Geographical Analysis ISSN 0016-7363

Neurofuzzy Modeling of Context–Contingent Proximity Relations

Xiaobai Yao,[1] Jean-Claude Thill[2]

[1]Department of Geography, University of Georgia, Athens, GA, [2]Department of Geography, University at Buffalo, The State University of New York, Amherst, NY

The notion of proximity is one of the foundational elements in humans' understanding and reasoning of the geographical environments. The perception and cognition of distances plays a significant role in many daily human activities. Yet, few studies have thus far provided context–contingent translation mechanisms between linguistic proximity descriptors (e.g., "near," "far") and metric distance measures. One problem with previous fuzzy logic proximity modeling studies is that they presume the form of the fuzzy membership functions of proximity relations. Another problem is that previous studies have fundamental weaknesses in considering context factors in proximity models. We argue that statistical approaches are ill suited to proximity modeling because of the inherently fuzzy nature of the relations between linguistic and metric distance measures. In this study, we propose a neurofuzzy system approach to solve this problem. The approach allows for the dynamic construction of context–contingent proximity models based on sample data. An empirical case study with human subject survey data is carried out to test the validity of the approach and to compare it with the previous statistical approach. Interpretation and prediction accuracy of the empirical study are discussed.

Introduction

The perception and cognition of distances play a significant role in many daily human activities. Its most overt impact is probably on cognitive tasks of navigation within a certain physical environment, where distance perception mediates people's orientation and positioning (Montello 1997). The notion of proximity, which is understood here to refer specifically to natural-language expressions (e.g., "near," "far") of people's psychological feelings of distances, also features very prominently in people's daily life. The last decade has witnessed tremendous technological innovations in transportation and communication, which have challenged

Correspondence: Xiaobai Yao, Department of Geography, University of Georgia, Athens, GA 30605
e-mail: xyao@uga.edu

Submitted: September 10, 2004. Revised version accepted: November 17, 2005.

us to reexamine critically the relationship between proximity expressions forged by perceptual filters and metric distances that objectively exist in the physical space surrounding us. In this information age, people are indeed empowered with greater abilities to overcome distance barriers. Modern communication and information technologies such as intelligent transportation systems, location-based computing and services (LBS), telematics, and other novel means of wireless communications may already have come together to reshape people's perception and cognition of distances.

A significant body of literature has accumulated so far on proximity relations, with perspectives from geography, cognitive science, linguistics, and others. While prior studies on proximity modeling range from the formalization of proximity relations in space (e.g., Frank 1992; Worboys 2001) to new measures of proximity (e.g., Bera and Claramunt 2003), the dynamical modeling of the relationship between proximity and metric distances is still rather deficient. The specification of the meanings of qualitative proximity measures is seen as a major challenge for the future of space-aware information systems, such as geographical information systems (GIS) and location-based information systems, which are grounded in the precept of metric information. To enable geospatial information systems to interpret people's qualitative distance measures, further work needs to be carried out to provide a translation mechanism between qualitative distance measures and the metric distance measure that is contingent upon the various contexts of proximity perception and sensitive to interpersonal perceptual differences. Proximity modeling of qualitative locations (Yao and Thill 2005b) is critical to a wide range of new applications of space-aware information technologies based on ubiquitous computing, interoperability, and real-time information (Goodchild 2000; Thill 2000; Karimi and Hammad 2004). For example, a translation mechanism between qualitative and quantitative measures is a key component of a human-centered interface that takes qualitative inputs of spatial relations. This is becoming increasingly relevant as GIS-based systems are seen almost ubiquitously, albeit in various disguises such as LBS, in-vehicle navigation systems, Web-based spatial search applications, and so forth. Other examples of the use of such a mapping mechanism include fuzzy spatial footprints to retrieve information from a digital library, automated navigation services, and the prompt interpretation of a linguistic description of a location by emergency response services. In addition, most users are not trained in either GIS or geography (Egenhofer and Mark 1995) and yet they do have commonsense knowledge of the spatial features and spatial relations in the environments (Kuipers 2004). Therefore, the capabilities of GIS to interpret linguistic spatial relation inputs are very important for the general public. In this article, we present a neurofuzzy modeling framework that supports the mapping between qualitative and metric distance measures while preserving the integrity of distinguishing properties of proximity relations.

The article proceeds as follows: the next section reviews prior studies of proximity modeling. The following section introduces the theoretical structure of the neurofuzzy approach in general and introduces its specifications in context–contingent proximity modeling. This section further discusses the merits of this approach in proximity modeling. In the penultimate section, we conduct a case study using the neurofuzzy approach, interpret the modeling results, and compare them with results of the ordered logit regression (OLR) model presented in Yao and Thill (2005a). Conclusions and future research needs are discussed in the final section.

