Generalizability of Artificial Neural Network Models in Ecological Applications:
Predicting Nest Occurrence and Breeding Success of the Red-winged Blackbird *Agelaius phoeniceus*

Uygar Özesmi, Can O. Tan, Stacy L. Özesmi*, and Raleigh J. Robertson

*Environmental Science Chair, Department of Environmental Engineering, Erciyes University, 38039 Kayseri Turkey
bDepartment of Biology, Middle East Technical University, 06531 Ankara Turkey
cDepartment of Biology, Queen’s University, Kingston, Ontario Canada

* Corresponding author. Environmental Science Chair, Department of Environmental Engineering, Erciyes University, 38039 Kayseri Turkey
Tel: 90 352 437 6748 Fax: 90 352 437 4404 email: stacy@erciyes.edu.tr

Abstract

Separate artificial neural network (ANN) models were developed from data in two geographical regions and years apart for a marsh-nesting bird, the red-winged blackbird *Agelaius phoeniceus*. Each model was independently tested on the spatially and temporally distinct data from the other region to determine how generalizable it was. The first model was developed to predict occurrence of nests in two wetlands on Lake Erie, Ohio in 1995 and 1996. The second model was developed to predict breeding success in
two marshes in Connecticut, USA in 1969 and 1970. Independent variables were vegetation durability, stem density, stem/nest height, distance to open water, distance to edge, and water depth. The nest occurrence model performance on the training data was at an average cross entropy, or concordance index (c-index), of 0.75. Within geographical region testing in two different wetlands resulted in c-indices of 0.66 and 0.53. The breeding success model performance was at a c-index of 0.75 on the training data and at c-indices of 0.47 and 0.53 for within region testing. When we tested the nest occurrence model on fledged nestling data we obtained c-indices of 0.69 and 0.47 in Clarkes Pond in 1969 and 1970 respectively, and 0.43 and 0.52 in All Saints Marsh in 1969 and 1970 respectively. When we tested the fledged nestling model on the nest occurrence data, we obtained c-indices of 0.70 and 0.41 in Stubble Patch in 1995 and 1996 respectively, and 0.54 and 0.55 for Darr in 1995 and 1996 respectively. With input variable relevances, sensitivity analyses and neural interpretation diagrams we were able to understand how the different models predicted nest occurrence and breeding success and compare their differences and similarities. Important variables for predicting nest site selection/breeding success in both models were vegetation durability and distance to open water. Both models also predicted increasing nest occurrence/breeding success with increasing water depth under the nest and increasing distance to edge. However, relationships for prediction differed in the models. Generalizability of the models was poor except when the marshes had similar values of important variables in the model, for example water depth. ANN models performed better than generalized linear models (GLM) on marshes with similar structures. Generalizability of the models did not differ in nest occurrence and breeding success data. Extensive testing also showed that the
GLMs were not necessarily more generalizable than ANNs. The results from this study suggest that ANN models make good definitions of a study system but are too specific to generalize well to other ecologically complex systems unless input variable distributions are very similar.

*Keywords*: Habitat models; General linear model; Spatial habitat selection; Habitat preference; Freshwater marsh; Wetlands; Model validation.

1. **Introduction**

Understanding bird habitat selection and predicting presence, absence or breeding success of a species in particular locations under certain environmental conditions in time and space is one of the key issues for wildlife management and conservation. The difficulty arises as most of the ecological processes and relationships, if not all, exhibit nonlinearity and are very complex in nature.

However, artificial neural networks (ANNs) have been shown to be a successful tool for interpretation and prediction under these conditions. The aim of this study is to test the generalizability of spatial habitat models of red-winged blackbirds (*Agelaius phoeniceus*) and to observe the commonalities between models to find out generalizable processes. Finally we examine whether nest occurrence models could predict breeding success and vice versa.
ANN models have been applied to many diverse ecological problems from phytoplankton production (Scardi, 1996, 2001) to community changes based on climatic inputs (Tan and Smeins, 1996) to animal damage to crops (Tourenq et al., 1999; Spitz and Lek, 1999). ANNs have been used to study the relationship between species and habitat variables for birds (Fielding, 1999b; Manel et al., 1999), cyanobacteria (Maier et al., 1998), fish (Baran et al., 1996; Reyjol et al., 2001), and macroinvertebrates (Hoang et al., 2001). However one goal for the use of ANNs in ecology, which has been elusive, has been the development of generalizable models that can be applied to different systems in different geographical regions. This is currently an active area of research. For example, Wilson and Recknagel (2001) developed generic ANN lake models to predict algal abundances in freshwater lakes. Scardi (2001) explained the use of a metamodeling procedure to generalize neural network models to other geographic regions when a large amount of data is not available for those regions.

In order to assess whether we have achieved our goal of generalizability it is necessary to have an independent test of the model (Fielding, 1999a). A model that has not been tested is only a definition of a system. It becomes a scientific pursuit, a hypothesis, when one starts testing it. An independent test turns a model into a hypothesis that can then be validated (Ayer, 1936). Hirzel et al. (2001) argued that comparing models of habitat suitability on one unique case is not enough. Therefore they created a virtual species and different scenarios to test their models. We agree that more than one independent test of a model is needed. For this reason, in this study we tested two ANN spatial habitat models for the red-winged blackbird, *Agelaius phoeniceus*, on data from
marshes in different geographical areas and years apart. This gave four completely
independent cases with which to test the generalizability of each model. We also
examine the input variable relevances, neural interpretation diagrams, and sensitivity
analyses of both models to understand how they are making predictions. Finally we
make multimodel inferences to suggest how to create a more general model.

One of our models was created to predict breeding success based on habitat
variables. The other model was originally developed to predict nest occurrence based on
habitat variables. When developing the nest occurrence model it was assumed that the
red-winged blackbirds selected nest locations in places with the highest habitat suitability
or quality and that by selecting this optimal habitat they would have high breeding
success. However other limiting factors such as intraspecific competition, interspecific
competition, and dispersal capability also influence habitat selection (Cody, 1981;
Burger, 1985; Cody, 1985). In addition some research has questioned the relationship of
habitat variables to breeding success. For example, in a study of savannah sparrow
Passerculus sandwichensis, Bedard and Lapointe (1983) found that reproductive success
depended on predators and catastrophic events such as flooding. Van Horne (1983)
stressed that density must be coupled with demographic data or some erroneous
conclusions on what constitutes optimal habitat could result. Nevertheless, many studies
on red-winged blackbirds have found a relationship between reproductive success and
habitat variables such as water depth, vegetation type and durability, nest height, and nest
location (i.e. Meanley and Webb, 1963; Goddard and Board, 1967; Holcomb and Twiest,
1968; Francis, 1971; Robertson, 1972; Holm, 1973; Caccamise, 1977; Weatherhead and
Robertson, 1977). Therefore we assumed in this research that when our model predicts a
high probability of nest occurrence, this corresponds to a location with high breeding success and vice versa.

