Environmental Management (2008) 42:789–802
DOI 10.1007/s00267-008-9186-3

Landscape Planning for Agricultural Nonpoint Source Pollution Reduction I: A Geographical Allocation Framework

Matthew W. Diebel · Jeffrey T. Maxted · Peter J. Nowak · M. Jake Vander Zanden

Received: 24 July 2007 / Accepted: 4 July 2008 / Published online: 14 August 2008

Abstract Agricultural nonpoint source pollution remains a persistent environmental problem, despite the large amount of money that has been spent on its abatement. At local scales, agricultural best management practices (BMPs) have been shown to be effective at reducing nutrient and sediment inputs to surface waters. However, these effects have rarely been found to act in concert to produce measurable, broad-scale improvements in water quality. We investigated potential causes for this failure through an effort to develop recommendations for the use of riparian buffers in addressing nonpoint source pollution in Wisconsin. We used frequency distributions of phosphorus pollution at two spatial scales (watershed and field), along with typical stream phosphorus (P) concentration variability, to simulate benefit/cost curves for four approaches to geographically allocating conservation effort. The approaches differ in two ways: (1) whether effort is aggregated within certain watersheds or distributed without regard to watershed boundaries (dispersed), and (2) whether effort is targeted toward the most highly P-polluting fields or is distributed randomly with regard to field-scale P pollution levels. In realistic implementation scenarios, the aggregated and targeted approach most efficiently improves water quality. For example, with effort on only 10% of a model landscape, 26% of the total P load is retained and 25% of watersheds significantly improve. Our results indicate that agricultural conservation can be more efficient if it accounts for the uneven spatial distribution of potential pollution sources and the cumulative aspects of environmental benefits.

Keywords Nonpoint source pollution · Watersheds · Riparian buffers · Prioritization · Threshold · Statistical simulation

Introduction

Agricultural runoff is the dominant source of sediment, phosphorus, and nitrogen pollution to aquatic ecosystems (USEPA 2000). Excess loading of nutrients in the form of nonpoint source (NPS) pollution has led to widespread eutrophication of lakes, rivers, and estuaries, and resultant ecological degradation and loss of ecosystem services (USEPA 1990; NRC 1992; USEPA 1996; Turner and Rabalais 1994; Burkart and James 1999; Carpenter and others 1998). Inputs of fine sediments degrade aquatic habitat (Waters 1995) and are among the most widespread pollutants to rivers and streams (USEPA 2000). Because the causes and extent of these effects are well understood, reducing agricultural NPS pollution and improving stream water quality has become an important priority of many state and federal resource management agencies (USEPA 2002).

Application of agricultural best management practices (BMPs), such as riparian buffers, can reduce inputs of nutrients and sediments to streams (e.g., Mendez and others 1999; Nerbonne and Vondracek 2001; Bishop and others 2005). To encourage their use, state resource management agencies and the U.S. Department of Agriculture (USDA) have developed programs that offer incentives to farmers to implement BMPs. For example, since 1987, the USDA Conservation Reserve Program has paid $29.7 billion to

M. W. Diebel (✉) · J. T. Maxted · M. J. Vander Zanden
Center for Limnology, University of Wisconsin - Madison, 680 North Park Street, Madison, WI 53706, USA
e-mail: mwdiebel@wisc.edu

P. J. Nowak
Gaylord Nelson Institute for Environmental Studies, University of Wisconsin - Madison, 550 North Park Street, Madison, WI 53706, USA

farmers to implement conservation practices (USDA 2006). Despite the large amount of money distributed through these voluntary programs, they have not generally produced measurable improvements in stream water quality (Wolf 1995; Meals 1996; Boesch and others 2001), prompting the search for ways to improve the effectiveness of pollution reduction programs.

The USDA has made it a priority to target agricultural conservation practices where they will "generate the most profound or widespread environmental benefits for a given [cost]" (Hansen and Hellerstein 2006). Theoretically, targeting should be particularly important, since available funds can only pay for conservation practices on a small fraction of agricultural lands. However, there are many ways to define environmental benefits, and many scales at which these benefits may be realized, some of which are more easily measured than others. These variables lead to program design questions, such as: Given limited financial resources, should a program aim to greatly improve water quality in a few streams or strive for modest improvements in water quality across a large watershed? Furthermore, given the political difficulties with unequal application of regulations or payments, how much additional benefit is likely to be realized by a targeted program compared to a voluntary program?

A recent policy-oriented research effort in Wisconsin provided an opportunity to explore these and other questions. In 2002, legislative efforts to revise NPS pollution regulations (Wisconsin Administrative Statute NR151) produced controversy among stakeholders surrounding the issue of riparian buffer implementation, leading to the establishment of a multi-stakeholder group called the Wisconsin Buffer Initiative (WBI) (UW-CALS 2005). The advisory committee of the WBI was composed of academic researchers, state and federal regulatory agencies, as well as agricultural and conservation groups, and was asked to design a set of recommendations based on the best available science concerning the use of riparian buffers to manage NPS pollution in Wisconsin. The WBI participants seized upon this as an opportunity to develop an innovative program that builds on recent conceptual and empirical advances in the understanding of NPS pollution and that considers buffers as only one option in a suite of available practices (UW-CALS 2005).

Here we present the first of a series of three articles deriving from our experience working with the Wisconsin Buffer Initiative. The goal of this article is to develop a geographical implementation framework for NPS pollution control. This framework is meant to guide the use of multiple, complementary BMPs, rather than solely riparian buffers. We begin by summarizing evidence that existing NPS pollution control programs have not produced significant, broad-scale water quality improvements, despite numerous studies showing that the management practices supported by these programs are effective at local scales. We suggest that this apparent contradiction can be explained by two primary factors. First, data from Wisconsin farms and streams indicate that sources and transport of NPS pollutants are distributed log-normally at multiple spatial scales across agricultural landscapes. Therefore, targeting pollution control toward the high ends of these distributions would likely result in greater environmental gains per unit of management effort. However, most government-sponsored NPS pollution control programs have additional objectives (e.g., reducing excess crop production) and politically-driven design elements (e.g., equal-opportunity, voluntary participation) that tend to limit targeting of BMPs according to their capacity for reducing NPS pollution. Second, seasonal and weather-driven variability in the concentrations of pollutants in receiving waters may make the detection of statistically-significant water quality changes difficult under common BMP application levels and monitoring approaches. In other words, pollution control effort is often too sparsely distributed across the landscape to make an appreciable difference in any one place.

We use statistical simulations to evaluate program efficiency gains that could be realized by geographically targeting and aggregating pollution control effort (i.e., extent of BMP implementation or monetary cost). Specifically, we estimate two types of cost/benefit curves (total pollution reduction and proportion of watersheds improved) for four geographical allocation approaches. Selection of a "best approach" depends on the relative importance of the two types of benefit, which involves a value judgment. However, in realistic implementation scenarios, one approach (aggregated/targeted) not only has the highest average benefit, but is also most suited to implementation in an adaptive management framework. To conduct these analyses, we had to assemble empirical data from multiple sources because no suitable unified dataset exists. Admittedly, this approach may make our results less generalizable than if the study was based on a single, multi-scale dataset collected with uniform methods across a variety of agricultural landscapes. To compensate, we use a sensitivity analysis to identify which model input uncertainties most strongly influence the results.

In the two following articles, we further develop the aggregated/targeted approach to NPS pollution control. First, Maxted and others (this issue) consider issues of watershed scale and delineation, and present a set of watersheds to serve as implementation and management units for the program. Second, Diebel and others (this issue) model broad-scale patterns of NPS phosphorus and sediment loading to streams, and provide a ranking of watersheds based on water quality restoration potential. Together, these articles provide a conceptual advance in the landscape-scale management of agricultural nonpoint source pollution.

Can Agricultural Conservation Practices Improve Stream Water Quality?

