	0.71590
Littoral	0.30329	0.12293	2.46724	0.01362
Littoral^2	-0.00650	0.00326	-1.99694	0.04583
Riverine	0.11590	0.04466	2.59495	0.00946
Riverine^2	-0.00156	0.00126	-1.24505	0.21311
Wet*Ag	0.00113	0.00048	2.33330	0.01963
Wet*Rice	-0.00067	0.00034	-1.96139	0.04983
Wet*Urban	0.00127	0.00118	1.08196	0.27927
Wet*Deep water	-0.00713	0.00817	-0.87225	0.38307
SDI	-0.50414	0.38028	-1.32570	0.18494
SEI	0.67971	0.54787	1.24064	0.21474
Precipitation	0.00739	0.03605	0.20511	0.83749
Latitude	0.09547	0.07236	1.31945	0.18702
Day of year	-0.00178	0.01609	-0.11075	0.91181
Average deg F	0.00319	0.02025	0.15757	0.87480
Average wind	0.00115	0.14606	0.00790	0.99370

Table 3.6 – First-order land use variables from version 7 of the negative binomial model, listed in order of decreasing effect size.

Variable	Coefficient	Standard error	t - ratio	p - value
Littoral	0.30329	0.12293	2.46724	0.01362
Riverine	0.11590	0.04466	2.59495	0.00946
Rice	0.05722	0.01562	3.66238	0.00025
Deep water	0.05023	0.14670	0.34241	0.73204
Palustrine	0.03318	0.02923	1.13526	0.25627
Pasture	0.02496	0.01408	1.77243	0.07632
Field crops	0.00377	0.01279	0.29440	0.76846
Vineyards	-0.00245	0.08940	-0.02739	0.97815
Urban	-0.03428	0.02622	-1.30736	0.19109
Orchards	-0.06000	0.01836	-3.26726	0.00109
Dairy	-0.07887	0.04875	-1.61794	0.10568

will generally have more open water than rice fields. However, rice fields are more likely to have water on them through much of the breeding season because they are irrigated. Thus, the (generally) larger effect sizes of the wetland variables, but lower "consistencies" (higher standard errors), can be explained by the naturally low precipitation in the Central Valley in the summer months.

Table 3.3 also shows that in the simpler models, where wetlands entered as a single aggregate type, mallards appeared to demonstrate a consistent affinity for wetlands in general. However, the best model, with wetlands disaggregated, suggested that the strong positive effects of littoral and riverine wetlands may have been responsible for much of that effect. The evidence in support of differential effects across wetlands types in this dataset was not overwhelming (w_{7N}/w_{6N} = 0.7324/0.2676 = 2.74), but the same reasoning used to interpret the differences in standard errors between the wetlands and rice variables can be used to interpret the differences across wetland types. The palustrine wetlands class includes a variety of wetland types, but consists mostly of small wetlands isolated from larger bodies of water (Cowardin et al. 1979). In the Central Valley, most palustrine wetlands will be seasonal, with standing water present during the rainy season (November through March) and for some time after, but usually not through the entirety of the breeding season (March through July). Littoral and riverine wetlands, on the other hand, are associated with permanent bodies of deep water – lakes and rivers – and so will sustain more reliable breeding conditions for mallards throughout the summer.

Because the evidence for differential effects across wetlands types was not strong, and because an estimate of the "average" effect of wetlands was required for the purpose of predicting impacts from wetlands restoration for the optimization applications presented in Chapter 6, I set the results from version 7 aside and focused on results from version 6 of the negative binomial model. Table 3.7 lists the parameter estimates, standard errors, t-ratios and p-values for version 6 of the negative binomial model, and Table 3.8 lists the first-order land use type variables, in decreasing order of their estimated effect on mallard abundance. The results were broadly consistent with version 7, but they were based on the restriction of no differential effects across wetland types.

Table 3.7 – Regression results for version 6 of the negative binomial model.

Variable	Coefficient estimate	Standard error	*t* - ratio	*p* - value
Constant	-1.70180	3.37834	-0.50374	0.61444
1998	0.21187	0.56478	0.37513	0.70757
1999	-0.05557	0.26080	-0.21307	0.83128
2000	-0.36322	0.27433	-1.32403	0.18549
Field crops	0.00256	0.01293	0.19830	0.84282
Field crops^2	0.00008	0.00012	0.65620	0.51170
Pasture	0.02185	0.01422	1.53695	0.12431
Pasture^2	-0.00010	0.00017	-0.60884	0.54263
Orchards	-0.05739	0.01895	-3.02870	0.00246
Orchards^2	0.00050	0.00020	2.45239	0.01419
Vineyards	-0.00457	0.09171	-0.04986	0.96023
Vineyards^2	-0.00079	0.00301	-0.26159	0.79364
Dairy	-0.08250	0.04886	-1.68836	0.09134
Dairy^2	0.00162	0.00201	0.80443	0.42115
Urban	-0.03133	0.02623	-1.19443	0.23231
Urban^2	0.00041	0.00030	1.35215	0.17633
Deep water	0.06947	0.13823	0.50255	0.61528
Deep water^2	-0.00121	0.00804	-0.15035	0.88049
Rice	0.05093	0.01557	3.27072	0.00107
Rice^2	-0.00032	0.00017	-1.89992	0.05744
Wetlands	0.11803	0.04680	2.52190	0.01167
Wetlands^2	-0.00085	0.00046	-1.85541	0.06354
Wet*Ag	0.00031	0.00068	0.45849	0.64660
Wet*Rice	-0.00115	0.00049	-2.37986	0.01732
Wet*Urban	0.00070	0.00119	0.59219	0.55373
Wet*Deep water	-0.00948	0.00523	-1.81411	0.06966
SDI	-0.45609	0.38670	-1.17944	0.23822
SEI	0.85791	0.52080	1.64730	0.09950
Precipitation	0.00434	0.03596	0.12073	0.90390
Latitude	0.08968	0.07438	1.20563	0.22796
Day of year	0.00003	0.01629	0.00155	0.99876
Average deg F	0.00285	0.02035	0.13988	0.88875
Average wind	-0.00078	0.14668	-0.00534	0.99574

Table 3.8 – First-order land use variables from version 6 of the negative binomial model, listed in order of decreasing effect size.

Variable	Coefficient	Standard error	t - ratio	p - value
Wetlands	0.11803	0.04680	2.52190	0.01167
Deep water	0.06947	0.13823	0.50255	0.61528
Rice	0.05093	0.01557	3.27072	0.00107
Pasture	0.02185	0.01422	1.53695	0.12431
Field crops	0.00256	0.01293	0.19830	0.84282
Vineyards	-0.00457	0.09171	-0.04986	0.96023
Urban	-0.03133	0.02623	-1.19443	0.23231
Orchards	-0.05739	0.01895	-3.02870	0.00246
Dairy	-0.08250	0.04886	-1.68836	0.09134

In the more general versions of the regression models, there were first-order, second-order, and interactive effects embedded within the exponential mean function. Because it is often difficult to interpret variables that enter regression models nonlinearly and multiple times (the coefficient estimate for the first-order term tells only part of the story), I calculated the overall marginal effects (the "marginals") to get a better idea of the influence of each land use variable. The marginal for variable k, M_k, was calculated as follows:

$$M_k = \frac{\sum_i \frac{\partial \hat{\mu}_i}{\partial x_{ik}}}{\sum_i y_i} \times 100 \qquad (3.5)$$

In equation 3.5, $\hat{\mu}_i$ is the predicted mean for observation i, x_{ik} is the percent of land in land use type k for observation i, and y_i is the count for observation i. The marginals are the expected percent changes in total mallard abundance from a one unit increase in each land use type across all route-stops. Because the units of the land use variables were in percent, the effect estimated by M_k was the expected percent change in the total count of mallards from (hypothetically) converting one percent of the area near BBS route-stops from natural uplands, the excluded land use type, to land use type *k*. I bootstrapped the regression model to estimate the standard errors of the marginals: the model was estimated 200 times, each time based on a different random draw from the raw dataset, with replacement and maintaining the original sample size. The estimated marginals, which are listed in Table 3.9, paint a slightly different picture of the effects of the various land use types than the coefficients on the first-order terms in Table 3.8. Wetlands is still seen to have the largest positive effect and orchards the largest (consistently) negative effect, but pasture appears to have a larger positive effect and rice appears to have a smaller effect than those suggested by the coefficients on their first-order terms.

Table 3.9 – Marginals from version 6 of the negative binomial model.

Variable	Average	Standard deviation	95% confidence interval
Wetlands	5.81	1.68	(1.638 , 9.929)
Pasture	4.30	1.59	(0.420 , 7.784)
Vineyards	2.70	3.06	(-6.739 , 8.710)
Field crops	2.31	1.49	(-1.041 , 6.857)
Rice	0.97	1.20	(-2.000 , 4.417)
Deep water	0.77	5.85	(-4.181 , 2.739)
Urban	-0.19	1.79	(-5.346 , 4.464)
Orchards	-4.36	1.09	(-7.968 , -1.727)
Dairy	-4.83	3.24	(-17.928 , 1.919)

3.2.2.4 Drawing out spatial effects from the regression results

One of the main objectives of this research was to investigate the importance of spatial effects. Towards that end, the results of version 6 of the negative binomial model were used to address the question of returns to scale with respect to the percent of land in wetlands. (This is one of the factors that determine the benefits of few large wetland patches vs. many small wetland patches.) The returns to the percent of land in wetlands is the expected change in average mallard abundance from a marginal increase in the percent of land in wetlands, and can be estimated by differentiating the exponential mean function with respect to the wetlands variable:

$$\frac{\partial \hat{\mu}}{\partial x_w} = \hat{\mu}\left(\hat{\beta}_w + 2\hat{\beta}_{w2} x_w + \hat{\beta}_{wa} x_a + \hat{\beta}_{wr} x_r + \hat{\beta}_{wu} x_u + \hat{\beta}_{wd} x_d\right) \tag{3.6}$$

In equation 3.6, the *x*'s are the percent of land in wetlands, agriculture, rice, urban, and deep water, and the $\hat{\beta}$'s are the associated parameter estimates. Note that the derivative in equation 3.6 appears in the expression for the marginal (equation 3.5), so the two concepts are closely related. The marginal for wetlands is the expected change in the total count of mallards after an incremental increase in wetlands at all route-stops; it is intended as an overall, or average effect. The derivative, on the other hand, focuses on a particular location. The regression model implies that returns to the percent cover in wetlands depends on the amount of other land use types nearby; the derivative can be different at every location.

Figure 3.5 shows a graph of $\frac{\partial \hat{\mu}}{\partial x_w}$ vs. x_w across the entire range 0 to 100%.

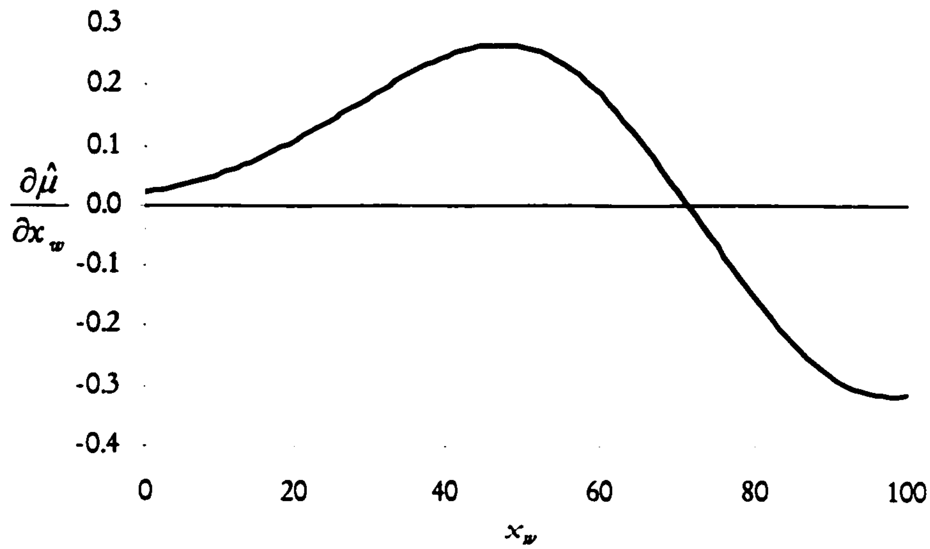


Figure 3.5 – The derivative of expected abundance with respect to the percent of land in wetlands.

To plot the graph in Figure 3.5, the derivative was calculated using the parameter estimates from version 6 of the negative binomial model, and the percent of land in non-wetland land use types were generated to satisfy the required adding-up condition – the sum of all x's had to equal 100%. As x_w ranged from 0% to 100%, the amounts of the other land use types were set so that the relative ratios were equal to the relative ratios at their mean values. For example, the average percent of urban land near the BBS route-stops was $\bar{x}_u = 4.9\%$ and the average percent of agricultural land was $\bar{x}_a = 59.6\%$, so $\bar{x}_u / \bar{x}_a = 0.082$. To generate the curve in Figure 3.5, I adjusted x_u and x_a as x_w varied from 0% to 100% in such a way that the ratio $x_u / x_a = 0.082$ was maintained. The other land use percents were calculated in a similar fashion and all were scaled so that their sum always equaled 100%.

Figure 3.5 suggests increasing returns from $x_w = 0$ to $x_w = 47\%$, still positive but decreasing returns up to $x_w = 71\%$, and negative returns for $x_w > 71\%$. This implies that small patches of wetlands scattered throughout the landscape, or more generally wetlands with a high edge to area ratio, should support the highest densities of breeding mallards. For example, multiple patches of 36 hectares each (71% of 50 hectares), distributed such that the overall mix of land use types includes 71% wetlands and 29% uplands, would maximize total expected abundance. This, however, is not the only way to design a landscape with the highest possible breeding mallard density; any landscape with 71% of each circular 50-hectare scene in wetlands would do. For example, larger wetlands with complex shapes, i.e. a sufficiently high edge to area ratio, could also satisfy this condition.

In Figure 3.6, the averages of the mallard counts are plotted against the percent of land in wetlands (panel A) and the percent in rice (panel B), using intervals of 10%. These simple plots support the general conclusion that mallards prefer a mix of wet and dry land use types. The graphs show that on average the highest mallard abundances were observed at route-stops with intermediate amounts of wetlands or rice nearby. There were relatively few observations with more than 50% of the land in wetlands, but taken together these plots are consistent with the implication of the negative binomial regression model shown in Figure 3.5.[23] These result are easily interpreted in light of the fact that breeding mallards require upland areas for nesting, while wetland and rice habitats are preferable for feeding and brood rearing.

[23] The plots in Figure 3.4 do not represent information independent of that used to construct Figure 3.4, of course. These plots are presented simply to show that the "clean" results implied by the regression model (which describes relationships between the dependent variable and each independent variable when all of the other independent variables are held constant) is suggested even in the "noisy" raw data itself.

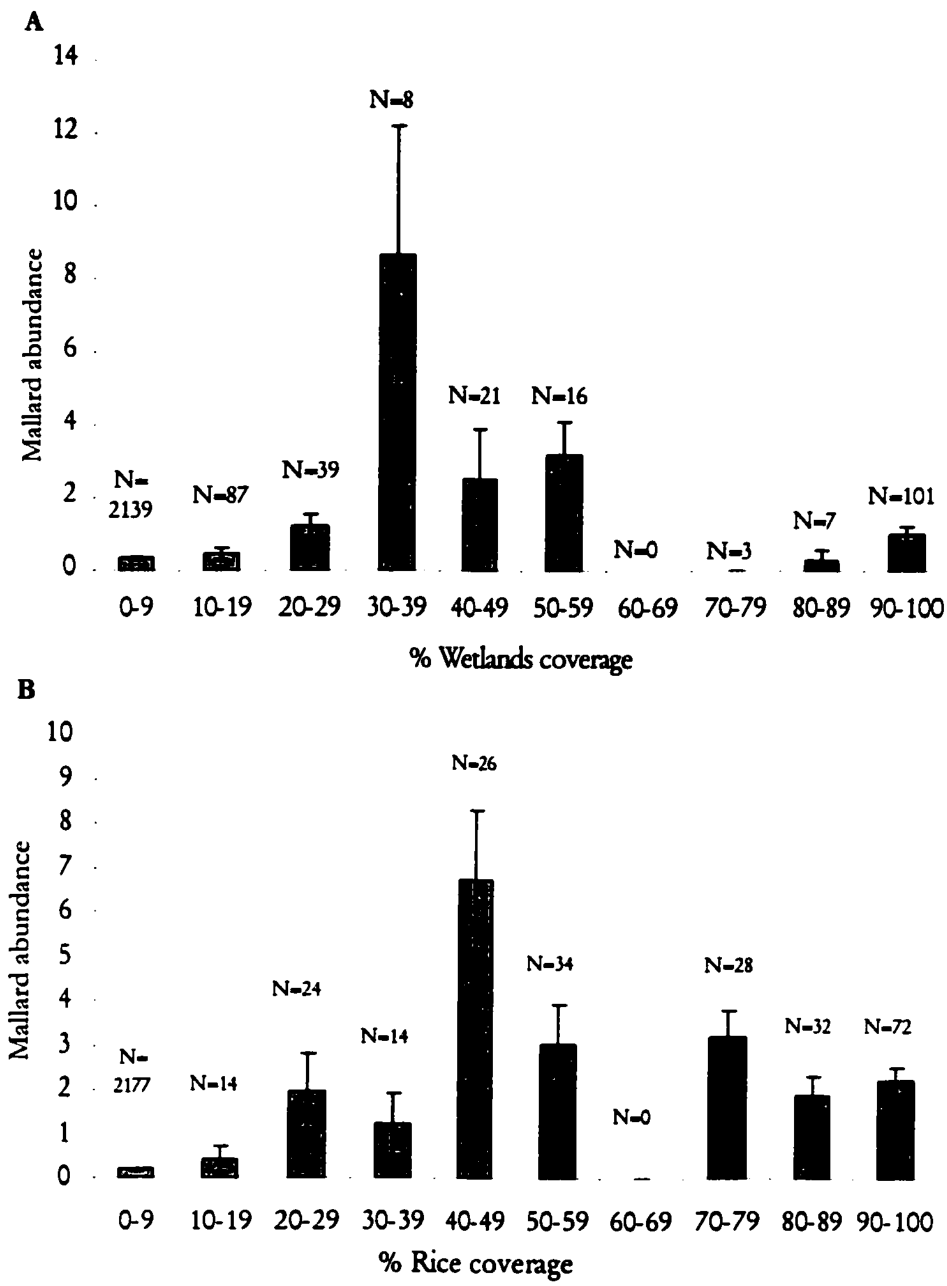


Figure 3.6 - The highest abundances occurred at route-stops with intermediate levels of wetlands or rice.

3.2.3 Management significance of the spatial effects

The regression results suggest that mallards prefer wet lands to dry in the breeding season, but that configuration also matters. But the regression results do not necessarily imply that the spatial effects are significant for management purposes. One way that the regression results could be significant for management is if the spatial effects were strong enough to make a spatially targeted wetlands restoration strategy significantly more effective than a non-targeted strategy. To measure the strength of the spatial effects from this management perspective, I simulated three wetlands restoration strategies. One was a spatially targeted strategy and the other two were not. The simulations were intended to represent a manager trying to decide how many hectares of wetlands to restore around each BBS route-stop in the Central Valley. The spatially targeted strategy (the optimal scenario) was simulated by solving the following optimization problem:

$$\underset{R_i}{Max} \left[\sum_i \exp\left(x_i' \hat{\beta}\right)\right] \tag{3.7a}$$

Subject to:

$$\sum_i \frac{R_i \times \pi(400m)^2}{10^6} \leq 1{,}563 \text{ hectares} \tag{3.7b}$$

$$R_i \leq \sum_k x_{i,k} \quad \forall i \tag{3.7c}$$

$$x_{i,k}' = x_{i,k}\left(1 - \frac{R_i}{\sum_k x_{i,k}}\right) \quad \forall k \tag{3.7d}$$

In equations 3.7a through 3.7d: i indexes the route-stops; the R_i's are the choice variables, the percent of the area around each route-stop restored to wetlands; x_i' is the modified vector of land use variables after restoration; and $\hat{\beta}$ is the vector of parameter estimates from version 6 of the negative binomial regression model. Expression 3.7a is the objective function, the total expected abundance of breeding mallards across all route-stops. The first constraint, equation 3.7b, limits the total area of wetlands restored to an arbitrarily set 50% increase, from the current 7.6% to 11.4% of the total area within 400 meters of the route-stops (1,563 hectares). The second set of constraints (of which there are 819, one for each route-stop in the Central Valley) requires that the amount of wetlands restored at each location be no greater than the total amount of the "restorable" land use types present initially, indexed by k. Urban land was assumed prohibitively expensive to purchase and restore, so they were excluded from the restorable types. Deep water and rice were also excluded. The final set of constraints, equation 3.7d, ensured that wetlands were restored proportionally from the restorable land use types. For example, if five hectares were restored to wetlands at a particular route-stop and there were 25 total hectares of the five restorable land use types initially, then 20% of each was converted to wetlands.

Two sub-optimal scenarios were simulated for comparison. These scenarios were intended to approximate possible outcomes from wetlands restoration strategies not based on a spatially explicit model of habitat selection. The first scenario was a relatively uniform distribution of restoration – a similar amount of restoration at each route-stop. The uniform scenario was simulated by solving the optimization problem with the added constraint that no more than 5.1% of the land at any route-stop could be restored to wetlands (5.1% just happened to be the figure that led to a total restored area of 1,563 hectares, the assumed 50% maximum increase). The resulting distribution of restoration activities was not perfectly uniform, however, because some route-stops initially had less than 5.1% of restorable land (at these locations the restored area was constrained by the initial availability of the restorable land use types), and because no minimum constraint was imposed. The second scenario used for comparison was a clumped distribution of restoration activities, which might result from a manager choosing the largest available sites, irrespective of their location. The clumped scenario was simulated by selecting route-stops at random and converting all restorable agriculture at each of these route-stops to wetlands until the 1,563-hectare limit was reached.

To account for the uncertainty in the parameter estimates, I performed each of the scenarios as a set of Monte Carlo trials. Under the optimal scenario, the hypothetical manager used (imperfect) information regarding the spatial habitat preferences of mallards from the negative binomial regression results. In every replication the manager used the same set of parameter estimates to predict the impacts of wetlands restoration, so the manager undertook the same levels of restoration activities every time. What changed across the replications was the "true beta," the vector of parameters the manager was trying to estimate with the regression model. The true beta's were defined for each replication by a random draw from the multivariate normal distribution implied by the estimated parameters and covariance matrix from the negative binomial regression results. Under both the uniform and randomly clumped scenarios, the hypothetical manager used no information regarding the spatial habitat preferences of mallards.

Table 3.10 lists the results from the three scenarios. All three scenarios yielded substantial increases in total expected abundance, but the spatially targeted strategy greatly outperformed the uniform and clumped strategies. Selecting restoration sites by accounting for spatial habitat preferences was predicted to increase mallard abundances much more than

Table 3.10 – Results from the restoration simulations.

Scenario	Expected total abundance
Initial	359.5 (45.0)[1]
Optimal	1,502 (1,160)[2]
Uniform	470.5 (58.5)[2]
Clumped	514.8 (87.9)[2]

Notes:

1. Standard deviation of results from 100 trials of resampling from the observed counts N times with replacement and summing them.

2. Standard deviation of results from 100 optimization trials, each trial using a draw from the implied distribution of parameter estimates from the regression results.

proportionally; a 50% increase in wetland area led to a 318% increase in total abundance. The sub-optimal scenarios were predicted to lead to less-than-proportional increases – 31% and 43% for uniform and clumped restoration respectively.[24] The predicted average increase under the optimal scenario was surprisingly large, but the results are at least consistent with the data in the following sense. The observed abundances ranged between zero (87% of the observations) and 31 (one observation), and less than 1% (19 of the 2,535 observations) were 10 or greater. The largest predicted expected abundance under optimal restoration was 19.79, which was within the range of the data, though well above the average. Approximately 27% of the route-stops (226 of 819) were predicted to have expected abundances of 10 or greater. These results suggest that there are many good opportunities

[24] Do not let the large standard deviation of the simulated optimal total expected abundance in Table 3.10 fool you. The variation was high, but when comparing the strategies a more relevant measure is the difference between the total expected abundances for each run, which is not shown in the table. In no runs did the uniform strategy outperform the optimal strategy, and in only one run did the clumped strategy outperform the optimal strategy.

for effective wetlands restoration in the Central Valley, and that the spatial component of habitat selection by mallards is substantial.

The simulation exercises were intended to measure the potential increase in effectiveness a decision maker could expect by using a spatially targeted approach over a merely opportunistic approach to selecting sites for wetlands restoration. If the regression models had suggested that mallard abundances were a simple function of the amount of each land use type, with no significant spatial effects, then there would have been no differences in the results from the three simulated scenarios. If mallard abundances were influenced only by the amount of each land use type present, then each of the simulated management scenarios would have predicted the same total abundance after restoration, because only the arrangement, not the total amount, of restoration differed across the scenarios. In Section 3.2.2.4, I argued that the regression results showed that spatial effects influence mallard abundances. The results from the simulations showed that these spatial effects could be significant for management purposes.

3.2.4 Do mallards look up from their ponds?

All results presented so far were based on models that assumed only nearby land use characteristics – within 400 meters – influenced mallard abundances. The 400-meter cutoff was based on the BBS data collection protocols; surveyors were instructed to count all birds they could identify within 400 meters of their location. But this only suggests that the 400-meter distance was a reasonable *minimum* distance within which to measure land use characteristics. If mallards used resources far from their nest sites, then they might judge site quality by more than just very local landscape characteristics.

To investigate this possibility, I estimated a set of three more negative binomial regression models, each an extension of the version 6 model discussed earlier. The first of these expanded models added land use characteristics between 400 and 700 meters from each route stop. The second added, on top of those, land use characteristics between 700 and 1000 meters from each route stop, and the third added land use characteristics 1000 to 2000 meters away.[25] My expectation was that the effects of the land use characteristics would rapidly diminish as the distance from the route-stops increased, and that any significant effects would be in the same direction as the nearby land use variables but smaller in magnitude.

The results, which are shown in Tables 3.11 and 3.12, did not conform to my expectations. Table 3.11 shows the Akaike model selection statistics for the baseline model, with only the 0-400 meter land use variables included, plus the three expanded models. The most general model, with three extra rings of land use characteristics included, scored best by the Akaike criteria, which implies that the extra land use variables added substantial explanatory power. Mallards apparently do look up from their pond, and they look at least two kilometers away. Table 3.12 shows the parameter estimates and standard errors for the most general model (for the land use variables only). The most general model explained more of the variation in the data than the more restricted versions, but the parameters were estimated much less precisely (which was likely a result of high colinearity between the spatially lagged land use variables). The *t*-statistics were much lower than in the previous models, and notably so for the 0-400 meter land use variables. The magnitudes of the effects did not diminish consistently with distance, and their signs often changed. This

---

[25] In the second and third versions, I had to exclude the second-order land use variables because the negative binomial model would not converge with all variables included, presumably because of excessive multicolinearity.

preliminary investigation into a spatially lagged model of species-habitat relations, then, yielded equivocal results. Land use characteristics at a distance do seem to matter, but the regression results do not reveal much about the structure of these effects.

Table 3.11 – Overall regression model results for the baseline model plus three models with spatially lagged land use variables. The model with all lagged land use variables has the largest Akaike weight

Variables included	$r^2_{y,\hat{\mu}}$	AIC	w_i
0-400 meters only	0.2888	2771.99	0.0000
With 400-700 meters	0.3599	2750.426	0.0111
With 700-1000 meters	0.3795	2747.65	0.0443
With 1000-2000 meters	0.371	2741.53	0.9447

<u>Table 3.12</u> – Coefficient estimates and *t*-statistics for the negative binomial model with three sets of spatially lagged land use variables. Virtually no variables were estimated with precision, and some of the key variables change sign with increasing distance from the survey location.

	0 - 400 meters		400 - 700 meters		700 - 1000 meters		1000 - 2000 meters	
Variable	$\hat{\beta}$	*t* - stat	$\hat{\beta}$	*t* - stat	$\hat{\beta}$	*t* - stat	$\hat{\beta}$	*t* - stat
Field crops	-0.03025	-1.20804	0.06301	1.76241	-0.01574	-0.50158	0.01477	0.84625
Field crops^2	0.00029	1.31077	-0.0003	-1.24417				
Pasture	-0.02951	-1.06751	0.09127	2.24163	0.00204	0.05189	-0.01266	-0.4657
Pasture^2	0.00022	0.75381	-0.0007	-1.36484				
Orchards	0.00256	0.0675	0.00259	0.05208	-0.036	-0.90239	0.00649	0.27617
Orchards^2	0.00000	-0.00547	0.00034	0.72573				
Vineyards	-0.01964	-0.13521	0.06416	0.46401	0.00528	0.04618	-0.07235	-0.99685
Vineyards^2	-0.00102	-0.24849	0.00022	0.09229				
Dairy	-0.13949	-1.71744	0.23711	2.29745	-0.01688	-0.22339	0.065	0.71444
Dairy^2	0.00421	1.12614	-0.00933	-1.23783				
Urban	-0.06171	-1.28564	0.07936	1.31072	0.00585	0.08656	-0.03977	-0.54685
Urban^2	0.00068	0.8713	-0.00068	-0.75996				
Deep water	-0.16491	-1.00282	0.06751	0.51924	0.18638	1.63304	-0.07579	-1.09488
Deep water^2	0.00822	1.01534	-0.00446	-1.45097				
Rice	0.05195	1.22431	-0.01344	-0.2705	0.00346	0.08728	0.01407	0.58176
Rice^2	-0.00055	-1.3678	0.00042	0.9983				
Wetlands	-0.03252	-0.32229	0.15643	1.4978	0.00049	0.01043	0.03406	1.06305
Wetlands^2	0.0003	0.2946	-0.00129	-1.19622				
Wet*Ag	0.00156	1.1683	-0.00106	-0.82629	-0.00036	-0.70379	0.00055	1.64211
Wet*Rice	-0.00048	-0.44433	-0.00007	-0.0481	-0.00046	-0.49875	0.00022	0.35663
Wet*Urban	0.00101	0.52774	-0.00357	-1.39528	0.00054	0.26553	-0.0022	-0.5635
Wet*Deep water	-0.00288	-0.39408	-0.00298	-0.3932	-0.0046	-0.87647	0.00434	0.90212
SDI	-0.1913	-0.27129	-0.02753	-0.03842	0.00033	0.01118	-0.23772	-0.43243
SEI	1.24161	1.52459	-1.11922	-1.09643	-0.10054	-0.81382	-0.2336	-0.18554

3.3 Regression models for other bird species

The BBS dataset contains counts of all resident breeding bird species in North America. This research focused on mallards, but similar analyses could be undertaken for many more species in the BBS dataset. Not all species of interest are amenable to the kind of statistical modeling used here, simply because many of the species that may be of most interest from a conservation perspective are rare, and therefore would not be observed frequently enough by BBS surveyors to support statistical modeling. However, some species will be observed frequently enough, and some of these could potentially be used as indicators for other species of concern. This section presents preliminary results for 21 other bird species, plus overall bird richness and diversity. These species were selected either because they depend on wetlands for at least part of their life cycle, or because they were observed frequently by BBS surveyors and therefore should be relatively easy to model.