2. Methods

2.1 Study areas

The study areas consisted of four marshes, two located in Sandusky Bay on Lake Erie, Ohio, USA and two located in Connecticut, in the northeastern USA. The data were from the years 1995 and 1996 and from the years 1969 and 1970 respectively.

2.1.1 Lake Erie marshes

The Lake Erie marshes were Stubble Patch and Darr. Data from these marshes, which included habitat variables and nest occurrence, was collected in 1995 and 1996 (Özesmi, 1996). These two marshes lie in the southwestern end of Lake Erie on Sandusky Bay and are part of Winous Point Wetlands, which consists of many wetland units divided by dikes. Stubble Patch was a 16.7 ha wetland that had been managed as a shallow emergent marsh since 1990. In 1995 it was dominated by *Typha angustifolia*, *Phalaris arundinacea*, and *Hibiscus palustris*. In 1996 *Hibiscus palustris* declined and *Sparganium eurycarpum* dominated large areas of the marsh as did *Phalaris arundinacea* and *Typha angustifolia*. The average water level for both years was about 40 cm in the vegetated areas. In 1995 the area with vegetation cover was 4.5 ha but in 1996 vegetation
cover increased to 7.3 ha. Open water covered more than half of the marsh in both years (Özesmi, 1996).

Darr was a 25.8 ha wetland dominated by *Typha angustifolia* and *Hibiscus palustris*. About one-half of that area was open water. The average water level in 1995 was 30.6 cm in the vegetated areas. However in 1996 Darr was managed as a shallow marsh and the average water level was 15.3 cm in the vegetated areas. From 1995 to 1996 the area of vegetation cover increased from 12.3 ha to 13.4 ha (Özesmi, 1996).

In 1995 a total of 85 red-winged blackbird nests were found in Darr and in 1996 42 red-winged blackbird nests were found. Darr had 5 marsh wren nests in 1995 but 28 marsh wren nests in 1996. Competition with marsh wren dramatically affected red-winged blackbird habitat selection as the marsh wren occupied the most central and deepest areas of the marsh (Özesmi and Özesmi, 1999). In Stubble Patch 30 and 49 red-winged blackbird nests were found in 1995 and 1996 respectively. Stubble Patch did not have any marsh wren nests. Thus the Lake Erie wetlands were much larger in size with a much lower density of red-winged blackbird nests than the Connecticut wetlands (Table 1).

In addition to the nest occurrence data, there was some information on breeding success available for Darr and Stubble Patch in 1996, although the nests were not followed until all the eggs were hatched and fledged. If at the time of the last nest check, the nests had eggs, or fledglings, they were scored as successful. Nests that were empty, showed signs of predation, or were knocked down were scored as unsuccessful.

### 2.1.2 Connecticut marshes
The two marshes in Connecticut were Clarkes Pond and All Saints Marsh. The data from these marshes, which included habitat information and breeding success, was collected in the years 1969 and 1970 (Robertson, 1972). Clarkes Pond was a constructed impoundment 4.65 ha in size on the Mill River in Hamden, Connecticut. The dominant emergent vegetation, which covered 1.92 ha of the marsh, was *Typha latifolia* and *Typha angustifolia*. Average depth of the water underneath nests in the emergent vegetation was about 42 cm. The remaining area, 2.73 ha, was open water with some water lily (*Nymphea*), pickerelweed (*Pontederia*) and arrowhead (*Sagittaria*). The surroundings of Clarkes Pond consisted of large areas of mixed deciduous woodlands, a pine plantation, a pasture, and a mowed field. Clarkes Pond had 202 and 167 red-winged blackbird nests in 1969 and 1970 respectively (Robertson, 1972).

All Saints Marsh was located northeast of New Haven, Connecticut about 15 km away from Clarkes Pond. Its area was 1.09 ha and it was covered with emergent vegetation consisting of *Typha latifolia* interspersed with small patches of open water. Dense areas of buttonbush (*Cephalanthus occidentalis*) as well as other bushes were scattered throughout the marsh, which were often used for nesting. The marsh seemed to be spring fed with an average water depth of 30-40 cm during the nesting season. All Saints Marsh was surrounded by extensive areas of mixed deciduous woodland on two sides and weedy fields in early stages of old field succession on two sides. All Saints Marsh had 108 red-winged blackbird nests in 1969 and 128 in 1970 (Robertson, 1972).

2.2 Artificial neural network analysis
The ANNs used were the feed-forward multi-layer perceptron with back-propagation for training, which were created using NevProp3 software (Goodman, 1996).

2.2.1 Lake Erie ANN nest occurrence model

The nest occurrence model was trained using 1995 data from the Lake Erie wetlands Stubble Patch and Darr. The independent variables were vegetation durability based on an ordinate scale between 0 and 100 (Özesmi and Mitsch, 1997), stem density (number of stems/m$^2$), stem height (cm above water), distance to open water (m), distance to edge (m), and water depth (cm). All inputs were standardized to a mean of zero with the units in standard deviations. The dependent variable was a binary index of nest occurrence.

For hidden units, we used a symmetric logistic function ranging from –0.5 to 0.5 and for output layer an asymmetric logistic ranging from 0 to 1. The output was a probability of nest occurrence between 0 and 1. The score threshold was set to 0.5, so that probability predictions of the network lower than 0.5 were classified as no nest (0) and above 0.5 as nest (1). The training algorithm used was gradient descent. The initial values of the random weights were randomly set between ±0.1. We used a learning rate of 0.01. We trained the model using Darr 1995 data. We used early stopping with Stubble Patch 1995 data as a holdout data set to determine when to stop training the ANN. We tried different configurations of architecture for the model but chose a single hidden layer with six hidden units for the final model because greater numbers of hidden
units and hidden layers did not give better results on the training data. The final model was tested on Stubble Patch and Darr data from 1996. More detailed information on how the nest occurrence model was developed can be found in Özesmi and Özesmi (1999).

