Improving Wetland Mitigation Site Identification Through Community Distribution Modeling and a Patch-Based Ranking Scheme

ARTICLE *in* WETLANDS · OCTOBER 2012

Impact Factor: 1.57 · DOI: 10.1007/s13157-012-0315-7

4 AUTHORS, INCLUDING:

Elizabeth A. Hunter
University of Georgia

Patrick A. Raney
Upper Susquehanna Coalition

James P. Gibbs
State University of New York College of Env...

Wetlands
DOI 10.1007/s13157-012-0315-7

ARTICLE

Improving Wetland Mitigation Site Identification Through Community Distribution Modeling and a Patch-Based Ranking Scheme

Elizabeth A. Hunter · Patrick A. Raney · James P. Gibbs · Donald J. Leopold

Received: 20 March 2012 / Accepted: 5 June 2012

Abstract Current wetland mitigation practices do not fully recover wetland function, often due to poor mitigation site selection. Improved mitigation site selection methods are needed to efficiently assess the suitability and quality of potential wetland mitigation sites at broad spatial scales. We present a novel application of maximum entropy-based predictive distribution models coupled with a patch-based ranking scheme to identify potential wetland mitigation sites and contrast their effectiveness relative to a conventional "expert opinion" model. We used hydrogeologic and landscape features and widely available wetland community distribution data to predict locations of wetlands that were previously unknown, destroyed, or biologically rare in the Upper Susquehanna River Basin in the northeastern United States. An "expert opinion" model predicted wetland occurrence based on topographic slope and soil type. Maximum entropy-based models predicted an independent sample of wetland locations well (Area Under the Curve=0.86–0.98; 92 % correct classification rate) and dramatically outperformed the "expert opinion" model (62 % correct classification rate). A patch-based ranking scheme, which incorporated additional influences on wetland quality, ranked sites with biologically important wetland plant communities highly among model-identified wetlands. We conclude that integration of maximum entropy-based predictive modeling and patch-based ranking can effectively identify high quality wetland mitigation sites.

Keywords Community distribution modeling · MaxEnt · Maximum entropy modeling · Patch-based ranking · Wetland mitigation · Wetland restoration

Introduction

Wetlands support a disproportionately large amount of Earth's biota (Dudgeon et al. 2006), and wetland dependent species are endangered by continued global wetland losses (Wilcove et al. 1998; Gibbs 2000; Bedford et al. 2001; Keddy 2010). Outright wetland protection is an important conservation strategy, yet additional measures must be applied to enhance degraded areas as well as increase wetland extent to restore the wetland biota (Gaston 2003; IUCN 2004). There is an urgent need for better integration of wetland restoration projects within broader wetland reserve networks to maximize benefits to biodiversity and ecological services rendered (Young 2000).

Frameworks and financial backing for wetland creation, protection, and restoration (hereafter, simply wetland restoration) in the United States are provided through the federal government's policy of no net loss and Section 404 Clean Water Act permits that require wetland mitigation for altered or destroyed "jurisdictional" wetlands (for overview: Mitsch and Wilson 1996; Mitsch and Gosselink 2000). Despite these legal structures and provisions, wetland restoration targets relevant to watershed functioning have not been established at large spatial scales for any region (Race and Fonseca 1996; Kershner 1997; Zedler 2003). A strategic planning approach is needed to improve the ecological and hydrological functioning of restored wetlands at the watershed to regional level (Margules and Pressey 2000; White

E. A. Hunter (✉) · P. A. Raney · J. P. Gibbs · D. J. Leopold
Department of Environmental and Forest Biology, State University of New York College of Environmental Science and Forestry,
1 Forestry Dr.,
Syracuse, NY 13210, USA
e-mail: ehunte01@syr.edu

E. A. Hunter · P. A. Raney
Upper Susquehanna Coalition,
183 Corporate Dr.,
Owego, NY 13827, USA

Published online: 10 July 2012

and Fennessy 2005). Restoration planning at the watershed scale may also vastly improve the quality of individual restoration projects for biodiversity by, for example, maximizing propagule dispersal probability, enhancing metapopulation dynamics, and reducing energy expenditure for wildlife moving among wetlands (Gibbs 2000; Margules and Pressey 2000). Much wetland restoration currently proceeds based on identifying restoration priorities with little information prior to field visitation (White and Fennessy 2005). As a result, restoration sites are often chosen on the basis of expert opinion, the evaluation of only a few landscape features (e.g., hydric soils maps), and the availability of parcels (Biebighauser 2007). This approach has resulted, in some cases, in failure of wetland mitigation schemes (Mitsch and Wilson 1996; Moreno-Mateos et al. 2012).

A long history of wetland drainage in the United States (e.g., Zedler 2003; Biebighauser 2007) has obscured many historical wetland locations. With only extant wetlands as references, restoration targets are often lacking. However, due to past drainage practices, restoration site identification must be based on presence-only datasets (that is, locations of extant wetlands). Therefore, regression-based modeling approaches requiring presence-absence data are impractical (Gotelli and Ellison 2004) and at a serious disadvantage to those operable with presence-only type data (Elith et al. 2006; Phillips et al. 2006).

Maximum entropy species distribution modeling (specifically the program MaxEnt) is a statistical modeling approach developed specifically for presence-only data (Phillips et al. 2006; Phillips and Dudík 2008). Recent analyses indicate that this modeling approach regularly outperforms both simpler approaches (like regression-based ones) as well as more sophisticated techniques such as genetic algorithms in predicting species distributions (Elith et al. 2006; Heikkinen et al. 2006). Despite these advantages, the maximum entropy approach has rarely been used to predict the occurrence of community types, such as wetlands (but see Deblauwe et al. 2008). Because wetland locations are largely determined by climate and environmental characteristics (e.g., geomorphology) (Brinson 1993; Mitsch and Gosselink 2000; Bedford et al. 2001; Keddy 2010), the landscape setting of both remaining and drained wetlands can potentially be predicted using presence-only datasets in the program MaxEnt. The use of MaxEnt could enhance predictive capabilities over "expert opinion" methods based on hydric soils (e.g. Van Lonkhuyzen et al. 2004) because MaxEnt is able to account for complex interactions among landscape features that determine the presence or absence of wetlands (Phillips and Dudík 2008).