I estimated versions 1, 2, 4, and 6 of the negative binomial model for each species and bird richness, and a linear regression model for bird diversity. The results are shown in Tables 3.13 and 3.14. The measures of fit listed in Table 3.13 suggest that mallards are among the easiest species to model, but the models explained a substantial amount of the variation in abundances for several other species as well. The squared correlations between the predicted and observed counts for black-crowned night heron (a non-game wetland species), ring-necked pheasant (another important game species), song sparrow (a land bird species of special concern in the valley[26]), spotted towhee (a ground-nesting land bird), and richness and diversity were all above 0.35 (the value for mallards was 0.29). Table 3.14 summarizes the qualitative results from the regression models.

[26] CDFG and PRBO. 2001. California Bird Species of Special Concern: Draft List and Solicitation of Input. http://www.prbo.org/BSSC/draftBSSClist.pdf.

Table 3.13 – Results from four versions of the negative binomial model applied to 22 species and bird richness, and a linear regression model applied to bird diversity, from the BBS database.

	Version 1		Version 2		Version 4		Version 6	
	p	r^2	p	r^2	p	r^2	p	r^2
American robin	0.0000	0.078	0.0000	0.092	0.0000	0.160	0.0000	0.226
Black-crowned night heron	0.0000	0.641	0.0000	0.893				
Brown-headed cowbird	0.0000	0.018	0.0000	0.034	0.0062	0.056	0.0000	0.119
Black phoebe	0.0000	0.029	0.0000	0.029	0.3784	0.059	0.0000	0.108
Barn swallow	0.0000	0.010	0.0000	0.014	0.0052	0.011	0.0000	0.042
Brewer's blackbird	0.0000	0.056	0.0000	0.057	0.0011	0.062	0.0000	0.142
European starling	0.0000	0.045	0.0000	0.055	0.0000	0.072	0.0000	0.082
Great blue heron	0.0000	0.103	0.0000	0.088	0.0167	0.104	0.0002	0.140
Great egret	0.0000	0.103	0.0000	0.054	0.3643	0.075	0.0000	0.158
Horned lark	0.0000	0.035						
Killdeer	0.0000	0.131	0.0000	0.131	0.0082	0.139	0.0000	0.204
Loggerhead shrike	0.0000	0.036	0.0000	0.077	0.9661	0.071	0.0001	0.103
Mallard	0.0000	0.145	0.0000	0.144	0.0000	0.220	0.0000	0.291
Mourning dove	0.0000	0.054	0.0000	0.066	0.0365	0.064	0.0000	0.089
Norhtern mockingbird	0.0000	0.041	0.0000	0.054	0.0000	0.075	0.0000	0.132
Nutall's woodpecker	0.0000	0.089	0.0002	0.106				
Rock dove	0.0000	0.007	0.0000	0.010	0.0015	0.006	0.0000	0.016
Ring-necked pheasant	0.0000	0.272	0.0000	0.368	0.0000	0.377	0.0000	0.398
Red-tailed hawk	0.0000	0.055	0.0000	0.064	0.1147	0.078	0.0032	0.093
Song sparrow	0.0000	0.422	0.0000	0.495	0.0025	0.504	0.0000	0.605
Spotted towhee	0.0000	0.164	0.0000	0.207	0.0032	0.242	0.0000	0.465
Tri-colored blackbird	0.0000	0.026	0.0000	0.012	0.0550	0.013		
Bird richness	0.0000	0.220	0.0000	0.262	0.0000	0.310	0.0000	0.476
Bird diversity	0.0000	0.153	0.0000	0.175	0.0000	0.219	0.0000	0.346

Notes:

The columns with heading p contain p-values associated with log likelihood ratio tests that compared the model with the next most restricted one. For example, the p-value of 0.0052 for the version 4 model for Barn swallows indicates that there was a 0.52% chance that the ratio of the log likelihood values for the version 4 and the version 2 models (the next most restricted one in this case) would be observed under the null hypothesis that the version 2 model was correct. The test was based on a Chi-squared distribution with N and J degrees of freedom, where N was the sample size and J was the number of restrictions that distinguish the version 2 and 4 models.

The columns with heading r^2 are the squared correlations between the observed counts and the predicted means.

The blank cells in the table indicate that the negative binomial model would not converge.

Table 3.14 -- Summary of qualitative results from four versions of the negative binomial regression model applied to 22 species in the BBS dataset. Column 3 shows the variables that were statistically significant at the 1% (bold), 5% (regular), and 10% (italics) levels.

Species	$r^2_{y,\hat{\mu}}$ Ver. 6	Statistically significant variables
American robin	0.226	VI: Orchards+, **SEI-** IV: Orchards+, SDI+, **SEI-** II: Ag+, **SEI-** I: Ag+, Urban+
Black-crowned night heron	0.893	II: Deep water+, **Rice-**, *Wetlands+*; I: **Ag-**, **Urban-**, Deep water+
Brown-headed cowbird	0.119	VI: SDI+ IV: Vineyards+, Rice+, SDI+ II: SDI+ I: Ag+, Deep water+, Wetlands+
Black phoebe	0.108	II: **SEI-** I: **Ag-**
Barn swallow	0.042	II: Ag+ I:Ag+, Rice+, Wetlands+
Brewers blackbird	0.142	IV: Field crops+, Vineyards+, Orchards+, Dairy+ II: Ag+, **Rice-**
European starling	0.082	VI: Dairy+, Urban+ IV: **Rice-** II: **Deep water-**, **Rice-**, **Wetlands-**, SDI+ I: **Rice-**, **Wetlands-**
Great blue heron	0.140	I: Deep water+, Rice+
Great egret	0.158	IV: Deep water+, Rice+ II: Deep water+, Rice+ I: **Ag-**, Deep water+, Rice+
Horned lark	0.035	
Killdeer	0.204	I: **Urban-**, Deep water+, Rice+, Wetlands+
Loggerhead shrike	0.103	VI: *Field crops*(+), SEI+ IV: Orchards-, SEI+ II: Ag+, **SDI-**, SEI+ I: **Ag-**, **Urban-**
Mallard	0.291	VI: **Orchards-**, Rice+, Wetlands+, **Wet*Rice-** IV: **Orchards-**, **Urban-**, Rice+, Wetlands+, **Wet*Rice-**, **Wet*Water-** II:**Urban-**, Rice+, Wetlands+, SEI+ I: Ag+, Rice+, Wetlands+
Mourning dove	0.089	VI: SEI- IV: **Pasture-** II: Ag-, SEI- I: **Ag-**, **Urban-**, **Rice-**, **Wetlands-**

Table 3.14 (continued) – Summary of qualitative results from four versions of the negative binomial regression model applied to 22 species in the BBS dataset. Column 3 shows the variables that were statistically significant at the 1% (bold), 5% (regular), and 10% (italics) levels.

Species	$r^2_{y,\hat{\mu}}$ Ver. 6	Statistically significant variables
Northern mockingbird	0.132	VI: Wet*Urban+ IV: Vineyards+, Dairy+, Wet*Urban+, SDI+, **SEI-** II: **Wetlands-**, SDI+, **SEI-** I: Urban+, **Rice-**, **Wetlands-**
Nutall's woodpecker	0.106	I: **Rice-**
Rock dove	0.016	I: Ag+
Ring-necked pheasant	0.398	VI: Pasture+, SEI+ IV: Pasture+, SEI+ II: **Ag-**, **Urban-**, Rice+, Wetlands+, SEI+ I: Ag+, Rice+, Wetlands+
Red-tailed hawk	0.093	II: **Urban-** I: **Ag-**, **Urban-**, **Rice-**, **Wetlands-**
Song sparrow	0.605	IV: Vineyards+ II: Ag+, Rice+ I: Ag+, Wetlands+
Spotted towhee	0.465	II: SDI+ I: Wetlands+
Tricolored blackbird	0.026	
Bird Richness	0.476	VI: SDI+ IV: **Urban-**, Wetlands+, **Wet*Ag-**, **Wet*Rice-**, **Wet*Water-**, SDI+ II: **Urban-**, Wetlands+, SDI+ I: Deep water+, Wetlands+, Ag+
Bird Diversity	0.346	VI: SDI+ IV: **Field crops-**, **Urban-**, **Wet*Water-**, SDI+, SEI- II: **Urban-**, SDI+, **SEI-** I: Ag+, **Rice-**, Wetlands+

3.4 An ideal free regression model

Regression models of species-habitat relationships can be very informative, but they do have their limitations. One important limitation is the fact that they cannot be used to predict changes in total population size over time. They are strictly applicable only to the question of how individuals in a population *of given size* will distribute themselves across the landscape. Using regression models of species counts to predict changes in total population size from changes in land use (which was implicit in the calculation of the marginals in Section 3.2.2.3, and explicit in the optimization exercise in Section 3.2.3) requires at least two fairly strong assumptions: (1) only the amount and arrangement of land use types limit the population, and (2) the population is in equilibrium. The first assumption will be immediately suspect for migratory species, which depend on habitat conditions in more than one region, which are often very far apart. The second assumption cannot be maintained if time series data indicate that the total population size is changing.[27] Without sufficient data on distribution and abundance collected over time, observed relationships between species and habitats alone can tell us little about population change or equilibrium population size. However, under certain assumptions, and with some extra information, regression results can be used to predict not only a species' distribution conditional on its current population size, but also its equilibrium distribution and population size, conditional only on observed habitat conditions.

In this section, I describe a model that supports a more ecologically satisfying interpretation of the parameters in a standard regression model, like those from the negative

[27] Average counts of mallards per BBS route-stop in the Central Valley were 0.4194, 0.5429, 0.4209, and 0.3390 for 1997 through 2000. I do not know if this indicates a stable or changing population, but it would in general be useful to have a means of relaxing the equilibrium assumption.

binomial model used extensively in this chapter. Using the model described below, the regression parameters can be interpreted as (combinations of) fundamental population and habitat parameters. I also discuss the extra information required to identify the parameters individually (there is no free lunch here).

The simple model described in this section rests on several important assumptions. The first assumption is that individuals will distribute themselves across the landscape according to an "ideal free distribution" (Fretwell and Lucas 1969, Harper 1982, Pulliam and Danielson 1991). This assumption implies that the share of resources each individual can procure will depend on the number of other individuals in the vicinity, and that each individual will occupy the best available site (the site where the individual can achieve the highest resource procurement rate, given the distribution of other individuals in the landscape). The highest quality sites will be occupied first; then, as the resource procurement rate decreases in those sites as more individuals occupy them, lower quality sites become attractive. Sites will be occupied in decreasing order of the potential resource procurement rate (which will change as more individuals arrive, or as individuals move around). The simplest case would be when the individuals occupying a site effectively share the available resources equally. In this case, the resource procurement rate would be inversely proportional to the density of individuals. The landscape-level implication of the ideal free assumption is that individuals will distribute themselves such that their realized resource procurement rates are equal across the landscape (Pulliam and Danielson 1991). In other words, the individuals in the population will distribute themselves across the landscape so that no arbitrage opportunities from relocating remain.

But that is getting ahead of the story. Let us begin by assuming that individuals attempt to maximize "fitness," which will be determined by the rate at which they can procure resources from the environment. Specifically, assume that the fitness of each individual in the vicinity of location *i*, F_i, is a function of the amount of preferred habitat and the number of individuals nearby:

$$F_i = \exp\left[a + bW_i + cW_i^2 + d\ln y_i\right] \tag{3.8}$$

W_i is the area of wetlands and y_i is the number of individuals in the vicinity of location *i*.[28] I assume that wetlands are the only viable habitat type, though this is for purposes of exposition only. (Note that at this point the ideal free assumption implies that all individuals near location *i* will have the same fitness; shortly I will extend this assumption to the entire population.)

From equation 3.8, the fitness when density equals one is:

$$F_i\big|_{y_i=1} = \exp\left[a + bW_i + cW_i^2\right] \tag{3.9}$$

The parameter *d* affects fitness by determining the intensity of interference between individuals. If $d = -1$, then interference will be proportional, i.e. $F_i = \frac{F_i|_{y_i=1}}{y_i}$. The individuals near location *i* will share the available resources equally. The second individual to arrive will appropriate half of the resources that the first individual was enjoying alone. If $-1 < d < 0$, then interference will be less than proportional, i.e. $\frac{F_i|_{y_i=1}}{y_i} < F_i < F_i|_{y_i=1}$. In this case the second individual to arrive will appropriate less than half of the resources that the

[28] I use the exponential function to facilitate a link with the standard count regression model, but it also turns out to be convenient for the explicit interpretation of fitness that I will introduce shortly, which must be non-negative.

first individual was using.[29] Equation 3.8 also implies that at any density, fitness will be greatest when:

$$\frac{\partial F_i}{\partial W_i} = 0 \Rightarrow W^* = \frac{-b}{2c} \tag{3.10}$$

W^* is the optimal density of wetlands in the landscape. Solving equation 3.8 for y_i gives the abundance at location i:

$$y_i = \exp\left[\frac{\ln F_i - a - bW_i - cW_i^2}{d}\right] \tag{3.11}$$

If individuals distribute themselves according to an ideal free distribution across the entire landscape, the fitness at all locations will be equal: $F_i = \bar{F} \quad \forall i$, which means:

$$y_i = \exp\left[\frac{\ln \bar{F} - a - bW_i - cW_i^2}{d}\right] \tag{3.12}$$

To make the link between the model as developed so far and the standard count regression model, I assume that equation 3.12 determines average abundance, and that actual abundances vary according to a Poisson distribution:

$$\Pr[y_i = y] = \frac{\exp(\mu_i)\mu_i^y}{y!}, \text{ where } \mu_i = \exp\left[\frac{\ln \bar{F} - a - bW_i - cW_i^2}{d}\right] \tag{3.13}$$

Equation 3.13 can be estimated by the same count regression techniques used extensively in Section 3.2.2. In light of equations 3.9 through 3.12, the parameters of the regression model can now be given a more ecologically satisfying interpretation. The regression model is:

$$\mu_i = \exp[\beta_0 + \beta_1 W_i + \beta_2 W_i^2] \tag{3.14}$$

which implies:

---

[29] This "resource sharing" interpretation is not the only possible one. The parameter d can be thought of more generally as determining the "degree to which individuals will interfere with each other," for whatever reason.

$$\beta_0 = \frac{\ln \bar{F} - a}{d} \tag{3.15}$$

$$\beta_1 = \frac{-b}{d} \tag{3.16}$$

$$\beta_2 = \frac{-c}{d} \tag{3.17}$$

Under the assumptions outlined above, the regression parameters can be interpreted as combinations of the population parameter d (which relates to the degree of density dependence exhibited by the species), and habitat parameters a, b, and c (which relate to habitat quality with respect to the species in question). There are five unknowns in equations 3.15 through 3.17, so extra information would be required to solve for all of the fundamental parameters individually. Two pieces of extra information that would be sufficient are: (1) the maximum fitness in wetland habitat, F^{max}, and (2) the maximum fitness in non-wetland habitat, F^0. With this extra information, combining equations 3.9 and 3.10 yields:

$$F^{\max} = \exp\left[a - \frac{b^2}{4c}\right] \tag{3.18}$$

and

$$a = \ln F^0 \tag{3.19}$$

Equations 3.15 through 3.19 can now be solved for all of the fundamental parameters:

$$b = \frac{4\beta_2\left(\ln F^0 - \ln F^{\max}\right)}{\beta_1} \tag{3.20}$$

$$c = \frac{4\beta_2^2\left(\ln F^0 - \ln F^{\max}\right)}{\beta_1^2} \tag{3.21}$$

$$d = \frac{4\beta_2\left(\ln F^{\max} - \ln F^{0}\right)}{\beta_1^2} \tag{3.22}$$

$$\bar{F} = \exp\left[\frac{4\beta_0\beta_2 \ln F^{0}\left(\ln F^{0} - \ln F^{\max}\right)}{\beta_1^2}\right] \tag{3.23}$$

Finally, if fitness is interpreted as a particular combination of survival probability and reproductive success, such that what is being maximized is the number of individuals each breeder contributes to the total population in the next year, then:

$$Y_{t+1} = \bar{F}_t Y_t \tag{3.24}$$

where $Y_t = \sum_i y_{it}$ is the total population size in year t. The estimated average fitness from equation 3.23 can then be used to predict the direction of population change. If $\bar{F}_t > 1$ then the population is currently below the carrying capacity and should increase; if $\bar{F}_t < 1$ then the population is currently above carrying capacity and should decrease. Using this definition of fitness the equilibrium population size can also be estimated. The population size will not change when $\bar{F}_t = 1$, so at equilibrium equation 3.12 gives:

$$y_i^{eq} = \exp\left[\frac{-a - bW_i - cW_i^2}{d}\right] \tag{3.25}$$

and

$$Y^{eq} = \sum_i y_i^{eq} \tag{3.26}$$

Equations 3.25 and 3.26 can be used to estimate the equilibrium spatial distribution and the equilibrium population size for the species, given the current distribution of land use. This can be done based on a cross-sectional analysis of species-habitat relationships, if the maximum fitness in wetlands and uplands can be estimated separately, and if the interpretation of fitness used above is accepted. The fact that no time series information is necessary follows directly from the assumptions regarding density dependence and the particular definition of fitness assumed. These are fairly strong assumptions, but strong assumptions are required to estimate carrying capacity based on cross-sectional data alone.[30]

To illustrate the use of the ideal free interpretation of the standard regression model, I estimated an alternative negative binomial regression model with wetlands and rice combined into one land use variable, "wet lands." This simplification was imposed mostly for convenience, but it is not entirely unjustified. According to the results presented in Section 3.2.2.3, these two land use types influence mallard abundances in (at least qualitatively) similar ways, and there is other evidence that rice can serve as a substitute for wetlands in the Central Valley, especially when little water is available in natural wetlands late in the breeding season (Central Valley Habitat Joint Venture Implementation Board 1990; McLandress et al. 1996; Elphick 2000). The coefficients on the first- and second-order wet lands variables were taken as β_1 and β_2. To estimate β_0, all other coefficients were multiplied by the mean values of their associated variables and summed. F^{max} was set to 3.0 and F^0 was set to 0.5. Ideally these parameters would be estimated by careful experiments or field observations, but that was beyond the scope of this project. These values seem plausible, however, considering the ranges of reproductive parameters in the literature (McLandress et al. 1996, Krapu and others 1997, Dzus and Clark 1998; Sheafer 1998). F^0 = 0.5 implies that with no wet lands in the landscape the mallard population would decrease by 50% every year, and F^{max} = 3.0 implies that the intrinsic growth rate in optimal wet land habitat is 300% per year.

[30] Another way to say this is that strong assumptions are implicit when regression results based on cross-sectional studies are used to make inferences about population level effects of landscape change.

Results from the alternative regression model are given in Table 3.15, and results from the ideal free model are given in Table 3.16 and Figure 3.7. The squared correlation between the observed counts and predicted means for the alternative regression model was 0.2692, which compared favorably with results for the other versions shown in Table 3.3. The AIC statistic was 2,805, which means that it would have ranked third on the list of a priori candidate models (refer back to Table 3.4). Table 3.16 gives the estimates of the fundamental parameters a, b, c, and d, plus the average fitness across the four years in this dataset, $\bar{F}$, and the optimal density of wetlands, W^*. Note that d was between 0 and –1, which implies that interference between individuals was something less than proportional. The total equilibrium abundance estimated using the ideal free model was 3,230 individuals. This was nearly ten times the average total abundance across the four years in the dataset, which was 360 individuals. The estimate of $\bar{F}$, which was well above one, also suggests that the population was far below carrying capacity between 1997 and 2000. These numerical results should be considered only suggestive however, because they are sensitive to the values of F^{max} and F^0. The point of this exercise was mainly to demonstrate the utility of the ideal free model. The model can be used to interpret coefficients estimated from regression models of species-habitat relationships in a more ecologically satisfying manner, and it can allow inference about the carrying capacity of the landscape for the species, even when no data on population change through time is available.

The limitations of the model include the fact that extra information regarding maximum fitness in wetland habitat and average fitness in non-wetland habitat is required. The researcher must also make explicit assumptions about what it is that individuals are trying to maximize, to facilitate estimation of maximum possible fitness in the different

Table 3.15 – Results from the modified negative binomial regression model, with wetlands and rice combined in a single "wet land" variable.

Variable	Coefficient estimate	Standard error	*t* - ratio	*p* - value
Constant	-0.72006	3.39052	-0.21200	0.83180
1998	0.33151	0.55481	0.59800	0.55020
1999	-0.04384	0.25678	-0.17100	0.86440
2000	-0.34600	0.27139	-1.27500	0.20230
Field crops	-0.00046	0.01158	-0.04000	0.96830
Field crops^2	0.00007	0.00011	0.58700	0.55740
Pasture	0.00790	0.01316	0.60000	0.54840
Pasture^2	0.00002	0.00016	0.13700	0.89060
Orchards	-0.05997	0.01613	-3.71800	0.00020
Orchards^2	0.00052	0.00018	2.89700	0.00380
Vineyards	0.00040	0.09600	0.00400	0.99670
Vineyards^2	-0.00086	0.00327	-0.26500	0.79130
Dairy	-0.08461	0.04739	-1.78500	0.07420
Dairy^2	0.00156	0.00194	0.80400	0.42140
Urban	-0.02189	0.01973	-1.10900	0.26720
Urban^2	0.00028	0.00024	1.16600	0.24370
Deep water	-0.01665	0.08994	-0.18500	0.85320
Deep water^2	0.00087	0.00515	0.16900	0.86610
Wet land	0.05792	0.01427	4.05800	0.00000
Wet land^2	-0.00037	0.00014	-2.66000	0.00780
SDI	-0.08256	0.34945	-0.23600	0.81320
SEI	0.84704	0.48612	1.74200	0.08140
Precipitation	0.00084	0.03474	0.02400	0.98070
Latitude	0.08169	0.07538	1.08400	0.27850
Day of year	-0.00519	0.01643	-0.31600	0.75200
Average deg F	0.00130	0.02078	0.06300	0.95010
Average wind	-0.04385	0.14811	-0.29600	0.76720

habitat types, and to specify the nature of population change across time periods. However, if the researcher is willing to make these assumptions, then the model offers the ability to estimate the equilibrium abundance and spatial distribution of individuals under current landscape conditions, and to predict changes from alterations in the landscape. If the un-adjusted parameter estimates from a standard regression model were used to predict total population change from changes in land use, the results would be biased. The bias can be computed by differentiating equations 3.7 and 3.20 with respect to W_i, which gives:

$$\frac{\partial y_i}{\partial W_i} = \frac{\partial y_i^{eq}}{\partial W_i} \exp\left(\frac{\ln \bar{F}}{d}\right) \tag{3.27}$$

According to equation 3.27, if $d < 0$ and the population were above carrying capacity ($\bar{F} < 1$), then the un-adjusted regression parameters would over-estimate the impact of an increase in wetlands. If the population were below carrying capacity ($\bar{F} > 1$), then the un-adjusted parameters would under-estimate the impact of an increase in wetlands. This makes intuitive sense. If one took a snap-shot of species-habitat relationships when the population was above carrying capacity, then it would appear that wetlands could support a higher density of individuals than they actually could at equilibrium, which would lead us to overestimate the expected increase in population size from wetlands restoration. But beyond providing this intuition, the model would allow one to calculate the magnitude of the bias. For example, if $d = -0.75$ and $\bar{F} = 2$, then $\frac{\partial y_i}{\partial W_i} = 0.39 \times \frac{\partial y_i^{eq}}{\partial W_i}$; if $\bar{F} = 0.5$, then

$\frac{\partial y_i}{\partial W_i} = 2.5 \times \frac{\partial y_i^{eq}}{\partial W_i}$.

Table 3.16 – Results from applying the ideal free model to the alternative regression model.

Parameter	Estimated value
F^0	0.5 (assumed)
F^{max}	3 (assumed)
β_0	-2.253
β_1	0.0579
β_2	-0.000365
a	-0.6931
b	0.0452
c	-0.000285
d	-0.781
$\bar{F}$	2.90
W^*	79.3

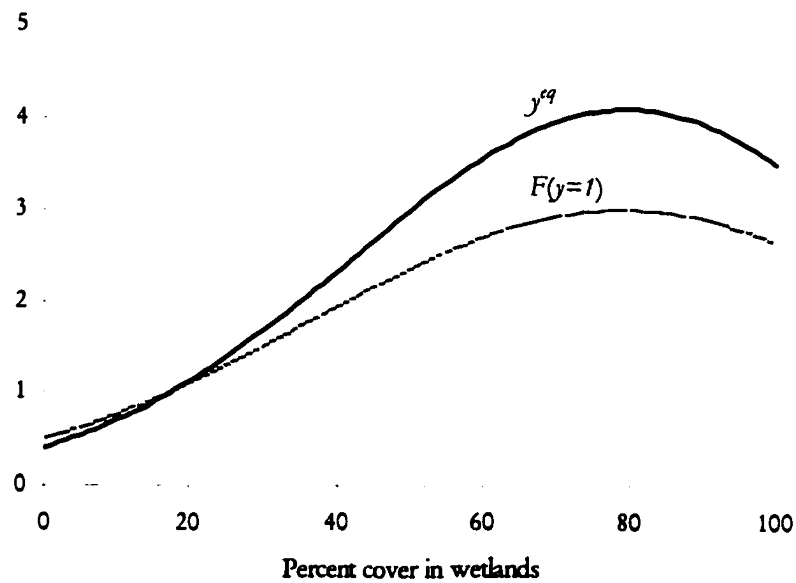


Figure 3.7 – Equilibrium mallard density and fitness when density equals one.

3.5 Conclusions

A lot of ground has been covered in this chapter, so before moving on to the water quality model in Chapter 4 it will be useful to review the main results. This chapter began with a description of a fairly comprehensive analysis of the relationships between breeding mallard abundances and land use in the Central Valley using count regression models. The results from these models showed the importance of "wet lands" – wetlands and rice in particular – for mallards in the breeding season. I interpreted some of the differences in the magnitudes of the parameter estimates and standard errors in light of the fact that some habitats (littoral wetlands, riverine wetlands, and rice) are more likely to be wet for a longer period in the breeding season than others (palustrine wetlands). The regression results also implied non-linear effects of the percent of land in wetlands on expected abundance, and I interpreted this non-linearity as an effect of land use configuration.

I used an optimization model based on the regression results to estimate the potential management significance of the spatial effects. The optimization model predicted that increasing the amount of wetlands near BBS route-stops in the Central Valley by 50% in a spatially targeted manner would result in a more than four-fold increase in expected abundances, while non-targeted strategies would result in increases of 30-40%. Next, I presented preliminary results for 21 other bird species, plus bird richness and diversity. These results suggested that the North American Breeding Bird Survey dataset could potentially provide useful information on habitat relationships for many other bird species, in addition to mallards. Finally, I described a model based on Fretwell and Lucas' "ideal free distribution" that could be used to interpret regression coefficients as combinations of fundamental population and habitat parameters. With some extra information and a few key assumptions, the model could also be used to predict the direction of population change and the equilibrium population size and distribution under current landscape conditions, and the expected change in the equilibrium population size and distribution under changed landscape conditions.

Chapter 4 – Wetlands and water quality

4.1 Pollution abatement in constructed and natural wetlands

In addition to habitat benefits, wetlands conservation is often motivated by the potential water quality benefits they can provide. Natural wetlands can help maintain water quality – and created and restored wetlands can help improve it – by removing pollutants from waters that flow through them, thereby preventing degradation of downstream water bodies (National Research Council 1992; Mitsch and Gosselink 1993; Lewis 2001). The ability of wetlands to remove nutrients, sediment, and other constituents that can impact water quality, and the mechanisms by which they do so, is well documented (Johnston 1991; Jennsen et al. 1994; Cronk 1996; Kadlec and Knight 1996; Gopal 1999; Knight et al. 1999; Tarutis et al. 1999). However, most of what is known about the performance of wetlands for water quality enhancement is based on studies of constructed wetlands, which are engineered specifically for the purpose of treating wastewater effluent. The factors that have the strongest influence on the ability of engineered wetlands to remove pollutants from inflowing waters are: (1) the area of the wetland, (2) the flow rate of water entering the wetland, and (3) the concentration of pollutant in the inflowing water (Kadlec and Knight 1996). Performance will be affected by other factors as well – such as the type of soil and vegetation in the wetland, the configuration of the wetland with respect to the direction of water flow, the variation in depth throughout the wetland, the temperature of the water, etc. – but it is the three factors listed above that are generally used to summarize the performance of existing wetlands and to design new constructed wetlands. The standard function used to describe the performance of constructed wetlands for wastewater treatment is the first-order removal rate equation, which can be written as:

$$r^i = Q_{in} C^i_{in} \left[1 - \exp\left(\frac{-K^i D}{h} \right) \right] \tag{4.1}$$

In equation 4.1, r^i is the mass of pollutant i removed from the inflowing water; Q_{in} is the flow rate of water into the wetland; C^i_{in} is the concentration of pollutant i in the inflowing water; h is the average depth of the wetland; D is the detention time, the average amount of time it takes for water to flow through the wetland; and K^i is the removal rate constant.[31] D is usually estimated by $(A \times h)/Q_{in}$, where A is the area of the wetland, and values for K have been estimated experimentally for a number of water quality constituents of concern (see Table 4.1). The first-order removal rate equation provides a general relationship between the size of a wetland, the flow rate of water and pollutant load presented to it, and the removal efficiency.

Table 4.1 – Removal rate constants for several important water quality constituents in treatment wetlands.

Constituent	K [meters/year]
Biological oxygen demand	34
Total suspended solids	1,000
Total nitrogen	22
Total phosphorus	12
Fecal coliform	75

From Kadlec and Knight (1996)

[31] The usual units of measurement for the variables that appear in the first-order removal rate equation are: r [g], Q [m³/yr], C [g/m³], D [yr], h [m], and K [g/m²/yr].

Natural and restored wetlands can also improve water quality (Whigham et al. 1988; Johnston et al. 1990; Chambers et al. 1993; Detenbek et al. 1993; Gilliam 1994; Weller et al. 1996), but they will be subject to much higher variability in loading rates and flow rates than are wetlands constructed for the purpose of treating wastewater from point sources. Furthermore, the structure of natural and restored wetlands will not necessarily be optimal for water treatment purposes, simply because they were not designed for that purpose. The degree to which any particular natural or restored wetland delivers water quality benefits will depend on a number of factors, including:

1. The size and shape of the wetland, which affect the amount of runoff that the wetland will intercept and the amount of time water is in contact with the soil and vegetation within it.
2. The nature of the landscape upstream of the wetland, which affects the concentration of pollutants in the inflowing water.
3. The nature of the landscape downstream of the wetland, which affects the fate of the pollutants in the water leaving the wetland.

These factors are not accounted for in the first-order removal rate equation because it was developed solely to explain changes in nutrient concentrations between the inlets and outlets of wetlands. A more comprehensive, landscape-level model would be required to address questions regarding the importance of the location of wetlands for the water quality benefits they will provide. This chapter describes one such model, developed for the Central Valley of California.