The Wisconsin Buffer Initiative initially sought to quantify the pollution reduction potential of riparian buffer strips and how this potential might vary with width and landscape setting. Literature reviews on the functions and performance of riparian buffers (Desbonnet and others 1994; Correll 1999; Wenger 1999) indicate that buffers can retain large proportions of the phosphorus (56%) and sediment (78%) that enters them in surface runoff, and that they can also be effective at removing nitrate (51%) dissolved in shallow groundwater (percentages are averages from 52 studies cited in Desbonnet and others 1994). Buffers are less effective at retaining: sediment, when channelized flow occurs (Dillaha and others 1989; Daniels and Gilliam 1996); phosphorus, when buffer soils become saturated with phosphorus (Osborne and Kovacic 1993); and nitrate, when subsurface flow paths do not intersect the root zone of riparian vegetation (Lowrance and others 1997). Thus, augmenting (and in some cases, replacing) riparian buffers with upland management practices such as reduced tillage and nutrient application or crop changes may better improve water quality in some situations.

Natural resource management agencies have long recognized the complimentary abilities of a wide array of BMPs and have therefore designed programs to fund the application of each practice where it is likely to be most effective (e.g., Hansen and Hellerstein 2006). Given the large amount of money that has been distributed to farmers through such programs, one might expect that water quality monitoring would detect significant improvements where BMPs have been implemented. However, this has often not been the case. We use the Wisconsin Priority Watershed Program (PWP) as a case study.

Between 1979 and 2006, the Wisconsin PWP provided approximately $201 million to counties and landowners to address land management activities that contribute to urban and rural runoff (Wisconsin Legislative Fiscal Bureau 2007). Funds were used to cost-share the implementation of BMPs by landowners in 87 "priority watersheds" (mean area = 369 km^2) where the need for NPS water pollution abatement was deemed most critical (Wisconsin Administrative Statute NR120). Funded practices included tillage and nutrient management changes, fencing to restrict animal access to streams, streambank shaping and reseeding, and structural barnyard improvements to reduce manure runoff. Participation in the program was voluntary, and for most of the program lifespan, there was no way to specifically induce owners of highly polluting lands to participate. In individual priority watersheds, an average of 38% of eligible landowners implemented one or more BMPs, although this figure ranged widely—from 4% to 82% (Wisconsin Department of Natural Resources, unpublished data). During an interim program evaluation in 1995, sampling-based water quality assessments for 20 of the priority watersheds found no statistically significant improvements, and it appeared that insufficient landowner participation was the reason (Wolf 1995). This finding is consistent with other studies of water quality (Davie and Lant 1994; Meals 1996; Boesch and others 2001) and biological (Wang and others 2002, 2006) response to agricultural BMPs. Subsequently, more intensive evaluations in a few of the priority watersheds found some improvements (Graczyk and others 2003; Corsi and others 2005), but outcomes in most watersheds have not been thoroughly evaluated.

In many cases, the effects of agricultural BMP implementation on stream water quality are never measured. For example, in a national database of river restoration efforts (Bernhardt and others 2005), only 13% of the 5141 "riparian management" project records included a monitoring component. However, even if all projects were monitored, there are several reasons why improvements would not be detected in many cases. First, agricultural BMPs are often implemented for reasons other than the improvement of water quality, including the creation of wildlife habitat and the reduction of excess crop production (USDA 2003). The locations that best serve these goals may not be the same ones that would best contribute to water quality improvement. Second, most programs, including the Wisconsin PWP and the largest federal programs (Conservation Reserve Program and Conservation Reserve Enhancement Program) are voluntary. Though posited on principles of fairness and accessibility, we contend that these arguably positive aspects of the program are also weaknesses, in that the program fails to ensure that the largest polluters participate. And third, partially as a result of the previous two factors, BMP implementation is typically dispersed geographically. While dispersed implementation may lead to small pollutant loading reductions to many water bodies, these reductions may not translate into statistically detectable water quality improvements if background variability in stream water pollutant concentrations (e.g., weather-related) obscures those changes. In the following sections, we examine how adjusting these program design elements can more efficiently cause water quality improvement.

Distribution of Nonpoint Source Pollution on Agricultural Landscapes

Natural and social scientists have traditionally emphasized the importance of characterizing the average behavior of a population, with the implicit assumption that aspects of the natural world that scientists are interested in can be approximated by a normal frequency distribution. In contrast, Limpert and others (2001) provided diverse examples of the natural world being characterized by log-normal probability distributions, and highlighted the important implications of this pattern. This idea of a "log-normal world" may be applicable to the problem of nonpoint source pollution on agricultural landscapes. Nowak and others (2006) proposed that among individual farm fields, both impacts of management practices and biophysical vulnerability to pollution follow log-normal frequency distributions. Because these factors interact to determine the pollution contribution of any one field (Wischmeier and Smith 1978), the resulting distribution of pollution will be even more strongly log-normal than either distribution on its own. Thus, the extreme situations at the tail of these frequency distributions have disproportionately large effects, and have the potential to drive pollution patterns at broad spatial scales.

We examined empirical phosphorus (P) pollution data from Wisconsin for evidence of this phenomenon at two spatial scales—within and among small watersheds. At the within-watershed scale, we used P index values (http://wpindex.soils.wisc.edu/) for 62 randomly-selected cropped fields (mean area = 5 ha) in a 20 km^2 watershed in southwestern Wisconsin (Hefty Creek, Green County) (L. Good, unpublished data). The Wisconsin P index uses soil P content, edaphic characteristics, topography, crop rotation, and management practices to estimate average annual per-area P loss across the rotation. Unlike most P indexes, which are uncalibrated P loss risk assessments (Sharpley and others 2003), the Wisconsin P index estimates actual losses, and has been validated with year-round runoff monitoring on 18 Wisconsin farm fields (Bundy and others 2008). The frequency distribution of the Hefty Creek P loss estimates is best fit by a log-normal probability distribution (Kolmogorov-Smirnov test, $p_{normal} = 0.03$, $p_{log\text{-}normal} = 0.74$), with a multiplicative standard deviation (s*) of 2.43 (Fig. 1a).

At the among-watershed scale, we used estimates of annual unit-area total P loads (kg/km^2/year) for 1005 small (20–100 km^2) watersheds in Wisconsin with at least a moderate amount of agricultural land (>30%) (Diebel and others, this issue). The frequency distribution of these P load estimates is also best fit by a log-normal probability distribution (Kolmogorov-Smirnov test, $p_{normal} = <0.001$, $p_{log\text{-}normal} = 0.02$), with s* = 2.32 (Fig. 1b). Assuming that within-watershed proportional P loss distributions in other Wisconsin watersheds are similar in shape to the distribution observed in the Hefty Creek watershed, the distribution of individual field P losses across Wisconsin is also log-normal, but with a higher s* (2.83) than either of the two distributions that form it (Fig. 1c is the joint probability of the distributions in Fig. 1a and b).