MaxEnt-predicted wetland sites can be further prioritized using simple abiotic criteria in a Geographic Information System (GIS). Patch-based rankings incorporate abiotic and biotic information that may influence the biological quality of a site but not necessarily the occurrence of a wetland. Including more information on site quality allows resource managers to efficiently identify locations with the highest potential for long-term restoration success. Patch-based ranking systems can be applied to large areas, tailored to fit a variety of resource management goals, and can be adapted as more information becomes available (Lee et al. 2001; Bayliss et al. 2003; Geneletti 2004). These advantages of patch-based ranking systems outweigh the disadvantages of subjectively weighting variable influence on site quality (Bayliss et al. 2003). Ideally, ranking schemes would incorporate information from direct field assessments of predicted suitable sites, but comprehensive field surveys are not always feasible, and evaluation prior to site visitation improves conservation planning efficiency (Margules and Pressey 2000). Variables that could contribute to wetland site quality, but not occurrence, include patch size, road density, and proximity to other wetlands. Reductions in patch area can compromise biodiversity (MacArthur and Wilson 1963) and water quality (Benayas et al. 2008). High road density has detrimental effects on wetland biota from road salt contamination and mortality from road crossings (e.g., Fahrig et al. 1995; Kaushal et al. 2005; Karraker et al. 2008). Existing large wetlands near restoration sites are likely biological sources of repopulation for nearby restoration sites (Nathan and Muller-Landau 2000), thus increasing the biodiversity value of such wetlands as mitigation sites.

We used maximum entropy distribution models and a patch-based ranking scheme to predict high quality wetland restoration sites in the Upper Susquehanna River Basin in New York State. The main objective of this study was to test the hypothesis that maximum entropy modeling is better at predicting hydrologically suitable wetland restoration sites and biologically important wetland communities than an "expert opinion" model based on landscape characteristics. Our secondary objective was to develop and demostrate application of an approach to integrate spatially explicit patch-based ranking schemes in conjunction with predicted wetland sites to prioritize resource management targets for large spatial scales.

Methods

Focal Region

Our study focused on the 16,186 km^2 New York State (NYS) portion of the Upper Susquehanna River Basin (USRB), which is located within the Appalachian Highlands ecoregion that serves as the headwaters for Chesapeake Bay (Fig. 1). Landcover in the region is ca. 50 % forest, 25 % agriculture, 5 % urban or developed, 5 % wetland, and 15 % remaining types (Homer et al. 2004). The USRB contains some of the thickest portions of the Marcellus Shale formation, and natural

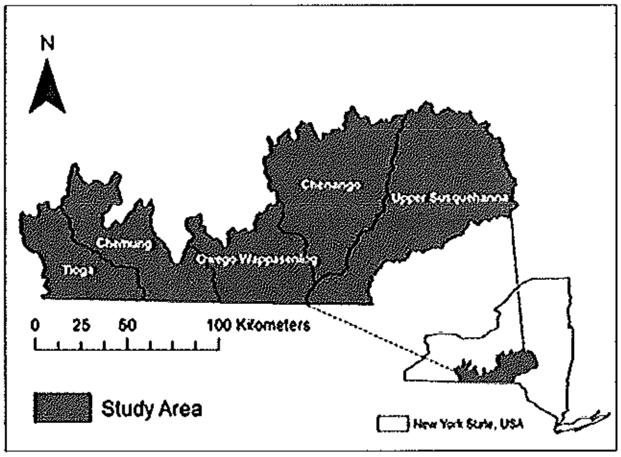


Fig. 1 Study site used to predict wetland occurrence based on National Wetland Inventory data in the subwatersheds of the New York State portion of the Upper Susquehanna River Basin (dark gray) in the northeastern United States

gas development appears imminent (Arthur et al. 2008), for which development of the required infrastructure would require dramatically increased wetland mitigation action (NYDEC 2011). Wetlands within the focal region are often isolated within fragmented landscapes of secondary forests in various stages of succession following agricultural land abandonment. Typical wetlands include: shallow and deep emergent marshes, riparian wetlands, bogs, poor to richly minerotrophic fens, and forested wetlands (Edinger et al. 2002). Of these wetland communities, fens in particular support a large number of the state protected plant species in NYS (e.g., Johnson and Leopold 1994; Bedford and Godwin 2003). Because their high biodiversity value could affect mitigation site value (Mitsch and Gosselink 2000), we modeled the occurrence of fens and bogs in addition to more common wetland communities.

Input Data

The National Wetland Inventory Program (NWI) database (USFWS 2010), which contains wetland occurrence polygons complete with habitat and hydrology descriptors, was used as a source of occurrences for common wetland types. The NWI database is the most comprehensive dataset covering the entire USRB and can be an accurate representation of wetland occurrence in predominantly forested areas, with accuracy above 90 % (Kudray and Gale 2000). Not all wetlands within a region are represented in the NWI database (Tiner 1997), but maximum entropy modeling is robust to errors of omission (Elith et al. 2006), making NWI data highly suitable for use in MaxEnt. The most typical wetlands within the region were palustrine (non-tidal, inland wetlands without flowing water), with emergent, forested, and scrub-shrub being the most common classes, representing 90 % of the wetland occurrences (Cowardin et al. 1979). Only the most common wetland types in the region (temporarily or seasonally flooded palustrine wetlands [NWI water regime modifiers A, C, and E]) were used to create the presence dataset.

Because wetlands differ biologically based on hydrogeologic setting (Godwin et al. 2002), we modeled the following wetland types separately: scrub-shrub, emergent, evergreen forested, and broad-leaved deciduous forested. Further division of the emergent and scrub-shrub wetlands to the subclass level (e.g., evergreen) are possible under the Cowardin et al. (1979) classification, but these wetland types were underrepresented in NWI data in the USRB. Additionally, the following wetland types of biological importance were also modeled: (1) polygons with the acidic designation ($n=59$) for "bog" locations and (2) medium to richly minerotrophic fen locations compiled from previous field visitations ($n=26$) and from the New York Natural Heritage Program database ($n=9$) (NYNHP 2011) for "fen" occurrences.