4.2 The Central Valley water quality model

This Central Valley water quality model is a spatially distributed hydrologic simulation model designed to estimate the amount of nitrogen and phosphorus that will enter rivers and streams from non-point source runoff. The model uses data on land use and agricultural practices specific to the Central Valley of California, but the approach and methods could be transferred to other regions as well. Chapter 6 describes an integrated optimization model for prioritizing sites for wetlands restoration, which combines the water quality model described in this chapter, the mallard habitat model described in Chapter 3, and estimates of the costs of wetlands restoration described in Chapter 5.

4.2.1 Model overview

In the water quality model, the Central Valley of California is represented as a regular grid of square cells, each of which is defined by its location, land use type, and soil type. The grid is made up of 1.48 million cells, each 200 meters on a side. Figure 4.1 shows a polygon and grid representation of a small watershed in the Central Valley. Panel A shows the location of the watershed; panel B shows the polygon representation of the watershed, which is the raw form of the land use data; and panel B shows the grid representation, which is used by the water quality model. The difference in resolution between panels B and C gives an indication of the information lost when the raw data was converted to grid form. At this 200-meter resolution most of the detail in the raw data was maintained, so the model has the potential to target restoration activities at the parcel level.[32]

Within-cell processes are modeled explicitly, and because the model was designed primarily for investigating the effects of landscape configuration on water quality, the flows of water and nutrients between cells are also modeled explicitly. The model uses a water balance approach to estimate flows into and out of each cell. Nitrogen and phosphorus fluxes are largely driven by water flows, but depend on other factors as well. The model

[32] For example, the average size of parcels offered for enrollment in the Wetlands Reserve Program in California in the year 2000 was approximately 450 acres, which would be represented by approximately 45 cells in the water quality model. This is sufficient to maintain the general shape of most parcels. However, this level of realism comes at a substantial computational cost, as will become apparent in Chapter 6 when the optimization model is presented.

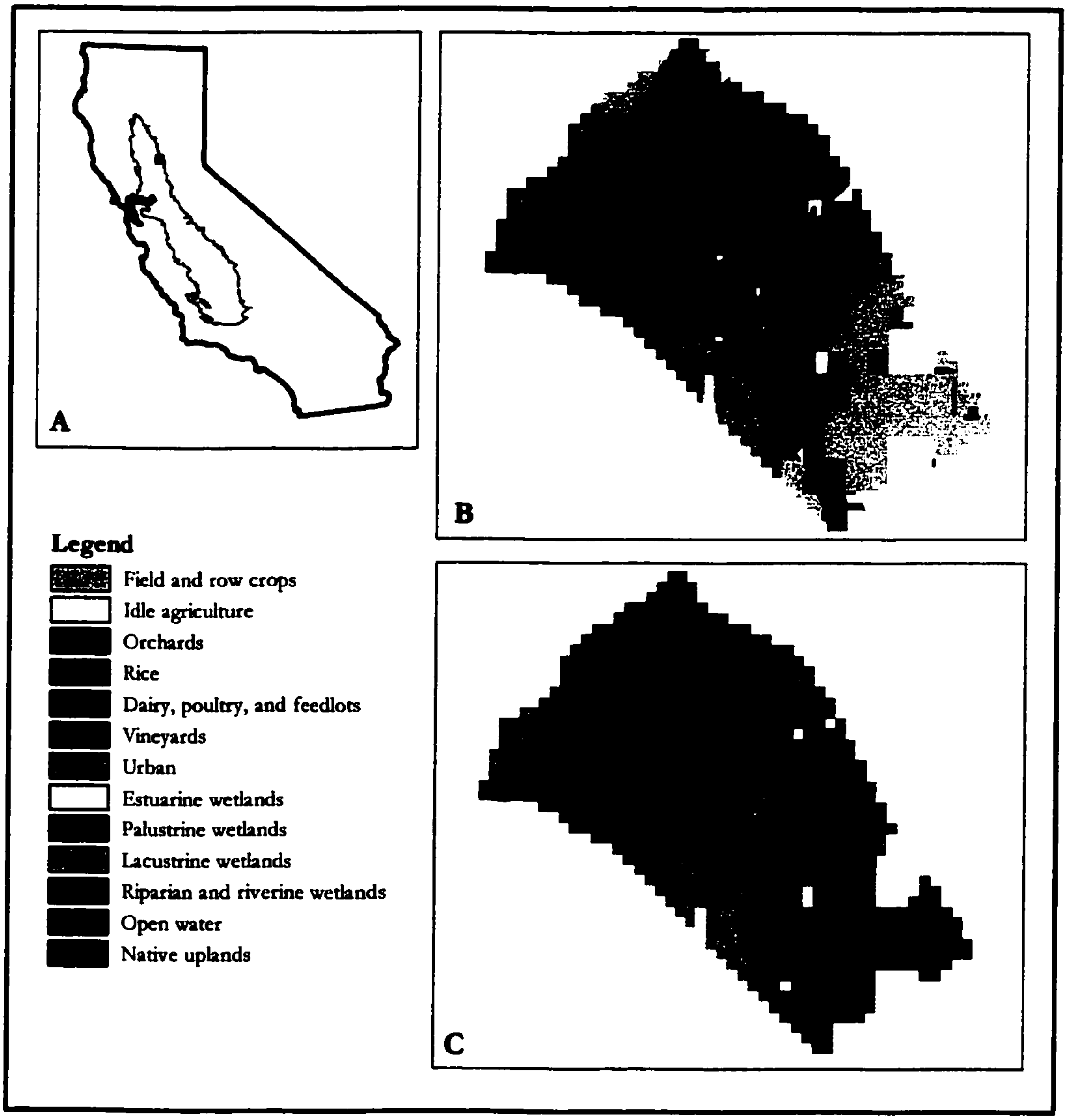


Figure 4.1 – Polygon and grid representation of the Central Valley for the water quality model. Panel A shows the location of the watershed in the Central Valley. Panel B shows the raw (polygon-based) land use data in the watershed. Panel C shows the grid representation of the watershed. The entire Central Valley covers 5.86 million hectares; the example watershed covers 4,148 hectares.

calculates the water and mass balance for each cell, and for the entire landscape, for each month in a representative year.

The model distinguishes between 29 types of land use, including ten types of agriculture and six types of urban land uses (see column 1 of Table 4.2). Because the focus of this research was on the effects of the landscape position of wetlands, not on site-specific characteristics that differentiate wetland types, the model treats all wetlands the same. The structure of the model does not preclude treating wetland types differently, however, and this is one feature that could be modified for future applications.

The foundation of the model is a set of water and mass balance equations, which are solved for each cell in each month. Figure 4.2 depicts the basic water balance for a representative cell. The water and mass balance equations differ across land use types, so before describing in detail the structure of the model I will describe the land use and other data required for its implementation.

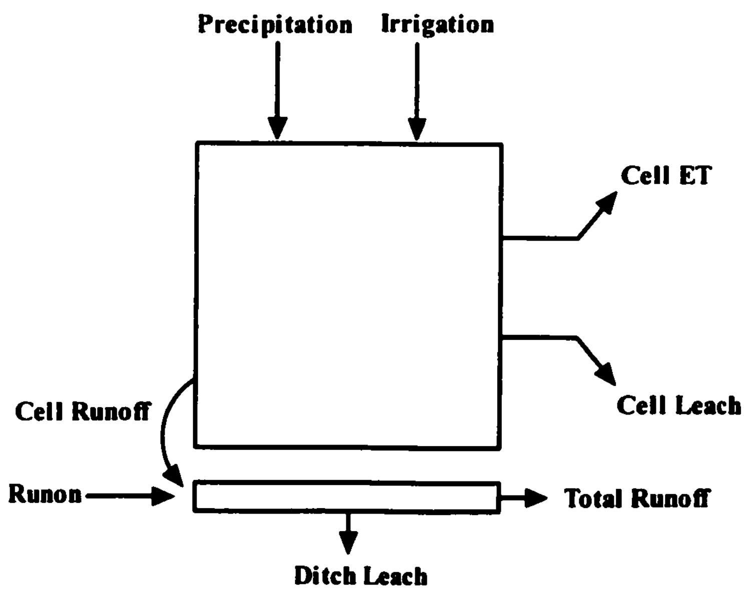


Figure 4.2 - Basic water balance for a representative cell.

4.2.2 The data

The data required for each cell include: average monthly rainfall, soil hydrologic group, population density, and land use type. Rainfall data came from the California Department of Water Resources.[33] Locations of the 28 stations in the Central Valley that collect precipitation data are shown in panel A of Figure 4.3. Data on total precipitation for each

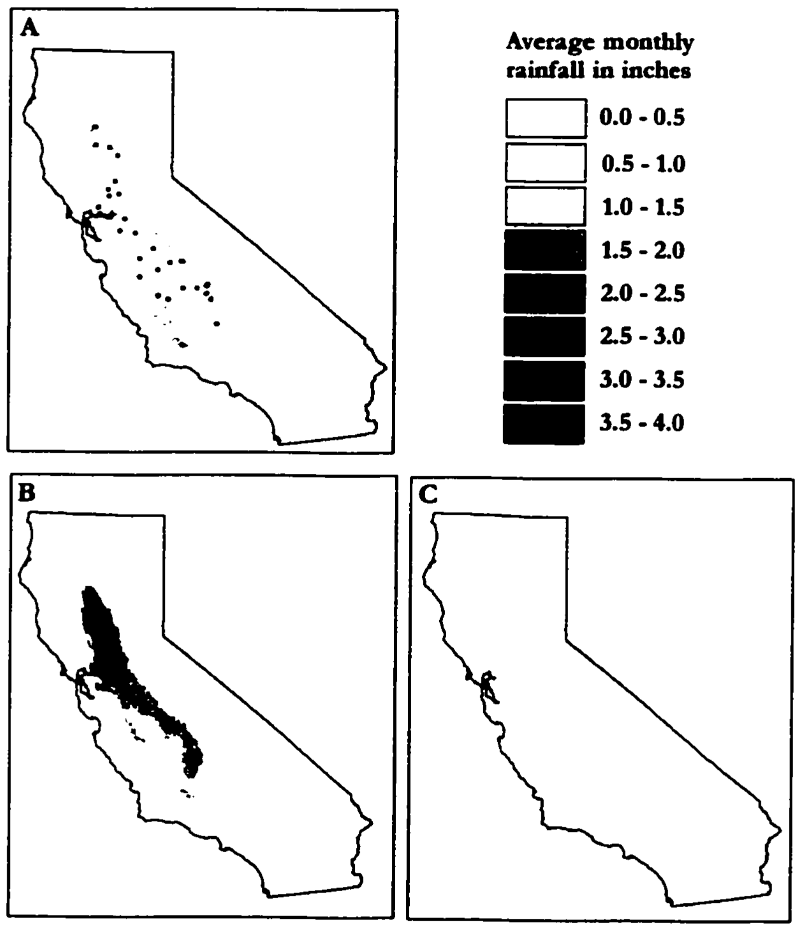


Figure 4.3 - Precipitation in the Central Valley. Panel A shows locations of the precipitation monitoring stations in the Central Valley, panel B shows interpolated average rainfall for December, and panel C shows interpolated average rainfall for May.

[33] http://cdec.water.ca.gov/

month at each station since 1965 were averaged, and the average monthly values for all cells in the grid were estimated using distance-weighted interpolation between all 28 stations. The average total annual rainfall across all 28 stations was 17.1 inches. The minimum total rainfall in the valley since 1965 was 9.3 inches in 1988-89, and the maximum was 30.7 inches in 1977-78. Panel B in Figure 4.3 shows interpolated average rainfall for December, and panel C shows interpolated average rainfall for May. These two panels give an indication of the temporal (within-year) and spatial variation of precipitation in the valley. The figure shows a strong north-south gradient, with the Sacramento Valley in the north receiving more precipitation than the San Joaquin Valley in the south throughout the year, and a less pronounced east-west gradient in the San Joaquin Valley in the winter, with the eastern half of the valley receiving more precipitation than the western half.

Data on soil characteristics came from STATSGO, a soils database maintained by the Natural Resources Conservation Service of the U.S. Department of Agriculture.[34] Soil characteristics relevant for the water quality model were determined by the hydrologic soil group (HSG) class, of which there are four: A, B, C, and D. A soil's placement in an HSG class depends on its permeability, with class A soils being the most permeable and class D soils being the least permeable (Corbitt 1990, pg. 7.25).

The water quality model uses information on the soil class for each cell to estimate how much precipitation and applied irrigation water will infiltrate into the groundwater and how much will run off the cell as surface flow. However, the structure of the STATSGO database is such that the HSG class for each cell could not be determined definitively. STATSGO data characteristics are associated with "mapunits," which are spatially defined regions on the order of 30,000 hectares each. Each cell in the water quality model grid is 4

[34] http://dbwww.essc.psu.edu/dbtop/doc/statsgo/statsgo_info.html#constraints

hectares, so there were many cells in each STATSGO mapunit. The soil characteristics assigned to each cell were determined by the percent of each HSG class in each mapunit. For example, if mapunit 500 had 50% class A soils, 25% class B soils, 15% class C soils, and 10% class D soils, then all cells within mapunit 500 were assigned that same mix of soils. Another way to interpret this is that for each cell in mapunit 500 there is a 50% chance it has class A soils, a 25% chance it has class B soils, and so on.

Data on population density came from the U.S. Census Bureau's TIGER/line files, which is a digital database of geographic features, including census tract boundaries and attributes, covering the entire United States.[35] The population density in each census tract was assigned to all coincident cells in the water quality model grid. The water quality model uses the population density estimates to calculate runoff from urban cells (using a model described in Section 4.2.3.3).

Data on land use came from the combined DWR and NWI dataset used for the habitat models in Chapter 3 (see Section 3.2.1). However, instead of the aggregated land use types used in the habitat models, the water quality model used a more detailed classification with 29 land use types. Table 4.2 shows the land use types used in the water quality model and the associated parameter values, which described in detail in the next section. Together, the estimates of average precipitation by month, percent of each hydrologic soil group class, population density, and land use type for each cell, provided the primary data on which the water quality model rests. These characteristics ultimately determine the total amount of water and nutrients applied to each cell in each month, and how much of each exit the cell by way of leaching to the groundwater or by way of surface runoff.

---

[35] See http://www.census.gov/geo/www/tiger/tigerua/ua_tgr2k.html, and coverages in e:/wetlands/masterarcviewproject.prj.

Table 4.2 – Land use types and associated parameters used in the Central Valley water quality model.

Land use type	Urban model parameters		Non-urban model parameters							Curve number description[2]
	α_N	α_P	AE[1]	N app. rate [lbs/acre/yr]	P app. rate [lbs/acre/yr]	CN_A	CN_B	CN_C	CN_D	
Citrus subtropical			0.75	113.14	57.71	67	78	85	89	Row crops, straight row, good hydrologic condition
Commercial	0.296	0.2498	0.20	0.00	0.00	98	98	98	98	Pavement & roofs, commercial & business areas
Deciduous fruits & nuts			0.75	93.00	49.33	67	78	85	89	Row crops, straight row, good hydrologic condition
Field crops			0.75	82.44	46.01	67	78	85	89	Row crops, straight row, good hydrologic condition
Grain and hay crops			0.75	40.67	20.33	67	78	85	89	Row crops, straight row, good hydrologic condition
Idle ag			0.75	0.00	0.00	76	85	90	93	Fallow, crop residue, poor condition
Industrial	0.277	0.0233	0.20	0.00	0.00	81	88	91	93	Urban districts, industrial
Lacustrine limnetic			0.60	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Lacustrine littoral			0.60	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Native barren			0.60	0.00	0.00	35	56	70	77	Brush, weed, grass mix, fair condition
Native misc			0.60	0.00	0.00	43	65	76	82	Woods-grass combination, fair condition
Native open water			0.20	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Native riparian			0.60	0.00	0.00	36	60	73	79	Woods, fair condition
Native vegetation			0.60	0.00	0.00	43	65	76	82	Woods-grass combination, fair condition
Palustrine			0.60	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Pasture			0.75	78.40	30.00	49	69	79	84	Pasture, grasslands or range, fair condition
Residential	0.131	0.011	0.20			77	85	90	92	Residential districts, average lot size 1/8 acre
Rice										Row crops, straight row, good hydrologic condition

Table 4.2 (continued) – Land use types and associated parameters used in the Central Valley water quality model.

Land use type	Urban model parameters		Non-urban model parameters							Curve number description
	α_N	α_P	AE[1]	N app. rate [lbs/acre]	P app. rate [lbs/acre]	CN_A	CN_B	CN_C	CN_D	
Riverine intermittent			0.20	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Riverine lower perennial			0.20	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Riverine tidal			0.20	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Riverine upper perennial			0.60	0.00	0.00	40	40	40	40	Low curve number used for open water or wetlands to reflect some storage
Semi-ag & incidental			0.75			59	74	82	86	Farmsteads
Truck nursery berry			0.75	173.50	53.22	67	78	85	89	Row crops, straight row, good hydrologic condition
Urban	0.1911	0.0719	0.20			98	98	98	98	Pavement & roofs, commercial & business areas
Urban landscape	0.0605	0.0033	0.60			49	69	79	84	Open space (lawns, parks, etc.), fair condition
Vacant urban	0.0605	0.0033	0.20			98	98	98	98	Pavement & roofs, commercial & business areas
Vineyards			0.75	67.00	55.00	67	78	85	89	Row crops, straight row, good hydrologic condition
Estuarine			0.60			40	40	40	40	Low curve number used for open water or wetlands to reflect some storage

Notes:

1. AE = Irrigation efficiency, the percent of applied irrigation water that is used (transpired) by the vegetation. According to Goldhamer and Snyder (1989) irrigation efficiencies range from 60-90% for most irrigation types. The model uses 0.75 for all agriculture types; i.e. the crop uses75% of applied water or rainfall, and the remainder is either leached to the groundwater or exits the patch as surface runoff. The model uses 0.60 for native vegetation and wetlands, and 0.20 for urban areas.

Information used to estimate the nutrient application rates for each agricultural land use type, which are shown in columns six and seven of Table 4.2, was taken from Owens et al. (1998), and Padgitt et al. (2000). Curve numbers and information used to estimate irrigation application rates for each agricultural land use type were taken from Goldhamer and Snyder (1989). Information used to estimate combined sewer and stormwater runoff and nutrient concentrations in runoff for urban land use types was taken from U.S. EPA (1977). The following sections describe the models that combine all of this information to estimate water and nutrient inputs and outputs for all cells in the model grid.

4.2.3 Water and mass balance calculations

The basic water balance shown in Figure 4.2 provides the template for water and mass balance calculations for all cells, but the details differ across agriculture, urban, natural upland, and wetland cells.

4.2.3.1 Water and mass balance for agriculture cells

The water and mass fluxes for agriculture cells are driven largely by applied irrigation water. The model assumes that farmers apply water according to recommendations made by the University of California Cooperative Extension (UC Extension) for irrigation scheduling (Goldhamer and Snyder 1989). The irrigation scheduling model is based on a water budget approach, which involves monitoring all of the additions and losses of water from a field to maintain a favorable soil water level (Ibid, pg. 23). The irrigation scheduling model can be summarized in five steps:

1. Estimate <u>reference evapotranspiration</u> (ET_0), the rate of evapotranspiration from a 4- to 7-inch-tall, unstressed, cool-season grass.
2. Estimate <u>crop water demand</u> by multiplying ET_0 by an appropriate correction factor (K_C), specific to the crop and region.

3. Estimate effective rainfall (R), the portion of rainfall that infiltrates the soil and contributes to satisfying crop water demand, using the Curve Number method (described below).
4. Estimate irrigation application efficiency (AE), the fraction of applied irrigation water that contributes to satisfying crop water demand. (The remainder either infiltrates below the root zone or leaves the field as surface runoff.)
5. Apply sufficient irrigation water (IA) to meet crop water demand, accounting for rainfall and expected losses through less-than-perfect irrigation applications.

The UC Extension recommends maintaining a water balance account on a daily basis in the growing season (based on weather data that can be obtained from the California Irrigation Management Information System, another UC Extension service). The Central Valley water quality model developed for this project estimated total monthly application rates by using average historical reference ET_0 values and a simple crop growth function for each agriculture type (Goldhamer and Snyder 1989, pgs. 30-31). The crop growth functions for annual and perennial crops are shown in Figure 4.4. ET_0 values differed across counties and months, and growth dates and associated K_C values differed across crop types and between the Sacramento and San Joaquin Valleys. All relevant values for these parameters were stored in tables that the water quality model referred to when calculating the water balance for each agriculture cell. The complete set of water balance equations for agriculture cells is:

$$S = (1000/CN) - 10 \tag{4.2a}$$

$$R = \min\{S, P\} \tag{4.2b}$$

$$S_r = S - R \tag{4.2c}$$

$$IA = \max\{0, (K_C ET_0 - R)/AE\} \tag{4.3d}$$

$$ET = K_C ET_0 \tag{4.2e}$$

$$\mathit{Runoff} = \max\{0, P - S\} + \max\{0, IA(1 - AE) - S_r\} \tag{4.2f}$$

$$\mathit{Leach} = \max\{0, P + IA - K_C ET_0 - \mathit{Runoff}\} \tag{4.2g}$$

Equations 4.2a through 4.2c constitute the Curve Number approach to estimating runoff and groundwater infiltration (Viessman et al. 1989, Corbitt 1990). The parameter *CN* in equation 4.2a has been estimated experimentally for a number of hydrologic soil class and cover type combinations by the Soil Conservation Service (now the Natural Resources Conservation Service) (Soil Conservation Service 1973). Values used for each land use type in the water quality model are shown in columns 8 through 11 in Table 4.2; descriptions of the conditions assumed for each land use type to assign the *CN* values are given in the last column of the table.

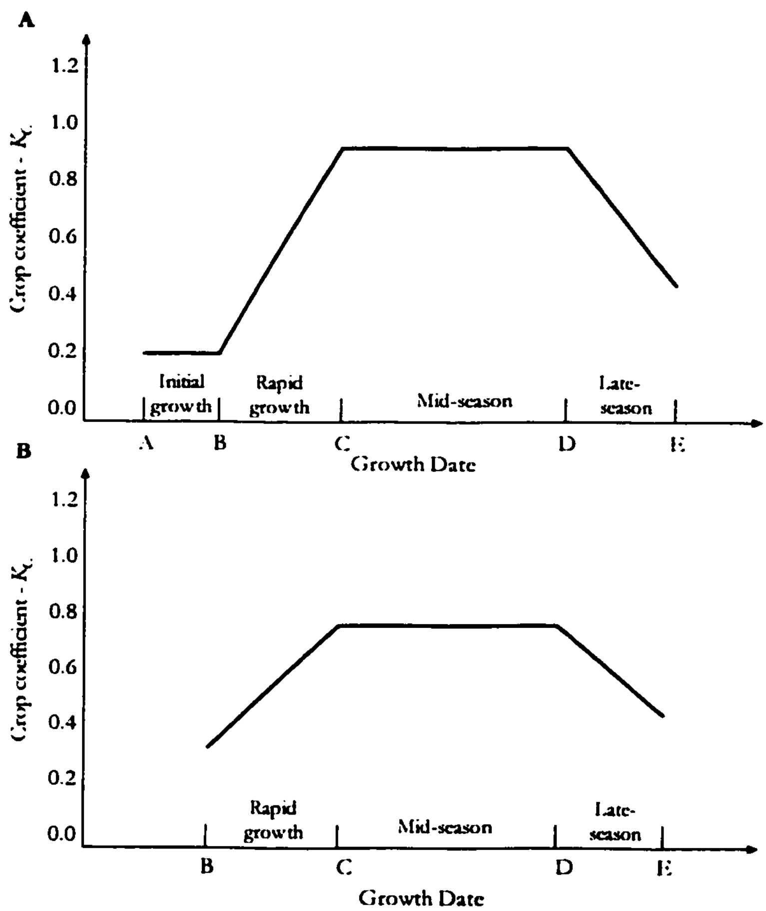


Figure 4.4 - Crop growth models for hypothetical annual crops (panel A) and perennial crops (panel B). From Goldhamer and Snyder (1989).

The model applies the water balance equations in sequence to each cell. Equation 4.2a is used first to estimate the maximum amount of water (measured in inches) that can infiltrate into the soil, S. Equation 4.2b is then used to estimate effective rainfall, R, which is the maximum of S and the average precipitation for the month, P. If the amount of precipitation is less than the infiltration capacity (if $P < S$), then all of the precipitation contributes to satisfying crop water demand, and none contributes to surface runoff. Equation 4.2c gives the residual infiltration capacity, S_r, the infiltration capacity that is not taken up by precipitation. This residual capacity is then available for (some of) the excess applied irrigation water to infiltrate. The model uses equation 4.2d to calculate the amount of irrigation water applied, IA, as per the recommendations for irrigation scheduling given above. No irrigation water is applied in a given month if the effective rainfall exceeds crop water demand. Finally, the last three equations determine the fate of the precipitation and irrigation water. Equation 4.2e says that the amount of water that exits a cell through evapotranspiration, ET, equals crop water demand (farmers are assumed to always supply their crops with sufficient water). The amount of water that exits the cell as surface runoff is estimated using equation 4.2f. Surface runoff could be some combination of water from precipitation, if the amount of precipitation is greater than the infiltration capacity (i.e. $P > S$), and irrigation water, if the amount of excess water applied is greater than the residual infiltration capacity (i.e. $IA(1 - AE) > S_r$). The amount of water that exits the cell through leaching to the groundwater is estimated using equation 4.2g. The amount of water that leaches below the root zone is the total amount applied less the amount that exits the cell through evapotranspiration or surface runoff.

The model performs the mass balance calculations for nutrient fluxes in agriculture cells in tandem with the water balance calculations. Farmers are assumed to apply fertilizers containing nitrogen and phosphorus at rates equal to the average annual application rates in the U.S. for each crop type (Padgitt et al. 2000). Crop nutrient use efficiencies, the fraction of applied nitrogen and phosphorus taken up by the growing plants, are assumed to be equal to the average nutrient utilization rates for major U.S. crops (U.S.D.A. 2000).[36] The total amount applied iss distributed across months according to the distribution of irrigation water applications, as determined by the crop growth models and the water balance calculations in equations 4.2a through 4.2g. Excess nitrogen (the amount not taken up by crops) exits each cell in proportion to the water fluxes; e.g. if 10% of the applied water exits a cell as leaching to the groundwater and 5% exits the cell as surface runoff, then the same fractions of excess nitrogen exit the cell by the same pathways.

Phosphorus is treated slightly differently than nitrogen in the model. Because phosphorus has low water solubility and easily becomes bound to soil particles, it is transported largely with suspended sediments (Soranno et al 1996; Carpenter et al. 1998). Because of its propensity to bind to soil particles, the water quality model assumes that no phosphorus infiltrates to the groundwater, and only some fraction of the excess phosphorus is carried off of a cell with surface runoff. The model uses the following equation to estimate the amount of excess phosphorus that exits the cell with surface runoff, P_{SR}:

$$P_{SR} = \frac{P_{Applied}\left(1-\gamma^{P}\right)Runoff}{Runoff + P_{1/2}} \tag{4.3}$$

In equation 4.3, $P_{Applied}$ is the amount of phosphorus applied to the crop, γ^{P} is the uptake efficiency, *Runoff* is the amount of water that exits the cell as surface runoff (estimated by equation 4.2f), and $\sqrt{P_{1/2}}$ is the runoff rate at which half of the excess phosphorus would be

[36] The average nitrogen use efficiency for corn, cotton, potatoes, and winter wheat for 1990 through 1997 was 65%. The average phosphorus use efficiency for corn, cotton, potatoes, soybeans, and winter wheat for 1990 through 1997 was 80%.

transported with surface runoff. Equation 4.3 implies that phosphorus transport (via sediment transport) initially increases with *Runoff* at an increasing rate, and eventually asymptotes to $P_{Applied}(1-\gamma^{P})$ at high flow rates. This functional form was chosen for its attractive qualitative characteristics: the fraction of phosphorus transported with surface flow is bounded between zero and one; the fraction transported increases rapidly at low flow rates and slowly at high flow rates; and only one parameter had to be estimated. The parameter $P_{1/2}$ was used to calibrate the phosphorus loading component of the model (as described later in Section 4.2.6.1).

4.2.3.2 Water and mass balance for natural upland cells

The water and mass balance calculations for natural upland cells differs slightly from those for the agriculture cells. No irrigation water is applied to natural upland cells; they receive water inputs only from precipitation and from runoff from other cells higher in the watershed that drained into them. No fertilizers are applied to natural upland cells either; they receive nitrogen and phosphorus inputs only from cells that drain into them. With these inputs, the water and mass balance calculations proceed as described for the agriculture cells (using equations 4.2a, b, e, f and g).

4.2.3.3 Water and mass balance for urban cells

The water and mass balance calculations for urban cells differ substantially from those for agriculture and natural upland cells in the model. The same water and mass balance principle is used for the urban cells, but instead of the equations based on crop water demands, the model uses equations developed by the U.S. EPA (1977) in a cross-sectional study of combined sewer overflows and stormwater discharges in 67 metropolitan areas across the United States. The complete set of water balance equations for urban cells is:

$$I = 9.6PD^{(0.573-0.0391\times\log(PD_d))} \tag{4.4a}$$

$$DS = 0.25 - \frac{0.1875I}{100} \tag{4.4b}$$

$$AR = \left(0.15 + \frac{0.75I}{100}\right)P - 5.345DS^{0.5957} \tag{4.4c}$$

$$DWF = 1.34PD \tag{4.4d}$$

$$Runoff_m = \frac{AR \times P_m}{\sum_m P_m} + \frac{DWF}{12} \tag{4.4e}$$

In equations 4.4a through 4.4e, I is the amount of impervious surface, PD is the population density, DS is the volume of detention storage, AR is the amount of annual surface runoff from precipitation, DWF is the annual dry weather flow, and $Runoff_m$ is the total amount of surface runoff in month m. Equations 4.4a through 4.4d were taken directly from U.S. EPA (1977); the only modification required for the Central Valley water quality model was equation 4.4e, which distributes the total annual runoff from precipitation across months in proportion to the temporal distribution of precipitation. The annual dry weather flow represents aggregate urban water demand, for residential, commercial, and industrial uses, and is assumed to be constant throughout the year. In the same EPA study, equations were also developed to estimate the amount of nitrogen and phosphorus in urban runoff. The general pollutant load function is:

$$M_{ij} = \alpha_{ij} P f(PD)\eta \tag{4.5}$$

In equation 4.5, M_{ij} is the pounds of pollutant j generated per acre of land use type i (industrial, commercial, or residential) per year; α_{ij} is a coefficient specific to each pollutant and land use type; P is the total annual precipitation; $f(PD) = 0.1412 + 0.218PD^{0.54}$ for residential areas, 1.0 for commercial areas, and 0.142 for industrial areas; and η is a function of the street sweeping interval (which is assumed to be equal for all urban cells in the model). Values of α for nitrogen and phosphorus used in the water quality model are shown in columns 3 and 4 of Table 4.2.