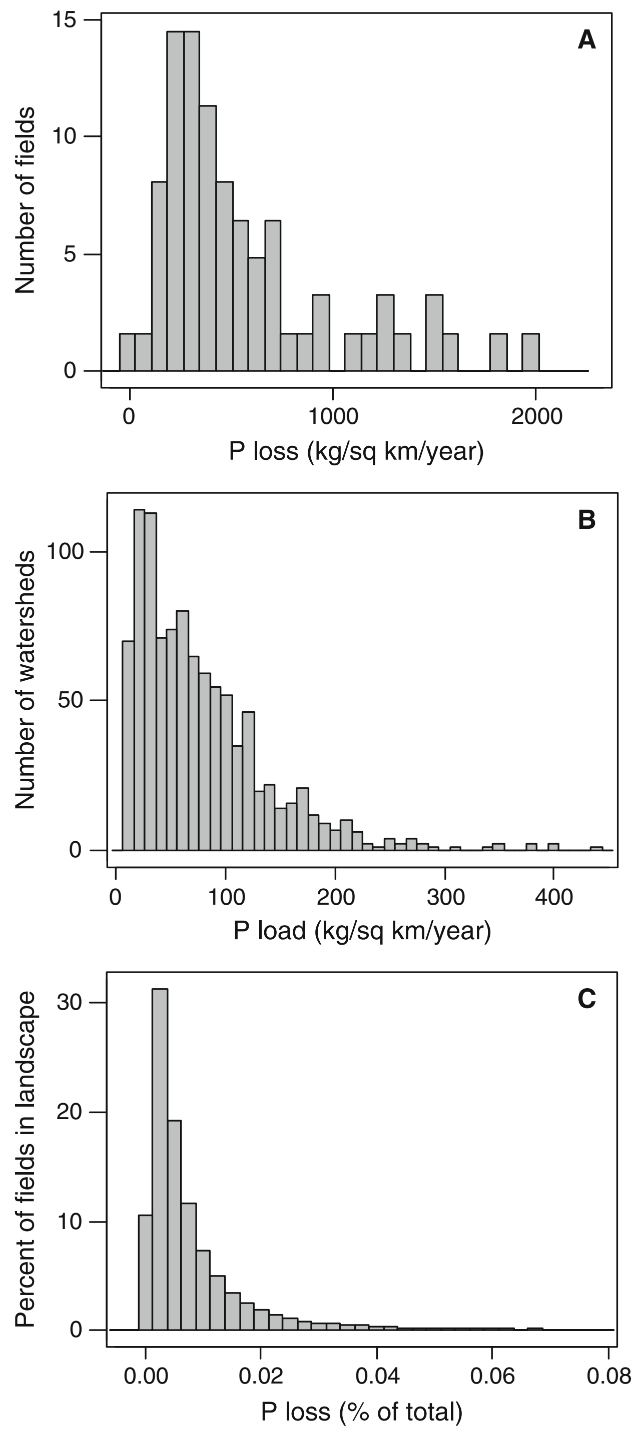

Fig. 1 Log-normal distribution of phosphorus pollution at multiple spatial scales: (**a**) P Index values from 62 fields in the Hefty Creek watershed, Green County, Wisconsin; (**b**) Modeled P loads from 1005 small agricultural watersheds in Wisconsin; (**c**) Modeled percent contribution of 10,000 fields to landscape P load (joint probability of distributions in a and b)

These distributions conflict with the predominant perception of NPS pollution as being diffusely distributed across the landscape. This new perspective shares some surprising commonalities with point source pollution: the problem (and the solution) originates from a limited number of locations on the landscape. An important difference is that it is easy to locate point sources, while identifying NPS "hotspots" can be very difficult. Nevertheless, if hotspots can be identified and managed appropriately, significant pollution reduction could be efficiently achieved. Later in this article, we use the above frequency distributions, which are the best available for Wisconsin, to parameterize a model that quantifies this increased efficiency.

Measuring the Effects of Best Management Practices

There are two primary approaches for measuring the benefits of NPS pollution reduction. Measured water quality change approaches use before-and-after water quality measurements to estimate the extent and statistical significance of water quality improvement in a water body whose watershed has been treated with BMPs. This approach sometimes also uses measurements on an unmanipulated watershed to control for extrinsic water quality influences such as weather. Modeled pollutant reduction approaches use empirical or mechanistic relationships between biophysical and management characteristics to estimate pollution loading from individual fields, farms, or subwatersheds before and after the implementation of BMPs. Model estimates for individual land units can then be aggregated up to estimate total pollutant load reduction in watersheds or political land units.

Government conservation programs often evaluate their programs with the modeled pollutant reduction approach. For example, the Wisconsin Priority Watershed Program (described above) used two models—BARNY (WDNR 1994a) and WINHUSLE (WDNR 1994b)—to estimate pollutant reductions resulting from BMP implementation. These models estimated that, relative to pre-implementation loadings, an average of 37% of phosphorus and 24% of sediment was prevented from entering surface waters in these watersheds (Holden and others 2006). As mentioned above, however, these estimated load reductions generally did not result in statistically significant water quality changes in receiving waters. One possible explanation for the incongruence of these two monitoring approaches is that the models overestimated the effectiveness of BMP implementation (Stuntebeck and Bannerman 1998). Water quality response to changes in land management may have also been delayed if nutrients or sediments accumulated as a result of past practices (detained) remain susceptible to transport by water (Carpenter 2005; Knox 2006). Another explanation is that the modeled reductions did occur, but that they were simply not detectable due to background variability in pollutant loads. This background variability is typically weather-driven, and in most water quality studies, it is seen as bothersome statistical noise, to be controlled for, if possible (Spooner and others 1987). Alternatively, it is possible to use the amount of background variability as a guide for setting target pollutant reductions. In other words, one could ask, "Given a certain variance in a water quality parameter, and given a realistic sampling regime for monitoring that parameter, how much would that parameter have to change before the intervention effect is likely to be statistically detected?" This amount of change could be translated into a minimum amount of program implementation effort that should be expended on any one watershed. In the following section, we use statistical simulations to address this question, coupled with a comparison of the relative efficiency of voluntary and targeted programs.

Methods

The Model Landscape

We compared four approaches for allocating conservation effort across a model agricultural landscape. For simplicity, we limited our evaluation of these approaches to their ability to reduce phosphorus (P) pollution at two scales: landscape and watershed. The P load from the landscape (L_l) is the sum of P loads (L_w) from 100 equally-sized, spatially-independent watersheds. Values of L_w were drawn randomly from the empirical distribution of the WBI watershed P loads (Fig. 1b). To simplify calculations involving percent reductions, watershed loads were converted to proportions of the landscape load. Thus, the standardized landscape load (λ_l) was set equal to 1, and the proportional watershed loads (λ_w) were calculated as:

$$\lambda_w = L_w / \sum_{w=1}^{100} L_w \qquad (1)$$

Each watershed is composed of 100 equally-sized fields. P loading from these fields is the only source of P to streams in the model landscape. In each watershed, P loads from fields (L_f) were assigned randomly selected values from the empirical distribution of the Hefty Creek fields (Fig. 1a). The proportional contribution of field f to the load from its watershed w (λ_f) is:

$$\lambda_f = L_f / \sum_{f=1}^{100} L_f \qquad (2)$$

and the proportional contribution of field f to the total landscape P load ($\lambda_{f,w}$) is:

$$\lambda_{f,w} = \lambda_f \cdot \lambda_w \qquad (3)$$

Thus,

$$\lambda_l = \sum_{w=1}^{100} \lambda_w = \sum_{f=1}^{10,000} \lambda_{f,w} = 1 \quad (4)$$

Each spatial allocation approach was then defined by the order in which conservation practices are applied to these 10,000 fields. The approaches differ in two ways: (1) whether effort is aggregated within certain watersheds or distributed without regard to watershed boundaries (dispersed), and (2) whether effort is targeted toward the most highly P-polluting fields or is distributed randomly with regard to field-scale P pollution levels. Simplified illustrations of these approaches are shown in Fig. 2.

Description of the Allocation Approaches

In the dispersed/random approach, fields are selected for BMP implementation in random order from among the 10,000 model fields. In the aggregated/random approach, fields are randomly selected from the watershed with the highest P load until the stop point (defined below) is reached. The same procedure is then applied to the watershed with the second highest P load, and so on. When BMP implementation has reached the stop point in all watersheds, the remaining fields are selected in random order. In the dispersed/targeted approach, fields are selected in descending order of percent contribution to the total landscape P load, regardless of watershed membership. And in the aggregated/targeted approach, fields are selected in descending order of percent P load contribution from the watershed with the highest P load until the stop point is reached. The same procedure is then applied to the watershed with the second highest P load, and so on. When BMP implementation has reached the stop point in all watersheds, the remaining fields are selected in descending order of percent contribution to the total landscape P load, regardless of watershed membership.