2.2.2 Clarkes Pond ANN breeding success model

In the breeding success model the independent variables were vegetation durability based on an ordinate scale between 0 and 100 (Özesmi and Mitsch, 1997), nest height (cm), distance to open water (m), distance to edge (m) and water depth (cm). We standardized all inputs to a mean of zero with the units in standard deviation. A binary index of whether or not any nestlings fledged was the dependent variable.

For hidden units, we used a symmetric logistic function ranging from −0.5 to 0.5 and for output layer an asymmetric logistic ranging from 0 to 1. Thus the output was a probability of fledglings between 0 and 1. The score threshold was set to 0.5, so that probability predictions of the network lower than 0.5 were classified as not fledged (0) and above 0.5 as fledged (1).

2.2.2.1 Training the model

When we varied the training algorithm for the networks we found that the corrected c-index was consistently higher for QuickProp than with pure gradient descent. Therefore we used the Quickprop algorithm, which is also faster than pure gradient descent. Quickprop increases the learn rate for each weight’s dimension if the global error decreases and assumes that the error surface curvature is locally
parabolic (quadratic), so that each change in the gradient and weight uniquely
determines the bottom of the current parabola (Goodman, 1996). Quickprop tries to
move towards that minimum.

We developed the model using data from one marsh, Clarkes Pond in 1969-
1970. In total 294 data points were available to train the model. We used
bootstrapping (150 with 66% of data) of the training data to determine the maximum
number of epochs to train the model. We stopped the training when corrected average
cross entropy (c-index) levelled off. C-index, also known as concordance index, is the
probability that the model will give a higher output probability to a case with an
actual output of 1 versus a case with an actual output of 0. If the c-index is 1, the
model predicts perfectly and if the c-index is 0.5 the model does not perform better
than a random model (Goodman, 1996). C-index also corresponds to the approximate
area under the Receiver Operating Characteristic (ROC) curve. In a ROC curve the
proportion of true-positive predictions (sensitivity) are plotted against the proportion
of false-positive predictions (1 minus specificity) for various score thresholds
(Fielding, 1999a).

We optimized the network parameters of learn rate, weight range, and
momentum. Several networks with different combinations of learn rate (ranging
between 0.1 and 0.001) and weight range (between 0.1 and 0.001) were run. After
optimizing the network architecture (see below), we again optimized the network
parameters. We chose 0.01 as the starting learn rate and 0.001 as the range in which
the initial random weights were chosen (± weight range).

2.2.2.2 ANN architecture
Although there are heuristic algorithms to determine the network architecture, we systematically explored architectures by running several different models: 1) a general linear model (GLM), (no hidden layer), 2) transformation only model (each input connected to only one hidden unit), 3) one hidden layer having different numbers of hidden units ranging between 2 and 40, 4) two hidden layer networks having 2, 3, 4 and 5 hidden neurons in each layer. All networks were run 5 times using the same predetermined random seeds, produced by a random number generator, to see the variation in the model performance due to different initial weights. All the models with one hidden layer performed better than the general linear model and transformation only model. The models with two hidden layers did not perform better than the models with one hidden layer, thus one hidden layer was chosen for the final network configuration. The accuracy of the one hidden layer networks first increased with increasing numbers of hidden units and then levelled off around 20 hidden units.

For good generalizability, it is necessary to have about 10 times as many training data points as there are weights in the network (Bishop, 1995). In our case, we had about 300 training input – output cases, so we should have approximately 30 connections in our network. We had 5 input variables and one output variable so this would mean about 4 hidden units arranged in the form of a single layer. Nevertheless, we obtained the best c-index with a network having 20 hidden units (c-index = 0.79), which gave 120 connections. For that reason, we tried to optimize the network with 4 hidden units (in terms of network parameter settings), and compare the model accuracies and sensitivity analyses both with the best network having 20 hidden units (120 connections) and the
best network having 4 hidden units (24 connections). Finally, we chose the network with 4 hidden units (c-index = 0.75) for the following reasons. First, given models with similar errors on the training sets, the simpler model is more likely to predict better on new cases, the dictum of Occam's razor. More formally a criteria such as Akaike Information criterion (AIC), which consists of an error term with a penalty for model complexity, could be used to choose a model (Burnham and Anderson, 2002; Bishop, 1995). This also would have resulted in choosing the model with 4 hidden units. Because increasing the number of estimated parameters (weights) in a model usually always increases its predictive ability, it is necessary to account for this increased model complexity when deciding which model to choose. Thus current literature is recommending the use of criteria to select a model when many models are created with the same dataset (Burnham and Anderson, 2002; McNally, 2000).

Second, we ran the two models with 4 and 20 hidden units with the same ten random starts to determine the variation in the c-index and corrected c-index (based on 150 bootstraps). For the model with 4 hidden units, the c-indices did not vary as much as with the 20 hidden units model (Table 2). We thought the larger variation in c-indices of the 20 hidden unit model might be an indication of overfit on the training data. Third, according to the sensitivity analyses the network with 4 hidden units seemed to capture the relationship between inputs and the output better; that is, it made more sense based on what is known about red-winged blackbird habitat selection.

2.2.2.3 Model test

The final model was tested using the data from All Saints Marsh from 1969 (n=101) and from 1970 (n=130). We compared the performance of the ANN with a
GLM, an ANN with no hidden units, which is basically a logistic regression model, on the training data from Clarkes Pond and on the All Saints Marsh test data.

2.2.3 Comparison of models

In this study the breeding success model, which had been trained on Clarkes Pond 1969-1970 data, was tested on the Lake Erie wetlands data from Darr and Stubble Patch in 1995 and 1996. The Lake Erie nest occurrence model was tested on the Connecticut wetlands data from Clarkes Pond and All Saints in 1969 and 1970. Because one model was developed to predict breeding success and the other nest occurrence, for this research we assumed that a high probability of nest occurrence corresponds to a high probability of breeding success and vice versa (for more discussion on this assumption see the introduction). In addition, since the Connecticut wetland variables did not include stem density, the average value of stem density from the Lake Erie wetlands was used when testing Connecticut wetlands data on the Lake Erie model. Note also that nest height was available for the Connecticut wetlands but that stem height was available for the Lake Erie wetlands. Stem heights were about 50 cm higher than nest heights on average. The assumption was made that stem height and nest height were correlated. Thus when the models were tested, nest heights were input instead of stem heights in the nest occurrence model and vice versa for the breeding success model.

In addition to the nest occurrence data from Darr and Stubble Patch, the information on breeding success from Darr and Stubble Patch in 1996 together with the habitat variables of vegetation durability, nest height, distance to open water, distance to
edge, and water depth under the nest was used as test data for the Clarkes Pond breeding success model.