4.2.3.4 Water and mass balance for wetland cells

The water and mass balance calculations for wetland cells are the same as those for natural upland cells, except a first-order nutrient attenuation function is included. As was the case for the natural upland cells, no irrigation water or nutrients are applied directly to the wetland cells; they receive water and nutrient inputs only from precipitation and runoff from other cells that drain into them. All nutrient removal is assumed to occur before leaching to the groundwater, so the amount of nitrogen in the surface runoff exiting a wetland cell, N_{SR}, is estimated as:

$$N_{SR} = \frac{(N_0 - r^N) Runoff}{Runoff + Leach} \tag{4.6}$$

and the amount of nitrogen leached to the ground water, N_L, is estimated as:

$$N_L = \frac{(N_0 - r^N) Leach}{Runoff + Leach} \tag{4.7}$$

In equations 4.6 and 4.7, N_0 is the amount of nitrogen entering the cell in runoff from other cells. Again, the model treats phosphorus differently because of its low water solubility. All phosphorus exits wetland cells by way of surface runoff: $P_{SR} = P_0 - r^P$, and $P_L = 0$. The model uses the first-order removal rate equation described at the beginning of this chapter (equation 4.1) to calculate r^N and r^P for all wetland cells. Average values of the removal rate constants, K, for nitrogen and phosphorus from the literature (see Section 4.1 and Table 4.1), were halved for use in the water quality model, to account for the likelihood that natural and restored wetlands would not perform as well as constructed wetlands engineered specifically for the purpose of treating wastewater.

4.2.4 Routing runoff through the landscape

The previous section described the water and mass balance calculations for each type of cell in the Central Valley water quality model. This section describes how the model routes surface runoff from each cell to other cells and eventually to rivers and streams. This may be the most important feature of the model. Many large-scale modeling studies of non-point source pollution from agriculture have used models that predict edge-of-field nutrient loads and then (implicitly) assumed that all runoff would make it to receiving waters downstream (e.g. Taylor et al. 1992; Coiner et al. 2001). More complex models have been developed (e.g. Skop and Sørensen 1998; Grunwald and Norton 2000), but the model presented in this chapter is simple enough to be applied to large landscapes at a fairly fine resolution with less demanding data requirements than the more complex models.

The water quality model does not keep track of water that leaches below the root zone from each cell; leached water is essentially treated as entering one large groundwater basin and staying there. To analyze impacts of non-point source pollution on groundwater quality, or the potential for wetlands to mitigate such impacts, a more detailed model of groundwater movement would be required.

The routing of surface runoff in the Central Valley water quality model can be thought of as "ditch flow." Referring back to the basic water balance pictured in Figure 4.2, alongside the representative cell there is assumed to be a ditch, which receives "Runon" from the left and "Cell Runoff" from the top. Runon is simply the "Total Runoff" from the ditch associated with the cell immediately uphill of the cell in question, and the direction of flow (what determines uphill and downhill) is defined by the shortest distance path from each cell to the closest river cell; see Figure 4.5. The shortest-distance-path routing algorithm was a necessary simplification. Ideally, the routing of surface runoff would be based on a detailed description of artificial channels and ditches and natural topographic relief in the study area. However, a description of the human-made drainage network in the valley was not available, and topographic relief in the valley is generally smaller than the resolution of the available digital elevation data. Without this more detailed information, the shortest-distance-path assumption provides the best description available of the likely direction of surface runoff in the study area.

The ditch-flow concept is used for agriculture and urban cells only. The practical implication of this convention for agriculture cells is that runoff from uphill cells do not contribute to precipitation and applied irrigation water, and the nitrogen and phosphorus in the runoff do not contribute to fertilizer applications for the purposes of plant growth. On- and between-farm drainage networks are assumed to be constructed to ensure that surface runoff does not flow from a field onto neighboring fields; it is shunted directly to ditches, to be carried to canals, and ultimately released to rivers and streams. However, the drainage network is not assumed to prevent runoff from an agricultural or urban cell from flowing onto downhill natural uplands or wetlands. If a natural upland or wetland cell is on the shortest-distance path from an agriculture or urban cell, it receives surface runoff from that cell on its way to the river.

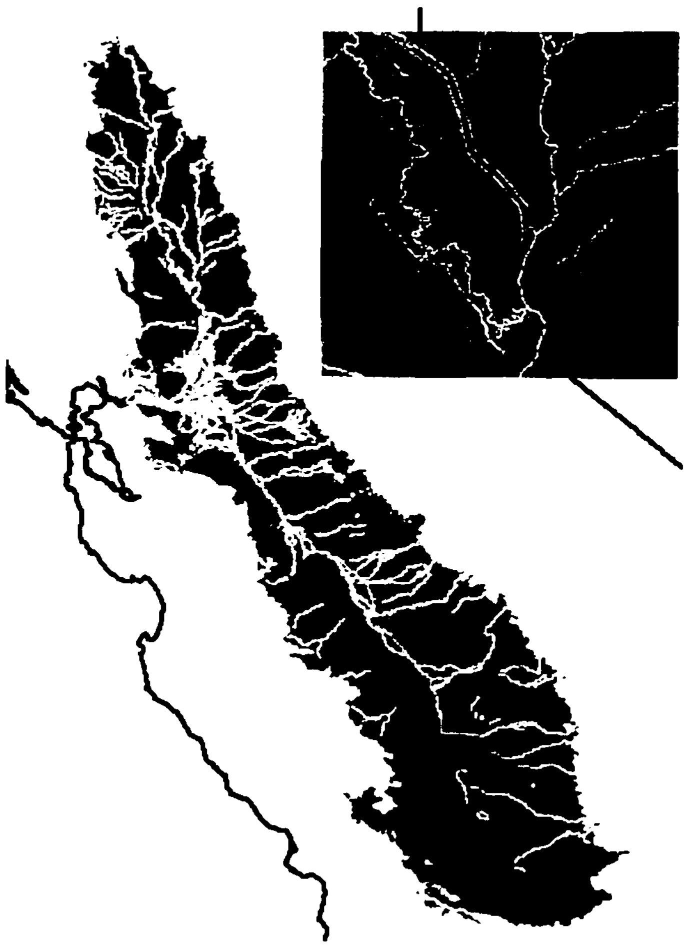


Figure 4.5 – Rivers in the Central Valley of California and the direction of surface runoff flow based on shortest-distance paths to the closest river cells. The close-up gives a closer view of the runoff surface.

Figure 4.6 shows a simplified depiction of the surface water runoff routing algorithm used in the water quality model. The shortest-distance path from cell 1, the urban cell, to the closest river cell passes through cells 2 through 5. The arrows pointing down into the cells represent water inputs from precipitation, urban water demand, or irrigation applications, depending on the land use type of the cell. The arrows at the bottom pointing out of the cells indicate the direction of flow for surface runoff exiting each cell. Surface runoff from

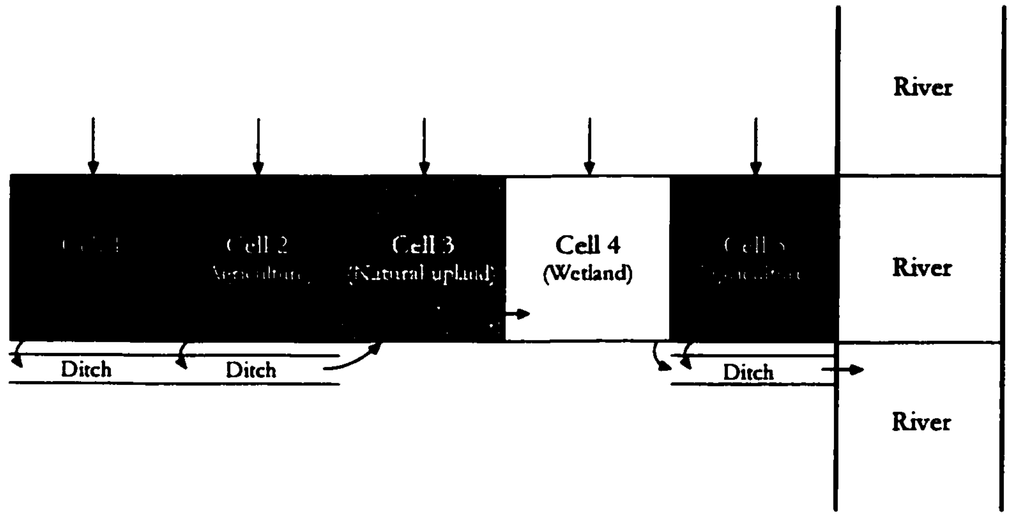


Figure 4.6 – Surface runoff routing in the water quality model.

cells 1 and 2 enter a ditch, which flows directly into cell 3, a natural upland cell. Surface runoff from cell 3 enters cell 4, a wetland cell. Surface runoff from cells 4 and 5 enter a ditch, which flows directly into the river. In short, the model assumes that in developed (urban or agricultural) areas an artificial drainage network is in place that shunts runoff to the closest un-developed area, be it a natural upland, wetland, or river. The most important implication of this assumption is that wetlands in the water quality model are fully connected, in a hydrologic sense, to the rest of the landscape. This means that they have the opportunity to mitigate non-point source pollution from all cells uphill from them. This is likely a generous assumption, and one that could be modified to analyze the effect of the degree of "hydrologic connectivity" on the overall impact that wetlands have on water quality.

Figure 4.6 also points out the importance of a wetland's location for its ability to deliver water quality benefits. All else being equal, a wetland's impact on water quality will increase with the average nutrient concentration in runoff from uphill cells and its proximity to a river cell. It should also be clear from Figure 4.6 (and equation 4.1) that a wetland's impact on water quality will decrease with the number of wetland cells downhill. To see this, consider two wetland cells in a row along a surface runoff flow path. According to the first-order removal rate equation, at any given flow rate a wetland will remove some fixed proportion of the nutrients from the water flowing through it. If the first wetland removes one half of the nutrients, then the second will remove one half of the half left over, or one quarter of the total. Thus the water quality model embodies decreasing returns to wetland proximity along runoff flow paths. This is different from the spatial structure of the habitat benefits, as suggested by the results of mallard models in Chapter 3, which implied initially increasing, then decreasing, then negative returns to contiguity. Furthermore, the direction of surface runoff flow does not affect habitat benefits at all.

Because the water and mass fluxes for any particular cell (or associated ditch) will depend on water and mass fluxes from all uphill cells, the model must perform the water balance calculations in the appropriate order. Before the total runoff from any particular cell can be calculated, the total runoff from all uphill cells must be known. This means that the water balance for the cells highest in the landscape, those that receive no runoff from other cells, must be calculated first. Then the water balance for the cells immediately downhill of those are calculated, and so on, until all runoff reaches the river. The model performs water balance calculations for each cell in decreasing order of their distance from the closest river cell. This convention ensures that when water balance calculations are performed for any given cell, the total runoff from all uphill cells has already been determined.

4.2.5 Results

4.2.5.1 Calibration and model fit

To calibrate the model and assess its accuracy, three parameters were adjusted to achieve the best possible fit between model predictions of total annual nitrogen and phosphorus loads and average historical loads for a number of river reaches in the Central Valley. Data used to calculate the average historical loads came from the USGS surface water database, which includes time-series data that describe stream levels, stream flow, reservoir and lake levels, surface-water quality, and rainfall from thousands of monitoring stations across the nation.[37] There were 194 USGS monitoring stations in the Central Valley, but only 9 and 22 river reaches had sufficient data to estimate average historical values for nitrogen and phosphorus loads, respectively. Data were considered sufficient if at the downstream station there was at least one record for average total nitrogen or total phosphorus concentration for every month of the year (though not necessarily in the same year).

The locations of the monitoring stations were used to delineate watersheds for predicting nutrient loads to the river reaches between the stations. The watersheds were defined by all cells that drain to the river reach between the downstream monitoring station and the upstream monitoring stations, according to the shortest-distance-path routing algorithm.[38] Figure 4.7 shows all watersheds used to calibrate the model, and a close-up view of one of the watersheds. I computed the average historical loads by applying mass

[37] http://waterdata.usgs.gov/nwis/sw

[38] Note that my use of the term "watershed" is not quite consistent with the way the term is used by hydrologists. Strictly speaking, a watershed includes all uplands that drain to a particular point in the river network. In lieu of defining a new term for this particular application, in this chapter and in Chapter 6 I use the term "watershed" to refer to only those uplands that drain to a particular river reach. I trust that the meaning will be clear from the context.

balance to the entire river reach in each watershed. Average historical loads were calculated as the difference between the total mass of nitrogen and phosphorus passing the downstream monitoring station and the total mass passing all upstream stations. The USGS

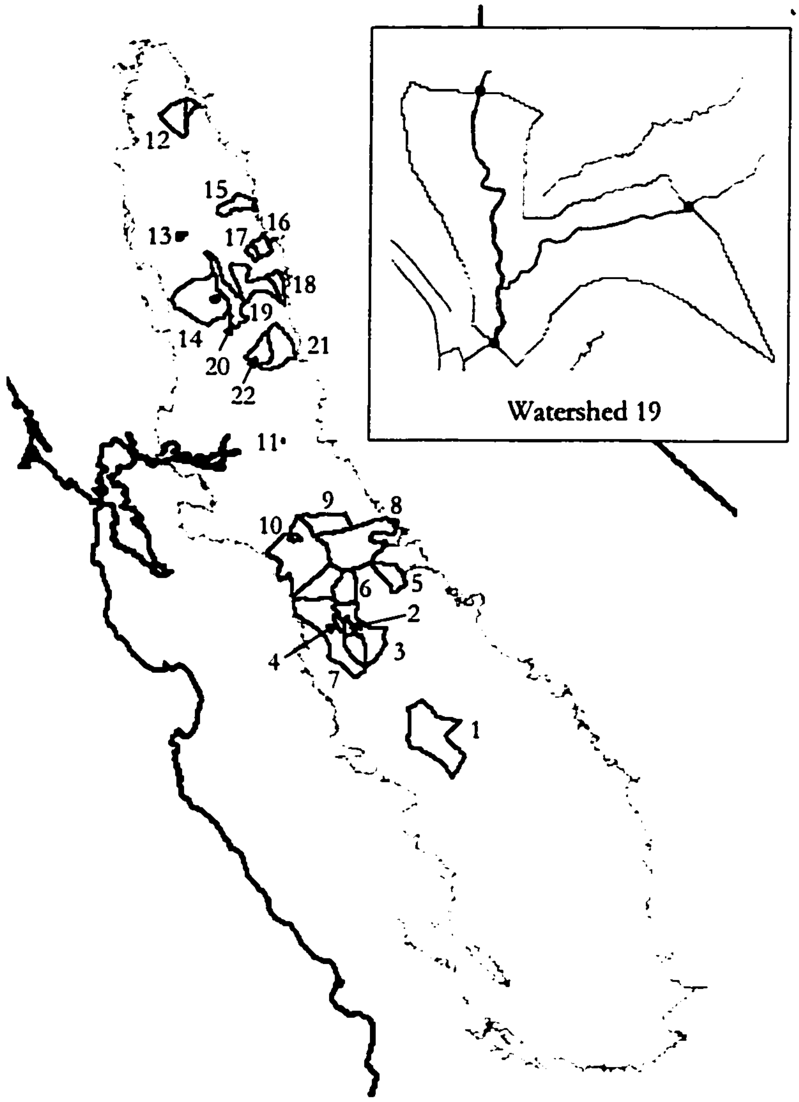


Figure 4.7 – Watersheds used to calibrate the Central Valley water quality model. The close-up shows the downstream and upstream monitoring stations for watershed 19.

surface water database contains measurements of average monthly stream flow and nutrient concentrations. The total mass of nutrient *j*, M^j, passing a station was computed using the following equation:

$$M^j = \sum_{m=1}^{12} C_m^j Q_m \tag{4.8}$$

In equation 4.8, *j* refers to total nitrogen or total phosphorus, *m* refers to months, C_m^j is the average concentration of nutrient *j* in month *m*, and Q_m is the average stream flow in month *m*. The average historical load, L^j, was computed by subtracting all M^j's for the upstream stations from the M^j for the downstream station. There were nine watersheds with sufficient data to compute M^N for the downstream station, and three of those had sufficient data to compute M^N for the upstream stations. There were 22 watersheds with sufficient data to compute M^P for the downstream station, and nine of those had sufficient data to compute M^P for the upstream stations. Missing values for upstream stations were imputed based on the average areal loading rate for each nutrient across the watersheds for which both upstream and downstream values were available. The average areal loading rate for total nitrogen in three watersheds was 2.90 kg/ha/yr,[39] and the average areal loading rate for total phosphorus in nine watersheds was 0.304 kg/ha/yr.[40]

The nitrogen utilization rate, γ^N, was adjusted to calibrate the model for nitrogen loads, and the surface runoff transport parameter, $P_{1/2}$, and the phosphorus utilization rate, γ^P (from equation 4.3), were adjusted to calibrate the model for phosphorus loads. Starting

---

[39] Compare to 4.88 from Peterjohn and Correll (1984), 19.5 from Kronvang et al. (1995), 22.23 from Skop and Sorensen (1998), 6.18 from Bhaduri et al. (2000), 20.5 from Brawley et al. (2000), and 5.07 from Castillo et al. (2000).

[40] Compare to 1.29 from Peterjohn and Correll (1984), 0.21 from Kronvang et al. (1995), 0.38 from Bhaduri et al. (2000), and 0.10 from Castillo et al. (2000).

values for the nutrient utilization rates were based on average values from nutrient budgets for major crops in the U.S. (U.S.D.A. 2000); $P_{1/2}$ was set at the arbitrary starting value of 50. The parameters were adjusted manually according to the overall fit of the model predictions to the average historical loads across the 22 watersheds. The fit of the model was assessed based on the intercept and slope coefficients and the R^2 value from a regression of the average historical loads on the loads predicted by the model. Perfect predictions would yield an $R^2 = 1.0$, an intercept equal to 0.0, and a slope equal to 1.0. The R^2 values and intercept estimates did not vary as much across trials as the slope estimates, so the best parameters were taken to be those that yielded slope estimates as close to 1.0 as possible. For phosphorus, $P_{1/2}$ was adjusted until the fit was best, given the starting value for γ^P, then γ^P was adjusted until the fit was best, given the calibrated value of $P_{1/2}$. Table 4.3 shows the results of the calibration. Combinations of parameter values are listed in the table in the order in which they were adjusted. The values of γ^N, $P_{1/2}$, and γ^P used to calculate baseline results for the entire Central Valley, which are reported in the next section, and for the optimization model described in Chapter 6, are shown in bold type in the final rows of the nitrogen and phosphorus calibration sections of Table 4.3. The calibrated value of γ^N was substantially higher than the starting value, but the calibrated value of γ^P was essentially unchanged.

Table 4.4 and Figure 4.8 show the final results from the calibrated model for the 22 watersheds. The average predicted areal loading rates for nitrogen and phosphorus were 1.64 and 0.086 kg/ha. Figure 4.8 shows that the fit of the model was better for nitrogen than for phosphorus, though there were far fewer observations of average historical loading rates for the former than for the latter. Also note that watershed 22 was not included in the regression for phosphorus shown in Figure 4.8; the average historical loading rate in that watershed was two orders of magnitude higher than that predicted by the model in all calibration trials, so it was considered an outlier.

Table 4.3 – Calibration of the water quality model.

Nitrogen calibration:			
γ^N	R^2	$L^N = a + b\hat{L}^N$	
		a	b
0.65	0.702	17,790	0.321
0.50	0.685	19,315	0.220
0.75	0.717	15,916	0.457
0.80	0.730	14,456	0.576
0.85	0.737	12,523	0.768
0.90	0.713	10,644	1.090

Phosphorus calibration:				
$P_{1/2}$	γ^P	R^2	$L^P = a + b\hat{L}^P$	
			a	b
50	0.80	0.352	696	1.529
25	0.80	0.338	761	0.783
37	0.80	0.347	721	1.141
32	0.80	0.344	735	0.992
32	0.85	0.344	733	1.322
32	0.75	0.344	736	0.793
32	0.81	0.344	734	1.044

Table 4.4 – Calibration results for the Central Valley water quality model.

Watershed	Area [ha]	M^N_1 [kg/yr]	M^N_2 [kg/yr]	L^N [kg/yr]	$\hat{L}^N$ [kg/yr]	M^P_1 [kg/yr]	M^P_2 [kg/yr]	L^P [kg/yr]	$\hat{L}^P$ [kg/yr]
1	75,596					0	2,130	2,130	7,045
2	4,656					0.0 [2]	922	922	4
3	37,268					2,804	3,492	688	1,390
4	3,572					0	449	449	0
5	17,196	0.0 [2]	828	828	5,163	0.0 [2]	42	42	385
6	21,556	0.0 [2]	20,040	20,040	5,374	0.0 [2]	1,135	1,135	203
7	68,228					3967	8,982	5,016	3,018
8	75,700					0.0 [2]	2,386	2,386	7,260
9	26,360					0.0 [2]	1,407	1,407	1,832
10	87,428	0	127,466	127,466	82,310	5,495	15,785	10,290	8,292
11	264	3,459.3 [1]	4,226	766	42	192.5 [1]	273	80	2
12	24,464					2,896.5 [1]	10,340	7,443	4,762
13	1,508	28,764.9 [1]	33,143	4,378	2,354	15,062.1 [1]	15,521	459	219
14	58,596	0.0 [2]	81,669	81,669	85,859	1,234.7 [1]	19,063	17,828	6,773
15	15,312					0.0 [2]	1,891	1,891	3,752
16	8,596					0	731	731	579
17	4,148					731	824	92	1,065
18	9,364					0.0 [2]	98	98	468
19	27,736					1,473	3,948	2,475	4,594
20	26,560	0	77,644	77,644	35,209	4,333	18,746	14,414	2,697
21	29,964	0.0 [2]	42,158	42,158	59,699	0.0 [2]	2,562	2,562	395
22	16,032	40,449	109,843	69,394	25,340	2,271	33,370	31,099	323

Notes: (1) Values imputed based on the average areal loading rate for watersheds with data for both upstream and downstream nutrient fluxes. (2) Assumed zero because the imputed value based on the average areal loading rate was negative.

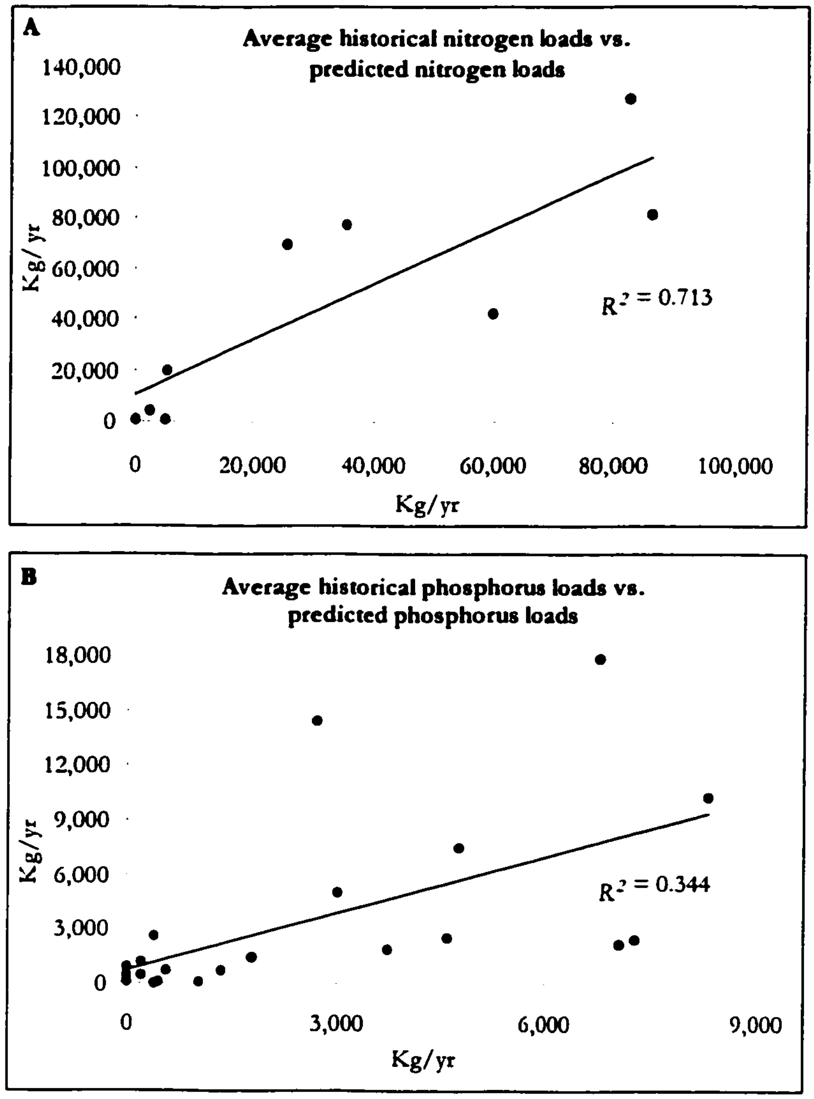


Figure 4.8 – Average historical nutrient loads vs. predicted loads from the calibrated model.

Errors in predictions from the water quality model could stem from a number of sources, and Monte Carlo and sensitivity analyses, like those performed on the hypothetical watershed in Chapter 2, could be used to assess the relative importance of various possible sources of error. However, the main goal of the calibration exercise was to ensure that the model generates predictions in a range consistent with average historical values. Figure 4.8

shows that it does, and that the model explains a fair amount of the variation in average historical loads across the 22 watersheds; therefore sensitivity analyses and further model refinements were left for future work. The next section presents baseline results from application of the water quality model to the entire Central Valley of California.

4.2.5.2 Baseline results for the Central Valley

I applied the water quality model to the entire Central Valley of California to generate baseline estimates of total annual nitrogen and phosphorus loads from surface runoff to all rivers and streams in the region. The structure of the model is such that the nutrient loads to any river cell (i.e. any 200-meter stretch of any river or stream) for each month could be computed separately. Such spatial and temporal detail could be useful for addressing localized water quality issues in the valley, but this research was not intended to address a particular problem area. The model was developed to provide a means of estimating the importance of the spatial distribution of wetlands on the water quality benefits they provide. The model will be put to just such a use in Chapter 6.

Table 4.5 lists baseline results for a number of key water and mass fluxes in the Central Valley as a whole. According to the model, approximately twice as much water is used in the valley for agriculture and urban uses than falls as precipitation.[41] Approximately 67% of the water inputs (in the form of precipitation, irrigation, and urban water demand) exits the landscape in the form of evapotranspiration, 25% exits as leaching to the groundwater, and 8% exits as runoff to rivers. Of the more than 350 million kilograms per year of nitrogen inputs, approximately 90% is taken up by crops for growth, 8.5% leaches to the groundwater, and 1.5% enters rivers by way of non-point source pollution in surface

[41] Most of the remainder (and much more in wet years) comes from runoff from the Sierra Nevada mountain range to the east. Some also comes from groundwater pumping.

runoff. Of the nearly 170 million kilograms of phosphorus inputs, approximately 81% is taken up by crops for growth, 18.8% is bound to soil particles and immobilized, and just over 0.2% enters rivers by way of non-point source pollution in surface runoff. The amount of nitrogen and phosphorus attenuated in wetlands is 9.1% and 6.1% of the total loads to rivers and streams from surface runoff. Figure 4.9 shows the variation in runoff, nutrient loads, and attenuation rates in wetlands across months. High runoff rates between November and March are due to precipitation, and high runoff rates in June and July are due to irrigation runoff. Nutrient loads peak in the growing season; they can be an order of magnitude higher in July than in November. Nutrient attenuation rates in wetlands vary much less drastically across months.

The annual nutrient attenuation rates in wetlands are similar in magnitude to the proportion of wetlands in the landscape, which is 6.9%. However, the existing wetlands are not necessarily in the locations that would yield the greatest water quality benefits possible. In Chapter 6 the question of optimal wetlands restoration in the Central Valley for water quality enhancement will be addressed specifically, using the water quality model described here. But first, Chapter 5 describes the estimates of wetlands restoration costs that are used in the integrated optimization model in Chapter 6.

Table 4.5 – Baseline results from the water quality model applied to the entire Central Valley.

	Total fluxes
	[1000 m3/yr]
Precipitation	18,379,348
Irrigation + urban water demand	37,202,691
Evapotranspiration	37,140,702
Leaching to groundwater	13,723,558
Runoff to rivers and streams	4,716,129
	[Kg/yr]
Nitrogen inputs	352,170,417
Nitrogen uptake	317,458,030
Nitrogen leaching	30,039,317
Nitrogen runoff	4,673,056
Phosphorus inputs	169,163,992
Phosphorus uptake	137,135,411
Phosphorus immobilized	31,617,146
Phosphorus runoff	411,433
Nitrogen attenuated in wetlands	424,870
Phosphorus attenuated in wetlands	25,239

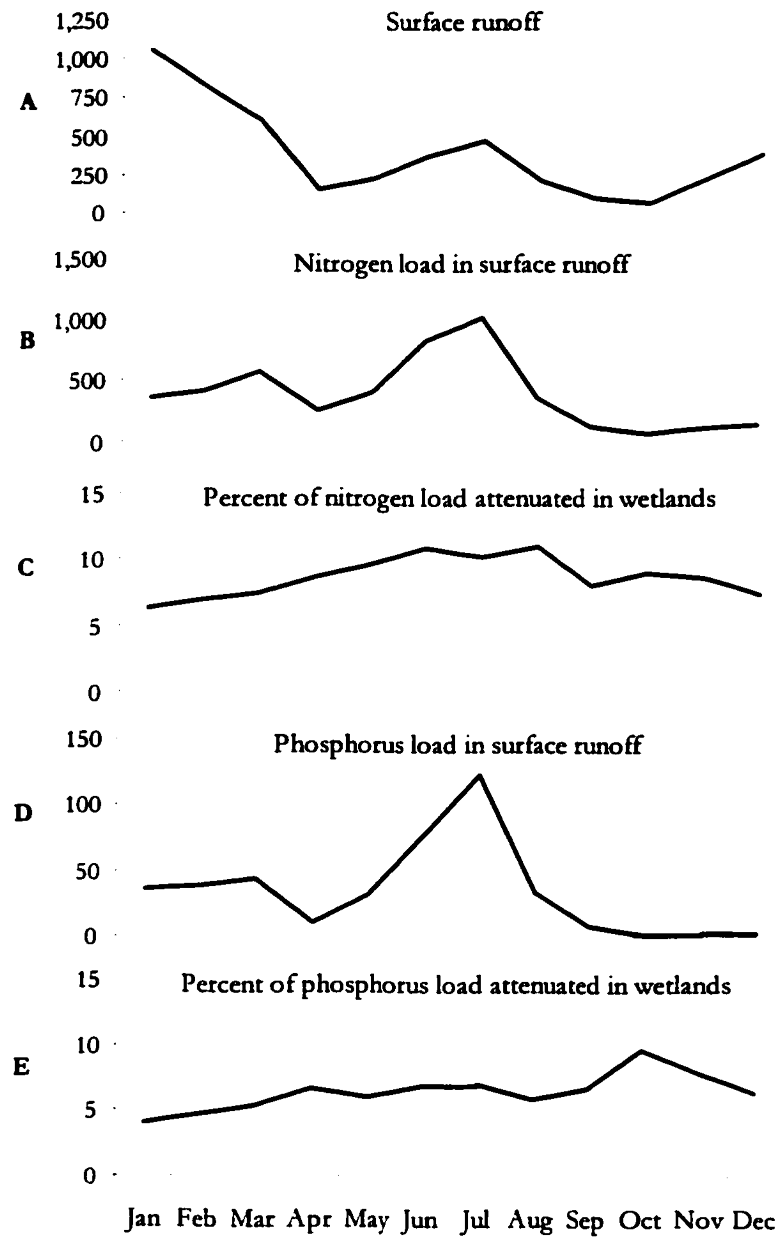


Figure 4.9 – Runoff, nutrient loads, and attenuation in wetlands across months in the Central Valley. The units are: surface runoff [10^6 m^3/yr]; nitrogen and phosphorus loads [1,000 kg/yr].