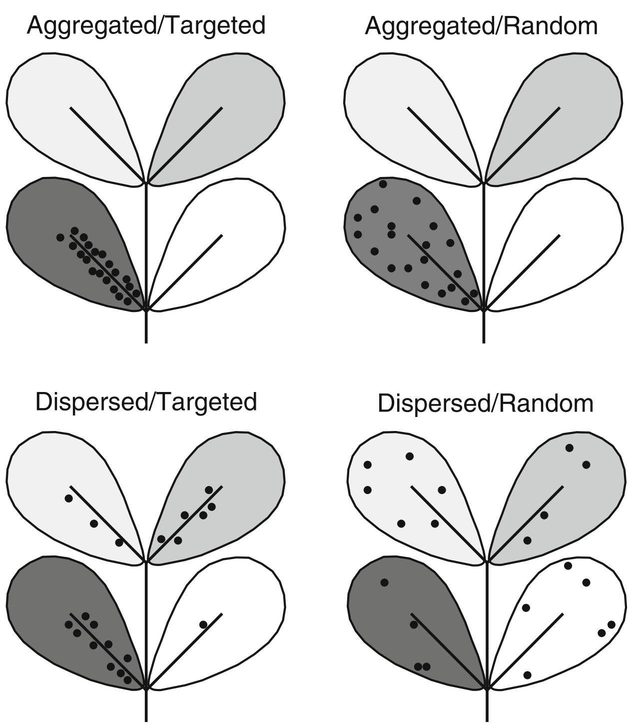

Fig. 2 Simplified illustrations of geographical allocation approaches. Black lines form a stream network; spatially-independent watersheds are shaded according to their contribution to NPS pollution (darker = higher); Black dots (20 in each approach) are locations of BMPs (for illustrative purposes, locations near streams reduce more pollution)

Description of the Benefit Indices

We compared these four approaches by implementing them on the model landscape and tabulating two benefit indices across the range of landscape BMP implementation levels (I_l, defined as the proportion of all 10,000 fields where BMPs are implemented). The first benefit index is the proportion of the total landscape P load that is eliminated (modeled pollutant reduction, $R_l = 1 - \lambda_l'$, where λ_l' is the landscape P load following BMP implementation). The second benefit index is the proportion of watersheds where a statistically significant reduction in the median stream water P concentration is observed (measured water quality change, $\bar{p}_w(R_f, I_w)$). This second index is an example of the type of nonlinear, threshold relationship between effort and benefit that has often been posited in the ecological and restoration science literature (e.g., Wang and others 1997; Carpenter and others 1999; Brazner and others 2004). When such a relationship exists, there is an optimal level of effort that maximizes efficiency (benefit/effort) (Statzner and others 1997), herein called the stop point. For both indices, we assumed that BMP implementation would reduce the P load from an individual field (R_f) by 75% (average P reduction potential estimated in Diebel and others [this issue]), and that reductions in P loading from fields would directly translate into the same percent reductions in stream P loads. Because the model calculates percent reductions, this assumption is reasonable even when in-stream processes remove P, as long as the rate of removal is proportional to the incoming load. However, the model will overestimate stream P reductions where detained P contributes appreciably to stream loads. We also assumed that the cost of BMP implementation is constant among fields. This is a reasonable assumption even when costs vary, as long as benefits are more variable than and uncorrelated with costs (Babcock and others 1997). Thus, BMP implementation level, effort, and cost are considered synonymous.

Statistical Simulations of Water Quality Change

We used statistical simulations to examine how the probability of detecting a statistically significant water quality improvement at the outlet of a watershed (p_w) depends on the BMP implementation level in that watershed (I_w, defined as the proportion of the 100 fields in each watershed where BMPs are implemented), and subsequently to estimate the stop points for the aggregated approaches. This process consisted of three steps: (1) Estimating $p_w(R_w)$, the relationship between p_w and the intervention level R_w (BMP-induced P reduction in watershed, as a proportion of the pre-BMP mean); (2) Using $p_w(R_w)$ and R_f (BMP-induced P reduction from individual fields; 75% in baseline scenario, but varied in sensitivity analysis) to calculate $p_w(R_f, I_w)$ for random and targeted approaches; (3) Calculating the stop points (I^*_{wr} and I^*_{wt}) as the values of I_w that maximize the benefit/effort ratio ($p_w[R_f, I_w]/I_w$) for the random and targeted approaches, respectively.

Step One

To detect the effect of an intervention on a variable parameter, such as stream P concentrations, the magnitude of the intervention must outweigh the variability of the parameter. We analyzed a separate empirical dataset to characterize the distribution of stream P variability in Wisconsin and subsequently to parameterize the model. This dataset consists of time series of P concentrations from 153 Wisconsin streams with more than 30% watershed agriculture (Robertson and others 2006) and provides the best available estimate of P variability in Wisconsin streams. Each time series consisted of six mid-monthly P concentration samples collected during the growing season (May–October) of 2001 or 2002. To simulate sample sets from a hypothetical pre-intervention monitoring program (6 monthly samples over 4 growing seasons), we fit log-normal frequency distributions to the sample values for each stream and then generated 24 random numbers from each distribution. For each stream, we then created post-intervention sample sets by multiplying each value in a second pre-intervention sample set by a range of R_w values (0–1, increments of 0.05). This procedure scaled the sample means and standard deviations proportionally, in accordance with the trend seen across the range of measured values ($\sigma = 0.58\mu - 0.004$, intercept not different from zero, $r^2 = 0.57$). We then conducted Wilcoxon rank-sum nonparametric tests to compare the median of the pre-intervention samples to the median of the post-intervention samples for each stream at each value of R_w (Graczyk and others 2003; Corsi and others 2005). We repeated these simulated experiments 20 times and then calculated the mean percentage of streams (equal to $p_w[R_w]$) where a significant ($p < 0.01$) difference was detected at each value of R_w. To calculate $p_w(R_w)$ for watersheds in the model landscape, a modified logistic function was fit to this relationship (line labeled "$n = 24$" in Fig. 3):

$$p_w = \frac{1.017}{1 + 59(e^{-12.8 \cdot R_w})} - 0.017 \tag{5}$$

Step Two

$R_w(I_w)$ was calculated as:

$$R_w = R_f \left(\sum_{f=1}^{100 \cdot I_w} \lambda_{f,w} \right) \tag{6}$$

where fields are selected in rank order (highest to lowest) of $\lambda_{f,w}$ in targeted approaches and in random order for random approaches. Substituting this function for R_w in equation 5 allows calculation of $p_w(R_f, I_w)$.

Step Three

The stop points (I^*_{wt} and I^*_{wr}) were identified as the maxima of $p_w(R_f, I_w)/I_w$ in the targeted and random approaches, respectively (Fig. 4).

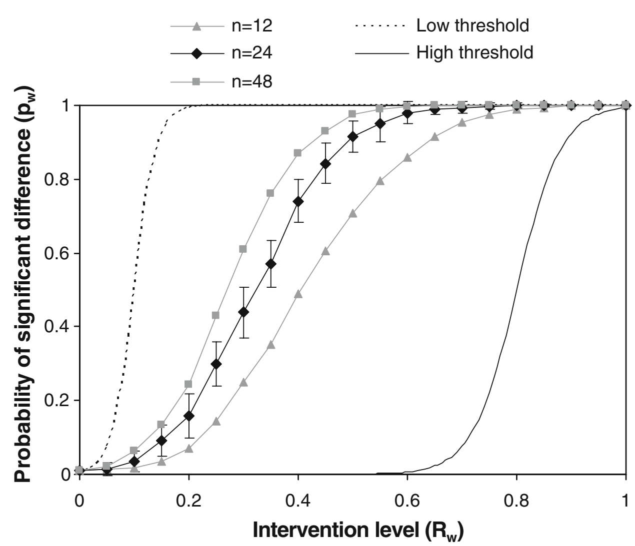

Fig. 3 Results of simulated Wilcoxon sign-rank tests. The probability (p_w) of detecting a significant difference ($p < 0.01$) in median phosphorus concentration is plotted against the proportional reduction in mean phosphorus concentration (intervention level, R_w) for three sampling scenarios ($n = 12$, 24, and 48 samples before and after intervention). On the $n = 24$ points, error bars are ± 1 standard deviation based on 20 simulated experiments (variability was similar for the other two sample sizes). Low and high threshold curves are hypothetical (see discussion)