Prior studies

Two properties of proximity relations are of utmost importance to the modeling of the correspondence between qualitative and quantitative distance measures. One characteristic is the dependence of the formation of proximity relations upon context. The other is the uncertain nature of proximity measures and therefore of the relationships between qualitative and metric distances.

It has been suggested that, while reasoning about proximity, human beings consider the context of the perceptual and cognitive task in hand. Discussion and analysis by Sharma, Flewelling, and Egenhofer (1994), Gahegan (1995), Hernandez, Clementini, and Felice (1995), and Worboys (1996, 2001) have highlighted the importance of *context* in qualitative spatial reasoning in general and in proximity modeling in particular. As Sharma, Flewelling, and Egenhofer (1994) pointed out, previous qualitative spatial reasoning models separated the metric properties of the data from the determination of qualitative magnitudes and events, which may be assessed differently depending on the *context*. This implies that a model that considers the relationship between the context and the metric properties of data is necessary. Context factors may vary among different types of distance perception experiences. For example, the context factors relevant in a distance of an international trip and that of a grocery shopping are quite different. Yao and Thill (2005a) classified context factors of proximity perception into objective and subjective factors. Subjective factors are those variables whose values change from person to person. Examples include the perceiver's levels of familiarity with the area and the route, the perceiver's time and financial budget, the perceiver's demographic characteristics, and socioeconomical characteristics. Objective factors are those context variables whose status is independent of the person who perceives the distance. Examples include scale of the study area, type of activity, reachability, intervening opportunity, transportation mode, traffic condition, and others.

Because a qualitative distance measure has inherent fuzziness (Robinson 1990), most prior studies have used fuzzy logic to model its relationship with metric distances (e.g., Zadeh 1965; Kaufmann 1975; Robinson 1990; Gahegan 1995; Worboys 2001; Wang 2003). Some of the recent and notable contributions are now briefly discussed. Robinson (1990, 2000) extends research by Freeman (1973), Robinson, Blaze, and Thoongs (1986), and Robinson, Frank, and Karimi (1987) with an automated and interactive method for acquisition and algebraic definition of fuzzy spatial relations using distance as the base variable. The concept of nearness is learned through a question-answering scheme. A multiperson concept is then constructed using various principles, including the agreement, global evidence, combined agreement and global evidence, and Zimmerman's method. Worboys (2001) captures the degree of nearness with a fuzzy-logic model of proximity spatial relations based on experimental data. The nearness neighborhood of an arbitrary place is defined by a simple linear membership function calibrated on majority responses in the experimental data. Neither Worboys' (2001)approach nor Robinson's (1990, 2000) attempts to generalize the mapping mechanism of qualitative to metric distances translation nor do they account for spatial context, however. Guesgen (2002a, b) and Guesgen and Albrecht (2000) also represent proximity relations (e.g., "close to") with explicit fuzzy sets. The shortcomings of the approach in these studies are discussed here. That an enumerative definition of fuzzy sets is used instead of an analytical definition inevitably renders computation time unacceptably large in most real-world applications and prevents generalization. In addition, the robustness of the approach is in doubt because no mechanism is advanced by the authors to estimate the membership functions and thus allow for context contingency.

In sum, what is missing in the previous fuzzy logic approaches for proximity modeling studies is a systematic way to find the membership function that captures another important characteristic of proximity relations, namely the nature of context contingency. In addition, most previous fuzzy approaches to proximity modeling presume the form of fuzzy membership functions. We maintain that one should not a priori assume the form of the fuzzy membership functions between qualitative proximity measures and the corresponding metric measures due to the sensitivity of the outcome of the translation process to model specification. We further argue that the membership function of the fuzzy relation is a function of the two distance measures and the context variables of the proximity measures.

Yao and Thill (2005a) proposed the statistical approach of OLR to construct and estimate a context–contingent proximity model. While this model was successful at predicting proximity given the corresponding metric distance and context variables, it has several drawbacks that limit its usefulness. The first limitation stems from the unproven assumption on the functional specification of the translation mechanism, and from the difficulty to account for uncertainty in this mechanism. Furthermore, ideally, a proximity model should allow for two-way mapping between qualitative and quantitative distance statements. In one direction, natural language distance is predicted given the metric distance and context variables; in the reverse direction, metric distances are predicted from natural language distances and context variables. The statistical approach cannot accomplish the second type of prediction. In this article, we propose a neurofuzzy proximity model and prove that this approach provides enormous potential to overcome the limitations of proximity modeling approaches advanced so far in the literature.