Chapter 5 – The Costs of Wetlands Restoration

Chapters 3 and 4 focused on two of the public benefits that wetlands can provide. Moving now to the other side of the benefit-cost equation, this chapter describes estimates of the private costs of wetlands restoration in the Central Valley of California. The cost of wetlands restoration on private land is made up of two components: (1) the benefits that would accrue to the landowner were the land to be used for some other purpose, and (2) the costs of converting the land from its present state back to, and maintaining it as, a wetland. The first component is the present value of the entire stream of expected future benefits from the land in its highest economic use. Absent significant distortions in the market, the value can be approximated by the market price of the land. The second component is the combined construction, operation, and maintenance costs of the wetland restoration itself. Conversion of wetlands to other uses generally involves removing vegetation, draining the land, and filling and leveling the site. Therefore, restoring wetlands generally requires excavation, reintroducing water to the site, and replanting native vegetation.[42]

5.1 Opportunity costs

I estimated land values in the Central Valley using county assessor data. County assessors are responsible for recording the market value of landholdings for the purposes of

[42] There is much discussion in the wetlands ecology literature about the potential for wetlands restoration success in general. Some researchers are very skeptical that restored wetlands can adequately replace converted ones, but others are more optimistic; for example, see: National Research Council (1992); Race and Fonseca (1996); Zedler (1996); and Mitsch and Wilson (1996). I ignore this issue here, but it could be incorporated into the integrated model presented in Chapter 6 in a crude way by adjusting relevant parameters to reflect the fact that restored wetlands may not provide as suitable habitat, or attenuate nutrients to the same degree, as natural wetlands.

determining tax rates for all landowners in their respective counties. Ideally, assessed values would be based on current market conditions and therefore accurately reflect the likely selling price of the property. However, in 1978 California voters passed Proposition 13, which limited to 2% per year the rate at which assessed property values could increase, unless the property changed ownership. The most recent change in ownership determines the Proposition 13 base year value, on top of which the 2% increase per year is added, unless the market value of the land is lower, in which case the property is temporarily reassessed to account for the drop in value.[43] The land values used in this study were not corrected for the effect of Proposition 13. This will likely bias all estimates of land values down, but unless the average rate at which properties change hands is very different across land use types in the study area, the relative values should not be much affected.[44]

I obtained county assessor data for 13 of the 20 counties in the Central Valley; data were not available for the remaining 7 counties. The land value data covers approximately 57% of the area of the valley; see Figure 5.1. Each county assessor uses their own land use categorization, and the land use types identified by each assessor generally do not match across counties. I aggregated each county assessor's categorization in such a way as to match, as closely as possible, the categorization in the DWR land use dataset (see Section 3.2.1). Some counties identified more land use types than are in the DWR categorization; for these counties land use types had to be aggregated. Other counties identified fewer land

---

[43] See http://www.co.sacramento.ca.us/assessor/assess-info/changes-in-ownership.html.

[44] If more accurate estimates of land values were required, assessed values could be adjusted by the average rate of inflation in the land market in the study area as follows: $\tilde{C}_{iT} = \dfrac{C_{iT} r^{T-t}}{0.02^{T-t}}$, where T is the current year, t, is the Proposition 13 base year for property i, $\tilde{C}_{iT}$ is the adjusted current property value, C_{iT} is the current assessed value, and r is the average rate of inflation in the land market between t and T. This would essentially remove the effect of the 2% per year increase in assessed value and replace it with the actual rate of inflation in land values, which should provide a better estimate of the current market value of the land.

use types; for these counties single values were used for multiple types in the DWR categorization. Missing values for land use types within counties were estimated using an equation of the form:

$$\hat{C}_{ij} = C_j^R \times \frac{\overline{C}_i}{\overline{C}^R} \tag{5.1}$$

In equation 5.1, $\hat{C}_{ij}$ is the estimated average per acre value for land use type *i* in county *j*, C_j^R is the average per acre value for residential parcels in county *j*, $\overline{C}_i$ is the average value for land use type *i* across all counties for which data were available, and $\overline{C}^R$ is the average value for residential parcels across all counties. Equation 5.1 was used for those counties where data were sufficient to estimate residential land values. For other counties the average per acre value for orchard lands was used in an equation similar to equation 5.1 to estimate the values for the missing land use types. Values for land use types in the seven counties where no data were available were estimated by the average values for all surrounding counties (or the nearest county).

Table 5.1 shows the estimates of average per acre land values for all counties in the Central Valley. There were many county-land use type combinations for which average values had to be imputed using the conventions described above. Nevertheless, the data were sufficient to capture the general nature of the variability in land values across the study area. The last column in Table 5.1 shows the average values for each land use type across all counties with data. The most valuable land use type, commercial, was over $500,000 per acre, more than two orders of magnitude greater than the least valuable type, rice, which was less than $5,000 per acre. Variation across counties was significant as well. In Solano County (and apparently Sacramento County), commercial land was worth over $1 million per acre, nearly twice the average across all counties. In Yuba and Madera counties, rice was worth approximately $2,100 and $1,700 per acre, less than half of the average across all counties.

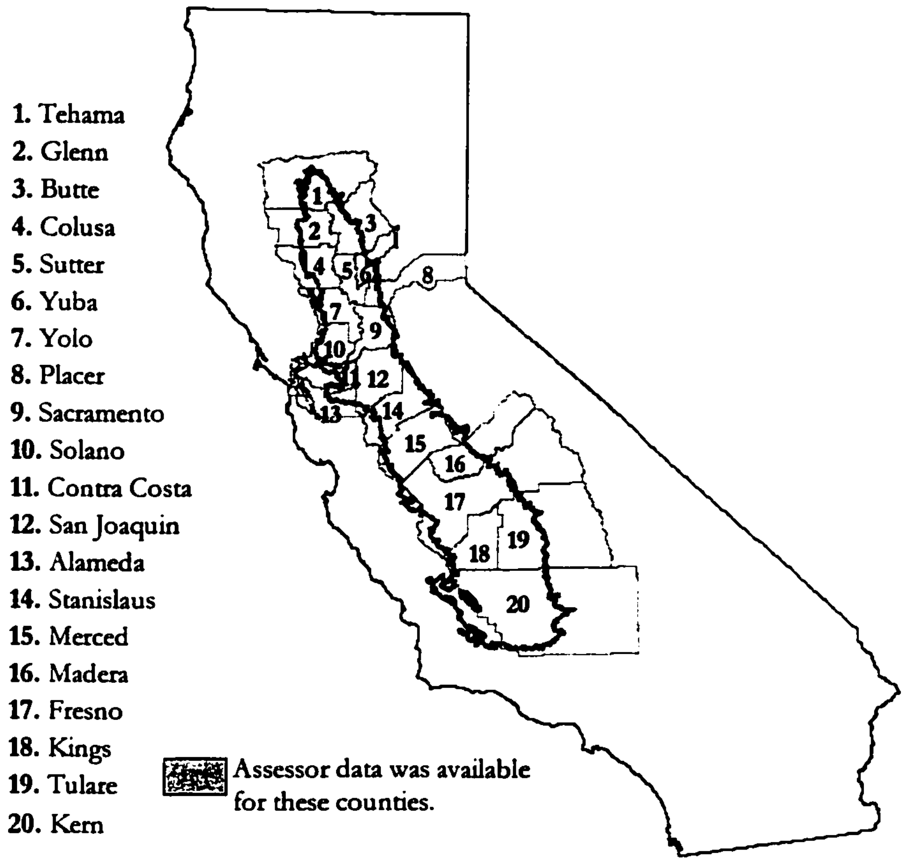


Figure 5.1 - Counties in the Central Valley and land value data availability.

Table 5.1 – Average land values in the Central Valley, by land use type and county [$/acre].

	Butte	Colusa	Contra Costa	Fresno	Glenn	Kern	Kings	Madera	Merced	Placer
Commercial	**558,440**	**594,181**	**970,470**	**448,177**	**270,442**	*448,177*	*448,177*	**346,061**	**324,099**	*598,223*
Industrial	**205,119**	**192,385**	**571,905**	**356,345**	**73,795**	*356,345*	*356,345*	**116,073**	**316,935**	*297,416*
Residential	**390,690**	**396,497**	**1,196,431**	444,344	**267,272**	*444,344*	*444,344*	**157,459**	**124,021**	*437,884*
Urban	387,296	393,053	1,186,037	440,484	264,950	*440,484*	*440,484*	156,091	122,944	*434,080*
Urban landscape	387,296	393,053	1,186,037	440,484	264,950	*440,484*	*440,484*	156,091	122,944	*434,080*
Vacant urban	**208,998**	165,842	**146,979**	185,855	111,791	*185,855*	*185,855*	65,860	51,874	*183,153*
Citrus subtropical	**23,194**	10,476	31,611	**11,237**	7,062	*11,237*	*11,237*	**8,400**	**8,269**	*11,569*
Deciduous fruits nuts	**11,435**	**19,002**	**9,270**	**9,993**	**8,192**	*9,993*	*9,993*	**7,635**	**8,269**	*11,328*
Field crops	**7,418**	5,893	**9,270**	**9,025**	3,972	*9,025*	*9,025*	**4,456**	**5,900**	*7,028*
Grain and hay crops	6,632	6,730	**9,270**	7,543	4,537	*7,543*	*7,543*	2,673	**5,900**	*7,433*
Idle ag	4,696	4,766	14,381	5,341	3,213	*5,341*	*5,341*	1,893	1,491	*5,263*
Pasture	**7,942**	**1,309**	**9,270**	**11,208**	**1,048**	*11,208*	*11,208*	**3,295**	**1,413**	*5,349*
Rice	**3,030**	**3,123**	**9,270**	4,763	**3,382**	*4,763*	*4,763*	1,688	**5,900**	*5,172*
Semi-ag & incidental	10,892	11,053	33,354	**13,484**	**5,682**	*13,484*	*13,484*	**8,359**	**18,350**	*14,940*
Truck nursery berry	5,489	5,570	16,809	6,243	3,755	*6,243*	*6,243*	2,212	1,742	*6,152*
Vineyards	**14,889**	8,691	26,227	**10,009**	5,859	*10,009*	*10,009*	**7,601**	**8,269**	*9,599*

Table 5.1 (continued) – Average land values in the Central Valley, by land use type and county [$/acre].

	Sacramento	San Joaquin	Solano	Stanislaus	Sutter	Tehema	Tulare	Yolo	Yuba	$\overline{C}$
Commercial	1,091,335	**859,418**	**1,114,022**	627,526	*562,604*	*414,441*	*448,177*	*727,173*	**140,730**	562,604
Industrial	538,061	**296,326**	**568,119**	309,389	*277,381*	*139,457*	*356,345*	*365,092*	**76,806**	277,381
Residential	**825,455**	437,373	**809,271**	474,643	*425,538*	*328,981*	*444,344*	*530,731*	**62,660**	425,538
Urban	818,284	433,573	802,240	470,519	*421,841*	*326,123*	*440,484*	*526,120*	62,116	421,841
Urban landscape	818,284	433,573	802,240	470,519	*421,841*	*326,123*	*440,484*	*526,120*	62,116	421,841
Vacant urban	345,261	182,939	338,492	198,527	*177,988*	*160,395*	*185,855*	*221,987*	26,209	177,988
Citrus subtropical	21,809	11,556	**5,115**	12,541	*11,243*	*15,128*	*11,237*	*9,956*	1,656	11,243
Deciduous fruits nuts	18,564	**9,836**	**5,115**	**10,674**	*9,570*	*9,813*	*9,993*	*9,774*	**5,849**	9,570
Field crops	12,268	**4,634**	**5,115**	**8,608**	*6,324*	*5,695*	*9,025*	*6,550*	**2,493**	6,324
Grain and hay crops	14,012	7,424	**5,115**	**8,608**	*7,223*	*5,584*	*7,543*	*6,853*	1,064	7,223
Idle ag	9,922	5,257	**5,115**	5,705	*5,115*	*3,954*	*5,341*	*5,226*	753	5,115
Pasture	9,474	**6,552**	**5,115**	5,448	*4,884*	*4,495*	*11,208*	*5,291*	**1,690**	4,884
Rice	8,847	4,688	**5,115**	5,087	*4,561*	*3,206*	*4,763*	*5,158*	**2,107**	4,561
Semi-ag & incidental	23,012	**19,038**	**5,115**	**14,930**	*11,863*	*8,287*	*13,484*	*12,484*	**9,946**	11,863
Truck nursery berry	11,597	**6,842**	**5,115**	6,668	*5,978*	*4,622*	*6,243*	*5,893*	880	5,978
Vineyards	18,095	**10,086**	**5,115**	10,404	*9,328*	*10,374*	*10,009*	*8,478*	1,374	9,328

Notes: Bold numbers correspond to land use types and counties for which data were available. Regular numbers correspond to values estimated by equation 5.1. Italicized numbers correspond to counties where no data were available. For these counties, average values for all nearby counties were used.

The land value estimates should be representative of the variation in opportunity costs between land use types across the entire study area, but they were not detailed enough to describe variations in costs within counties. Other researchers have used more detailed, parcel level data to analyze variations in property values due to site characteristics and landscape characteristics, including proximity to wetlands. Key results from two such studies are worth briefly reviewing here because they have implications for spatial effects on the cost side of the benefits-cost equation.

Geoghegan et al. (1997) used a hedonic framework to estimate implicit prices for landscape configuration in the Washington D.C. metropolitan area. In addition to the house and property characteristics usually found in standard hedonic models (lot area, number of bathrooms, living area, lake view, age of the house, etc.), they included several measures of landscape heterogeneity at two scales: a diversity index, a fragmentation index, and percent open space, each measured within a 0.1 km buffer and a 1 km buffer. None of the landscape heterogeneity indices were significant individually in the baseline model, except percent open space. However, a varying parameters model was also estimated, to assess differences in the effects of landscape heterogeneity across spatial scales. The parameters associated with the landscape variables were allowed to vary linearly and quadratically with distance, and in this model most of the landscape variables were statistically significant. Even though Geoghegan et al. (1997) did not focus on wetlands, their results are still relevant because they suggest that people have preferences for particular configurations of land use in the vicinity of their homes, and that the intensity of these preferences can be inferred through differences in housing prices.

Doss and Taff (1996) did focus on wetlands. They investigated preferences for proximity to four different types of wetlands – forested, scrub-shrub, emergent vegetation, and open water wetlands – in one county in Minnesota. County assessor data (without the Proposition 13 problem) and NWI data for Ramsey County were used in a hedonic model, which included measures of the distance to each wetland type and distance squared, along with a standard set of housing price variables. The results suggested that a lake view was worth $46,000 on average, and that being 200 meters closer to three of the four wetland types was worth up to $2,900.[45]

The results of these studies suggest that people have preferences for landscape configuration in general, and for proximity to wetlands in particular. Therefore, a general model of land use decision-making would account for the endogeneity of the costs as well as the benefits of wetlands conservation. Restoring a wetland in a particular location could have the effect of increasing nearby property values, thereby making future restoration in the area more expensive. The wetlands restoration case studies described in Chapter 6 focused on parcels of agricultural land fairly distant from urban areas, so ignoring this issue for the present study was not likely to significantly affect the results. Nevertheless, this is an issue that warrants future research.

5.2 The costs of constructing restored wetlands

I estimated the costs of constructing restored wetlands in the Central Valley using cost projections for potential Wetlands Reserve Program (WRP) sites in California for the year 2000. The WRP is a voluntary federal program, run by the Natural Resources Conservation Service of the U.S. Department of Agriculture, which offers landowners the opportunity to receive payments for restoring and maintaining wetlands on their property. In California, the program operates by soliciting offers from landowners (mostly farmers) for parcels to be

[45] Proximity to forested wetlands was apparently a disamenity. The authors did not speculate on the reason for this result.

enrolled in the program, then selecting from the offered parcels those that are: (1) expected to have a high probability of restoration success, (2) expected to deliver high levels of "biological benefits" – mostly for migratory birds and threatened and endangered species, and (3) relatively inexpensive to purchase an easement for and restore.[46] All of the factors used in the site selection process must be estimated before embarking on the projects, so the cost projections are based on engineering estimates and past experience from the program. There were 87 sites offered for enrollment in the program in the year 2000. Figure 5.2 shows a plot of the projected costs versus the area of each site, and the results of a regression using 83 of the points (four outliers were excluded). The estimated regression equation was:

$$\text{Projectedcost}[\$] = 14{,}057 + 205.2 \times (\text{Parcelacres}) \quad R^2 = 0.96 \tag{5.2}$$

The intercept in equation 5.2 can be interpreted as a fixed cost, which must be paid no matter the size of the site, and the slope as a (per acre) variable cost.

There is one important component of wetlands restoration costs that was not accounted for here: the cost of securing rights to, and delivering, sufficient water to wetlands restoration sites in the valley. Wetlands in the Central Valley require approximately 5.75 ft of water per year to maintain normal functions, and largely due to the cost of water the operation and maintenance costs of private wetlands in the region average between $75-150 per acre per year (CVPIA 2000). At an interest rate of 5%, that is a net present value of $1,500-3,000 per acre. If the cost of water were constant throughout the study area, then these figures could just be added to the variable costs in equation 5.2, but in fact water costs vary across water districts in the valley. However, there is evidence to suggest that in California water is quantity constrained, so that the crucial issue for wetlands restoration will

[46] See the Appendix for a discussion of the California WRP site selection process.

generally be the availability of water, not the price of water (Kanazawa 1993; Moore and Dinar 1995). There is also evidence to suggest that that the availability of water and price differentials (e.g., as a consequence of federal water subsidies) are capitalized into land values (Huffaker and Gardner 1986). Therefore, adding water costs to the estimates of land values based on county assessor data could double-count the cost of water. The specifics of the water costs could be important for case-specific management applications, but that extension was left for future work.

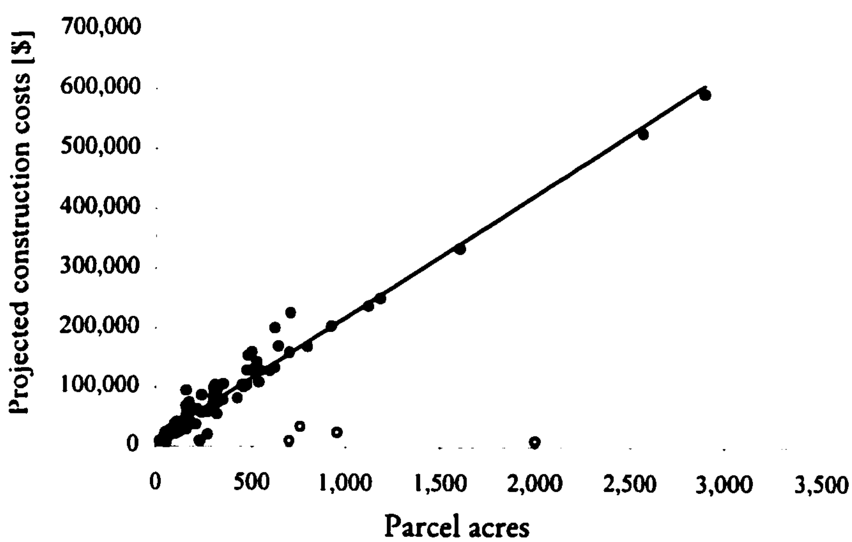


Figure 5.2 – Projected restoration costs versus area for sites offered for inclusion in the WRP in California in the year 2000. Empty circles are outliers, which were excluded from the regression.

Chapter 6 – An Integrated Model for Prioritizing Wetlands Restoration Activities

This chapter describes an integrated optimization model for prioritizing sites for wetlands restoration, and presents results from applications of the model to two case studies in the Central Valley of California. To develop the integrated model, the production function for habitat benefits from Chapter 3, the production function for water quality benefits from Chapter 4, and the estimates of restoration costs from Chapter 5, were all combined into a numerical optimization framework, similar to the one described in Chapter 2. Before describing the integrated model and presenting the results, it will be useful to briefly revisit the standard reserve site selection problem introduced in Chapter 1, to highlight some of the contributions of this research.

6.1 The reserve site selection problem revisited

The reserve site selection literature demonstrates that in a world of limited conservation resources systematic approaches to site selection problems can have important implications for environmental policies. However, to date the reserve site selection literature has largely ignored a number of important features of the general problem of targeting conservation activities cost-effectively. Four important features that have often been ignored in reserve site selection applications, but were addressed in this research, are: (1) variation in costs across sites, (2) land use changes, (3) spatial effects, and (4) tradeoffs between multiple objectives.

Recall that the standard reserve site selection problem is to maximize the number of species protected in a reserve network that can contain only a limited number of sites. Because the constraint in the standard problem is on the number of sites (or area) that can be included in the network, the implicit assumption is that all sites are of equal cost (or equal cost per area). Several researchers have pointed out that more effective methods would account for differences in costs across sites and maximize species protection subject to a budget constraint. In a reserve site selection application for endangered species in the continental United States, Ando et al. (1998) estimated that the cost of protecting half of the species in their dataset in a budget constrained reserve network would be approximately one third the cost of a site constrained network. Polasky et al. (2001a) estimated that the cost of protecting up to 85% of the terrestrial vertebrates in Oregon in a budget constrained reserve network would be less than 10% of the cost of a site constrained network.

In addition to ignoring variations in costs across sites, most reserve site selection applications have ignored the effects of land use changes on species persistence. Land use changes are apparently presumed imminent outside of the reserve network, but these changes are left unspecified.[47] The only management option considered is the establishment of a "reserve" on the selected sites, which would entail protecting the sites in their current condition and prohibiting any incompatible uses on them. Because the species are assumed present, no changes on the sites are envisaged aside from their legal status – thus the "reserve" site selection problem, not the "management" site selection problem or the "restoration" site selection problem. A more general model would allow for enhancement

[47] The motivation of the problem seems to be based on the idea that all land outside of the reserve network will be completely unsuitable as habitat (at least eventually). Those portions of the landscape that cannot be protected are assumed to contribute nothing to species persistence. This simplification has some appeal in that it seems to be a prudent use of the "precautionary principle." On the other hand, it could be grossly inefficient because it would promote the use of scarce conservation resources to ensure the protection, wholly within reserves, of species that could find suitable habitat both inside and outside of reserves.

or restoration, where the selected sites could be modified to provide habitat for species that do not currently occur there. An even more general model would account for both endogenous land use changes (influenced by the manager directly) and exogenous land use changes (influenced by forces outside of the manager's control). The model described in this chapter incorporates the former generalization; the latter was left for future work.

As discussed in Section 1.4, most reserve site selection applications also have ignored spatial interactions that could affect species persistence. In lieu of models of species-habitat relationships, many reserve site selection applications have relied on large-scale species range maps, and a species was considered protected if it occurred on at least some threshold number of sites in the network (refer back to expressions 1.2a-c). As Polasky et al. (2001) pointed out, species persistence is generally a complex function of the amount and type of land set aside, as well as its spatial configuration. In the ecology literature, spatially sophisticated treatments of species behavior, population dynamics, and habitat preferences are common (see Tilman and Kereiva (1997) for a survey), but spatially explicit models have yet to find wide use in applications of terrestrial reserve site selection.[48] If spatial interactions do affect species abundances, then the benefits of including a site in the network would be a function of the size and location of other protected sites, because these would affect the configuration of the surrounding landscape. Therefore, the benefits of management on each site would be endogenous with respect to the decisions to manage other sites. Accounting for this endogeneity requires spatial models of species-habitat relationships, in addition to an explicit treatment of land use changes. The model presented in this chapter can account for such spatial effects.

[48] They are not so uncommon in research on marine reserves. In particular, Sanchirico and Wilen (1999 and 2001) used spatial bioeconomic models to investigate questions of marine reserve design.

Another feature that most reserve site selection applications have ignored is the possibility that protected land will deliver public benefits other than those associated with the protection of species. Beyond merely defining benefits in terms of the number of species included in the reserve network, Polasky et al. (2001) discussed the possibility of attaching weights to each species to represent their relative values to society, and the possibility of using a measure of taxonomic diversity as the objective. However, there has been virtually no consideration in the literature of other types of public benefits from protected areas, such as opportunities for recreation, amenity values provided by proximity to natural areas, water quality maintenance, or flood control benefits, to name but a few possibilities. And because multiple benefits are rarely considered, tradeoffs between competing objectives are not often addressed in the reserve site selection literature (but see Calkin et al. 2002). The model presented in this chapter accounts for both habitat and water quality benefits of wetlands restoration, and it can assess tradeoffs between the two.

To sum up, this chapter addresses a site selection problem where (1) the variation in costs across sites were accounted for, (2) sites were restored instead of protected, so land use changes from management decisions were modeled directly, (3) spatially explicit models of the production functions for ecosystem services were used in the objective function, and (4) two classes of environmental benefits were considered and the tradeoffs between them assessed.

6.2 The integrated model

As in Chapter 2, the modeling approach taken here presumes a manager charged with the task of choosing sites for wetlands restoration with a limited budget. The manager would like to provide both habitat and water quality benefits, but does not know their relative values and therefore wants to consider the set of solutions that maximize all possible combinations of weights on the two benefits. Chapter 2 discussed the utility of the production possibility frontier, and I will use that concept in this chapter as well.

The integrated optimization model used to investigate tradeoffs between habitat and water quality benefits of wetlands and to prioritize sites for restoration in the Central Valley, can be summarized as follows:

$$\underset{x}{Max}\left[W_H f_H(x) + W_W f_W(x)\right] \tag{6.1a}$$

Subject to:

$$f_C(x) \leq Budget \tag{6.1b}$$

I will refer to this as "the wetlands restoration problem." In expression 6.1a, x is a vector of binary choice variables where x_i is 1 if site i is restored and 0 otherwise; $f_H(x)$ is the expected habitat improvement if wetlands are restored in locations represented by x, which is based on the regression model of habitat selection by breeding mallards described in Chapter 3 (version 6 of the negative binomial model); and $f_W(x)$ is the expected water quality improvement via reductions in nutrient loads, which is based on the spatially distributed hydrologic simulation model described in Chapter 4 (using calibrated parameter values shown in Table 4.3). W_H and W_W are weights applied to the ecosystem services from wetlands, and in expression 6.1b, $f_C(x)$ is the total cost, which is based on the estimates of wetlands restoration costs described in Chapter 5.[49]

[49] One limitation of the model in its current form is that it does not include information on the distribution of hydric soils in the study area. In practice, wetlands restoration activities are usually limited to areas that have hydric soils because these are the areas that were most likely wetlands in the past. Hydric soils are common in the Central Valley, but without information on their exact distribution the site selection model described here could identify restoration sites that are not on hydric soils. This is one of the first extension that should be made to the model.

6.3 Solving the wetlands restoration problem

The integrated optimization model is based on the grid described in Chapter 4 for the water quality model, which means that in the extreme case there would be 1.48 million decision variables; each cell could be considered for wetlands restoration independently. Due to the large number of decision variables, the complexity of the production functions for ecosystem services, and the fact that restoration benefits are endogenous, solving this extreme case is not feasible. Operations researchers have developed a number of heuristics for large combinatorial optimization problems, but their performance generally depends on the nature of the problem to which they are applied (Reeves 1993, Michalewicz and Fogel 2000). The best that can be done for the wetlands restoration problem is to generate good candidate solutions by using a heuristic designed to account for the endogeneity of benefits as much as possible.[50] I used a simple heuristic to solve the wetlands restoration problem, but one that took advantage of the structure of the production functions for ecosystem services to make the problem tractable. The heuristic used to solve the optimization problems can be summarized as follows:

> Step 1: Calculate the baseline habitat and water quality conditions in the study area.
> Step 2: Simulate restoration of each site independently and measure benefits and costs.
> Step 3: Select the site with the highest benefit-cost ratio.
> Step 4: Re-calculate the benefits and costs of all sites that interact with the last site chosen.
> Step 5: Repeat Steps 3 and 4 until the budget is exhausted.

This is a greedy algorithm, and it is similar to the iterative algorithm used in the stylized wetlands restoration scenario in Chapter 2. Under certain conditions an algorithm that

---

[50] From now on I will refer to "generating good candidate solutions" to the problem as "solving" the problem. It should be understood, however, that better solutions may exist, and future work on more efficient algorithms for these types of problems is warranted.

excludes Step 4 would guarantee the globally optimal solution, but one of those conditions is that the benefits and costs of restoring the sites must be independent, which is not the case for the wetlands restoration problem. The reason that Step 4 is necessary is that the benefits and costs of restoring any particular cell are endogenous with respect to the decisions to restore other cells. If the heuristic described above is thought of as a walk down the marginal benefits curve, then the endogeneity due to spatial effects means that the marginal benefits curve can not be traced out by calculating the benefits and costs of restoring each cell alone. All relevant combinations must be considered. The key is to consider *only* the relevant combinations. By taking advantage of the spatial structure of the production functions for ecosystem services, the heuristic focuses in each iteration on only those sites that can possibly be affected by other sites chosen for restoration. A closer look at each production function will make this clear.

In the water quality production function, wetlands attenuate nitrogen from surface water runoff as it flows through them from land higher in the watershed on its way to the river, so the amount of nitrogen that a wetland attenuates depends on whether or not another wetland is restored uphill from it. The wetland higher in the watershed would intercept and attenuate some of the nitrogen in the runoff before it reached the wetland lower in the watershed. Two wetlands in a row, along a single overland flow path, do not do twice the work of a single wetland. In Step 2 of the heuristic, benefits are computed as if no other sites will be restored, which means that after the first site is selected the benefits for (some) other sites must be re-calculated to account for the change in the landscape. To make the heuristic as efficient as possible, the watershed was divided into its constituent drainsheds, which are those sets of upland cells that drain to the river through a single cell. Figure 6.1 shows the drainsheds for watershed 17. In Step 4 of the heuristic, only the sites

Figure 6.1 - The drainsheds in watershed 17.

that share the same drainshed as the site selected for restoration in the previous iteration must be re-calculated.