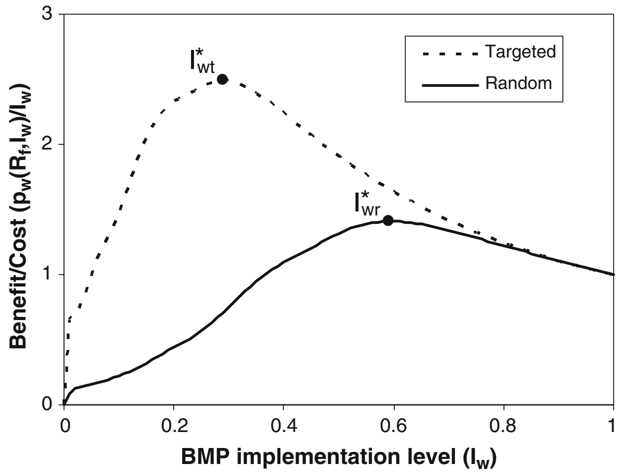

Fig. 4 Estimation of the stop points (I^*_{wr} and I^*_{wt}) for aggregated approaches in the baseline scenario. Benefit is equal to the probability of detecting a significant difference in stream water phosphorus at the outlet of a watershed (p_w). Cost is the proportion of fields with BMPs (I_w). The stop point is the maximum of each curve, where net benefit is greatest (targeted = 29%, random = 59%)

Calculation of the Benefit Indices

Using the field selection algorithms defined above, the modeled pollutant reduction benefit index R_l was calculated as:

$$R_l = R_f\left(\sum_{f=1}^{10{,}000\cdot I_l} \lambda_{f,w}\right) \tag{7}$$

and the measured water quality change benefit index $\bar{p}_w(R_f,I_w)$ was calculated as:

$$\bar{p}_w(R_f,I_w) = \frac{\sum_{w=1}^{100} p_w(R_f,I_w)}{100} \tag{8}$$

where the $p_w(R_f,I_w)$ are calculated separately for each watershed using equations 5 and 6. $\bar{p}_w(R_f,I_w)$is equivalent to the proportion of watersheds where a significant P reduction is detected.

Sensitivity Analysis

We tested the sensitivity of the relative performance of the allocation approaches to parameter values and assumptions that differ from the baseline scenario described above (Table 1). First, BMP-induced P reduction (R_f) was varied from 40% to 90% to reflect uncertainty in both the theoretical and operational performance of these practices (Gitau and others 2005). Second, the frequency distributions of P losses at both watershed and field scales was changed from log-normal to normal (with the same means as the empirical log-normal distributions, and standard

Table 1 Performance of four allocation approaches according to two benefit indices under nine scenarios (rows)

Scenario parameters				Stop points		Measured water quality change index				Modeled pollutant reduction index			
R_f	# of samples	Watershed P distribution	Field P distribution	Targeted approaches	Random approaches	Dispersed random	Aggregated random	Dispersed targeted	Aggregated targeted	Dispersed random	Aggregated random	Dispersed targeted	Aggregated targeted
0.90	24	log-normal	log-normal	0.22	0.49	0.03	0.17	0.21	0.33	0.09	0.22	0.36	0.30
0.75	48	log-normal	log-normal	0.21	0.50	0.04	0.16	0.21	0.32	0.08	0.19	0.30	0.25
0.75	*24*	*log-normal*	*log-normal*	*0.29*	*0.59*	*0.02*	*0.14*	*0.17*	*0.25*	*0.08*	*0.20*	*0.30*	*0.26*
0.75	12	log-normal	log-normal	0.35	0.74	0.01	0.11	0.13	0.17	0.08	0.22	0.30	0.27
0.60	24	log-normal	log-normal	0.40	0.74	0.02	0.11	0.13	0.17	0.06	0.17	0.24	0.22
0.40	24	log-normal	log-normal	0.55	1.00	0.01	0.07	0.07	0.08	0.04	0.13	0.16	0.14
0.75	24	log-normal	normal	0.46	0.59	0.02	0.14	0.12	0.17	0.08	0.20	0.26	0.22
0.75	24	normal	log-normal	0.29	0.59	0.02	0.14	0.18	0.25	0.08	0.11	0.21	0.18
0.75	24	normal	normal	0.46	0.59	0.02	0.14	0.12	0.17	0.08	0.11	0.14	0.13

For each scenario, r_f is the proportion of P removed from an individual field and "# of samples" refers to the number of P measurements used to monitor water quality change. The values in italics are the baseline scenario. Stop point values are the watershed-scale BMP implementation levels (I^*_{wt} and I^*_{wr}) that maximize the efficiency of the aggregated approaches. As a benchmark to facilitate comparison of approaches, benefit index values (proportions) are reported for a 10% landscape-scale BMP implementation level deviations equal to mean/3). This change was intended to address uncertainty about whether the empirical distributions represent the true distributions for Wisconsin, and to examine how the allocation approaches would perform in more uniform agricultural landscapes. Third, we varied the number of pre- and post-intervention water samples to represent shorter (12 samples) and longer (48 samples) monitoring periods. We use the results of this sensitivity analysis to identify influential variables and important uncertainties.

Results

Based on the simulated experiments, the probability of observing a statistically significant P reduction in Wisconsin agricultural streams is related to the intervention level by a sigmoid (threshold) function (Fig. 3). In any one stream, the threshold is abrupt, but because P concentration variance is not consistent or predictable among streams, the cumulative probability density function for a randomly chosen stream is more gradual. Based on this function, the optimal levels of effort in a watershed—the stop points—are 59% of randomly-chosen fields (I^*_{wr}) or 29% of targeted fields (I^*_{wt}) in the baseline scenario (Fig. 4). The stop points are sensitive to variation in all model parameters; targeted stop points ranged from 21% to 55%, and random stop points from 49% to 100%, among scenarios (Table 1). For a given scenario, the targeted stop point was about half of the random stop point, except when the field-level P loss distribution was normal, where it was about three-fourths.

The four allocation approaches perform differently according to our two evaluation methods (Fig. 5). We use cumulative net benefit (benefit/effort) as a common evaluation criterion. This criterion is different than instantaneous efficiency (marginal change in benefit per unit effort), which is an appropriate criterion for selecting sites in iterative site selection algorithms (Hyman and Leibowitz 2000), but which is not useful for evaluating the relative performance of the approaches presented here. The relevant sections of the benefit/effort curves (Fig. 5) are at low conservation effort because most state and federal programs charged with controlling NPS pollution do not have sufficient funds to implement BMPs on a substantial proportion of the landscape. For example, the Wisconsin PWP helped pay for BMPs on approximately 10% of farms in the state during its period of activity (Wisconsin Department of Natural Resources, unpublished data).

Based on the modeled pollutant reduction index (Fig. 5a), the dispersed/targeted approach creates the most benefit at all levels of effort. The aggregated/targeted approach is next best, and at low and high effort performs nearly as well as dispersed/targeted. The aggregated/random approach performs well at low effort, but becomes equivalent to the dispersed/random approach once the stop point has been reached in all watersheds. The relative performance of the allocation approaches was only sensitive to the distribution of P losses. Aggregation and targeting are particularly important when among-watershed and among-field P losses, respectively, are log-normally distributed (Table 1).

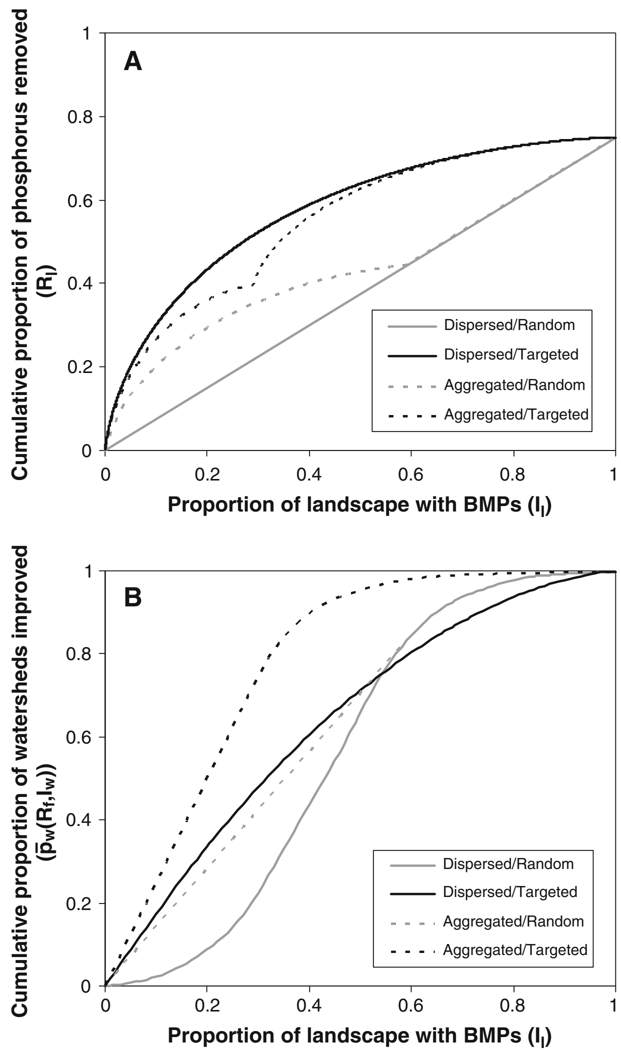

Fig. 5 Benefit/cost curves for the four geographical allocation approaches in the baseline scenario (see Table 1). (**a**) Benefit is the reduction in P pollution from entire model landscape (R_l); (**b**) Benefit is the proportion of watersheds where a statistically significant difference is observed in stream water P concentration ($\bar{p}(R_f, I_w)$)