To predict wetland occurrence we used five environmental variables: elevation, slope, aspect, surface geology, and soil type, all scaled to 10 by 10 m cells (Table 1). Primary bedrock types (e.g., shale) were used to represent surficial geology. Over 1,000 soil types were present in the USRB, and each of the 16 counties in the USRB had hundreds of soil types represented, with few types common among counties. Soil types were reclassified into 15 primary classes based on majority soil composition listed in the SSURGO database (NRCS 2010): alluvial, channery loam, channery silt loam, gravelly loam, gravelly sandy loam, gravelly silt

Table 1 Description of source data for environmental variables used to predict wetland occurrence based on National Wetland Inventory data in the Upper Susquehanna River Basin in New York State

Variable	Source	Scale	Datum	Units
Elevation (DEM)	National Elevation Dataset (Gesch et al. 2002)	1:24,000	NAD 27	Meters
Slope	Calculated from DEM in ArcGIS	1:24,000	NAD 27	Percent
Aspect	Calculated from DEM in ArcGIS	1:24,000	NAD 27	Degrees
Soils	SSURGO (NRCS 2010)	Various, typically 1:24,000	NAD 83	Categorical
Bedrock	USGS (Nicholson et al. 2006)	1:2,500,000	WGS 84	Categorical

Each variable was used as a predictive layer in the program MaxEnt loam, gravelly or stony, loam, muck or peat, pits (severely anthropogenically disturbed soils), sand, sandy loam, silt loam, silty clay loam, and water. Classifying soils in this way improved conformity among counties and facilitated wetland occurrence prediction.

Modeling Assumptions

MaxEnt assumes all presence locations are of the same type (usually a species) and there is little variability in the species' niche preference for particular locations (Phillips et al. 2006). Although our inputs were wetland community types rather than species, the environmental conditions required to produce wetlands and their types appear well enough defined biologically to overcome this limitation (see Brinson 1993; Godwin et al. 2002; White and Fennessy 2005). Second, MaxEnt assumes that environmental conditions at extant wetland locations represent the fundamental niche for a particular type of wetland. It is possible that through wetland losses due to drainage for agriculture, some highly productive areas are no longer represented in present-day wetland occurrences. Given that our modeling efforts are explicitly intended to identify potential mitigation sites, and highly productive agricultural lands are unlikely to be available for purchase as mitigation sites, the underrepresentation of highly productive areas is of reduced importance. Third, an unbiased sample of wetlands is assumed, as MaxEnt has limited ability to adjust for differences in sampling effort (Elith et al. 2006). Our use of remotely-sensed wetland presence data (NWI) ensured an unbiased sample. Modeling limitations were collectively addressed by using extensive wetland presence coverages to perform validation studies.

Model Calibration

Wetland presence locations were modeled in MaxEnt version 3.3.3 using all feature types (linear, quadratic, product, threshold, and hinge), which allows modeling of complex relationships between wetland presence and environmental covariates (Phillips and Dudík 2008). We used the logistic output format, which gives the closest approximation to probability of occurrence. All other MaxEnt settings were left at their default values. To create a sample of presence locations for MaxEnt modeling, random points ≥50 m apart were generated within NWI wetland polygons. From this occurrence set, 1,000 points were randomly selected for each of the four main wetland types (emergent, deciduous forested, evergreen forested, and scrub-shrub) for use as model training sites. One hundred additional points were withheld from this training dataset for each wetland type to use for model validation. The model output produced spatially explicit occurrence probability maps for the USRB and associated goodness-of-fit statistics (area under the Receiver Operating Characteristic curve or AUC).

Probability maps created by MaxEnt for each wetland type were used to designate high probability wetland patches (hereafter patches). Commonly used metrics for determining presence thresholds for predicted occurrence probabilities (Jimenez-Valuerde and Lobo 2007), such as maximum training sensitivity plus specificity (MTSS), were not suitable because their cut-off values were too low for our datasets (average MTSS=0.165) and predicted too much area to be useful for site selection. We determined the suitable probability threshold using three metrics:

1) Average patch size—small patches provide too few ecological functions, while large areas can be prohibitively expensive to purchase. Therefore, we targeted sites of intermediate size (about 5 ha on average, the minimum size threshold for wetlands under New York State jurisdictional protection), which allows for prediction of both small and large patch sizes.
2) Percent of predicted area overlapping NWI—although the prediction of area not designated by NWI is desirable, thresholds with higher percentages of NWI area within the total predicted area indicate better predictive capabilities (Hosmer and Lemeshow 2000). This measure is similar to AUC in that it accounts for errors of omission (NWI area not predicted) and commission (predicted area that is not NWI), but differs from AUC in that it accounts for the total area predicted.
3) Geographical evenness—calculated similarly to species diversity evenness, geographical evenness is calculated as: $\left(\sum_{1}^{HUC} A\right)/Max(A)$ where HUC is the number of 8-digit Hydrologic Unit codes (Seaber et al. 1987) present in the USRB (5 HUCs), and A is the predicted area within each HUC at the given probability threshold. Evenness values range from 1 (area only in one HUC) to 5 (same amount of predicted area within each HUC). We selected geographical evenness thresholds when they predicted at least some area in all HUCs (mitigation sites needed in all HUCs) but evenness did not substantially exceed that of current NWI (evenness=2.9).

A probability threshold that optimized these metrics across all wetlands types was selected. Pixels with probabilities greater than the threshold were selected, buffered by 25 m, and joined into single polygons (i.e. "patches"). Patches for different wetland types were dissolved together to create a "MaxEnt model" for patch-based ranking and comparison with the expert model.

Patches were ranked on the basis of four criteria that could affect their value as a mitigation site. Criteria considered here include: patch area (A), designation as significant biological community (B), distance to nearest road (R), and distance to nearest large remaining wetlands (W: NWI wetlands ≥5 ha). Ranking criteria presented here were not intended to be exhaustive, nor is it implied that these are the best criteria to consider when evaluating potential wetland mitigation sites. In practice, specific resource management objectives will ultimately dictate appropriate assessment criteria and weighting. Our aim was to explore the utility of integration of patch ranking schemes with MaxEnt model outputs to identify high quality mitigation sites.