An even more efficient shortcut was available for the production function for habitat quality. The integrated optimization model uses version 6 of the negative binomial regression model from Chapter 3 to calculate the average mallard abundance, $\hat{\mu}$, within 400 meters of the center of each cell. The percent of each land use type is calculated in the neighborhood of the cell as shown in Figure 6.2. The average abundance on the cell is estimated as $\frac{\hat{\mu} \times 200^2}{\pi \times 400^2}$, and $f_H(x)$ was calculated as the sum of the average abundances over all cells in the study area. Given the structure of the mallard model, the habitat benefits of restoring a particular site are endogenous with respect to the decisions to restore nearby sites - the abundance on any given cell can be affected only if it or another cell within 400 meters is restored to wetlands. So, in Step 4 of the heuristic only those sites within 400 meters of the site selected in the previous iteration must be re-calculated. Because the model considers both ecosystem services simultaneously, the sites that require re-calculating in each iteration include those in the drainshed and any extra sites not in the drainshed but still within 400 meters of the last site selected.

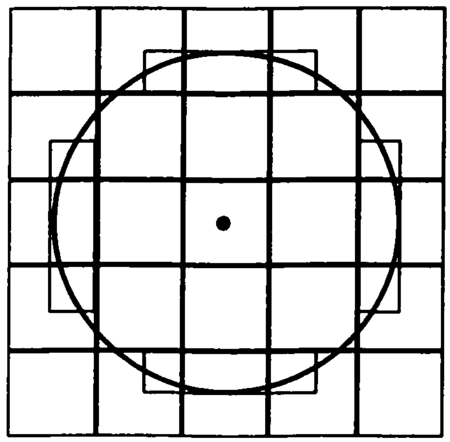


Figure 6.2 - The grid neighborhood used to calculate the percent of all land use types within 400 meters of the center of each cell.

6.4 Two case studies

I applied the integrated model to two case studies of wetlands restoration in the Central Valley. The first case study considered wetlands restoration in four small watersheds. Each watershed was treated separately, but within each watershed all dry land agriculture cells were treated as "sites," i.e. considered available for purchase and restoration to wetlands. The second case study considered wetlands restoration in the entire Central Valley, but only a limited number of sites (contiguous groups of cells in this case) were treated as available for purchase and restoration to wetlands. The optimization problem was solved for each case study to determine the optimal locations for wetlands restoration and the levels of each ecosystem service expected from the modified landscapes. The following sections describe the case studies in more detail and present results from each.

6.4.1 The watersheds case study

In the first case study the integrated model was applied to four of the watersheds used to calibrate the water quality model (refer back to Section 4.2.5.1 and Figure 4.7). Watersheds 13, 16, 17, and 18 were chosen because they were the largest watersheds for which the model could be solved in a reasonable amount of time. The baseline land use configurations for the four watersheds are shown in Figure 6.3. To make comparisons across watersheds as meaningful as possible, the budget for each was set proportional to the area of the watershed; specifically: $Budget = \$[50 \times \text{Hectares in watershed}]$. I investigated the effects of changing the budget in watershed 18.

The results for watershed 13 are shown in Table 6.1 and Figure 6.4. Table 6.1 lists the baseline conditions in the watershed, the restoration budget, and the results from solving the optimization problem at six different levels of W_W and W_H. Other combinations of weights did not produce different results. Figure 6.4 shows the PPF for the watershed and the restoration activities associated with each point on the PPF (the numbers next to the points that make up the PPF correspond to the numbered panels that show the cells selected for restoration). There were only three points on the PPF, and they were all close to the origin, which suggests that the few opportunities for wetlands restoration in the watershed would lead to only small improvements in water quality and habitat quality. Given \$68,200 for wetlands restoration, the greatest possible improvement in water quality is an 11% decrease in nitrogen loads from surface runoff and the greatest possible improvement in habitat quality is a 3% increase in total mallard abundance in the breeding season. Note that restoration activity always occurred on pasture cells, which were the least expensive dry land agriculture cells in the watershed. The solution that maximized mallard abundance consisted of the two pasture cells that were farthest from orchards (recall the negative effect of orchards on mallard abundances from the regression models in Chapter 3). The solution that maximized water quality consists of the two pasture cells that were closest to the river, and the intermediate solution consisted of one cell far from orchards and one cell close to the river. The results for this watershed show clearly that the model generates solutions consistent with the spatial effects built into the production functions for ecosystem services, and that differences in costs exert a strong influence on the results as well.

The results for watershed 16 are shown in Table 6.2 and Figure 6.5. There were more opportunities for wetlands restoration in this watershed than in watershed 13, especially for water quality, and there were stark tradeoffs between improvements in the two ecosystem services. Given $413,800 for wetlands restoration, the greatest possible improvement in water quality is a 45% decrease in nitrogen loads from surface runoff and the greatest possible improvement in habitat quality is a 21% increase in total mallard abundance in the breeding season. The solutions that resulted in large improvements in water quality (points 2, 3, and 4) yielded small improvements in habitat quality, and the solution that resulted in the largest improvement in habitat quality (point 1) yielded virtually no improvement in water quality. There was also no overlap spatially between the solutions that maximized habitat quality and water quality. The solution that maximized habitat quality was completely in a pasture, in the upper left portion of the watershed, and near rice cells, maintaining a mix of wet and dry land in the vicinity. The solutions that delivered large improvements in water quality were on the other side of the river – the side with most of the contributing area in the watershed. The best cells for improving water quality bordered the river, but the solutions that delivered large improvements in water quality also included some pasture cells higher in the watershed because they were relatively inexpensive.

The results for watershed 17 are shown in Table 6.3 and Figure 6.6. In this watershed there were relatively large gains in habitat quality possible, and lower gains in water quality possible, which was the reverse of the situation in watershed 16. Given $198,200 for wetlands restoration, the greatest possible improvement in water quality is a 20% decrease in nitrogen loads from surface runoff and the greatest possible improvement in habitat quality is a 51% increase in total mallard abundance in the breeding season. Again, the solution that maximized habitat quality consisted entirely of pasture cells near rice cells, maintaining a mix of wet and dry land in the vicinity, and again the solution that maximized water quality consisted of some cells near the river and some pasture cells higher in the watershed.

The results for watershed 18 are shown in Table 6.4 and Figure 6.7. Given $457,600 for wetlands restoration, the greatest possible improvement in water quality is a 35% decrease in nitrogen loads from surface runoff and the greatest possible improvement in habitat quality is a 26% increase in total mallard abundance in the breeding season. In this watershed the PPF looked more like a smoothly concave text book version of the curve, but there was still a pronounced gap between the solutions that resulted in large water quality improvements and those that resulted in large habitat quality improvements. Unlike in watersheds 16 and 17, in this watershed the solutions were very close to each other spatially, and the solution that maximized water quality was on the side of the river with the smallest area. This was because the larger side was dominated by natural uplands and therefore contributed a lower nutrient load than the smaller side of the watershed. Restoration again occurred mostly on pasture cells, but there was a surprisingly large amount of variation in benefits possible from different configurations of restoration activities within the single pasture.

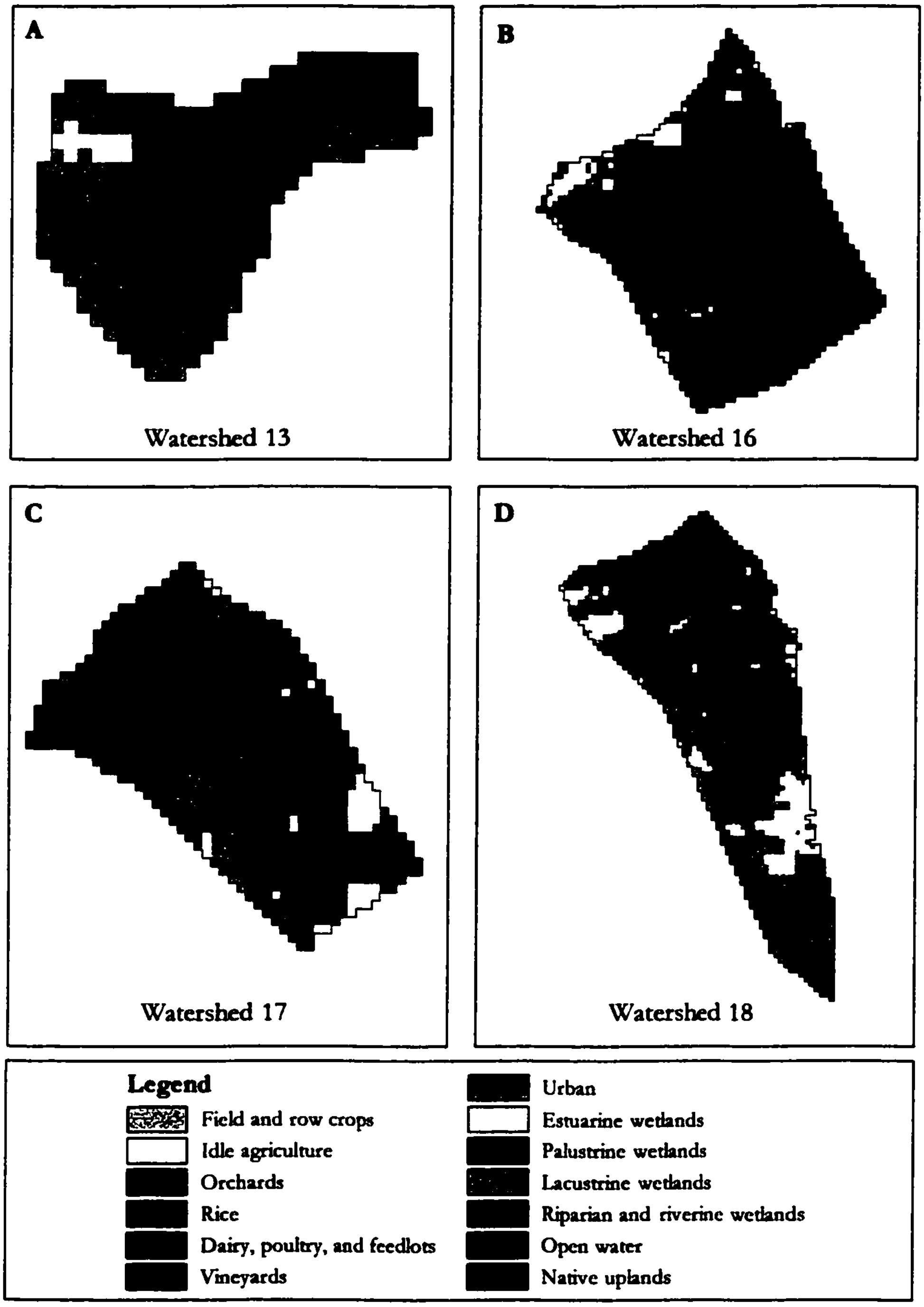


Figure 6.3 – Baseline land use configurations in the four watersheds in which the integrated model was solved.

Table 6.1 – Results from watershed 13.

WATERSHED AREA: 1,364 ha

INITIAL CONDITIONS:

% Urban	27.32
% Agriculture	44.30
% Rice	0.27
% Native uplands	3.45
% Open water	0.00
% Wetlands	24.67

Baseline nitrogen load	2354.0 kg/yr
Baseline mallard abundance	80.3 individuals

BUDGET: $68,200

W_W	W_H	Predicted nitrogen load reduction	Predicted mallard abundance increase
		[kg/yr]	[individuals]
0.0000	1.0000	44.5	2.3
0.0110	0.9890	44.5	2.3
0.0115	0.9885	215.9	1.1
0.0210	0.9790	215.9	1.1
0.0215	0.9785	257.9	0.22
1.0000	0.9785	257.9	0.22

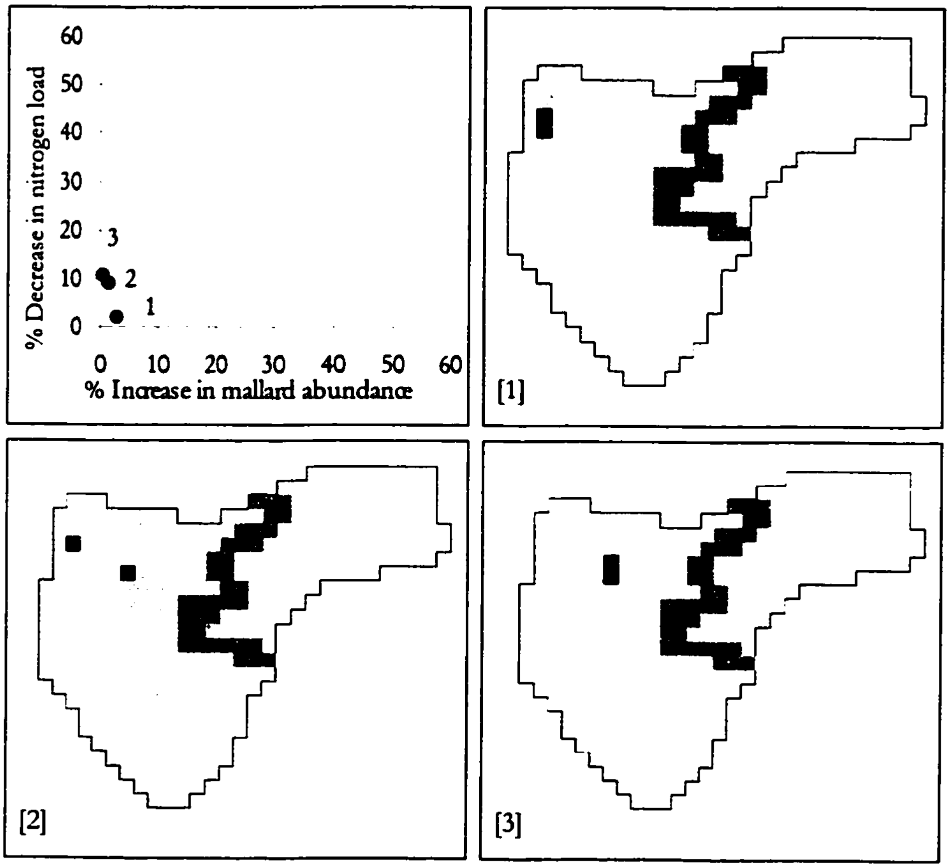


Figure 6.4 – The PPF and associated restoration activities for watershed 13.

Table 6.2 – Results from watershed 16.

WATERSHED AREA: 8,276 ha

INITIAL CONDITIONS:

% Urban	5.8
% Agriculture	17.7
% Rice	10.0
% Native uplands	63.6
% Open water	2.0
% Wetlands	0.8

Baseline nitrogen load	5,117.7 kg/yr
Baseline mallard abundance	133.6 individuals

BUDGET: $413,800

W_W	W_H	Predicted nitrogen load reduction [kg/yr]	Predicted mallard abundance increase [individuals]
0.0000	1.0000	54.9	27.7
0.0050	0.9950	54.9	27.7
0.0100	0.9900	1,905.8	4.37
0.0150	0.9850	2,182.5	2.65
0.0200	0.9800	2,309.7	0.527
0.0250	0.9750	2,309.7	0.527
0.0750	0.9250	2,310.5	0.527
1.0000	0.0000	2,310.5	0.527

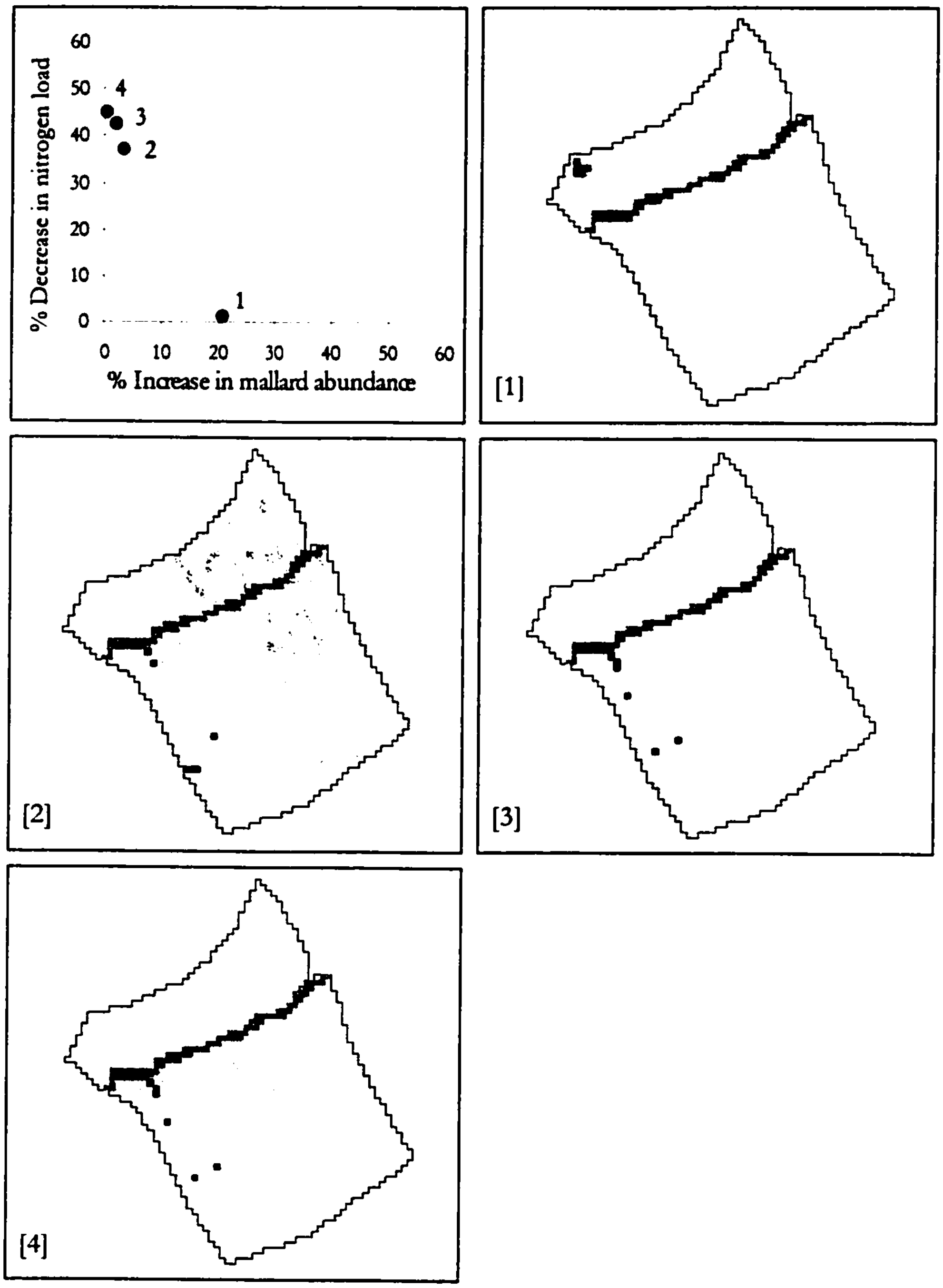


Figure 6.5 – The PPF and associated restoration activities for watershed 16.

Table 6.3 – Results from watershed 17.

WATERSHED AREA: 3,964 ha

INITIAL CONDITIONS:

% Urban	11.5
% Agriculture	51.5
% Rice	16.3
% Native uplands	17.6
% Open water	0.5
% Wetlands	2.7

Baseline nitrogen load	9,370.6 kg/yr
Baseline mallard abundance	62.9 individuals

BUDGET: $198,200

W_W	W_H	Predicted nitrogen load reduction [kg/yr]	Predicted mallard abundance increase [individuals]
0.0000	1.0000	162.6	31.66
0.0050	0.9950	162.6	31.66
0.0075	0.9925	178.5	28.19
0.0100	0.9900	651.6	23.39
0.0250	0.9750	651.6	23.39
0.0400	0.9600	1,083.3	14.73
0.0500	0.9500	1,795.0	3.07
0.1000	0.9000	1,795.0	3.07
0.1500	0.8500	1,799.4	2.50
0.1250	0.8750	1,799.4	2.50
1.0000	0.0000	1,858.0	0.01

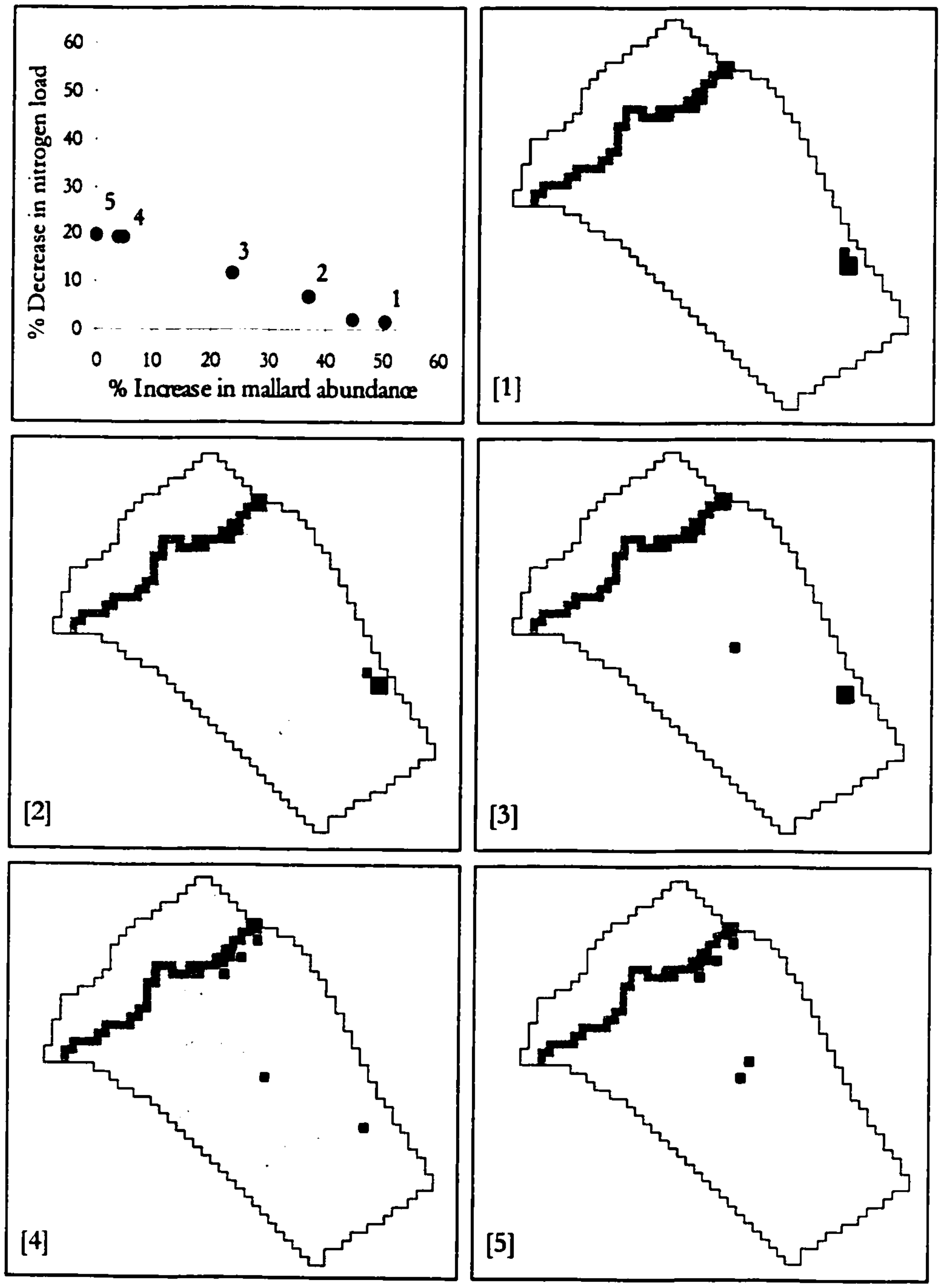


Figure 6.6 – The PPF and associated restoration activities for watershed 17.

Table 6.4 – Results from watershed 18.

WATERSHED AREA: 9,152 ha

INITIAL CONDITIONS:

% Urban	5.6
% Agriculture	18.3
% Rice	0.0
% Native uplands	0.3
% Open water	2.9
% Wetlands	69.2

Baseline nitrogen load	4,507.2 kg/yr
Baseline mallard abundance	149.2 individuals

BUDGET: $457,600

W_W	W_H	Predicted nitrogen load reduction [kg/yr]	Predicted mallard abundance increase [individuals]
0.0000	1.0000	41.07	39.67
0.0010	0.9990	41.07	39.67
0.0020	0.9980	196.82	39.45
0.0031	0.9970	196.82	39.45
0.0031	0.9969	1,289.93	23.54
0.0150	0.9850	1,312.12	23.79
0.0200	0.9800	1,343.85	22.09
0.0400	0.9600	1,347.55	22.09
0.0500	0.9500	1,386.88	17.74
0.1250	0.8750	1,510.39	12.82
0.2000	0.8000	1,505.09	13.98
0.2500	0.7500	1,524.63	8.83
0.3000	0.7000	1,524.63	8.83
0.3750	0.6250	1,515.83	12.62
0.5000	0.5000	1,515.83	12.62
1.0000	0.0000	1,517.12	10.60

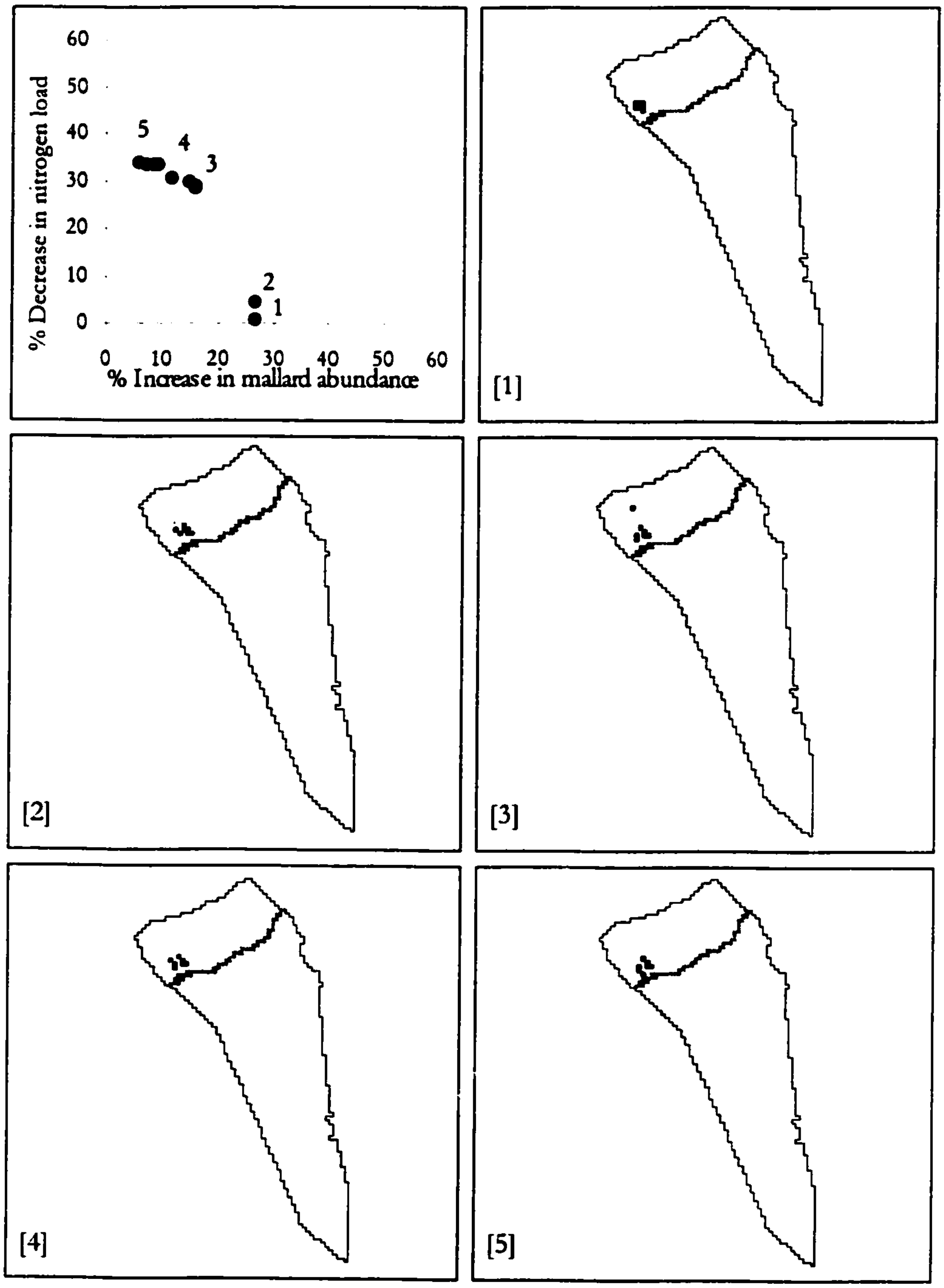


Figure 6.7 – The PPF and associated restoration activities for watershed 18.

As a final exercise for the watersheds case study, I varied the restoration budget in watershed 18 to assess its effect on the PPF. Increasing the budget should shift the PPF out and decreasing it should shift it in, but it is not clear a priori how far it will shift and whether or not some portions of the curve will shift more than others. The increments of the shifts and the shapes of the shifted PPFs will provide an indication of the "returns to scale" in the watershed, where "scale" is with respect to the total funds available for wetlands restoration activities in general. Two more sets of optimization problems were solved for watershed 18, one with a budget of $915,200 and one with a budget of $228,800 (twice and half that used to generate the results in Table 6.4 and Figure 6.7). The results of this exercise are given in Figure 6.8, which shows that water quality improvements initially increased rapidly then flattened out, while habitat quality improvements remained high for a greater range of the

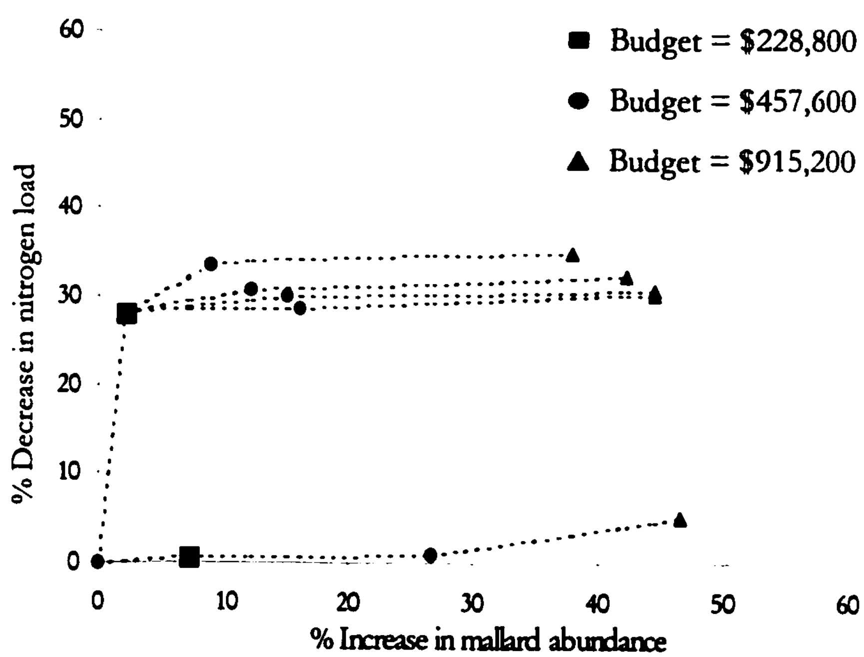


Figure 6.8 – Shift of the PPF with a change in the restoration budget.

budget. The results suggest that up to a 35% reduction in nitrogen loads from surface runoff can be achieved by restoring wetlands in a few strategic locations, but achieving removal rates greater than 35% will be difficult.