Based on the measured water quality change index (Fig. 5b), the aggregated/targeted approach creates the most benefit at all levels of effort, particularly with high BMP performance and log-normal P loss distributions at the among-field scale (Table 1). The dispersed/targeted approach performs slightly better than the aggregated/random approach up to moderate effort, in part because of the high P reduction efficiency of targeting, and in part because of de facto aggregation of highly polluting fields. The dispersed/random approach performs particularly poorly at low levels of effort, but becomes comparable to other approaches at higher levels, where most large pollution sources are addressed and most watersheds have high levels of BMP implementation.

The dispersed/random approach has the least structure of all the approaches, and therefore can be considered a baseline to which others can be compared. It is clear that adding a targeting *or* an aggregating component to this baseline approach improves program efficiency according to both evaluation methods. Targeting, by itself, improves performance more than aggregating, particularly for landscape-scale P reduction. However, targeting without aggregating would be impractical to implement, since it would require estimating pollutant loss from all fields in the landscape. While we did not formally evaluate the impact of assessment costs on program efficiency, these costs are likely to be substantial and offset the benefits of this approach.

Combined targeting *and* aggregating of agricultural BMPs can most efficiently cause measurable stream water quality changes. The aggregated/targeted approach is most efficient overall; for example, with effort on only 10% of the landscape in the baseline scenario, 26% of the total agricultural NPS P load would be eliminated and 25% of watersheds would significantly improve. The aggregated/targeted approach is a streamlined approximation of an iterative site selection approach, where the benefit derived from each site is dependent on which sites have already been selected (Hyman and Leibowitz 2000). The hierarchical spatial organization of potential benefit in agricultural landscapes allows this streamlined method to function much like a true iterative approach, but without the necessity of re-evaluating site-scale instantaneous efficiency at each step. Furthermore, the hierarchical method for identifying potential sites for management action can reduce assessment costs.

The sensitivity analysis indicates that the superiority of the *aggregated/targeted* approach is robust to uncertainties in model parameters and to assumption choices. With BMPs on 10% of the landscape, the aggregated/targeted approach had the highest average (mean of benefit indices) efficiency for all scenarios (Table 1). Increasing the number of water quality samples decreased the stop points and increased the likelihood of observing significant differences at a given level of effort. However, these changes were proportional among approaches. Increasing the assumed effectiveness of BMPs (R_f) exaggerated differences among approaches, particularly according to the measured water quality change index. In particular, with R_f at 40%, differences between all the approaches, except dispersed/random, were minimal. The effect of changing P loss distributions from log-normal to normal was different for the two benefit indices. When done at the among-field scale, it reduced the relative ability of targeting to improve water quality, whereas doing so at the among-watershed scale most strongly reduced the ability of both targeting and aggregating to reduce landscape-scale P loss. Together, these findings suggest that targeting and aggregating NPS pollution reduction will be most useful in highly variable agricultural landscapes and when BMPs are effective.

Discussion

Thresholds

Targeting agricultural conservation effort toward the largest sources of pollution can dramatically improve the efficiency of landscape-scale pollution reduction (Fig. 5a) (see also Hansen and Hellerstein 2006). However, this approach fails to consider the value of cumulative benefits at discrete, relatively small scales. The relationship between nutrient levels and the ecological condition of running waters may often be nonlinear (e.g., Dodds and others 2002; Brazner and others 2004; Wang and others 2007). Moreover, different ecological response variables (e.g., community structures of fish, invertebrates, and periphyton) may exhibit different relationships with nutrient levels (Wang and others 2007). Ideally, managers could identify thresholds for all variables of interest and allocate management effort according to the position of different ecosystems relative to these thresholds (Statzner and others 1997).

The statistical threshold we use in this article is probably not equivalent to all relevant ecological thresholds. Sigmoid curves are often used to represent threshold relationships between a response (y) and a predictor (x), and can be defined by their mean (μ) and steepness (s) through the equation $y = 1/(1 + \exp(-(x - \mu) \cdot s))$. The statistical P reduction threshold has a moderate mean (0.32) and a moderate steepness (4.3). The "low threshold" in Fig. 3 ($\mu = 0.1, s = 5$) is a hypothetical relationship where the response variable increases quickly at relatively small intervention levels. For example, acute stream water hypoxia is often caused by manure spills originating from one or a few farms in a watershed (Tegtmeier and Duffy 2004). If these farms can be identified, a relatively minor intervention (in terms of the mean watershed-wide P loss) could result in a large increase in the probability of eliminating hypoxic events. The "high threshold" in Fig. 3 ($\mu = 0.8,\ s = 20$) represents a relationship where the response variable does not increase until relatively high intervention levels. For example, most agricultural streams carry P loads that are much higher than they would be without human influence (i.e., their reference condition) (Robertson and Saad 2003). Achieving reference P loads or concentrations in most streams would require a higher reduction in P inputs than would simply detecting a significant reduction in P.

It is clear that there are many possible thresholds in the relationship between conservation effort and ecological response; the statistical threshold we use in this article is only one example. We chose this threshold primarily because it was quantifiable from available data, but also because achieving statistically significant water quality improvements is useful, in that it will facilitate calibration of and provide empirical support for modeled pollutant reduction approaches. Using thresholds to allocate environmental management effort will generally result in aggregation of effort (unless the threshold is very low). In addition to creating cumulative benefits more efficiently, aggregation can improve management outcomes in two ways. First, public perception of the performance of conservation programs is likely to benefit from success stories from specific, identifiable places rather than the calculated sum of tiny, incremental changes sparsely distributed over a broad region. And second, aggregation of effort into a discrete set of geographic units facilitates experimentation and adaptive management.

Adaptive Management of Nonpoint Source Pollution

Based on the above evaluation, we recommend an aggregated/targeted approach to the allocation of landscape-scale NPS pollution control effort. In brief, implementation of this approach follows a four-step process: (1) Delineate a set of watersheds whose size maximizes overall program utility (e.g., Maxted and others, this issue); (2) Rank these watersheds according to their water quality restoration potential (e.g., Diebel and others, this issue); (3) Use field-scale pollutant loss models (e.g., Phosphorus Index) to rank fields within a watershed according to their contribution to the total pollutant load; (4) Beginning with the top-ranked watershed, implement BMPs on the top 29% (the stop point for the aggregated/targeted approach) of fields in each watershed.

Each component in this approach is based on models of how NPS pollution is distributed, how it can be mitigated, and what mitigation outcomes are likely to occur. As with any model-based management approach, its real-world performance may be limited by simplifying design decisions. For example, spatial relationships among landscape units (terrestrial and aquatic features acting as sources and sinks for pollutants) were not accounted for in this analysis and may have a significant influence on net pollutant export (Alexander and others 2008). Our analysis assumes consistent application of management goals among landscape units, where in reality, goals may vary according to the degree of impairment relative to designated uses (e.g., Total Maximum Daily Loads; USEPA 1999). In addition, the cost of land management is strongly dependent on institutional arrangements, which can vary both among and within management jurisdictions. Therefore, an overarching factor in the design of this approach was that it be amenable to experimentation and adaptation.