To create a comprehensive wetland resource coverage for ranking, non-overlapping portions of MaxEnt-predicted wetland patches were merged with NWI polygons, and MaxEnt-designated fen and bog occurrences were transferred to the resulting wetland polygons. The ranking criteria were divided by respective sample maxima to produce indices on a 0–1 scale. Variability in patch size required A to be log transformed, values<0 were set to "0.0001" for conformity to a 0–1 scale. Fens and bogs received a B value of 1 (all other wetlands received 0). The following formula was used to rank patch quality: $A+B+R+(1-W)/Max(A+B+R+(1-W))$; due to limited information on the relative contribution of each variable to wetland restoration site quality, all four criteria were weighted equally.

Model Validation

We created an "expert" model based on two landscape features (slope and soil type) to be evaluated alongside the MaxEnt model. Outputs were models based on informal knowledge experts typically apply to identify areas suitable for wetland mitigation: areas with low slopes (<1 %) and soils high in organic content (muck, silt loam, and loam), which largely represent designated hydric soils (NRCS 2010) for the area.

The probability threshold metrics previously described were used to evaluate the efficacy of MaxEnt versus expert models. Additionally, the ability of the models to identify wetland hydrologic conditions was assessed using aerial imagery in a GIS. Using 1-m resolution color satellite photographs taken in 2009, the presence or absence of hydrological indicators intersecting randomly selected patches for the MaxEnt (n=100) and expert models (n=100) was tabulated. Assessment categories considered for this verification were: intersection with NWI, visible standing water, or a lack of hydrological indicators. To determine whether non-NWI predicted patches were suitable wetland locations, visible standing water was only recorded when patches were not already designated as NWI. Either intersection of NWI or visible standing water was considered to indicate the presence of wetland hydrology. Using the tabulated hydrology presence-absence results from the aerial indicator survey, a Fisher's exact test was performed to assess the independence of wetland detection ability of the MaxEnt and expert modeling approaches (Sokal and Rohlf 1995). To test the efficacy of the patch-based ranking, we calculated the average ranking for all patches, and for the ten wetland areas within the USRB deemed to be biologically important by the New York Natural Heritage Program (NYNHP 2011). We then compared the average ranking of the NYNHP sites to the average ranking of all sites. All GIS work was performed in ArcMap 9.3.1 (ESRI 2009).

Results

Model Calibration

MaxEnt performed well at predicting the withheld NWI sites with AUC values ranging from 0.86 to 0.98, which indicate "excellent" (AUC=0.8–0.9) to "outstanding" (AUC>0.9) model performance (Hosmer and Lemeshow 2000). The most important environmental variables for predicting wetland occurrences were lower slopes and soils with some organic content (Table 2). Highest wetland occurrence probabilities generally occurred within the northeast portion of the USRB, where the greatest density of NWI wetlands also occurred (Fig. 2).

The MaxEnt probability threshold that optimized patch estimation across the three assessment metrics (average patch size, percent NWI, and geographical evenness) was determined to be 0.75 (Fig. 3). However, all metrics could

Table 2 Percent contribution of environmental variables to the prediction of wetland community type occurrence in the Upper Susquehanna River Basin in New York State. For each wetland type, the most important variable in determining occurrence is in bold

Wetland type	Slope	Soil type	Geology	Elevation	Aspect
Emergent	43.7	**44.2**	6.1	5.2	0.7
Forested—deciduous	41.2	**49.0**	5.3	3.5	1.0
Forested—evergreen	**52.8**	34.5	5.7	5.7	1.4
Scrub-shrub	**48.0**	39.9	7.6	4.4	0.2
Fen	6.6	**47.4**	21.7	1.3	23.0
Bog	40.6	**44.1**	10.0	4.7	0.7

Models of wetland occurrence for each type were created in the program MaxEnt

Variable units: slope (percent), soil type (categorical), geology (categorical), elevation (meters), aspect (degrees)

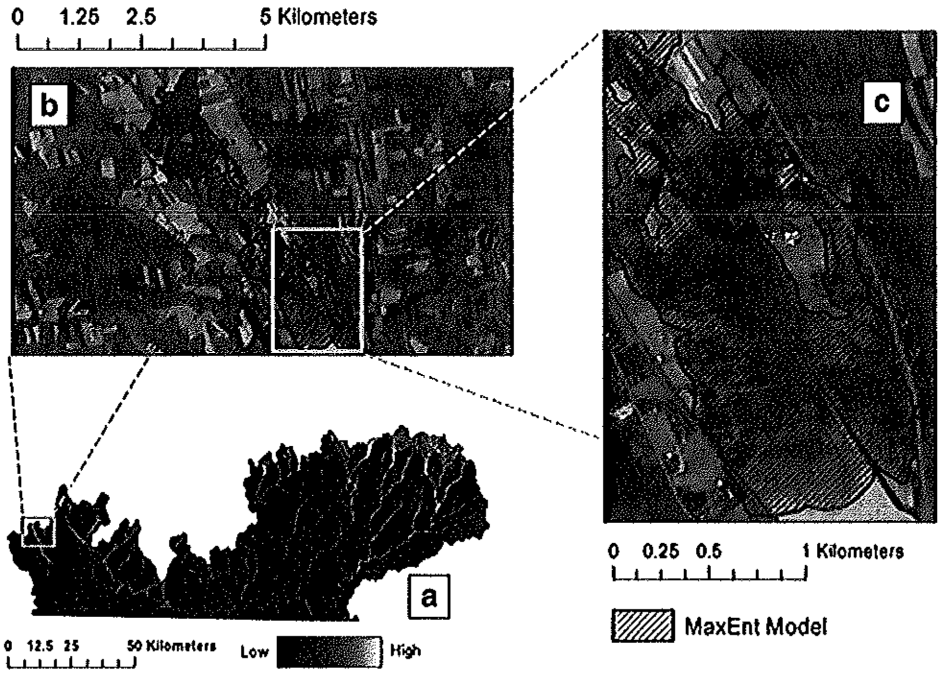


Fig. 2 Map of the MaxEnt predicted probability of occurrence for emergent wetlands in the Upper Susquehanna River Basin in New York State, based on occurrence data from the National Wetland Inventory. Lighter shades indicate a higher probability of occurrence from 0 to 1. The MaxEnt model (b, c) of dissolved polygons for all predicted wetland types with >75 % occurrence probability, where (b) depicts a 490 ha partially drained wetland complex and (c) shows a drained portion of the wetland complex that was not delineated by NWI

not be satisfied across all wetland types. Average patch size was closer to the target 5.0 ha for more selective (higher) thresholds, although very high thresholds (>0.8) produced very small average patch sizes (<2.0 ha). Percent overlap with NWI was higher for more selective thresholds. Geographical evenness was optimized for a middle range of thresholds. The most selective thresholds predicted areas only in one or two HUCs, which produced very low evenness values, and the least selective thresholds had evenness values that were much higher than the target NWI value of 2.9. Using the 0.75 probability threshold, and dissolving across all wetland types, produced a "MaxEnt model" of overall wetland predictions with 22,808 patches. The mean patch size was 4.1 ha but with large variability (± 21.3 SD, range 0.3 ha to 1,530 ha).