6.4.2 The WRP case study

In the second case study the integrated model was applied to the entire Central Valley. In this case, however, the decision variables did not correspond to individual cells. Instead, a set of "sites" (varying numbers of contiguous agriculture cells) was constructed randomly, in an effort to mimic as closely as possible the set of parcels offered for inclusion in the WRP in California in the year 2000. In that year, 83 of the 87 parcels offered for inclusion in the WRP were in counties that overlap the Central Valley. Detailed information on the locations of the offered parcels was not available. However, the number of parcels offered in each county and the area of each parcel was known. To simulate the set of parcels offered for inclusion in the WRP, I chose 83 sets of contiguous agriculture cells at random according to the county-level distribution of the parcels offered in 2000. The number and the size range of the parcels offered in 2000 were replicated for each county. For example, in Tulare County ten parcels were offered, the smallest of which was 20 hectares and the largest 453 hectares. Therefore, I chose 10 sets of contiguous agriculture cells totaling between 20 and 453 hectares at random from Tulare County and treated them as offerings to be considered by a hypothetical WRP manager. Sites were constructed in a similar manner for other counties, and the entire set served as the basis of the optimization problem. The wetlands restoration problem was then solved for a range of W_H and W_W to delineate the production possibilities frontier, as in the watersheds case study. In this case the decision variable x_i, where $i = 1,2,\ldots 83$, referred to whether or not site i (each of which consisted of multiple contiguous cells) was chosen for inclusion in the WRP. This process was repeated multiple times to generate a distribution of possible outcomes, all loosely analogous to the situation in California in the year 2000.

I used the same optimization heuristic for the WRP case study as for the watersheds case study, but in this case finding the globally optimal solution was virtually guaranteed because the offered sites generally only interacted with each other weakly. The water quality benefits of restoring a particular site could only be affected by restoration of other sites that happened to be in the same drainshed, an unlikely occurrence given the small number of sites considered and the large study area. Similarly, the habitat benefits of restoring a particular site could only be affected by restoration of other sites that happened to be within 400 meters. As a result, Step 4 of the optimization algorithm was greatly simplified in this scenario.

Table 6.5 presents summary output for 50 repetitions of the WRP scenario. For simplicity I focused only on the endpoints of the PPFs – the maximum habitat and water quality benefits attainable and the water quality and habitat benefits associated with those solutions. The results were quite consistent across the repetitions, in spite of the wide latitude inherent in the randomization algorithm for defining the locations of the sites. In aggregate, the results in Table 6.5 suggest that at current funding levels (about $11,000,000 per year), the WRP in California could restore approximately 1,000 hectares of wetlands, or increase the abundance of mallards in the breeding season by approximately 650 individuals, or decrease nitrogen loads to rivers and streams from surface runoff by approximately

12,000 kg per year.[51] However, not all of these improvements can be made simultaneously; there are significant tradeoffs to consider.

In all repetitions the area of wetlands restored to maximize water quality was less than the area restored to maximize habitat quality, a result that can be attributed to the strong spatial effects embodied in the production function for water quality. The structure of the production function for water quality was such that the benefits of restoring wetlands in very specific locations – close to the river and in those drainsheds with a large amount of contributing area – were much higher than other locations. This was due to the

Table 6.5 – Results from the WRP case study.

	Average	Coefficient of Variation
	[Kg/yr]	
Baseline nitrogen load	4,673,056	NA
W^{max}	12,377	0.226
$W \mid H^{max}$	3,691	0.281
$W \mid Area^{max}$	7,889	0.327
	[individuals]	
Baseline mallard abundance	77,000	NA
H^{max}	655.5	0.111
$H \mid W^{max}$	324.1	0.358
$H \mid Area^{max}$	454.7	0.168
	[Hectares]	
Baseline wetland area	317,908	NA
$Area^{max}$	1,016	0.075
$Area \mid W^{max}$	702.2	0.130
$Area \mid H^{max}$	867.6	0.110

heterogeneity in the spatial distribution of surface runoff and associated pollutant loads across the landscape. In the water quality model nonpoint source pollution is diffuse, but far

[51] The numbers in Table 6.5 are likely underestimates because most WRP participants accept easement payments that are well below the market value of their land. This is because the landowner retains the title to the land and the right to use it in ways compatible with maintaining it as a wetland. Therefore, the landowner does not forego all of the private benefits of the land.

from uniform. The production function for habitat quality, on the other hand, implies that mallard abundances are less influenced by the spatial arrangement of wetlands. Because only cells within 400 meters of each other interacted, there were many arrangements of restored cells that could lead to similar levels of habitat benefits. This can be understood by imagining the production functions for habitat and water quality as general functions of the amount (A) and configuration (C) of wetlands: $H = h(A, C)$ and $W = w(A, C)$. The marginal rate of substitution between the two "inputs," $-\frac{dA}{dC}$, is larger for the water quality production function than the habitat quality production function.

The average nitrogen attenuation rate on restored cells points to the increase in effectiveness possible from spatial targeting. The average attenuation rate in the restored wetlands in the WRP case study was approximately 18.1 kg/ha, which was much higher than the average of 1.2 kg/yr in existing wetlands (according to baseline outputs from the water quality model), but much lower than the 68.9 kg/ha in restored wetlands from the watersheds case study. In the only other cost-effectiveness study of the water quality benefits of wetlands restoration that I am aware of, Ribaudo et al. (2001) assumed an average attenuation rate of 200 kg/ha in an analysis in the Mississippi Basin. The results presented here imply that the average rate will depend crucially on the spatial arrangement of the restored wetlands. The difference between the average attenuation rates in restored wetlands in the watersheds case study and the WRP case study arises because the manager in the watershed scenario had a completely free hand to spatially target restoration activities; all cells were treated as available for purchase and restoration. In the WRP scenario the manager was constrained by the initial set of offered sites.

The average increase in abundance of mallards on restored sites in the WRP case study was 0.76 individuals/ha, and for the watersheds case study it was 1.1 individuals/ha. Again, because the marginal rate of substitution between wetland area and configuration was lower for the habitat function, it follows that there would be lower gains in effectiveness from spatial targeting possible, and therefore a less pronounced difference between the two case studies.

The watersheds and WRP case studies were not intended to answer definitively the question of where wetlands should be restored in the Central Valley (though they were a start in that direction). They were intended to estimate, in a general way, the potential gains in effectiveness possible form a spatially targeted approach to selecting sites for wetlands restoration, and the magnitude of the tradeoffs between water quality and habitat quality improvements that managers should expect to face. The results from both the watersheds case study and the WRP case study suggest that significant gains in effectiveness could be possible through a spatially targeted approach, and that there could be significant tradeoffs between objectives to consider. Wetlands policies (in the aggregate) and wetlands management activities (in particular instances) could be more effectively evaluated and designed with these issues in mind, and with the tools described in this dissertation at hand.

Chapter 7 - Conclusions

The research described in this dissertation was predicated on the notion that the goal – or at least one important goal – of public policy related to wetlands is to maximize the net benefits they deliver to society. From that perspective, measuring the public and private benefits and costs of alternative wetlands conservation strategies is an important endeavor. Our ability to measure these benefits and costs depends largely on our understanding of the determinants of wetland ecosystem services, which is where the research described in this dissertation fits into the overall public policy debate regarding wetlands conservation.

This dissertation described several semi-independent pieces of research, but all were executed with the two broad objectives first mentioned in Chapter 1 in mind: (1) to further our understanding of the role that landscape configuration plays in the provision of ecosystem services from wetlands, and (2) to enhance our practical ability to account for spatial effects and tradeoffs between competing environmental objectives when evaluating, designing, and implementing wetlands policies. Section 7.1 reviews some of the main results from the research, and Section 7.2 concludes the dissertation with a brief discussion of potential improvements to suggest directions for future work.

7.1 Summary of the main results

1. *A set of integrated optimization models was developed that can provide a useful framework for analyzing and prioritizing wetlands restoration activities.*

Chapter 2 presented a general numerical framework for analyzing land use decisions when spatial effects and multiple objectives are important. The framework provides a useful way to organize our thinking about land use decision-making, and it provides a means to measure the importance of spatial effects and to compare the effectiveness of different site selection strategies. Results from the optimization exercises in Chapter 2 suggested that site selection heuristics that account for spatial effects could deliver near-optimal solutions to site selection problems. This is important because optimizing algorithms will not be feasible for many real-world problems. Results from the cases studies in Chapter 6 showed that the framework could be implemented in large landscapes using readily available data. It should also be clear that this approach requires a multi-disciplinary effort. The framework provides a rational way to combine models and methods from ecology, hydrology, economics, and operations research, to more effectively analyze and design wetlands conservation strategies.

2. *Regression models using GIS land use data and bird abundance data from the North American Breeding Bird Survey suggested that breeding mallards in the Central Valley of California exhibit a preference for areas with a mix of wet and dry land use types.*

Results from the regression models of mallard habitat selection in Chapter 3 showed that in the Central Valley mallards have a preference for locations with a mix of wet (wetlands or rice) and dry land use types in the breeding season. Too few wetlands are sub-optimal, but too many are sub-optimal as well. Specifically, the results suggest that an optimal landscape for breeding mallards would consist of 71% of each 50-hectare scene in wetlands. This result is consistent with the fact that mallards require upland areas for nesting and wetlands for foraging and brood rearing.

3. *Using regression models to predict changes in total population size from changes in land use is valid only under special circumstances. If individuals are distributed according to an ideal free distribution, then the regression results can be used, along with some extra information, to predict changes in carrying capacity from changes in land use.*

Using regression models to predict changes in total population size from changes in the landscape is not strictly valid because the estimated parameters are conditional on the current population size. If data are collected when the population is not at carrying capacity, then inferences regarding the effects of wetlands restoration based on such data will be biased. In Section 3.4, I described a supplemental model that allows a more ecologically coherent interpretation of the parameters in a standard regression model of species abundances, and with some extra information and a few key assumptions the model could be used to estimate the equilibrium population size that the landscape can support. The model must ultimately be judged on its predictive capabilities, but the fact that it offers a more theoretically consistent means of predicting population-level impacts of land use change certainly warrants further research along these lines.

4. *A large-scale hydrologic simulation model that uses an explicit runoff routing algorithm was developed. The model can estimate nutrient loads to rivers and streams from non-point source runoff, and can be used to support the prioritization of wetlands restoration activities for improving water quality.*

The hydrologic simulation model described in Chapter 4 can be used to estimate average nutrient loading rates to each 200-meter stretch of all rivers and streams in the Central Valley, and to predict changes in loading rates from wetlands restoration activities anywhere in the valley. According to the baseline results 4,670,000 kg of nitrogen and 410,000 kg of phosphorus enter rivers and streams from non-point source surface runoff, and wetlands attenuate approximately 425,000 kg of nitrogen and 25,200 kg of phosphorus in the Central Valley in an average year. The Central Valley water quality model represents a compromise between simpler models that estimate edge-of-field nutrient loads only, and more complex models that require detailed data that is often unavailable.

5. *Differences in restoration costs can exert a strong influence on the solution to the wetlands restoration problem.*

Chapter 5 presented estimates of wetlands restoration costs in the Central Valley. County assessor data was used to estimate the opportunity costs of wetlands restoration, based on average per acre values for each land use type in each county. Projected costs for potential WRP projects in California in the year 2000 were used to estimate wetlands construction costs. The estimated land values varied substantially across land use types and across counties, but the land value data were not detailed enough to characterize costs at the parcel level. Results from the watersheds case study presented in Chapter 6 suggested that restoration costs could have a strong influence on the optimal configuration of wetlands restoration activities. The solutions to the wetlands restoration problem often consisted largely of pasture cells, which were the least expensive to restore, even if those cells were not otherwise ideal for delivering habitat or water quality benefits. Because both the nature of the benefits and the costs were seen to have a large influence on the solutions, analyses that focus exclusively on the benefits side of the equation would be incomplete. Many reserve site selection applications have ignored the variation in costs of setting aside nature reserves. A few researchers have shown that such an omission can lead to substantially less effective conservation prescriptions, and the results of the present research support their findings.

6. *Tradeoffs between habitat and water quality benefits from wetlands restoration can be severe.*

Results from the watersheds case study and the WRP case study in Chapter 6 suggested that decision-makers could face significant tradeoffs between competing environmental objectives when designing wetlands conservation policies. The results from the watersheds case study suggested that the tradeoffs could be especially severe when a manager has a completely free hand to target restoration activities in a watershed. The solutions that maximized water quality benefits delivered very little habitat benefits, and vice versa. The results from the WRP case study suggested that the tradeoffs could be less severe, but still substantial, when a manager must choose from a limited set of sites offered for enrollment in an easement program. The solutions that maximized water quality delivered about half of the maximum possible habitat benefits, and solutions that maximized habitat benefits delivered less than one third of the maximum possible water quality benefits. This research is among the first applications of reserve site selection methods to focus on tradeoffs between competing environmental objectives.

7. *The integrated optimization model can be used to estimate the levels of environmental benefits possible from wetlands conservation programs (such as the WRP) under different levels of funding.*

The results from the WRP case study suggested that at current funding levels the Wetlands Reserve Program in California could decrease nitrogen loads to rivers and streams in the Central Valley by approximately 12,000 kg per year, or increase mallard abundances in the breeding season by approximately 650 individuals per year. The model could also be used to estimate the expected levels of environmental benefits from alternative levels of funding. The figures presented in Chapter 6 likely underestimate the benefits that the WRP can deliver, however, because the costs of purchasing easements is generally lower than the cost of purchasing land outright, which is what the costs estimates based on county assessor data used in the model imply. Nevertheless, this research is among the first to estimate actual levels of ecosystem services from wetlands. The standard approach in the wetlands assessment literature uses relative values and indices of functions, which cannot be readily converted to actual levels.

8. *A spatially targeted approach to wetlands conservation can deliver substantially higher levels of environmental benefits than non-targeted approaches.*

The simulation exercises presented in Chapter 2 showed that when spatial effects are important a spatially targeted approach to prioritizing wetlands restoration activities could deliver higher levels of environmental benefits than a non-targeted approach. The optimization exercise in Chapter 4 showed that the spatial habitat preferences of mallards in the breeding season in the Central Valley could be sufficient to support a spatially targeted approach to restoring wetlands for waterfowl habitat in the region. A simulated 50% increase in the area of wetlands near BBS route-stops was predicted to lead to an increase in total mallard counts of over 300%. The results from the WRP case study in Chapter 6 suggested that a spatially targeted approach to wetlands restoration in the Central Valley – even when a manager must choose from a limited set of sites – could yield an approximately 55% greater decrease in nitrogen load from non-point source runoff and an approximately 45% greater increase in mallard abundances, compared to an approach that maximizes the total area of wetlands restored. The results from the watersheds case study also suggested that there could be significant differences in benefits from restoration activities that are spatially very similar. In watershed 18 in particular, all solutions to the wetlands restoration problem consisted of cells within a single pasture, but the differences in water quality and habitat benefits between the alternative solutions were substantial.

In all, this research has demonstrated that spatial effects can have a strong influence on the level of ecosystem services that wetlands deliver to society, and that there can be significant tradeoffs to consider when designing and implementing wetlands policies. But beyond merely calling attention to these issues, the models developed in this research can be used to measure the magnitude of spatial effects and tradeoffs and ultimately to make more effective environmental policy decisions in their presence.

7.2 Potential improvements and directions for future research

Most of the limitations of the models were discussed in the relevant chapters, but it will be useful to review the major ones here to suggest directions for future work. First, the regression models in Chapter 3 are limited in the scope of their applicability. They cannot be used alone to make predictions about population-level effects of land use changes. However, the model sketched out in Section 3.4 has the potential to ameliorate this shortcoming. Another limitation of the regression models in Chapter 3 was their inability to distinguish between effects of land use near and far. Better model specifications may produce better results, but due to the inherently high level of spatial autocorrelation it will always be difficult to separate these effects. Better data (e.g., from radio telemetry studies of species movements) may be required to address the more difficult questions of spatial habitat preferences of birds and other species in a satisfactory way. Another limitation of the mallard model is its exclusive focus on the breeding season. A more complete model of mallard population dynamics would account for annual migration between, and the conditions on, both the breeding and wintering grounds.

The water quality model could be improved by addressing the following limitations: (1) the paucity of data for calibration, (2) the disparate and mismatched nature of the data used to parameterize the model, and (3) the simplified functional forms used as the foundation of the model. These limitations have been listed in decreasing order of importance; the first improvement that should be made is to gather more data for calibration, and only after the calibration data and input data are sufficiently improved should the model be made any more complex. A limitation that was not discussed in detail in Chapter 4, and one that could go under the third category listed above, is the purely static nature of the model. One of the dynamic processes that is ignored, and one that pertains to the ability of wetlands to perform water quality services in particular, is the possibility that wetlands could become effectively saturated with nitrogen or phosphorus (or other water quality constituents of concern), such that over time their ability to attenuate these pollutants would diminish. The water quality model described in Chapter 4 could provide the foundation for addressing issues of this nature, where the timing as well as the location of management decisions are important, but substantial extensions of the model would be required.

The cost estimates used for this research could be improved by collecting more data to fill the gaps in estimated land values. Data on land values were available for only 13 of the 20 counties in the Central Valley, and many land use types in those 13 counties were not recorded separately by county assessors. The estimates of restoration costs were also incomplete because they excluded the cost of delivering water to the sites. Another limitation of the costs data was that they were not detailed enough to estimate a hedonic model to describe the influence of wetland proximity on land values, which would be necessary to account for the potentially endogenous nature of the costs of wetlands restoration.

Large problems are difficult to solve with the numerical optimization framework used for this research, so heuristics tailored specifically to the problem at hand must be employed. Much of the research described in this dissertation was aimed at developing useful heuristics for the particular problem of wetlands restoration in agricultural landscapes. Nevertheless, even the simple optimization heuristics used here were able to delineate sets of restoration sites that yielded much larger improvements in habitat and water quality than the set that yielded the largest increase in wetland area. Therefore, analyses of the effectiveness of spatial targeting and tradeoffs between different objectives need not be hampered by an inability to apply optimizing algorithms. The integrated optimization model used for the case studies in Chapter 6 could be improved by making it faster and easier to implement. It could be made faster by decreasing the resolution of the model grid, from 200-meter cells to

400-meter cells for example. This would decrease the resolution at which restoration activities could be targeted, but it would allow case studies in larger watersheds, and it would allow more sites to be considered when prioritizing restoration activities throughout the valley.

References

Adamus PR., Clairain, EJ Jr., Smith RD, Young RE. 1987. Wetland Evaluation Technique, Volume II: Methodology. Vicksburg, MS: U.S. Army Corps of Engineers, Waterways Experiment Station.

Adamus RP, Stockwell LT, Clarain EJ Jr., Morrow ME, Rozas LE, Smith RD. 1991. Wetland Evaluation Technique (WET), Volume I: Literature Review and Evaluation Rationale. Wtlands Research Program Technical Report WRP-DE-2. Vicksburg, MS: U.S. Army Corps of Engineers, Waterways Experiment Station.

Anderson DR, Burnham KP, Thompson WL. 2000. Null hypothesis testing: problems, prevalence, and an alternative. *Journal of Wildlife Management* 64(4):912-923.

Ando A, Camm J, Polasky S, Solow A. 1998. Species distributions, land values, and efficient conservation. *Science* 279:2126-2128.

Austin JE, Sklebar HT, Guntenspergen GR, Buhl TK. 2000. Effects of roadside transect width on waterfowl and wetland estimates. *Wetlands* 20(4):660-670.

Bhaduri B, Harbor J, Engel B, Grove M. 2000. Assessing watershed-scale, long-term hydrologic impacts of land-use change using a GIS-NPS model. *Environmental Management* 26(6):643-658.

Bockstael NE. 1996. Modeling economics and ecology: the importance of a spatial perspective. *American Journal of Agricultural Economics* 78:1168-1180.

Bolger DT, Scott TA, Rottenberry JT. 1997. Breeding bird abundance in an urbanizing landscape in coastal southern California. *Conservation Biology* 11:406-421.

Bosch DD, Theurer FD, Bingner RL, Felton G, Chaubey I. 1998. Evaluation of the AnnAGNPS water quality model. Presentation at the 1998 ASAE Annual International Meeting, Orlando, Florida.

Boulinier T, Nichols JD, Hines JE, Sauer JR., Flather CH, Pollock KH. 2001. Forest fragmentation and bird community dynamics: inference at regional scales. *Ecology* 82(4):1159-1169.

Brawley WW, Collins G, Kremer JN, Sham C, Valiela I. 2000. A time-dependent model of nitrogen loading to estuaries from coastal watersheds. *Journal of Environmental Quality* 29:1448-61.

Brinson MM. 1993. A hydrogeomorphic classification for wetlands. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Brinson MM, Rheinhardt RD. 1996. The role of reference wetlands in functional assessment and mitigation. *Ecological Applications* 6(1):69-76.

Brinson MM, Rheinhardt RD. 1998. Wetland functions and relations to societal values. In Southern Forested Wetlands: Ecology and Management. Eds: Messina MG, Conner WH. CRC Press, Boca Raton, FL.

Brown CD, Hollis JM. 1996. SWAT: A semi-empirical model to predict concentrations of pesticides entering surface waters from agricultural land. *Pesticide Science* 47(1):41-50.

Burnham KP, Anderson DR. 1998. Model selection and inference: a practical information theoretic approach. Springer-Verlag, New York, New York, USA.

California Department of Fish and Game. 1999. CWHR Version 7.0 computer program. Sacramento, CA.

Calkin D, Montgomery CA, Schumaker NH, Polasky S, Arthur JL, Nalle DJ. 2002. Developing a production possibility set of wildlife species persistence and timber harvest value using simulated annealing. *Canadian Journal of Forest Research* [in press].

Cameron AC, Trivedi PK. 1998. Regression analysis of count data. Cambridge, Cambridge University Press.

Camm JD, Polasky S, Solow A, Csuti B. 1996. A note on optimal algorithms for reserve site selection. *Biological Conservation* 78: 353-355.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications* 8(3):558-568.

Castillo MM, Allan JD, Brunzell S. 2000. Nutrient concentrations and discharges in a midwestern agricultural catchment. *Journal of Environmental Quality* 29:1142-51.

Central Valley Habitat Joint Venture Implementation Board. 1990. Central Valley Habitat Joint Venture implementation plan: a component of the North American Waterfowl Management Plan. U.S. Fish and Wildlife Service, Sacramento, CA. 102pp.

Cedfeldt PT, Watzin MC, Richardson BD. 2000. Using GIS to identify functionally significant wetlands in the Northeastern United States. *Environmental Management* 26(1):13-24.

Chambers JM, Wrigley TJ, McComb AJ. 1993. The potential use of wetlands to reduce phosphorus export from agricultural catchments. *Fertilizer Research* 36(2):157-164.

Clark CW, Mangel M. 2000. Dynamic state variable models in ecology: methods and applications. New York, Oxford University Press.

Coiner C, Wu J, Polasky S. 2001. Economic and environmental implications of alternative landscape designs in the Walnut Creek Watershed if Iowa. *Ecological Economics* 38:119-139.

Corbitt RA. 1990. Standard Handbook of Environmental Engineering. McGraw Hill, New York, NY.

Cowardin LM, Carter V, LaRoe ET. 1979. Classification of Wetlands and Deepwater Habitats in the United States. U.S. Fish and Wildlife Service. Washington D.C. FWS/OBS 79-31. 103 pp.

Cowardin LM, Shaffer TL, Arnold PM. 1995. Evaluations of duck habitat and estimation of duck population sizes with a remote-sensing-based system. National Biological Service, Biological Science Report 2. 26 p.

Cronk JK. 1996. Constructed wetlands to treat wastewater from dairy and swine operations: A review. *Agriculture, Ecosystems, and Environment* 58(2-3):97-114.

Csuti B, Polasky S, Williams PH, Pressey RL, Camm JD, Kershaw M, Kiester AR, Downs B, Hamilton R, Huso M, Sahr K. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80:83-97.

CVPIA. 2000. Central Valley Wetlands Water Supply Investigations, CVPIA 3406 (d)(6)(A,B). A report to Congress, Final Report, December 2000.

Dahl TE. 1990. Wetlands losses in the United States 1780's to 1980's. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. Jamestown, ND: Northern Prairie Wildlife Research Center Home Page. http://www.npwrc.usgs.gov/resource/othrdata/wetloss/wetloss.htm (Version 16JUL97).

Dale VH. 1997. The relationship between land-use change and climate change. *Ecological Applications* 7(3):753-769.

Davis FW, Stoms DM, Church RL, Okin WJ, Johnson KN. 1996. Selecting biodiversity management areas. *In*: Sierra Nevada Ecosystem Project: Final report to Congress. Davis, University of California, Centers for Water and Wildland Resources.

Detenbek NE, Johnston CA, Niemi GJ. 1993. Wetland effects on lake water quality in the Minneapolis-St. Paul Metropolitan area. *Landscape Ecology* 8(1):39-61.

Diamond JM. 1975. The island dilemma: lessons of modern biogeographic studies for the design of nature reserves. *Biological Conservation* 7:129-46.

Doss CR., Taff SJ. 1996. The influence of wetland type and wetland proximity on residential property values. *Journal of Agricultural and Resource Economics* 21(1):120-129.

Dunning JB, Stewart DJ, Danielson BJ, Noon BR, Root TL, Lamberson RH, Stevens EE. 1995. Spatially explicit population models: current forms and future uses. *Ecological Applications* 5:3-11.

Dzus EH, Clark RG. 1998. Brood survival and recruitment of mallards in relation to wetlands density and hatching date. *The Auk* 115(2):311-318.

Edwards AL, Weakley AS. 2001. Population biology and management of rare plants in depression wetlands of the southeastern Coastal Plain, USA. *Natural Areas Journal* 21(1):12-35.

Elphick CS. 2000. Functional equivalency between rice fields and seminatural wetland habitats. *Conservation Biology* 14(1):181-191.

Flather CH, Sauer JR. 1996. Using landscape ecology to test hypotheses about large-scale abundance patterns in migratory birds. *Ecology* 77:28-35.

Fretwell SD, Lucas HLJ. 1969. On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica* 19:16-36.

Geoghegan J, Wainger LA, Bockstael NE. 1997. Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. *Ecological Economics* 23:251-264.

Gibbs JP. 1993. Importance of small wetlands for the persistence of local populations of wetland-associated animals. *Wetlands* 13:25-31.

Gilliam JW. 1994. Riparian wetlands and water quality. *Journal of Environmental Quality* 23(5):896-900.

Goldhamer DA, Snyder RL. 1989. Irrigation scheduling: a guide for efficient on-farm water management. University of California, Division of Agriculture and Natural Resources, publication 21454.

Gopal B. 1999. Natural and constructed wetlands for wastewater treatment: potentials and problems. *Water Science and Technology* 40(3):27-35.

Grunwald S, Norton LD. 2000. Calibration and validation of a non-point source pollution model. *Agricultural Water Management* 45:17-39.

Hanski I. 1999. Metapopulation Ecology. Oxford University Press, Oxford, UK.

Harper DGC. 1982. Competitive foraging in mallards: 'ideal free' ducks. *Animal Behavior* 30:575-584.

Heitmeyer ME, Connelly DP, Pederson RL. 1989. The Central, Imperial, and Coachella valleys of California. In: Habitat management for migrating and wintering waterfowl in North America. Smith LM, Pederson RL, Kiminski RM, editors. Lubbock, TX: Texas Technical University Press.

Herkert JR. 1998. The influence of the CRP on grasshopper sparrow population trends in the mid-continental United States. *Wildlife Society Bulletin* 26:227-231.

Hof J, Flather CH. 1996. Accounting for connectivity and spatial correlation in the optimal placement of wildlife habitat. *Ecological Modelling* 88:143-155.

Hof J, Bevers M. 1998. Spatial optimization for managed ecosystems. Columbia University Press, New York.

Hof J, Sieg CH, Bevers M. 1999. Spatial and temporal optimization in habitat placement for a threatened plant: the case of the western prairie fringed orchid. *Ecological Modelling* 115:61-75.

Hubbell SP. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton, NJ.

Huffaker RG, Gardner BD. 1986. The distribution of economic rents arising from subsidized water when land is leased. *American Journal of Agricultural Economics* 68(2): 306-12.

Hyman JB, Leibowitz SG. 2000. A general framework for prioritizing land units for ecological protection and restoration. *Environmental Management* 25(1):23-35.

Jenssen PD, Maehlum T, Roseth R, Braskerud B, Syversen N, Njos A, Krogstad T. 1994. The potential of natural ecosystem self-purifying measures for controlling nutrient inputs. *Marine Pollution Bulletin* 29(6-12):420-425.

Johnston CA, Detenbeck NE, Niemi GJ. 1990. The cumulative effects of wetlands on stream water quality and quantity: a landscape approach. *Biogeochemistry* 10:105-141.

Johnston CA. 1991. Sediment and nutrient retention by freshwater wetlands: Effects on surface water quality. *Critical Reviews in Environmental Control* 21(5-6):491-565.

Knight RL, Kadlec RH, Ohlendorf HM. 1999. The use of treatment wetlands for petroleum industry effluents. *Environmental Science and Technology* 33(7):973-980.

Kadlec RH, Knight RL. 1996. Treatment Wetlands. Lewis Publishers, Boca Raton, Florida.

Kanazawa M. 1993. Pricing subsidies and economic efficiency: the U.S. Bureau of Reclamation. *Journal of Law and Economics.* 36(1):205-34.

King DM, Wainger LA, Bartoldus CC, Wakeley JS. 2000. Expanding Wetland Assessment Procedures: Linking Indices of Wetland Function with Services and Values. ERDC/EL TR-00-17, US Army Engineer Research and Development Center, Vicksburg, MS.

Kirkland GL, Ostfeld RS. 1999. Factors influencing variation among states in the number of federally listed mammals in the United States. *Journal of Mammalogy* 80(3):711-719.

Knox JC. 2001. Agricultural influence on landscape sensitivity in the Upper Mississippi River Valley. *Catena* 42(2-4):193-224.

Koenig WD. 1998. Spatial autocorrelation in California land birds. *Conservation Biology* 12:612-620.

Krapu GL, Greenwood RJ, Dwyer CP, Kraft KM, Cowardin LM. 1997. Wetland use, settling patterns, and recruitment in mallards. *Journal of Wildlife Management* 61(3):736-746.

Kronvang B, Grant R, Larsen SE, Svendsen LM, Kristensen P. 1995. Non-point-source nutrient losses to the aquatic environment in Denmark: impact of agriculture. *Marine and Freshwater Research* 46(1):167-77.

Krysanova V, Mueller-Wohlfeil D, Becker A. 1998. Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecological Modelling* 106:261-289.

Lewis WM, Jr. 2001. Wetlands Explained: Wetland Science, Policy, and Politics in America. Oxford University Press, New York.

Lomolino MV. 1994. An evaluation of alternative strategies for building networks of nature reserves. *Biological Conservation* 69(3):243-9.

MacArthur RH, Wilson EO. 1967. The Theory of Island Biogeography. Princeton University press, Princeton, New Jersey.