Adaptive management means treating policies as testable hypotheses (Lee 1993). In our approach, comparably-sized, spatially-independent watersheds serve as experimental units (Maxted and others, this issue). Spatial independence facilitates statistical analysis, and because these watersheds are smaller than those commonly used in other landscape-scale management programs (e.g., USGS 1994; WDNR 2005), more of them can be funded, providing better opportunities for evaluation. Many existing "hydrologic unit" delineations (e.g., USGS 1994) contain true watersheds (the entire area upslope of a point), downstream segments of larger watersheds, and collections of several adjacent smaller watersheds. In watershed-based NPS pollution reduction programs, it is prudent to allocate effort to true watersheds, so that all potential pollution sources are accounted for, and so that changes in water quality can be monitored at a single point (Griffith and others 1999).

Models for targeting conservation effort among (Diebel and others, this issue) and within (Phosphorus Index) watersheds provide predictions of outcomes that can be compared with real outcomes. To evaluate model predictions, implementation should purposely deviate from the approach recommended above. This deviation should be focused on only one or two parameters at a time and should be structured to address critical uncertainties. For example, the stop point (29% of fields) identified by the simulation model signifies the amount of effort that should be expended in one watershed before moving on to the next watershed. With adaptive management, the amount of effort should be varied among watersheds to assess the accuracy of this estimated threshold and to assess whether relevant ecological responses correspond with statistical thresholds.

Despite strong theoretical support for the utility of adaptive management (Holling 1978), putting it into practice within state and national environmental policy has been challenging (Ruhl 2006). Policies that result in inconsistent application of regulations or incentives are often unpopular. Furthermore, landowners who are accustomed to environmental regulations that seem to be based in scientific certainty may balk at the idea of their land being subject to an experiment. And from an administrative standpoint, the continuous monitoring and adjustment that is at the heart of adaptive management is at odds with the "command-and-control" paradigm (Holling and Meffe

1996) that has come to dominate natural resource policy. Thus, true adaptive management of agricultural NPS pollution will require a willingness by landowners to see the broad landscape context of their actions, as well as a willingness by creators of administrative law to authorize agencies to adapt to new information.

The four-step approach outlined above is a spatially hierarchical method for targeting NPS pollution control, from broad regions such as states, where policy is enacted; to watersheds, where water quality is relevant; to farms and fields, where management is practiced. Our landscape simulation results show that targeting at multiple spatial scales is necessary to optimize program efficiency. In two companion articles, we fully develop a set of watersheds (Maxted and others, this issue) and a method for evaluating their restoration potential (Diebel and others, this issue). Our approach is most specific at this broad spatial scale because these issues had not previously been addressed in the literature. Targeting NPS pollution reduction within watersheds can be accomplished with a number of existing models (e.g., Phosphorus Index [Sharpley and others 2003], WINHUSLE [WDNR 1994b], BARNY [WDNR 1994a]). Because program implementation at this level is typically carried out by local (e.g., county) conservation staff, model choice can be flexible to accommodate different technical capabilities and preferences. Alternative approaches could then be evaluated in a more holistic version of adaptive management, where administrative efficiency is monitored along with ecosystem response.

The use of agricultural conservation practices has no doubt benefited the environment. However, much environmental degradation is still caused by agriculture, and the benefits of conservation have been difficult to measure. Because the size of the problem greatly exceeds the amount of resources available to address it, policies have emphasized implementation—using available funds to get as many conservation practices on the ground as possible. Opportunities to learn from experience have been lost because of this emphasis. In the long term, implementation-oriented programs may create fewer environmental benefits than if they had incorporated prioritization, monitoring, and adaptation from the beginning. Our approach acknowledges that conservation practices can only be implemented on a limited portion of the landscape at any one time. It then capitalizes on nonlinear relationships between effort and benefit to prioritize aggregated effort in small watersheds. Successful observation of watershed-scale responses in these watershed management units should be viewed as a first goal in a more protracted mission to reduce nonpoint source pollution.

Acknowledgments This research was supported by the USDA Natural Resources Conservation Service and the Wisconsin Department of Natural Resources. The authors thank the Wisconsin Buffer Initiative Advisory Committee for contributing ideas and feedback throughout this project, Laura Ward Good for providing the phosphorus index data, Dale Robertson for providing the stream phosphorus data, and Brian Weidel for coding the statistical simulations. Emily Stanley and Chris Solomon provided comments that substantially improved the manuscript.

References

Alexander RB, Smith RA, Schwarz GE, Boyer EW, Nolan JV, Brakebill JW (2008) Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin. Environmental Science and Technology 42:822–830

Babcock BA, Lakshminarayan PG, Wu JJ, Zilberman D (1997) Targeting tools for the purchase of environmental amenities. Land Economics 73:325–339

Bernhardt ES, Palmer MA, Allan JD, Alexander G, Barnas K, Brooks S, Carr J, Clayton S, Dahm C, Follstad-Shah J, Galat D, Gloss S, Goodwin P, Hart D, Hassett B, Jenkinson R, Katz S, Kondolf GM, Lake PS, Lave R, Meyer JL, O'Donnell TK, Pagano L, Powell B, Sudduth E (2005) Synthesizing US river restoration efforts. Science 308:636–637

Bishop PL, Hively WD, Stedinger JR, Rafferty MR, Lojpersberger JL, Bloomfield JA (2005) Multivariate analysis of paired watershed data to evaluate agricultural best management practice effects on stream water phosphorus. Journal of Environmental Quality 34:1087–1101

Boesch DF, Brinsfield RB, Magnien RE (2001) Chesapeake Bay eutrophication: scientific understanding, ecosystem restoration, and challenges for agriculture. Journal of Environmental Quality 30:303–320

Brazner JC, Tanner DK, Detenbeck NE, Batterman SL, Stark SL, Jagger LA, Snarski VM (2004) Landscape character and fish assemblage structure and function in western Lake Superior streams: General relationships and identification of thresholds. Environmental Management 33:855–875

Bundy LG, Mallarino AP, Good LW (2008) Field-scale tools for reducing nutrient losses to water resources. Pages 159–170 in Final Report: Gulf Hypoxia and Local Water Quality Concerns Workshop. September 26–28, 2005, Ames, Iowa. American Society of Agricultural and Biological Engineers, St. Joseph, Michigan

Burkart MR, James DE (1999) Agricultural-nitrogen contributions to hypoxia in the Gulf of Mexico. Journal of Environmental Quality 28:850–859

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH (1998) Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8:559–568

Carpenter SR, Ludwig D, Brock WA (1999) Management of eutrophication for lakes subject to potentially irreversible change. Ecological Applications 9:751–771

Carpenter SR (2005) Eutrophication of aquatic ecosystems: Bistability and soil phosphorus. Proceedings of the National Academy of Sciences 102:10002–10005

Correll DL (1999) Vegetated stream riparian zones: their effects on stream nutrients, sediments, and toxic substances. Smithsonian Environmental Research Center, Edgewater, MD

Corsi SR, Walker JF, Wang L, Horwatich JA, Bannerman RT (2005) Effects of best-management practices in Otter Creek in the Sheboygan River Priority Watershed, Wisconsin, 1990–2002. U.S. Geological Survey Scientific Investigations Report 2005–5009, Middleton, WI

Daniels RB, Gilliam JW (1996) Sediment and chemical load reduction by grass and riparian filters. Soil Science Society of America Journal 60:246–251

Davie DK, Lant CL (1994) The effect of CRP enrollment on sediment loads in two southern Illinois streams. Journal of Soil and Water Conservation 49:407–412

Desbonnet A, Pogue P, Lee V, Wolf N (1994) Vegetated buffers in the coastal zone: a summary review and bibliography. University of Rhode Island, Providence, RI

Diebel MW, Maxted JT, Han S, Robertson DM, Vander Zanden MJ (2008) Landscape planning for agricultural nonpoint source pollution reduction III: assessing phosphorus and sediment reduction potential. Environmental Management, this issue

Dillaha TA, Reneau RB, Mostaghimi S, Lee D (1989) Vegetative filter strips for agricultural nonpoint source pollution control. Transactions of the ASAE 32:513–519

Dodds WK, Smith VH, Lohman K (2002) Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Canadian Journal of Fisheries and Aquatic Sciences 59:865–874