Input variable relevances and neural interpretation diagrams (NIDs) (Özesmi and Özesmi, 1999), as well as sensitivity analyses (Lek et al., 1996; Scardi, 1996; Recknagel et al., 1997) were used to understand the models’ predictions. The input variable relevance is the sum square of weights for one input variable divided by the sum square of weights for all input variables and thus it shows the contribution of a variable to the model. In a NID, the weights of the connections between units are represented by pixel weights of lines scaled to the relative values of the weights in the ANN. Black lines represent positive signals and gray lines represent negative signals. By looking at the weights and signs of the connections in a NID it is possible to see the importance of variables in the model and the interactions between variables. Sensitivity analyses are done by varying the values of one input variable at a time while holding the other input variables at constant values (we used average values) and plotting the values against probability of breeding success.

3. Results and discussion

3.1 Within geographic region model results

3.1.1 Lake Erie ANN and GLM nest occurrence model
The Lake Erie ANN nest occurrence model performance on the training data, which was Stubble Patch and Darr data for 1995, was at a c-index of 0.75 (Table 3). When the model was tested on the 1996 data for Stubble Patch and Darr the resulting c-indices were 0.66 and 0.53 respectively. C-indices of the Lake Erie GLM model on the test data were 0.67 and 0.57 for Stubble Patch and Darr respectively. The Lake Erie ANN nest occurrence model did not perform well on Darr in 1996 because of the presence of marsh wren. More details on the Lake Erie models as well as a marsh wren model and a marsh wren and red-winged blackbird interaction model can be found in Özesmi and Özesmi (1999).

3.1.2 Clarkes Pond ANN and GLM breeding success model

The Clarkes Pond ANN model of breeding success, which was defined as at least one nestling fledged, had a c-index of 0.75 for the training data of Clarkes Pond in 1969 and 1970 (Table 3). The corrected c-index, calculated using 150 bootstraps, levelled off at 0.66 after 70 epochs. This corrected c-index should reflect how generalizable the model is. The Clarkes Pond breeding success ANN model was tested independently on All Saints Marsh. All Saints Marsh was quite different from Clarkes Pond (Table 1). Clarkes Pond was an open water pond surrounded by *Typha* vegetation where the red-winged blackbirds nested. All Saints Marsh was a shallow emergent marsh dominated by *Typha* with scattered shrubs throughout. All Saints Marsh had a lower water depth on average than Clarkes Pond. In addition, All Saints Marsh was about
one-fourth of the size of Clarkes Pond. All Saints Marsh did not have any large areas
of open water in either 1969 or 1970. Therefore the distance to open water variable
for All Saints Marsh was input in two different ways: 1) as the greatest distance in
Clarkes Pond (112 m) and 2) as the mean distance to open water in the Clarkes Pond
data. These two different ways to input distance to open water in All Saints Marsh
gave similar results. A separate model developed without the distance to open water
variable as input did not give better results as well.

In addition to these structural differences, the density of nests was greater in
All Saints Marsh than Clarkes Pond, 108.3 versus 96.1 nests/ha per year. The
percentage of successful nests was also greater in All Saints Marsh for both years,
64.4% versus 47.6% average for both years (Robertson, 1972).

When this model was tested on All Saints Marsh for 1969 and 1970, the
within region testing resulted in c-indices of 0.47 and 0.53 respectively. The Clarkes
Pond GLM model had a c-index of 0.63 on the training data and c-indices of 0.53 and
0.53 for All Saints Marsh in 1969 and 1970 respectively. Thus both the artificial
neural network model and general linear model both did not perform better than a
random model.

One reason for the poor model performance is probably the structural
differences of the marshes (different vegetation types, different vegetation/open water
ratios, different water depths, different sizes). In addition, in these marshes predation
was a significant factor in breeding success. In particular, Clarkes Pond was under
very high predation pressure with 45% and 47.3% of the nests predated in 1969 and
1970 respectively (Robertson, 1972). All Saints Marsh had 11.1% and 16.4 % of the
nests predated in 1969 and 1970 respectively. Racoon (*Procyon lotor*) was responsible for the most nest predations, accounting for 54 of 80 nests predated in the nestling phase in these marshes. Other nest predators were likely birds or water snakes. Long-billed marsh wrens (*Cistothorus palustris*) were present in Clarkes Pond in 1969 but there was no evidence that these birds destroyed red-winged blackbird nests.

Predation seemed to dominate over other environmental factors (the habitat input variables) for determining breeding success. However, environmental variables affect predation. It is more difficult for a predator to get to a nest in deeper water, in the center of a marsh, and surrounded by open water. For example, predation decreased with increasing water depth under the nest. Although predators are influenced by environmental variables included in the model, complete reliance on environmental variables to account for predation is not sufficient and the inclusion of predators into the model can improve performance (Özesmi and Özesmi, 1999).

### 3.2 Model generalizability test results

#### 3.2.1 Lake Erie nest occurrence ANN and GLM models

The Lake Erie ANN nest occurrence model tested on the Connecticut breeding success data resulted in c-indices of 0.69 and 0.47 for Clarkes Pond in 1969 and 1970 respectively, and 0.43 and 0.52 for All Saints Marsh in 1969 and 1970 respectively (Table 3). Thus the Lake Erie ANN nest occurrence model performed well on Clarkes
Pond 1969 data but no better than a random model on the other data. Clarkes Pond 1969 data were the most similar to the training data for the Lake Erie model, showing closer distances to open water and deeper water for higher breeding success (Figures 1a-b). These variables had high relevances in the Lake Erie ANN model.

The Lake Erie GLM model tested on Clarkes Pond resulted in c-indices of 0.66 and 0.48 for 1969 and 1970 respectively. All Saints Marsh in 1969 and 1970 had c-indices of 0.38 and 0.48 respectively. The GLM model performed very poorly on all marshes and years except Clarkes Pond in 1969. This model relied heavily on distance to open water (Table 4). Similar to the training data, Darr and Stubble Patch in 1995, which had higher nest occurrence close to open water, Clarkes Pond in 1969 had higher breeding success for closer distances to open water (Figure 1a). However, Clarkes Pond in 1970 showed the opposite and All Saints Marsh did not have large areas of open water in either year.