Model Validation

Model validation measures indicated that MaxEnt generally outperformed the expert model (Fig. 4). The two models performed equally well in average patch area (MaxEnt= 4.1 ha, expert=4.7 ha) compared to the target of 5 ha. Models were also similar in their abilities to reach the target NWI evenness (2.9), with MaxEnt being less even (2.0) and expert being more even (3.9) than desired but both within the range of acceptability. The models differed in their abilities to accurately predict current NWI area. The MaxEnt model had 76 % overlap with NWI, whereas the expert model had only 20 % overlap, indicating a higher error rate of commission for the expert model. Additionally, the MaxEnt model predicted 92 % of NWI area, whereas the expert model only predicted 62 %, indicating a higher error rate of omission for the expert model.

Aerial imagery verification of model performance indicated that the MaxEnt model was more likely to predict areas with suitable wetland hydrology than the expert model (78 % and 57 % respectively, p=0.002, Fisher's exact test, Fig. 5). Average patch rank for all patches was 0.40 (±0.14 SD), while average patch rank for heritage sites was 0.79 (±0.13 SD) (Fig. 6). Nine of the ten NYNHP sites ranked in the top ranking quartile.

Discussion

Model Performance

Using readily available wetland occurrence and landscape data, we have implemented a novel method for predicting and ranking >22,000 suitable wetland restoration sites in a region covering >16,000 km^2. Our methodology identified the physical setting of wetlands, irrespective of drainage, and performed better than commonly used expert opinion models based on typically invoked selection criteria of topographic slope and hydric soils (e.g. Van Lonkhuyzen et al. 2004). This method identified previously unknown wetlands and drained areas, identifying favorable locations of suitable wetland hydrology for potential successful wetland restoration (Biebighauser 2007).

The MaxEnt models performed well at predicting wetland occurrence, using both standard AUC measures and aerial imagery verification, indicating that the environmental variables used for prediction were sufficient for making accurate predictions (Elith et al. 2006). Soil type and topographic slope had substantially greater influence over the probability of

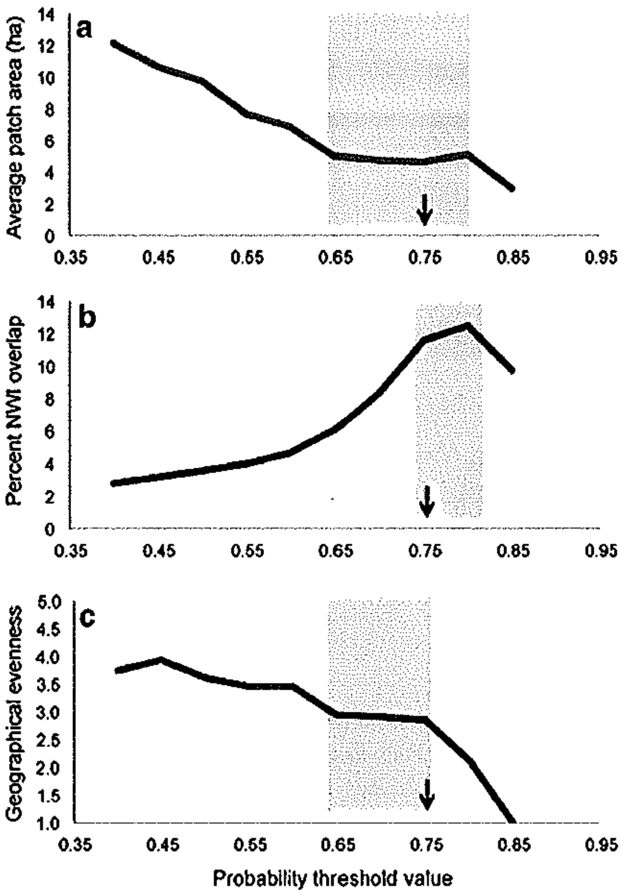


Fig. 3 Examination of the dependence of three suitability metrics on MaxEnt probability thresholds for deciduous forested wetlands in the Upper Susquehanna River Basin in New York State. Suitability metrics: (a) average patch area with a target of 5.0 ha, (b) percent of the predicted area that overlaps with National Wetland Inventory (NWI) polygons with higher percentages being more desirable, and (c) geographical evenness with better values being closer to NWI evenness (2.9). Gray boxes indicate the acceptable range for each metric, and arrows indicate the 0.75 probability threshold that best met the metric targets for all wetland types

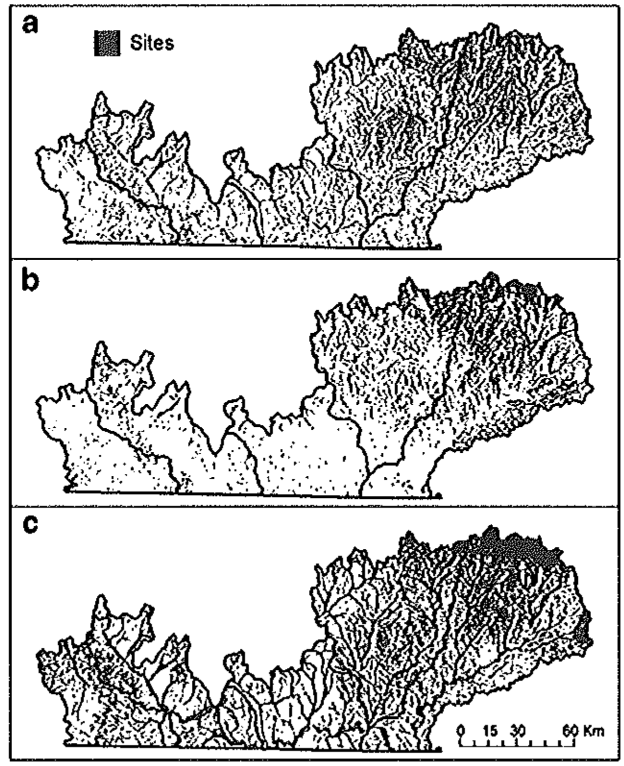


Fig. 4 Distributions (gray lines) for: (a) National Wetland Inventory polygons used as occurrence records for MaxEnt model training, (b) MaxEnt predicted polygons (all wetland types, probability threshold: 0.75), and (c) polygons predicted by the expert model. Black lines delineate the 8-digit Hydrologic Unit codes (Seaber et al. 1987) within the Upper Susquehanna River Basin in New York State