Gitau MW, Gburek WJ, Jarrett AR (2005) A tool for estimating best management practice effectiveness for phosphorus pollution control. Journal of Soil and Water Conservation 60:1–10

Graczyk DJ, Walker JF, Horwatich JA, Bannerman R (2003) Effects of best-management practices in the Black Earth Creek Priority Watershed, Wisconsin, 1984–98. U.S. Geological Survey Water-Resources Investigations Report 03-4163, Middleton, WI

Griffith GE, Omernik JM, Woods AJ (1999) Ecoregions, watersheds, basins, and HUCs: how state and federal agencies frame water quality. Journal of Soil and Water Conservation 54:666–677

Hansen L, Hellerstein D (2006) Better targeting, better outcomes. U.S. Department of Agriculture, Economic Research Service. Economic Brief No. 2

Holden C, Castelnuovo R, Fallat C, Patel D, Modaff K, Knapp K (2006) Wisconsin land and water conservation: Annual progress report (2005). Publication No. ARM pub 163. Available online at http://dnr.wi.gov/org/water/wm/nps/pdf/2005jointannualreport.pdf

Holling CS (ed) (1978) Adaptive environmental assessment and management. Wiley-Interscience, New York

Holling CS, Meffe GK (1996) Command and control and the pathology of natural resource management. Conservation Biology 10:328–337

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

Knox JC (2006) Floodplain sedimentation in the Upper Mississippi Valley: natural versus human accelerated. Geomorphology 79:286–310

Lee KN (1993) Compass and gyroscope: integrating science and politics for the environment. Island Press, Washington, DC, p 243

Limpert E, Stahel WA, Abbt M (2001) Log-normal distributions across the sciences: keys and clues. Bioscience 51:341–352

Lowrance R, Altier LS, Newbold DJ, Schabel RR, Groffman PM, Denver JM, Correll DL, Gilliam JW, Robinson JL, Brinsfield RB, Staver KW, Lucas W, Todd AH (1997) Water quality functions of riparian forest buffers in Chesapeake Bay watersheds. Environmental Management 21:687–712

Maxted JT, Diebel MW, Vander Zanden MJ (2008) Landscape planning for agricultural nonpoint source pollution reduction II: balancing watershed size, number of watersheds, and implementation effort. Environmental Management, this issue

Meals DW (1996) Watershed-scale response to agricultural diffuse pollution control programs in Vermont, USA. Water Science and Technology 33:197–204

Mendez A, Dillaha TA, Mostaghini S (1999) Sediment and nitrogen transport in grass filter strips. Journal of the American Water Resources Association 35:867–875

Nerbonne BA, Vondracek B (2001) Effects of local land use on physical habitat, benthic macroinvertebrates, and fish in the Whitewater River, Minnesota, USA. Environmental Management 28:87–99

Nowak P, Bowen S, Cabot PE (2006) Disproportionality as a framework for linking social and biophysical systems. Society and Natural Resources 19:153–173

NRC (National Research Council) (1992) Restoration of aquatic ecosystems: science, technology and public policy. National Academy Press, Washington, DC

Osborne LL, Kovacic DA (1993) Riparian vegetated buffer strips in water-quality restoration and stream management. Freshwater Biology 29:243–258

Robertson D, Saad D (2003) Environmental water-quality zones for streams: a regional classification system. Environmental Management 31:581–602

Robertson DM, Graczyk DJ, Garrison PJ, Wang L, LaLiberte G, Bannerman R (2006) Nutrient concentrations and their relations to the biotic integrity of wadeable streams in Wisconsin. U.S. Geological Survey Professional Paper 1722, Middleton, WI

Ruhl JB (2006) Regulation by adaptive management-is it possible? Minnesota Journal of Law, Science, and Technology 7:21–57

Sharpley AN, Weld JL, Beegle DB, Kleinman PJA, Gburek WJ, Moore PA, Mullins G (2003) Development of phosphorus indices for nutrient management planning strategies in the United States. Journal of Soil and Water Conservation 58:137–152

Spooner J, Jamieson CJ, Maas RP, Smolen MD (1987) Determining statistically significant changes in water pollutant concentrations. Lake and Reservoir Management 3:195–201

Statzner B, Capra H, Higler LWG, Roux AL (1997) Focusing environmental management budgets on nonlinear system responses-potential for significant improvements to freshwater ecosystems. Freshwater Biology 37:463–472

Stuntebeck TD, Bannerman R (1998) Effectiveness of barnyard best management practices in Wisconsin. U.S. Geological Survey Fact Sheet FS-051-98

Tegtmeier EM, Duffy MD (2004) External costs of agricultural production in the United States. International Journal of Agricultural Sustainability 2:1–20

Turner RE, Rabalais NN (1994) Coastal eutrophication near the Mississippi River delta. Nature 368:619–621

USDA (US Department of Agriculture) (2003) ASCS handbook: conservation reserve program. 1-CRP (Rev. 1). USDA. Agricultural Stabilization and Conservation Service, Washington, DC

USDA (US Department of Agriculture) (2006) Conservation reserve program: summary and enrollment statistics–2005. USDA. Farm Service Agency, Washington, DC

USEPA (US Environmental Protection Agency) (1990) National water quality inventory: 1988 Report to Congress. Office of Water, Washington, DC

USEPA (US Environmental Protection Agency) (1996) Environmental indicators of water quality in the United States. Office of Water. EPA 841/R-96/002

USEPA (US Environmental Protection Agency) (1999) Protocol for developing nutrient TMDLs, First Edition. Office of Water. EPA 841/B-99/007

USEPA (US Environmental Protection Agency) (2000) Water quality conditions in the United States: a profile from the 1998 National water quality inventory report to congress. EPA 841/F-00/006

USEPA (US Environmental Protection Agency) (2002) A review of statewide watershed management approaches. Office of Water

USGS (US Geological Survey) (1994) 1:250,000-scale hydrologic units of the United States. Available online at http://water.usgs.gov/GIS/metadata/usgswrd/XML/huc250k.xml

UW-CALS (University of Wisconsin-Madison, College of Agricultural and Life Sciences) (2005) Wisconsin Buffer Initiative final report. Available online at http://www.drs.wisc.edu/wbi/

Wang LZ, Lyons J, Kanehl P, Gatti R (1997) Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. Fisheries 22:6–12

Wang LZ, Lyons J, Kanehl P (2002) Effects of watershed best management practices on habitat and fish in Wisconsin streams. Journal of the American Water Resources Association 38:663–680

Wang LZ, Lyons J, Kanehl P (2006) Habitat and fish responses to multiple agricultural best management practices in a warm water stream. Journal of the American Water Resources Association 42:1047–1062

Wang LZ, Robertson DM, Garrison PJ (2007) Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implication to nutrient criteria development. Environmental Management 39:194–212

Waters TF (1995) Sediment in streams: sources, biological effects and control. American Fisheries Society, Bethesda, MD

Wenger S (1999) A review of the scientific literature on riparian buffer width, extent and vegetation. Literature review University of Georgia Institute of Ecology. Office of Public Service & Outreach, Athens, GA

Wischmeier WH, Smith DD (1978) Predicting rainfall erosion losses-a guide to conservation planning. USDA Handbook 537, Washington, DC

WDNR (1994a) (Wisconsin Department of Natural Resources). BARNY 2.2-The Wisconsin barnyard runoff model, inventory instructions and user's manual. Report WR-285-91

WDNR (1994b) (Wisconsin Department of Natural Resources). WINHUSLE 1.4.4–Model documentation and user's manual. Report WR-294-91

WDNR (2005) (Wisconsin Department of Natural Resources). Hydrography of Wisconsin (GIS data). Available online at http://dnr.wi.gov/maps/gis/datahydro.html

Wisconsin Legislative Fiscal Bureau (2007) Nonpoint source and water pollution abatement and soil conservation programs. Informational Paper 66. Available online at http://www.legis.state.wi.us/lfb/Informationalpapers/66.pdf

Wolf A (1995) Rural nonpoint source pollution control in Wisconsin: the limits of a voluntary program. Water Resources Bulletin 31:1009–1022