3.2.2 Clarkes Pond breeding success ANN and GLM models

The Clarkes Pond ANN breeding success model tested on the Lake Erie nest occurrence data gave c-indices of 0.70 and 0.41 for Stubble Patch in 1995 and 1996 respectively (Table 3). For Darr the c-indices were 0.54 and 0.55 in 1995 and 1996 respectively. The Clarkes Pond ANN breeding success model performed well on Stubble Patch 1995 data but no better than a random model on the other data. One reason for the good performance on Stubble Patch in 1995 may have been because in this year Stubble Patch had the most open water and thus would have been most similar to Clarkes Pond.
(Table 1) in terms of vegetation/open water ratio. The mixture of vegetation cover and open water is known to be an important factor in red-winged blackbird nest site selection (Orians, 1980).

The Clarkes Pond ANN breeding success model and the Lake Erie nest occurrence ANN model are predicting differently. The Lake Erie model predicts high nesting probability in areas near to open water and in the center of the marsh while the Clarkes Pond model probabilities seem to be based on water depth together with other variables.

When the Clarkes Pond ANN breeding success model was tested on Darr and Stubble Patch breeding success data from 1996 c-indices of 0.64 and 0.52 respectively were obtained (Table 3). The Clarkes Pond GLM breeding success model tested on the Lake Erie nest occurrence data resulted in c-indices of 0.60 and 0.60 on Darr in 1995 and 1996, and 0.62 and 0.71 on Stubble Patch in 1995 and 1996 (Table 3). Thus the Clarkes Pond GLM breeding success model performed better than the ANN model, except for on Stubble Patch in 1995. This GLM model relied heavily on water depth to predict breeding success (Table 4). The water depth of Clarkes Pond was most similar to Stubble Patch, especially in 1996 (Figure 1b).

The Clarkes Pond GLM breeding success model tested on Darr and Stubble Patch 1996 breeding success data gave c-indices of 0.57 and 0.57 respectively. Next we made informal multimodel inferences using input variable relevances, sensitivity analyses, and neural interpretation diagrams.
Input variable relevances for both models are shown in Table 4. The variables with the highest relevances for the Lake Erie ANN nest occurrence model were distance to open water, vegetation durability, stem density and distance to edge (Özesmi and Özesmi, 1999). For the Clarkes Pond breeding success model, the highest input variable relevances were nest height, water depth, distance to open water, and vegetation durability. Both models are in agreement that two of the four most important variables for red-winged blackbird nesting and fledging probability are distance to open water and vegetation durability. The importance of vegetation structure for red-winged blackbirds has previously been noted (Goddard and Board, 1967; Holm, 1973; Burger, 1985). Murkin et al. (1989) noted that red-winged blackbirds seemed to prefer a certain distance to open water for nest placement.

According to the average values for distance to open water, and also water depth, the Lake Erie model data (Stubble Patch and Darr 1995) were most similar to Clarkes Pond 1969 data (Figures 1a-b). Distance to open water had a high relevance in the Lake Erie model. Another reason the nest occurrence model may have performed better on Clarkes Pond 1969 data is that marsh wrens were present in that year but not in 1970 or in All Saints Marsh (Table 1). Five marsh wrens were also present in Darr in 1995. Darr and Stubble Patch 1995 data were used to train the Lake Erie ANN nest occurrence model.

According to mean values for water depth, Clarkes Pond data from 1969 and 1970, which were used to develop the Clarkes Pond breeding success model, were closest
to the Stubble Patch 1995 data. As water depth was one of the two most relevant
variables in the Clarkes Pond breeding success, this may be one reason why the Clarkes
Pond breeding success model performed the best on Stubble Patch 1995 data.

3.4 Sensitivity analysis

For the Lake Erie nest occurrence model, nesting probability increased with
increasing vegetation durability, increasing water depth, and increasing distance to edge
(Özesmi and Özesmi, 1999). Nesting probability increased with increasing stem height
to about 130 cm and then decreased. Nesting probability increased with increasing
stem density to about 160 stems/m² and then decreased. Nesting probability decreased
with increasing distance to open water.

For the Clarkes Pond breeding success model, sensitivity analyses indicated that
breeding success increased with increasing water depth and increasing distance to edge
(Figure 2a-e). Breeding success generally decreased with increasing nest height and
increasing distance to open water. Breeding success was high with lower vegetation
durability and high vegetation durability while vegetation with mid-range durability had
low breeding success.

Comparing the sensitivity analyses curves, both the Lake Erie nest occurrence
model and the Clarkes Pond breeding success model predict increased nesting and
breeding success respectively with increasing water depth and increasing distance to
edge. Weatherhead and Robertson (1977), with their weighted linear nest score that
significantly correlated with breeding success for a marsh in Ontario, Canada, gave
higher scores to nests with deeper water underneath. Brown and Goertz (1978),
Goddard and Board (1967) and Robertson (1972) also report increased breeding
success with greater water depth under the nest.

Differences in the models are that the Lake Erie nest occurrence model predicts
increased nesting with increasing vegetation durability while the Clarkes Pond breeding
success model predicts high breeding success in low durability vegetation and high
durability vegetation with a decrease in breeding success in middle range of vegetation
durability. The preference of red-winged blackbirds for vegetation of high durability
has been noted before (Robertson, 1972; Albers, 1978; Berstein and McLean, 1980;
Özesmi and Mitsch; 1997). These studies are in agreement with the Lake Erie model.
However, Weatherhead and Robertson (1977) scored less durable grasses higher than
*Typha* but *Typha* higher than sedges for an Ontario, Canada marsh. This is more in
agreement with the Clarkes Pond model.

Another difference is that the Lake Erie nest occurrence model predicts an
optimum stem height while the Clarkes Pond breeding success model predicts
increasing breeding success with decreasing nest height. Some studies showed that
nesting success increased with increasing nest height (Meanley and Webb, 1963;
Holcomb and Twiest, 1968) and some with decreasing nest height (Goddard and Board,
1967). The nest score of Weatherhead and Robertson (1977) gave a higher score to
nests that were lower, which is in agreement with the Clarkes Pond model. Brown and
Goertz (1978) found that breeding success was the highest in nests 1.2-2.4 m high with
lower and higher nests less successful. Caccamise (1977) showed increasing breeding
success with nests up to 1.6-1.79 m high and then decreasing success. These two
studies are more in agreement with the Lake Erie model. The relationship between nest height and breeding success is not a simple one but may also be related to vegetation type or durability (Francis, 1971) and other environmental factors, such as depth of water under the nest. Brown and Goertz (1978) noted that red-winged blackbird nests tended to be lower over open water. In addition, the difference between the Lake Erie model and Clarkes Pond model may also be caused by measurement of different variables, stem height versus nest height. Stowers et al. (1968) reported that red-winged blackbird average nest heights were different in different vegetation types.