wetland occurrence than other environmental features (Table 2); this result, however, is not surprising given that wetlands typically occur at poorly drained lower slope positions where soils are periodically inundated (Keddy 2010). Although these two features played a large part in wetland prediction, it is notable that a model based *only* on slope and soil type (the expert model) did not perform as well at predicting NWI wetlands. The expert model had higher error rates of both omission and commission in predicting NWI area than the MaxEnt model. The higher error rates indicate that the expert model was not as specific to wetland areas as the MaxEnt model, and was likely predicting a large amount of area that would not be suitable for wetland establishment (Fig. 4). Comparison against the expert model provides strong evidence that the MaxEnt model, through its synthesis of interactions between multiple landscape features (Elith et al. 2006), performs better at predicting wetland occurrences than currently employed methods. Given strong predictive performance based on several indicators, wetland occurrence prediction through MaxEnt appears to have overcome the constraints typical of presence-only modeling. Because the models were trained using a densely sampled landscape (through remotely sampled NWI locations), the prediction of wetland occurrence was largely interpolative, and MaxEnt performs better at interpolation than extrapolation (Peterson et al. 2007; Elith et al. 2010). The patch-based ranking scheme also performed well, as indicated by the high ranks received by biologically important sites compared to average patch ranks.

Limitations and Advantages

Several limitations and caveats to our approach deserve mention. First, although the use of remotely sensed NWI data for wetland presence locations allowed for comprehensive coverage throughout the USRB, the NWI categorizations by wetland type may be somewhat unreliable as they may ignore wetland succession; thus, it is possible that not all wetlands were properly classified (Kudray and Gale 2000). Second, the aerial assessment of wetland patches should be interpreted with caution as it

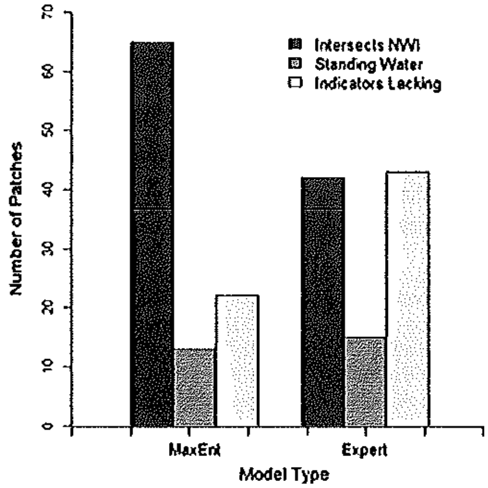
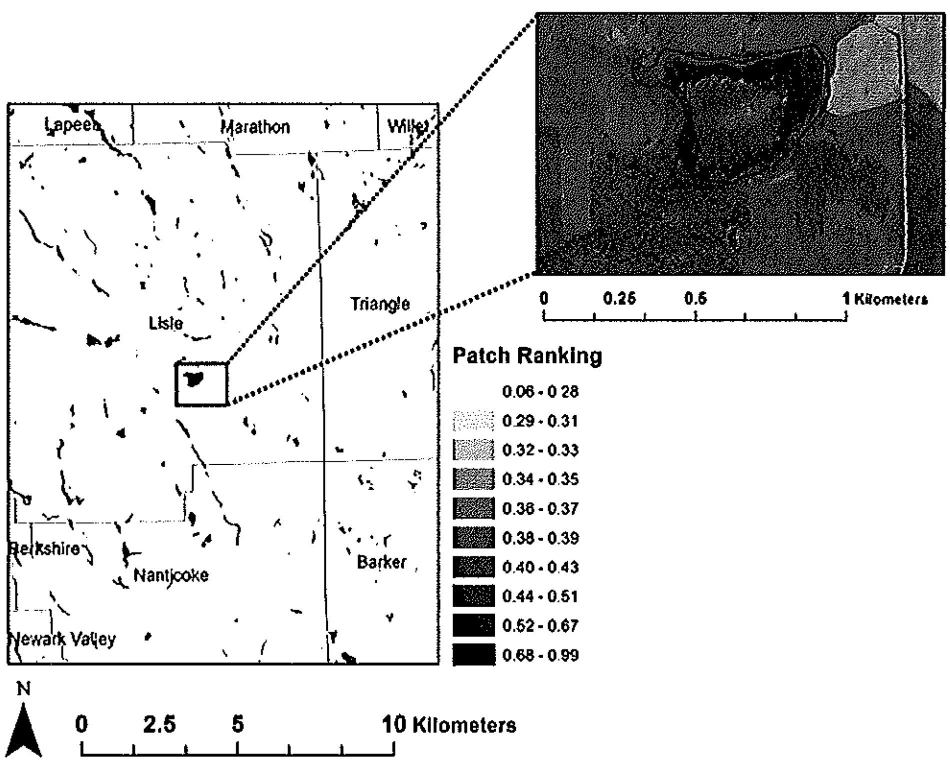


Fig. 5 Comparison of wetland detection ability for MaxEnt model and expert model patches (n=100 respectively) using 1 m resolution color aerial imagery in the Upper Susquehanna River Basin in New York State, with higher numbers of patches intersecting with National Wetland Inventory (NWI) polygons or having visible standing water being more desirable. Patches crossing the outline of NWI wetlands were recorded as "Intersects NWI". Standing water was recorded only for patches not intersecting NWI, and patches neither intersecting NWI nor displaying visible standing water were counted as "Indicators lacking", indicating poorer predictive performance

is possible that some locations that were previously drained would not produce remotely visible hydrological indicators of a former wetland (Biebighauser 2007), which would indicate that MaxEnt actually performed somewhat better than indicated by this assessment. Third, our selection of a probability threshold for delineating wetland patches was suited to the mitigation needs within the USRB. Other scenarios may require the use of different probability threshold selection criteria, thus potentially changing the outcome of the modeling procedure. Criteria for probability threshold selection should be carefully considered before beginning the modeling process, but our criteria (patch size, overlap with NWI, and geographical evenness) should be widely applicable to other wetland mitigation efforts. Last, model outputs would be strengthened by field visitation of predicted sites to verify the existence of wetland or drained wetland conditions. However, we emphasize that this modeling procedure was explicitly designed as a pre-screening tool to better focus field visitations that are already a typical component of final mitigation site selection.