Manel S, Dias JM, Ormerod SJ. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: acase study with a Himalayan river bird. *Ecological Modelling* 120:337-347.

Margules CR, Pressey RL. 2000. Systematic conservation planning. *Nature* 405:243-253.

Marshall E, Haight R, Homans FR. 1998. Incorporating environmental uncertainty into species management decisions: Kirtland's Warbler habitat management as a case study. *Conservation Biology* 12:975-985.

Martello S, Toth P. 1990. Knapsack Problems: Algorithms and Computer Implementations. John Wiley and Sons, New York.

McAllister LS, Peniston BE, Leibowitz SG, Abbruzzese B, Hyman JB. 2000. A synoptic assessment for prioritizing wetland restoration efforts to optimize flood attenuation. *Wetlands* 20(1):70-83.

McLandress MR, Yarris GS, Perkins AEH, Connelly DP, Raveling DG. 1996. Nesting biology of mallards in California. *Journal of Wildlife Management* 60(1):94-107.

Michalewicz Z, Fogel DB. 2000. How to Solve It: Modern Heuristics. Springer-Verlag Press, Berlin.

Mitsch WJ, Gosselink JG. 1993. Wetlands. 2nd Edition. Van Nostrand Reinhold, New York.

Mitsch WJ, Wilson RF. 1996. Improving the success of wetland creation and restoration with know-how, time, and self-design. *Ecological Applications* 6(1):77-83.

Moilanen A, Smith AT, Hanski I. 1998. Long-term dynamics in a metapopulation of the American pika. *American Naturalist* 152:530-542.

Moore MR, Dinar A. 1995. Water and land as quantity-rationed inputs in California agriculture: empirical tests and water policy implications." *Land Economics* 71(4):445-61.

Mount JF. 1995. California Rivers and Streams. University of California Press, Berkeley, CA.

Murdock NA. 1994. Rare and endangered plants and animals of southern Appalachian wetlands. *Water, Air, and Soil Pollution* 77(3-4):385-405.

National Research Council. 1992. Restoration of Aquatic Ecosystems. National Academy Press, Washington, D.C.

National Research Council. 1995. Wetlands: Characteristics and Boundaries. National Academy Press, Washington, D.C.

Nevo A, Garcia L. 1996. Spatial optimization of wildlife habitat. *Ecological Modelling* 91:271-281.

Newbold SC. 2002. Integrated modeling for watershed management: multiple objectives and spatial effects. *The Journal of the American Water Resources Association* 38(2):341-353.

Noss RF, O'Connell MA, Murphy DD. 1997. The Science of Conservation Planning: Habitat Conservation Under the Endangered Species Act. Island Press, Washington, D.C.

O'Connor RJ, Dunn E, Johnson DH, Jones SL, Petit D, Pollock K, Smith CR, Trapp JL, Welling E. 2000. A programmatic review of the North American Breeding Bird Survey: report of a peer review panel.

O'Neill MP, Schmidt JC, Dobrowolski JP, Hawkins CP, Neale CMU. 1997. Identifying sites for riparian wetland restoration: application of a model to the Upper Arkansas River basin. Restoration Ecology 5(4):85-102.

Owens LB, Van Keuren RW, Edwards WM. 1998. Budgets of non-nitrogen nutrients in a high fertility pasture system. *Agriculture, Ecosystems and Environment* 70:7-18.

Padgitt M, Newton D, Penn R, Sandretto C. 2000. Production practices for major crops in U.S agriculture, 1990-97. Resource economics division, Economic Research Service, USDA; http://www.ers.usda.gov/publications/sb969/sb969c.pdf.

Pearce J, Ferrier S. 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling* 128:127-147.

Peterjohn WT, Correll DL. 1984. Nutrient dynamics in an agricultural watershed: Observations on the role of a riparian forest. *Ecology* 65(5):1466-75.

Polasky S, Camm JD, Solow AR, Csuti B, White D, Ding R. 2000. Choosing reserve networks with incomplete species information. *Biological Conservation* 94(1):1-10.

Polasky S, Camm J, Garber-Yonts B. 2001a. Selecting biological reserves cost-effectively: an application to terrestrial vertebrate conservation in Oregon. Land Economics 77(1):66-78.

Polasky S, Csuti B, Vossler CA, Meyers SM. 2001b. A comparison of taxonomic distinctness versus richness as criteria for setting conservation priorities for North American birds. *Biological Conservation* 97(1):99-105.

Polasky S, Solow AR. 2001. The value of information in reserve site selection. *Biodiversity and Conservation* 10(7):1051-1058.

Possingham HP, Andelman SJ, Noon BR, Trombulak S, Pulliam HR. 2001. Making smart conservation decisions. *In* Conservation Biology: research priorities for the next decade. Editors Soule ME, Orians GH. Washington, D.C.: Island Press.

Prendergast JR, Quinn RM, Lawton JH. 1999. The gaps between theory and practice in selecting nature reserves. *Conservation Biology* 13(3):484-92.

Pressey R L, Possingham HP, Margules CR. 1996. Optimality in reserve selection algorithms: when does it matter and how much? *Biological Conservation* 76:259-267.

Pressey RL, Possingham HP, Day JR. 1997. Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. *Biological Conservation* 80:207-219.

Pulliam HR, Danielson BJ. 1991. Sources, sinks, and habitat selection: A landscape perspective on population dynamics. *The American Naturalist* 137:S50-S66.

Pulliam HR, Dunning JB. 1994. Demographic processes: population dynamics on heterogeneous landscapes. *In* Principles of conservation biology. Editors Meffe GK, Carroll CR. Sunderland, MA: Sinauer Associates, Inc.

Race MS, Fonseca MS. 1996. Fixing compensatory mitigation: what will it take? *Ecological Applications* 6(1):94-101.

Reeves CR (editor). 1993. Modern Heuristic Techniques for Combinatorial Problems. Halsted Press, New York.

Rheinhardt RD, Brinson MM, Farley PM. 1997. Applying wetland reference data to functional assessment, mitigation, and restoration. *Wetlands* 17(2):195-215.

Ribaudo MO, Heimlich R, Classen R, Peters M. 2001. Least-cost management of nonpoint source pollution: source reduction versus interception strategies for controlling nitrogen loss in the Mississippi Basin. *Ecological Economics* 37:183-197.

Russell GD, Hawkins CP, O'Neill MP. 1997. The role of GIS in selecting sites for riparian restoration based on hydrology and land use. *Restoration Ecology* 5(4):56-68.

Sanchirico JN, Wilen JE. 1999. Bioeconomics of spatial exploitation in a patchy environment. *Journal of Environmental Economics and Management* 37:129-150.

Sanchirico JN, Wilen JE. 2001. A bioeconomic model of marine reserve creation. *Journal of Environmental Economics and Management* 42:257-276.

Saunders DA, Hobbs RJ, Margules CR. 1991. Biological consequences of ecosystem fragmentation: a review. *Conservation Biology* 5(1):18-32.

Schumaker NH. 1998. A User's Guide to the Patch Model. EPA/600/R-98/135. U.S. Environmental Protection Agency, Environmental Research Laboratory, Corvallis, Oregon.

Sheafer SE. 1998. Recruitment models for mallards in eastern North America. *The Auk* 115(4):988-997.

Shuford WD, Page GW, Kjelmyr JE. 1998. Patterns and dynamics of shorebird use of California's Central Valley. *Condor* 100(2):227-244.

Simberloff D. 1988. The contribution of population and community biology to conservation science. *Annual Review of Ecology an Systematics* 19:473-511.

Simberloff D, Abele LG. 1982. Refuge design and island biogeographic theory: effects of fragmentation. *The American Naturalist* 120(1):41-9.

Skop E, Sørensen PB. 1998. GIS-based modelling of solute fluxes at the catchment scale: a case study of the agricultural contribution to the riverine nitrogen loading in the Vejle Fjord catchment, Denmark. *Ecological Modelling* 106(2-3):291-310.

Soil Conservation Service. 1973. A Method for Estimating Volume and Rate of Runoff in Small Watersheds. Technical paper no. 149, USDA-SCS, Washington, D.C.

Soranno PA, Hubler SL, Carpenter SR. 1996. Phosphorus loads to surface waters: a simple model to account for spatial pattern of land use. *Ecological Applications* 6(3):865-878.

Soulé ME, Simberloff D. 1986. What do genetics and ecology tell us about the design of nature reserves? *Biological Conservation* 35:19-40.

Stokland JN. 1997. Representativeness and efficiency of bird and insect conservation in Norwegian boreal forest reserves. *Conservation Biology* 11:101-111.

Szentadrasi S, Polasky S, Berrens R, Leonard J. 1995. Conserving biological diversity and the Conservation Reserve Program. *Growth and Change* 26:383-404.

Tarutis WJ, Stark LR, Williams FM. 1999. Sizing and performance estimation of coal mine drainage wetlands. *Ecological Engineering* 12:352-372.

Taylor ML, Adams RM, Miller SF. 1992. Farm-level response to agricultural effluent control strategies: the case of the Willamette Valley. *Journal of Agricultural and Resource Economics* 17(1):173-185.

Tilman D, Kareiva P, editors. 1997. Spatial ecology: the role of space in population dynamics and interspecific interactions. Princeton, NJ: Princeton University Press. 368 p.

Underhill LG. 1994. Optimal and suboptimal reserve selection algorithms. *Biological Conservation* 70:85-87.

U.S.D.A. 2000. Agricultural Resources and Environmental Indicators, 2000. Economic Research Service, Natural Resources and Environment Division.

U.S. EPA. 1977. Nationwide Evaluation of Combined Sewer Overflows and Urban Stormwater Discharges, Volume 1: Executive summary 1977. Municipal Environmental Research Laboratory, Office of Research and Development, U.S. EPA, Cincinnati, OH.

U.S. EPA. 1995. America's Wetlands: Our Vital Link Between Land and Water. EPA843-K-95-001.

van Langevelde F, Schotman A, Claassen F, Sparenburg G. 2000. Competing land use in the reserve site selection problem. *Landscape Ecology* 15:43-256.

Veech JA. 2000. Choice of species-area functions affects identification of hotspots. *Conservation Biology*. 14(1):140-147.

Viessman W Jr., Lewis GL, Knapp JW. 1989. Introduction to Hydrology. 3rd Edition. Harper Collins Publishers, New York.

Walpole SC, Sinden JA. 1997. BCA and GIS: integration of economic and environmental indicators to aid land management decisions. *Ecological Economics* 23:45-57.

Weller CM, Watzin MC, Wang D. 1996. Role of wetlands in reducing phosphorus loading to surface water in eight watersheds in the Lake Champlain Basin. *Environmental Management* 20(5):731-739.

Whigham DF, Chitterling C, Palmer B. 1988. Impacts of freshwater wetlands on water quality: a landscape perspective. *Environmental Management* 12(5):663-671.

Williams JC, ReVelle CS. 1997. Applying mathematical programming to reserve selection. *Environmental Modeling and Assessment* 2:167-175.

Williams P, Gibbons D, Margules C, Rebelo A, Humphries C, Pressey R. 1996. A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology* 10:155-174.

Wright J, ReVelle C, Cohon J. 1983. A multiobjective integer programming model for the land acquisition problem. *Regional Science and Urban Economics* 13:31-53.

Zedler JB. 1996. Ecological issues in wetland mitigation: an introduction to the forum. Ecological Applications 6(1):33-37.

Appendix – Benefits and costs, optimization, and the Wetlands Reserve Program

The Wetlands Reserve Program (WRP) is a voluntary federal program run by the Natural Resources Conservation Service of the U.S. Department of Agriculture. The WRP offers landowners the opportunity to receive payments for restoring and maintaining wetlands on their property.[52] Implementation of the WRP is the responsibility of the states, and in California it generally involves selecting from among parcels offered for enrollment in the program those that are expected to yield the greatest "biological benefits" possible. The expected biological benefits of purchasing and easement for and restoring an offered parcel to wetlands are estimated during field visits by WRP staff with the aid of a "biological ranking criteria worksheet" (see Figure A.1[53]). In brief, the WRP site selection strategy works as follows:

1. A raw biological score is calculated for each site being considered for inclusion in the WRP, based on the biological ranking criteria worksheet.
2. To account for differences in costs of the land, the raw biological score is adjusted by how much it differs from $3,000/acre. It is adjusted up if it is less expensive, and down if it is more expensive. This is then the biological score.
3. Contributions from partnerships are taken into account by adjusting the biological score by the percent decrease in total NRCS costs that would result from partnership contributions. This is then the adjusted score.
4. Sites are chosen in decreasing order of the adjusted score until the budget for the year is exhausted.

If the goal of the program is to maximize the biological benefits from wetlands restoration, and if the raw biological score is assumed to be the best available indicator of the expected biological benefits from restoring each parcel, then the site selection strategy

[52] http://www.wi.nrcs.usda.gov/programs/wrp.asp
[53] The biological ranking criteria worksheet reproduced here is not exactly the same as the one currently in use in California. It has recently been changed slightly, but it is still very similar to the one in Figure A.1.

should be designed to choose the set of sites that will yield the greatest possible total biological benefits, subject to the limited annual budget. The current strategy is apparently intended to do this, by choosing sites that have high biological benefits but low costs. However, it turns out that the strategy described below would perform better than the strategy currently in use. The alternative strategy would work as follows:

1. Calculate the raw biological score for each prospective site.
2. Calculate the expected costs of restoring each prospective site (accounting here for partnership contributions).
3. Divide the raw biological score by the expected cost to get a biological benefit-cost ratio.
4. Choose sites in decreasing order of the biological benefit-cost ratio until the budget is exhausted.

To estimate how much better the alternative strategy would perform compared to the current strategy, I applied it to data on all 87 parcels offered for inclusion in the WRP in California in the year 2000, which are listed in Table A.1 in increasing order of their final rank based on the current selection strategy.[54] The budget for the WRP in California in the year 2000 was approximately $10 million.[55]

In the year 2000 the first 15 parcels listed in Table A.1 were selected. The sum of the raw biological scores for the first 15 parcels, i.e. the "total biological benefit," was 2,799. If parcels were chosen in decreasing order of the adjusted score, then the total biological benefit would have been 1,677. If the parcels were chosen in decreasing order of the correct adjusted score (based on an adjustment to the biological score, as called for by the scoring

[54] To maintain confidentiality, all information associated with each offered parcel that could potentially be used to identify the landowners has been omitted from the table.

[55] There is some confusion regarding how sites were actually selected in 2000. According to data on which Table A1.1 was based, parcels were selected in increasing order of their final rank (column 1). However, there is a discrepancy between the rank and the adjusted score (column 11), which is the result of the currently used site selection strategy as described in the main text. This suggests that some other (unknown to me) criteria went into determining the final rank. Furthermore, the adjusted score was apparently calculated incorrectly. Contrary to what the scoring strategy calls for, it appears to have been based on an adjustment to the raw biological score, instead of the adjusted biological score. However, this confusion does not affect the conclusions I draw here regarding the effectiveness of the current site selection strategy.

strategy), then the total biological benefit would have been 2,072. If the parcels were chosen in decreasing order of the biological benefit-cost ratio, as suggested by the alternative site selection strategy outlined above, then the total biological benefit would have been 6,714. The alternative strategy substantially outperforms (any of the plausible versions of) the current strategy. In fact, a randomization experiment suggested that a purely random site selection strategy would be nearly as effective on average as the current strategy. I ranked the offered parcels randomly 100 times, and each time chose parcels in decreasing order of the random rank until the budget was exhausted. The average total biological benefit from these trials was 2,445, and the standard deviation was 524. In 91 of the trials the total biological benefit was greater than 2,072, which is what the total biological benefits in the year 2000 would have been if parcels were chosen based on the correct adjusted score.[56]

The WRP is perhaps a prototypical example of a government program designed to provide incentives to landowners to take marginal farmland out of production and restore it to natural habitat. If the goals of the WRP and other state and federal programs like it are to deliver the maximum possible level of environmental benefits, then thoughtful design of the strategy used to select sites for enrollment in the programs is crucial for cost-effectiveness.

[56] I brought these results to the attention of an acquaintance at NRCS who is involved in implementation of the WRP in California. His response was that the raw biological score is not intended as a measure of expected biological benefits alone. It also includes measures of expected restoration success, based on the current and past hydrologic conditions at the site, among other things. Nevertheless, the point I make here regarding the sub-optimal performance of site selection strategies that are not based on a rationally constructed benefit-cost ratio remains valid.

Figure A.1 – The California Wetlands Reserve Program biological ranking criteria worksheet.

RATING SCORE ____________

WETLAND RESERVE PROGRAM
CALIFORNIA RANKING CRITERIA WORKSHEET

Client Name:______________________________ Date:______________

Address:______________________________ Application No.__________

Evaluators:______________________________

Check WRP Option: _____ Perpetual _____ 30-Year _____ 10-Year Restoration Agreement

Estimated Percentage of Hydric Soils in Offered Acres ________

Acreage Offered for Restoration Total Acres ____________

	Present Land Use Type (ac)	Proposed Land Use Types (ac)
Ag Type	______	______
Wetlands	______	______
Woody Riparian	______	______
Upland	______	______

SUMMARY OF RATING FACTOR SCORES

	Maximum	Points
FACTOR 1. RESTORED HYDROLOGY (Present Condition)	25	______
FACTOR 2. RESTORED HYDROLOGY (Future Condition)	25	______
FACTOR 3. SPECIES UTILIZATION	30	______
FACTOR 4. HABITAT DIVERSITY	20	______
FACTOR 5. GEOGRAPHIC LOCATION	15	______
FACTOR 6. ADJACENT PROTECTED HABITAT	15	______
FACTOR 7. SURROUNDING HABITAT	15	______
FACTOR 8. RESTORED HYDROLOGY: Soil Factors	40	______
FACTOR 9. SIZE OF OFFERING	15	______
SUBTOTAL SCORE	200	______
COST RANKING SCORE	40	______
PARTNERSHIP POINTS	20	______
TOTAL	260	______

______________________________ ______________
Signature - FWS Representative Date

______________________________ ______________
Signature - Technical Team Leader Date

______________________________ ______________
Signature - District Conservationist **Date**

______________________________ ______________
Landowner Review **Date**

Figure A.1 (continued)

1. RESTORED NATURAL HYDROLOGY (Present Condition)

*Hydrology Functions Absent (Example: PC, does not Pond or Flood for 15 Consecutive Days	(25 Points)
*Hydrology Functions Degraded (Example: FW or Wetlands Farmed Under Natural Conditions, still Ponds or Floods for 15 days or more)	(20 Points)
Cropland that has been restored to wetland voluntarily, will be scored according to its prior condition.	
*Formerly PC	(25 Points)
*Formerly FW or Wetlands Farmed Under Natural Conditions	(20 Points)

NOTE: If more than one situation, making up 25% or more of the offered acreage, calculate the weighted average.

Total ________
(max - 25 pts)

Field Notes: (describe the natural and supplemental hydrology, type and source of water, and dependability of supplemental sources).

2. HYDROLOGY RESTORATION (Expected Future Condition)

*High Probability of Restoration on at least 50% or Supplemental Water readily available during the growing season. Vernal Pool Complex > 15%.	(25 Points)
*High Probability of Restoration on 25-50% or Supplemental Water available during growing season	(20 Points)
*Probability of Restoration on < 25%	(10 Points)

NOTE: If more than one situation, making up 25% or more of the offered acreage, calculate the weighted average.

Total __________
(max - 25 pts)

Field Notes: (describe restorable hydrology, ponding/flooding probability, source and reliability of supplemental water).

3. SPECIES UTILIZATION

-Listed Federal & State T&E Species	5 points each
-Species of Concern: Proposed/Candidate T&E Species	3 points each

Total __________
(max - 30 pts)

SCORE	SPECIES	HABITAT	SEASON

Figure A.1 (continued)

4. HABITAT DIVERSITY

Number of Habitat Elements Present After Restoration for each selected Wetland Type

All	(15 points)
All but One	(10 points)
All but Two	(5 points)
Three or More Absent	(1 point)

Wetland Types*/
Habitat Elements**

_____ Forested/Scrub Shrub
____Open Water (>3 ft deep)
____Submergents
____Shrubs/Trees
____Associated Uplands

____Vernal Pools
____Mud Flat
____Submergents
____Emergents
____Associated Uplands

_____ Seasonal Herbaceous
____Mud Flat
____Open Water
____Emergents
____Shrubs/Trees***
____Associated Uplands

____Semi-permanent Herbaceous
____Mud Flat
____Open Water
____Submergents
____Trees/Shrubs***
____Associated Uplands

_____Coastal/Tidal
____Mudflats
____Open Water
____Submergents***
____Emergents***
____Shrubs/Trees***
____Associated Uplands

FOOTNOTES:
* Check appropriate wetland type evaluated and the Habitat Elements present.

** The Habitat Elements listed above must occupy at least 10% of the specific associated Wetland Type to be considered present.

*** Considered in this Wetland Type only if present historically and will occupy at least 5% of the restored wetland.

NOTE: ADD 5 ADDITIONAL POINTS FOR > 1 WETLAND TYPE

Total __________
(maximum - 20 points)

Field Notes:

Figure A.1 (continued)

5. GEOGRAPHIC LOCATION BASED ON JOINT VENTURE GOALS

(Check appropriate basin)

____American Basin...	(15 points)
____Great Basin/Intermountain Valleys	
____Tulare Basin	
____Sutter Basin	
____Butte Basin...	(10 points)
____Yolo Basin	
____Delta Basin	
____North Coast...	(5 points)
____Central Coast	
____San Joaquin Basin	
____Colusa Basin	
____Southern California.....................................	(1 point)

Total ______
(max - 15 pts)

Field Notes:____

____

____

6. ADJACENT TO PROTECTED AREAS OR WITHIN CRITICALLY LIMITED HABITAT.

*Adjacent to Existing Easement, Refuge or other Protected Area...	(15 points)
*Protected Habitat within 1 mile...	(10 points)
*Protected Habitat greater than 1 mile...	(5 points)
*Critically Limited: Greater than 10 miles...	(15 points)

Total ______
(max - 15 pts)

Field Notes: Document type and size of protected area and distance from Proposed Esaement.

TYPE OF PROTECTED AREA	DISTANCE
______________________________	__________
______________________________	__________

Figure A.1 (continued)

7. SURROUNDING HABITAT DIRECTLY ADJACENT TO THE PROPOSED EASEMENT

Adjacent Habitats

* 3 Types, or Wetlands, making up greater than 75% of the adjacent land use types.	(15 points)
* 2 Types, or Wetlands making up greater than 50% of the adjacent land use types.	(10 points)
* 1 Type, or Wetlands, making up greater than 25% of the adjacent land use types.	(5 points)
* 1 Type, or Wetlands, making up less than 25% of the adjacent land use types.	(1 point)

_____ Grassland　　_____ Grain Crops

_____ Woodland　　_____ Irrigated Pasture

_____ Brush/Scrubland　　_____ Woody Riparian

_____ Wetlands

Total _______
(max - 15 pts)

Field Notes: ____

8. RESTORED NATURAL HYDROLOGY/SOIL CHARACTERISTICS
(Potential to restore water features based on soil characteristics)

***Flooding Potential: Temporary Inundation by Flowing Water**	
Frequent (>50 events in 100 years)	(15 points)
Occasional (5-50 events in 100 years)	(5 points)
Rare (1-5 events in 100 years)	(1 point)
***Ponding: (Determined by Permeability)**	
Very Slow	
Slow	(15 points)
Moderate	(10 points)
Moderately Rapid	(5 points)
Excessive	(3 points)
	(1 point)
***Saturation: (Depth to Water Table)**	
0 to 1 foot	
1 to 3 feet	(10 points)
Greater than 3 feet	(5 points)
	(1 point)

NOTE: Refer to local soil survey data for specific categories related to flooding, permeability and depth to water table. Points for a high water table may be considered even though the water table is seasonal.

Total _______
(max - 40 pts)

Field Notes: ____

Figure A.1 (continued)

9. SIZE OF EASEMENT	
100 to 500 acres	(15 points)
0 to 100 acres, or 500 to 1000 acres	(10 points)
Greater than 1000 acres	(5 points)
	Total ________
	(maximum - 15 points)

Table A.1 (continued)

Rank	Raw Bio Score	Biological Score	Acres	Easement Cost [$]	Restoration cost [$]	Other Costs [$]	Partner Contrib. [$]	Total NRCS Costs [$]	Cost Decrease	Adjusted Score	Correct Adjusted Score	Benefit-cost ratio
				(NRCS Estimate)	(NRCS Estimate)	(NRCS Estimate)	(proposed)		(percent)			
26	188	170.3	320	352,000	64,000	10,000	0	426,000		188	170	0.000441
27	188	170.3	320	352,000	64,000	10,000	0	426,000		188	170	0.000441
28	188	170.3	318	349,800	63,600	10,000	0	423,400		188	170	0.000444
29	188	170.1	237	260,700	47,400	10,000	0	318,100		188	170	0.000591
30	188	170.4	476	523,600	95,200	10,000	0	628,800		188	170	0.000299
31	180	158.8	425	595,000	72,475	10,000	46,238	631,237	7	187	166	0.000285
32	187	157.1	300	600,000	65,100	10,000	0	675,100		187	157	0.000277
33	185	167.3	294	323,400	58,800	10,000		392,200		185	167	0.000472
34	185	167.3	320	352,000	64,000	10,000	0	426,000		185	167	0.000434
35	185	154.8	525	1,050,000	133,875	10,000	0	1,193,875		185	155	0.000155
36	183	176.0	958	480,000	10,000	15,000	0	505,000		183	176	0.000362
37	175	144.8	476	952,000	119,000	10,000	95,200	985,800	9	181	154	0.000178
38	180	162.4	456	501,600	91,200	10,000	0	602,800		180	162	0.000299
39	180	162.4	463	509,300	92,600	10,000	0	611,900		180	162	0.000294
40	180	162.3	300	330,000	60,000	10,000	0	400,000		180	162	0.000450
41	179	160.9	101	116,200	11,615	10,000	0	137,815		179	161	0.001299
42	179	161.3	320	352,000	64,000	10,000	0	426,000		179	161	0.000420
43	179	161.3	318	349,800	63,600	10,000	0	423,400		179	161	0.000423
44	177	146.6	115	230,000	23,000	10,000	0	263,000		177	147	0.000673
45	177	147.3	300	600,000	60,000	10,000	0	670,000		177	147	0.000264
46	176	145.5	215	430,000	53,750	10,000	0	493,750		176	145	0.000356
47	176	143.4	50	100,000	12,500	10,000	0	122,500		176	143	0.001437
48	175	169.7	55	13,750	0	8,000	0	21,750		175	170	0.008046
49	175	168.7	270	108,000	10,800	10,000	0	128,800		175	169	0.001359
50	175	155.1	210	276,000	27,600	10,000	0	313,600		175	155	0.000558

Table A.1 (continued)

Rank	Raw Bio Score	Biological Score	Acres	Easement Cost [$] (NRCS Estimate)	Restoration cost [$] (NRCS Estimate)	Other Costs [$] (NRCS Estimate)	Partner Contrib. [$] (proposed)	Total NRCS Costs [$]	Cost Decrease (percent)	Adjusted Score	Correct Adjusted Score	Benefit-cost ratio
51	175	157.3	347	381,700	69,400	10,000	0	461,100		175	157	0.000380
52	173	147.1	270	468,000	48,600	10,000	0	526,600		173	147	0.000329
53	161	133.8	229	458,000	0	10,000	57,250	410,750	12	173	146	0.000392
54	173	155.3	320	352,000	64,000	10,000	0	426,000		173	155	0.000406
55	151	121.6	800	1,600,000	160,000	10,000	400,000	1,370,000	23	171	144	0.000110
56	171	160.2	160	101,280	19,200	10,000	0	130,480		171	160	0.001311
57	170	139.6	300	600,000	75,000	10,000	0	685,000		170	140	0.000248
58	169	138.9	500	1,000,000	120,000	10,000	0	1,130,000		169	139	0.000150
59	169	156.1	320	256,000	45,120	10,000	0	311,120		169	156	0.000543
60	148	128.0	2560	3,328,000	512,000	15,000		3,855,000		168	128	0.000038
61	167	137.6	154	308,000	22,330	10,000	0	340,330		167	138	0.000491
62	166	146.3	60	66,000	15,000	8,000	0	89,000		166	146	0.001865
63	165	134.4	350	700,000	91,000	15,000		806,000		165	134	0.000205
64	165	160.4	760	228,000	26,660	8,000	0	262,660		165	160	0.000628
65	165	146.1	75	78,000	18,750	10,000	0	106,750		165	146	0.001546
66	145	115.2	238	476,000	47,600	10,000	0	533,600	0	165	115	0.000272
67	162	130.1	100	200,000	30,000	10,000	0	240,000		162	130	0.000675
68	162	129.4	55	110,000	15,000	10,000	0	135,000		162	129	0.001200
69	159	129.6	1182	2,364,000	236,400	15,000	0	2,615,400		159	130	0.000061
70	159	140.3	1600	1,920,000	320,000	15,000	0	2,255,000		159	140	0.000071
71	159	146.1	320	256,000	45,120	10,000	0	311,120		159	146	0.000511
72	153	120.4	20	40,000	4,000	5,000	1,250	47,750	3	156	123	0.003204
73	155	135.9	450	540,000	90,000	15,000	0	645,000		155	136	0.000240
74	154	147.2	700	350,000	0	10,000	0	360,000		154	147	0.000428
75	153	145.0	2000	1,200,000	0	10,000	0	1,210,000		153	145	0.000126

Table A.1 (continued)

Rank	Raw Bio Score	Biological Score	Acres	Easement Cost [$]	Restoration cost [$]	Other Costs [$]	Partner Contrib. [$]	Total NRCS Costs [$]	Cost Decrease	Adjusted Score	Correct Adjusted Score	Benefit-cost ratio
				(NRCS Estimate)	(NRCS Estimate)	(NRCS Estimate)	(proposed)		(percent)			
76	153	122.9	640	1,280,000	160,000	10,000	0	1,450,000		153	123	0.000106
77	153	136.9	50	37,500	15,000	8,000	0	60,500		153	137	0.002529
78	149	118.1	80	160,000	16,000	10,000	0	186,000		149	118	0.000801
79	148	117.1	705	1,410,000	211,500	15,000	0	1,636,500		148	117	0.000090
80	147	118.6	160	272,000	54,400	15,000	0	341,400		147	119	0.000431
81	147	117.7	192	384,000	28,800	10,000	0	422,800		147	118	0.000348
82	147	128.6	164	164,000	52,808	10,000	0	226,808		147	129	0.000648
83	143	113.1	100	200,000	15,000	10,000	0	225,000		143	113	0.000636
84	140	132.4	320	108,000	64,000	10,000	0	182,000		140	132	0.000769
85	136	105.6	122	244,000	24,400	10,000	10,000	268,400	4	140	109	0.000507
86	138	119.3	120	144,000	14,400	10,000	0	168,400		138	119	0.000819
87	130	115.2	1120	1,008,000	224,000	15,000	0	1,247,000		130	115	0.000104
Totals:			36319	49,078,755	6,787,409	894,000	2,905,211	53,854,953				