The third difference is that the Lake Erie model predicts decreasing nest occurrence with increasing distance to open water. The relationship between breeding success and distance to open water is more complicated in the Clarkes Pond breeding success model but this model also predicts decreasing breeding success when the distance to open water increases beyond about 45 m.

Stem density was not an input variable in the Clarkes Pond breeding success model. The Lake Erie model showed an optimal stem density at about the midrange of the maximum stem density found in these wetlands. Many authors have stated that red-winged blackbirds prefer dense vegetation (i.e. Albers, 1978; Short, 1985). However, Holm (1973) found that nests in less dense vegetation fledged more young per nest and Weatherhead and Robertson (1977) found that more open nests were more successful than those in dense cover. Since sensitivity analysis cannot provide direct interpretation of these complex interactions it was necessary to examine the Neural Interpretation Diagrams of the models.
3.5 Neural Interpretation Diagrams (NIDs)

The NID of the Clarkes Pond breeding success ANN model shows the interaction of variables with each other (Figure 3). Breeding success is higher in nests if the vegetation is not durable but the nest height is high and the water depth is deep (unit 6). Fledging probability decreases with high distance to open water, high nest height and low water depth (7). Breeding success increases if the vegetation is durable and the distance to open water is high (8). If the vegetation durability, nest height, distance to open water, and water depth are all high, fledging probability decreases unless the distance to edge is high (9).

The neural interpretation diagrams (NIDs) of the two models are similar in that nesting/breeding success probability always decreases with increasing distance from open water, unless water depth is high (Özesmi and Özesmi, 1999). Both models also show that increasing vegetation durability increases nest occurrence/breeding success. However in the Clarkes Pond model, fledging probability also increases with low vegetation durability if the nest is high up and in deep water. Other rules in the models are also different. For example, in the Lake Erie model, nesting probability increases with increasing stem density and height unless vegetation durability, distance to open water and distance to edge increase. In the Clarkes Pond model, fledging probability decreases if nest height, vegetation durability, distance to open water, and water depth are all high. In the Lake Erie model, as distance to open water decreases nesting probability increases but only in areas where stem density and height are enough to support a nest.
3.6 Model generalizability

These two models agree on some of the important variables and have some of the same relationships to predict nest occurrence and breeding success. Both models agree on the importance of vegetation durability and distance to open water for nest occurrence/breeding success. The two models also both predict increased nest occurrence/breeding success with increasing depth of water under the nest and increasing distance to edge. However, the models are also using different relationships to predict nest occurrence and breeding success. Therefore the tests of generalizability usually gave poor results. These results could indicate that either the Lake Erie ANN and GLM nest occurrence model and the Clarkes Pond breeding success model are not generalizable to very different geographic locations and/or marshes of different size and structure, or that nest occurrence is a poor predictor of breeding success.

Model generalizability was poor except when the marshes had similar values of high relevance input variables, for example water depth. The artificial neural network models developed were not generalizable to marshes with different sizes and structures. Red-winged blackbirds are highly flexible in their habitat requirements and occupy a wide range of sites from uplands to marshes over a broad geographic area (Orians, 1980). For such an adaptable species, it may be difficult to generalize to different marshes and different geographical regions. However this research suggests that to create a generalizable model, which can be applied in many different marshes in different
locations, we would use the important habitat variables and relationships that are in
common for both of these models. Therefore wide testing provides wider understanding.

In our research we found that the Clarkes Pond breeding success model predicts
nest occurrence as well as the Lake Erie nest occurrence model predicts breeding
success (Table 3). However, previous research has indicated that nest occurrence is not
always a good predictor of breeding success. Robertson (1972) concluded that habitat
selection by red-winged blackbirds is probably a combination of site tenacity of adults
breeding for the second or greater time and selection for the “optimal” breeding habitat
by first time breeders. Thus nests could be located in both optimal and suboptimal
habitats. In addition, Weatherhead and Robertson (1977) found that breeding success
was correlated positively with habitat quality and negatively with female breeding
density within territories. Thus females are nesting in territories where their chance of
reproductive success is lower than it might have been in another territory. They
postulated that females were cueing in not only on habitat variables but were choosing
good mates and thus enhancing their ultimate fitness rather than their immediate
reproductive success. However Holm (1973) asserted that females were only choosing
good territories and that was why nesting density was high within some territories.
Similarly using the Lake Erie data set we found out that incorporating clustering of
nests into a spatial autocorrelation regression model and a spatial data mining tool
"Predicting Locations Using Map Similarity (PLUMS)" increased prediction accuracy
significantly (Chawla et. al., 2001).

4. Conclusion
The ANN models that were tested were very specific to certain conditions of the marshes they were trained on. However we can expect them to perform well when applied to wetlands with similar values of high relevance input variables, given that there are no other confounding effects, such as the presence of interspecific competitors.

When developing an ANN model, conditions to which the model is applicable should be specified exactly. Then the model should be tested only on independent data sets that meet those conditions. However, if we want an ANN model that is the most generalizable we need to develop our model on the “average marsh” that has a small number of special cases. The GLM models developed relied heavily on one input variable and performed better on other wetlands if that variable in those wetlands was also similarly related to nest occurrence/breeding success. However the extensive testing showed that the GLMs were not necessarily more generalizable than the ANN models.

The ANN models developed are not well suited for predicting and generalizing to other marshes and locations. However the ANNs are better than the GLMs for defining a system and interpreting the factors that govern nest occurrence/breeding success in a certain location. In addition, the ANN models developed in different marshes and different locations agree on some of the habitat variables and relationships for predicting nest occurrence/breeding success. Therefore we suggest a more generalizable model may be created with these variables and relationships. The process of modelling complex evolutionary and ecological systems humbled us as modellers reminding us that all encompassing relatively simple ecological models, such as ANNs, might not be possible. Yet the modelling
process is revealing, allowing us to understand these systems and their peculiarities better. The problem of generalizability in ANN modelling continues to be trapped within the tension of the local and universal in ecology.

Acknowledgements

We would like to thank Inci Togan for bringing us together to work on modelling and her constant encouragement. Significant computer time was allocated for our research by Erciyes University, Mechanical Engineering Department.

References

Albers, P.H. 1978. Habitat selection by breeding red-winged blackbirds. Wilson Bull. 90(4), 619-634.

Ayer, A.J. 1936. Language, Truth and Logic. (1990 ed. Penguin Books Ltd., London) Original Publisher Victor Gollancz.