Despite these limitations, our approach represents an improvement over existing methods for identifying suitable wetland mitigation sites. Other approaches have used similar data in spatially explicit models, but they rely almost entirely on subjective variable weighting (Van Lonkhuyzen et al. 2004; White and Fennessy 2005) or more complex data and calculations that may not be accessible to managers (Palmeri and Trepel 2002). Modeling wetland occurrence in MaxEnt reduces subjectivity and is also an easy modeling procedure once GIS data are compiled. Notably, the MaxEnt model was a useful tool for finding previously un-identified wetland locations through its detection of areas with visible signs of

Fig. 6 Sample ranking of potential wetland mitigation areas, predicted using the program MaxEnt, in the Upper Susquehanna River Basin in New York State (*left*), inset of a highly ranked bog (*right*). Darker shades indicate higher rankings. Townships are delineated



hydrology not previously delineated by NWI (Fig. 5). Data used in all aspects of the occurrence estimation and patch ranking are widely available, making our approach easy to implement for mitigation and conservation practitioners who need to set watershed scale (or larger) conservation and restoration priorities.

Modeling the presence of wetland communities instead of individual species within those wetland communities has a number of advantages over single species distribution modeling for identifying potentially valuable mitigation sites. First, modeling communities may allow for the generation of larger presence datasets with which to train MaxEnt, resulting in better fitting models (Ferrier and Guisan 2006). Second, modeling communities avoids problems with the stochasticity in dispersal and survival that often determine whether a species is present in a given location (Phillips et al. 2006). Third, specific ecological communities are more broadly distributed than individual species, particularly rare species that are narrowly distributed due to human disturbances (Elith et al. 2006; Ferrier and Guisan 2006). As a result, modeling communities rather than species could identify a greater number of quality protection and restoration targets, thus providing a coarse filtered approach to conservation (*sensu* Hunter et al. 1988).

Conclusions

Although frameworks exist for the protection of wetlands and mitigation of wetland losses in the United States, mitigation site failure continues to lead to biodiversity and ecosystem function loss (Race and Fonseca 1996; Moreno-Mateos et al. 2012). Improper establishment of hydrological settings has been repeatedly identified as the primary cause of wetland mitigation site failure (Mitsch and Wilson 1996). Complicating the task of establishing proper hydrological conditions is that exact locations and ecosystem types of drained wetlands (e.g., fens, marshes) are often unknown, and for rare communities only a relatively few locations remain (Nekola and Bruelheide 2004). Because of these historical losses, techniques that work well with presence-only data may be of increasing utility for predicting the occurrence of rare ecological communities (see Phillips and Dudík 2008), and subsequent protection and restoration efforts. Presence-only modeling is especially pertinent for wetland occurrence prediction because of the highly coupled relationship between wetlands and hydrogeological features in the landscape (Brinson 1993; Godwin et al. 2002). The ability to identify where wetlands likely once occurred can take much guesswork out of restoration, and our approach can be easily applied to other wetland restoration efforts at watershed scales. Ranking systems are also easily updated as additional abiotic and biotic information related to wetland quality become available (Lee et al. 2001; Bayliss et al. 2003). Although modeling and patch-based ranking are not without their limitations, integration of these tools allows for a very large number of locations to be systematically identified and screened for current status and pending threats. Conservation approaches that are able to efficiently identify and prioritize wetland mitigation sites are likely to result in improved conservation and restoration activities.

Acknowledgments This work was supported by EPA Wetland development grant number CD97225309 as a joint project between The Upper Susquehanna Coalition and the Department of Environmental and Forest Biology at the State University of New York College of Environmental Science and Forestry. The authors would like to thank Jim Curatolo and Melissa Yearick for their inputs throughout the project and the New York Natural Heritage Program for sharing spatial data. Two anonymous reviewers' comments greatly improved the manuscript.

References

Arthur DJ, Bohm B, Layne M (2008) Hydraulic fracturing considerations for natural gas wells of the Marcellus shale. Technical Report 1:1–16

Bayliss J, Helyar A, Lee JT, Thompson S (2003) A multi-criteria targeting approach to neutral grassland conservation. Journal of Environmental Management 67:145–160

Bedford BL, Godwin KS (2003) Fens of the United Sates: distribution, characteristics, and scientific connection versus legal isolation. Wetlands 23:608–629

Bedford BL, Leopold DJ, Gibbs JP (2001) Wetland ecosystems. Encyclopedia of Biodiversity 5:781–804

Benayas JR, Bullock JM, Newton AC (2008) Creating woodland islets to reconcile ecological restoration, conservation, and agricultural land use. Frontiers in Ecology and the Environment 6:329–336

Biebighauser TR (2007) Wetland drainage, restoration, and repair. University of Kentucky Press, Lexington

Brinson M (1993) Hydrogeomorphic classification for wetlands. Technical Report WRP-DE-4:1–79

Cowardin LM, Carter V, Golet FC, Laroe ET (1979) Classification of wetlands and deepwater habitats of the United States. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C

Deblauwe V, Barbier N, Couteron P, Lejeune O, Bogaert J (2008) The global biogeography of semi-arid periodic vegetation patterns. Global Ecology and Biogeography 17:715–723

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler DJ, Lévêque C, Naiman RJ, Prieur-Richard A, Soto D, Stiassny MLJ, Sullivan CA (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews 81:163–182

Edinger GJ, Evans DJ, Bebbauer S, Howard TM, Hunt DM, Olivero AM (2002) Ecological communities of New York State. A revised and expanded edition of Carol Reschke's Ecological Communities of New York State. (Draft for review). New York Natural Heritage Program, Albany, NY