Baran, P., Lek, S., Delacoste, M., Belaud, A., 1996. Stochastic models that predict trout population density or biomass on a mesohabitat scale. Hydrobiol. 337, 1–9.

Bedard, J., LaPointe, G. 1984. The savannah sparrow territorial system: Can habitat features be related to breeding success? Can. J. Zool. 62, 1819-1828.

Bernstein, N. P., McLean E. B. 1980. Nesting of red-winged blackbirds in cattails and common reed grass in Mentor Marsh. Ohio J. of Sci. 80,15-19.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Oxford, p 482.
Brown, B.T., Goertz, J.W., 1978. Reproduction and nest site selection by red-winged blackbirds in North Louisiana. Wilson Bull. 90 (2), 261-270.

Burger, J. 1985. Habitat selection in temperate marsh-nesting birds. In: Cody, M.L., (Ed.), Habitat Selection in Birds, Academic Press, San Diego, CA, pp 253-281.

Burnham, K. P., Anderson, D. R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, Springer, New York, NY, 488 pp.

Caccamise, D. F. 1977. Breeding success and nest site characteristics of red-winged blackbird. Wilson Bull. 89 (3), 396-403.

Chawla, S., Shekahar, S., Wu, W., Özesmi, U. 2001. Modelling spatial dependencies for mining geospatial data. In: Miller, H.J., Han, J. (Eds.), Geographic Data Mining and Knowledge Discovery, Taylor and Francis, London, pp. 131-159.

Cody, M. L. 1981. Habitat selection in birds: The roles of vegetation structure, competitors, and productivity. Bioscience 31 (2), 107-113.

Cody, M. L. 1985. An introduction to habitat selection in birds. In: Cody, M.L., (Ed.), Habitat Selection in Birds, Academic Press, San Diego, CA, pp. 4-59.

Fielding, A.H., 1999a. How should accuracy be measured? In: Fielding AH, editor. Machine Learning Methods for Ecological Applications. Kluwer, Dordrecht, p 209-223.

Fielding, A.H., 1999b. An introduction to machine learning methods. In: Fielding AH, editor. Machine Learning Methods for Ecological Applications. Kluwer, Dordrecht, p 1-35.

Francis, W. J. 1971. An evaluation of reported reproductive success in red-winged blackbirds. Wilson Bull. 83 (2):178-185.
Goddard, S.V., Board, V.V., 1967. Reproductive success of red-winged blackbirds in North Central Oklahoma. Wilson Bull. 79, 283–289.

Goodman, P.H., 1996. NevProp software, version 3. University of Nevada, Reno, NV. http://brain.unr.edu/index.php

Hirzel, A.H., Helfer, V., Metral, F., 2001. Assessing habitat-suitability models with a virtual species. Ecol. Model. 145, 111-121.

Hoang, H., Recknagel, F., Marshall, J., Choy, S., 2001. Predictive modelling of macroinvertebrate assemblages for stream habitat assessments in Queensland (Australia). Ecol. Model. 195, 195–206.

Holcomb, L. C., Twiest, G. 1968. Ecological factors affecting nest building in Red-winged Blackbirds. Bird-Banding 39, 14-22.

Holm, C. H. 1973. Breeding sex ratios, territoriality, and reproductive success in the red-winged blackbird (Agelaius phoeniceus). Ecology 54 (2), 356-365.

Lek, S., Belaud, A., Baran, P., Dimopoulos, I., Delacoste, M., 1996. Role of some environmental variables in trout abundance models using neural networks. Aquat. Living Resour. 9, 23–29.

Maier, H.R., Dandy, G.C., Burch, M.D., 1998. Use of artificial neural networks for modelling cyanobacteria Anabaena spp. in the River Murray, South Australia. Ecol. Model. 105 (2-3), 257–272.

Manel, S., Dias, J.-M., Ormerod, S.J., 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. Ecol. Model. 120, 337–347.
MacNally, R., 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between – and reconciliation of – ‘predictive’ and ‘explanatory’ models. Biodiversity and Conservation 9, 655–671.

Meanley, B., Webb, J. S. 1963. Nesting ecology and reproductive rate of the red-winged blackbird in tidal marshes of the upper Chesapeake Bay region. Chesapeake Sci. 4 (2), 90-100.

Murkin, H. R., Kaminski, R. M., Titman, R. D. 1989. Responses by nesting red-winged blackbirds to manipulated cattail habitat. In: Sharitz, R.R., Gibbons, J.W., (Eds), Freshwater Wetlands and Wildlife, USDOE Office of Scientific and Technical Information, Oak Ridge, Tennessee.

Orians, G. H. 1980. Some adaptations of marsh-nesting blackbirds. Princeton Univ. Press, Princeton, NJ, 295 pp.

Özesmi, U., 1996. A spatial habitat model for the marsh-breeding red-winged blackbird (Agelaius phoeniceus) in coastal Lake Erie wetlands. M.S. Thesis, The Ohio State University, Columbus, Ohio, 106 pp.

Özesmi, U., Mitsch, W.J., 1997. A spatial habitat model for the marsh-breeding red-winged blackbird (Agelaius phoeniceus) in coastal Lake Erie wetlands. Ecol. Model. 101, 139–152.

Özesmi, S.L., Özesmi, U., 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. Ecol. Model. 116, 15–31.

Recknagel, F., French, M., Harkonen, P., Yabunaka, K., 1997. Artificial neural network approach for modelling and prediction of algal blooms. Ecol. Model. 96, 11–28.
Reyjol, Y., Lim, P., Belaud, A., Lek, S., 2001. Modelling of microhabitat used by fish in natural and regulated flows in the river Garonne (France). Ecol. Model. 146, 131–142.

Robertson, R.J., 1972. Optimal niche space of the red-winged blackbird (*Agelaius phoeniceus*), (I) Nesting success in marsh and upland habitat. Can. J. Zool. 50, 247–263.

Scardi, M., 1996. Artificial neural networks as empirical models for estimating phytoplankton production. Mar. Ecol. Prog. Ser. 139, 289–299.

Scardi, M., 2001. Advances in neural network of phytoplankton primary production. Ecol. Model. 146, 33–45.

Scardi, M., Harding, L.W.J., 1999. Developing an empirical model of phytoplankton primary production: a neural network case study. Ecol. Model. 120, 213–223.

Short, H. L., 1985. Habitat Suitability Index Models: Red-winged Blackbird. U. S. Fish and Wildl. Serv. Biol. Rep. 82 (10.95)

Spitz, F., Lek, S., 1999. Environmental impact prediction using neural network modelling. An example in wildlife damage. Journal of Applied Ecology 36, 317-326.