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2010) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43–57
ESRI (2009) ArcMap version 9.3.1. Environmental Systems Research Institute, Redlands
Fahrig L, Pedlar JH, Pope SE, Taylor PD, Wegner JF (1995) Effect of road traffic on amphibian density. Biological Conservation 73:177–182
Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. Journal of Applied Ecology 43:393–404
Gaston KJ (2003) The structure and dynamics of geographic ranges. Oxford University Press, Oxford
Geneletti D (2004) A GIS-based decision support system to identify nature conservation priorities in an alpine valley. Land Use Policy 21:149–160
Gesch D, Oimen M, Greenlee S, Nelson C, Steuck M, Tyler D (2002) The national elevation dataset. Photogrammetric Engineering and Remote Sensing 68:5–11
Gibbs JP (2000) Wetland loss and biodiversity conservation. Conservation Biology 14:314–317
Godwin KS, Shallenberger JP, Leopold DJ, Bedford BL (2002) Linking landscape properties to local hydrogeologic gradients and plant species occurrence in minerotrophic fens of New York State, USA: a hydrogeologic setting (HGS) framework. Wetlands 22:722–737
Gotelli NJ, Ellison AM (2004) A primer of ecological statistics. Sinauer Associated Inc., Sunderland
Heikkinen RK, Luoto M, Araujo MB, Virkkala R, Thuiller W, Sykes MT (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. Progress in Physical Geography 30:751–777
Homer C, Huang C, Yang L, Wylie B, Coan M (2004) Development of a 2001 national land cover database for the United States. Photogrammetric Engineering and Remote Sensing 70:829–840
Hosmer DW, Lemeshow S (2000) Applied logistic regression. Wiley, New York
Hunter ML, Jacobson GL, Web T (1988) Paleoecology and the coarse-filter approach to maintaining biological diversity. Conservation Biology 2:375–385
IUCN (2004) 2004 IUCN Red list of threatened species. A global species assessment. IUCN, Gland, Switzerland and Cambridge, UK
Jimenez-Valuerde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either-or presence-absence. Acta Oecologica 31:361–369
Johnson AM, Leopold DJ (1994) Vascular plant species richness and rarity across a minerotrophic gradient in wetlands of St. Lawrence County, New York, USA. Biodiversity and Conservation 3:606–627
Karraker NE, Gibbs JP, Vonesh JR (2008) Impacts of road deicing salt on the demography of vernal pool-breeding amphibians. Ecological Applications 18:724–734
Kaushal SS, Groffman PM, Likens GE, Belt KT, Stack WP, Kelly VR, Band LE, Fisher GT (2005) Increased salinization of fresh water in the northeastern United States. Proceedings of the National Academy of Sciences of the United States of America 102:13517–13520
Keddy PA (2010) Wetland ecology, principles and conservation. Cambridge University Press, New York
Kershner JL (1997) Setting riparian/aquatic restoration objectives within a watershed context. Restoration Ecology 5:15–24
Kudray GM, Gale MR (2000) Evaluation of national wetland inventory maps in a heavily forested region in the upper great lakes. Wetlands 20:581–587
Lee JT, Woddy SJ, Thompson S (2001) Targeting sites for conservation: using a patch-based ranking scheme to assess conservation potential. Journal of Environmental Management 61:367–380
MacArthur RH, Wilson EO (1963) An equilibrium theory of insular zoogeography. Evolution 17:373–387
Margules CR, Pressey RL (2000) Systematic conservation planning. Nature 405:243–253
Mitsch WJ, Gosselink JG (2000) Wetlands, 3rd edn. Wiley, Hoboken
Mitsch WJ, Wilson RF (1996) Improving the success of wetland creation and restoration with know-how, time and self-design. Ecological Applications 6:77–83
Moreno-Mateos D, Power ME, Comin FA, Yockteng R (2012) Structural and functional loss in restored wetland ecosystems. PLoS Biology 10:e1001247
Nathan R, Muller-Landau HC (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. Trends in Ecology & Evolution 15:278–285
Natural Resources Conservation Service (NRCS) United States Department of Agriculture (2010) Soil Survey Geographic (SSURGO) Database for New York. Available online at http://soildatamart.nrcs.usda.gov Accessed 01/15/2011
Nekola JC, Bruelheide H (2004) Vascular plant compositional gradients within and between Iowa fens. Journal of Vegetation Science 15:771–780
Nicholson, SW, Dicken CL, Horton JD, Foose MP, Mueller JAL, Hon R (2006) Preliminary integrated geologic map databases for the United States. United States Geological Survey
NYDEC (2011) Revised draft SGEIS on the oil, gas and solution mining regulatory program (September 2011). Well permit issuance for horizontal drilling and high-volume hydraulic fracturing in the marcellus shale and other low-permeability gas reservoirs
NYNHP (2011) Rare species and community occurrences, biodiversity databases, element occurrence record digital data set. New York Natural Heritage Program, Albany
Palmeri L, Trepel M (2002) A GIS-based score system for siting and sizing of created or restored wetlands: two case studies. Water Resources Management 16:307–328
Peterson AT, Papeş M, Eaton M (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography 30:550–560
Phillips SJ, Dudík M (2008) Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. Ecography 31:161–175
Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231–259
Race MS, Fonseca MS (1996) Fixing compensatory mitigation: what will it take? Ecological Applications 6:94–101
Seaber PR, Kapinos FP, Knapp GL (1987) Hydrologic units maps. United States Geological Survey Water Supply Paper 2294:1–63
Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. Freeman, New York
Tiner RW (1997) NWI maps: what they tell us. National Wetlands Newsletter 19:7–12
USFWS (2010) National wetlands inventory. 1:25,000:1
Van Lonkhuyzen RA, LaGory KE, Kuiper JA (2004) Modeling the suitability of potential wetland mitigation sites with a geographic information system. Environmental Management 33:368–375
White D, Fennessy S (2005) Modeling the suitability of wetland restoration potential at the watershed scale. Ecological Engineering 24:359–377
Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Quantifying threats to imperiled species in the United States. BioScience 48:607–615
Young TP (2000) Restoration ecology and conservation biology. Biological Conservation 92:73–83
Zedler JB (2003) Wetlands at your service: reducing impacts of agriculture at the watershed scale. Frontiers in Ecology and the Environment 1:65–72