Stowers, J.F., Harke, D.T., Stickley, A.R., Jr., 1968. Vegetation used for nesting by the red-winged blackbird in Florida. Wilson Bull. 80 (3), 320-324.

Tan, S.S., Smeins, F.E., 1996. Predicting grassland community changes with an artificial neural network model. Ecol. Model. 84, 91–97.

Tourenq, C., Aulagnier, S., Mesleard, F., Durieux, L., Johnson, A., Gonzalez, G., Lek, S., 1999. Use of artificial neural networks for predicting rice crop damage by greater flamingos in the Camargue, France. Ecol. Model. 120, 349–358.
Van Horne, B., 1983. Density as a misleading indicator of habitat quality. J. of Wildlife Man. 47, 813-901.

Weatherhead, P.J., Robertson, R.J., 1977. Harem size, territory quality, and reproductive success in the red-winged blackbird (*Agelaius phoeniceus*). Can. J. Zool. 55, 1261–1267.

Wilson, H., Recknagel, F., 2001. Towards a generic artificial neural network model for dynamic predictions of algal abundance in freshwater lakes. Ecol. Model. 146, 69–84.
Table 1. Comparison of the marshes by size, vegetation, water depth, number of red-winged blackbird (RWB) nests and number of marsh wrens nests.

| Marsh Year Location | Total area (ha) | Vegetated area (ha) | Dominant vegetation | Avg water depth (cm) | RWB Nests | Marsh wrens Nests |
|---------------------|----------------|---------------------|---------------------|----------------------|-----------|------------------|
| Darr 1995 Lake Erie | 25.8           | 12.3                | *Typha, Hibiscus*   | 31                   | 85        | 5                |
| Darr 1996 Lake Erie | 25.8           | 13.4                | *Typha, Hibiscus*   | 15                   | 42        | 28               |
| Stubble Patch 1995 Lake Erie | 16.7        | 4.5                | *Typha, Phalaris, Hibiscus Sparganium* | 40 | 30 | 0 |
| Stubble Patch 1996 Lake Erie | 16.7        | 7.3                | *Typha, Phalaris*   | 40                   | 49        | 0                |
| All Saints 1969 Connecticut | 1.1          | 1.1                | *Typha, bushes*     | 30-40                | 202       | 0                |
| All Saints 1970 Connecticut | 1.1          | 1.1                | *Typha, bushes*     | 30-40                | 167       | 0                |
| Clarkes Pond 1969 Connecticut | 4.7          | 1.9                | *Typha*             | 42                   | 108       | some             |
| Clarkes Pond 1970 Connecticut | 4.7          | 1.9                | *Typha*             | 42                   | 128       | 0                |
Table 2. Variation in model accuracy caused by using 10 different random seeds.

|                | 4 hidden units | 20 hidden units |
|----------------|----------------|-----------------|
|                | c-index        | Corrected c-index | c-index  | Corrected c-index |
| Mean           | 0.77           | 0.70             | 0.82     | 0.74               |
| Std Dev        | 0.01           | 0.01             | 0.03     | 0.03               |
| Maximum        | 0.79           | 0.71             | 0.87     | 0.79               |
| Minimum        | 0.76           | 0.68             | 0.78     | 0.69               |
Table 3. Model performances (c-index) on training and independent test data sets.

Underlined values are model performances on training data.

|                          | Clarkes Pond GLM Breed. success Model | Clarkes Pond ANN Breed. success Model | Lake Erie GLM Nest occur. Model | Lake Erie ANN Nest occur. Model |
|--------------------------|--------------------------------------|---------------------------------------|-------------------------------|-------------------------------|
| Clarkes Pond 1969        | 0.63                                 | 0.75                                  | 0.66                          | 0.69                          |
| Clarkes Pond 1970        | 0.63                                 | 0.75                                  | 0.48                          | 0.47                          |
| All Saints Marsh 1969    | 0.53                                 | 0.47                                  | 0.38                          | 0.43                          |
| All Saints Marsh 1970    | 0.53                                 | 0.53                                  | 0.48                          | 0.52                          |
| Darr 1995                | 0.60                                 | 0.54                                  | 0.71                          | 0.75                          |
| Darr 1996                | Nest occurrence                      |                                       | 0.57                          | 0.53                          |
|                          | Breeding success                      |                                       | 0.57                          | 0.64                          |
| Stubble Patch 1995       | 0.62                                 | 0.70                                  | 0.71                          | 0.75                          |
| Stubble Patch 1996       | Nest occurrence                      |                                       | 0.67                          | 0.67                          |
|                          | Breeding Success                      |                                       | 0.57                          | 0.52                          |
| Average on test data     | 0.59                                 | 0.55                                  | 0.54                          | 0.55                          |
| Min on test data         | 0.53                                 | 0.41                                  | 0.38                          | 0.43                          |
| Max on test data         | 0.71                                 | 0.70                                  | 0.67                          | 0.69                          |
Table 4. Relevances of input variables for the Clarkes Pond breeding success and the Lake Erie nest occurrence GLM and ANN models.

| Independent Input Variable | Clarkes Pond GLM Model | Clarkes Pond ANN Model | Lake Erie GLM Model | Lake Erie ANN Model |
|----------------------------|------------------------|------------------------|---------------------|---------------------|
| Vegetation Durability      | 5.8%                   | 11.2%                  | 6.4%                | 19.2%               |
| Nest Height/Stem Height    | 9.9%                   | 48.2%                  | 0.9%                | 4.6%                |
| Distance to Open Water     | 0.2%                   | 13.9%                  | 88.2%               | 45.1%               |
| Distance to Edge           | 6.5%                   | 2.1%                   | 2.9%                | 11.0%               |
| Water Depth                | 77.6%                  | 24.7%                  | 1.0%                | 6.5%                |
Figure 1a-b
Figure 2a-e
Figure 3
Figure Captions

Figure 1a-b. Mean and standard errors of distance to open water (m) and water depth (cm) for the different marshes. Shaded bars indicate nests/fledged, white bars indicate no nests/not fledged. In figure (a) distance to open water at All Saints Marsh was set to a very high number since there was no open water during 1969-70.

Figure 2a-e. Sensitivity analyses for the Clarkes Pond 1969-70 ANN breeding success model.

Figures 3. Neural Interpretation Diagram (NID) for Clarkes Pond 1969-70 ANN breeding success model. The thickness of each connection is proportional to its relevance. Black lines represent positive relationships and gray lines represent negative relationships.