Neurofuzzy inference for proximity modeling

Fuzzy logic and artificial neural networks (ANN) are two prominent techniques in contemporary geographical information science (GIScience). Both techniques have been widely applied in GIScience research and applications. Examples of the use of ANN include knowledge discovery and data mining (e.g., Fischer and Abrahart 2000; Miller and Han 2001), remote-sensing data processing (e.g., Fiset et al. 1998), simulation (e.g., Lloyd 1994; Li and Yeh 2002), and spatial analysis and problem solving (e.g., Kathmann 1993; Thill and Mozolin 2000). Fuzzy set theory has been applied to various areas of GIScience including spatial analysis (e.g., Davis and Keller 1997; Ferrier and Wadge 1997), modeling spatial relations (Worboys 2001; Guesgen 2002a, b), spatial cognition and knowledge representation (Thill and Sui 1993; Brown, Groves, and Gedeon 2003), and others. More recently, it has been suggested that fuzzy set theory and the principles of neurocomputing can advantageously be brought together for geospatial modeling. A notable example is given by Gopal, Liu, and Woodcock (2001), who followed this line of thought and developed a fuzzy ARTMAP neural network for information processing and visualization. Fuzzyneural networks have been proved to be powerful supervised classifiers (e.g., Gopal, Liu, and Woodcock 2001; Odhiambo et al. 2004). This study proposes a neurofuzzy approach that integrates the principles and operations of fuzzy set theory with ANN to map mathematically natural-language proximity descriptors onto metric distance measures, and vice versa.

Fuzzy relationship between natural language and metric distances

Fuzzy sets and fuzzy logic were first introduced in 1965 by Zadeh as a means to model the ambiguity of natural language. Since then, a large body of literature has blossomed around fuzzy sets and fuzzy logic in an incredibly wide range of areas (Nguyen and Sugeno 1998; Fischer 2000). As a superset of conventional (Boolean) logic, fuzzy logic handles the concept of partial truth—values between ''completely true'' and ''completely false'' or between 0 and 1.

When it is applied to proximity representation, the traditional crisp set theory has to give an exclusive demarcation of the metric distance continuum into the linguistic distance measures. The boundary distance values setting apart linguistic statements are used like magic numbers, assuming a much more critical role than they have in reality. The crisp set theory model cannot capture the gradual change of influence of metric distances along the distance continuum on people's inclination to form linguistic distance terms. Therefore, the crisp set theory is not appropriate for proximity modeling.

Owing to the fuzzy nature of a statement on proximity expressed in natural language (linguistic distance measure), we expect to map each linguistic distance measurement to multiple metric distances. A linguistic distance measure, say, *near*, implies different trip distances for different people. It also means different trip distances by the same person in different scenarios. For example, when one is looking for a *nearby* grocery store, both a store 1 mile away and another store 1.1 miles away may be viewed as *nearby*. When metric distance increases, the likelihood of the distance being phrased as ''near'' decreases and that of being phrased as ''far'' increases. They all increase or decrease continuously. In other words, a person's inclination to phrase a distance as a certain linguistic distance measure (such as *near*) changes gradually along the metric distance continuum. Because we have a discrete number of linguistic expressions of proximity (near/nearby/close, far, etc.), the issue boils down to identifying a mapping of continuous measures to discrete measures. This situation is well suited for fuzzy relationship modeling.

Let $R(L, M)$ denote the relation set of linguistic distance measures and metric distance measures, where L is the linguistic distance measure, and M is the metric distance measure. If D_L and D_M are the domains of the linguistic and metric distance measurements, respectively, and μ_R is the domain of fuzzy membership grades of fuzzy relation R, then R can be defined as follows:

$$\mu_R : D_L \times D_M \rightarrow U, \; U = [0, \; 1] \tag{1}$$

The fuzzy membership grade of linguistic distance measure L_i being paired with metric distance measure M_j, symbolized as μ_{ij}, is a function of L_i, M_j, and of the context variables. Equation (2) shows a general form of the membership function

$$\mu_{ij} = f(L_i, M_j, C_1, \ldots, C_n) \tag{2}$$

where C_1 through C_n are context variables.

The question remains to determine the membership function f. Previous proximity modeling efforts based on fuzzy logic made a priori assumptions on the form and parametric specification of the function. It is our contention that this is unnecessarily restrictive and reduces the external validity of the resulting mappings. In the following section, we will present the neurofuzzy approach to estimate the form and parameters of the function from sample data.

Neurofuzzy proximity modeling

Neurofuzzy systems

A neurofuzzy system is a hybrid system that integrates fuzzy logic and neural networks. To identify a suitable fuzzy system for a given problem, membership functions and a rule base typically consisting of IF–THEN fuzzy rules must be specified. This can be done by prior knowledge (to predefine a membership function), by learning from sample data, or by both of these approaches (Nauck and Kruse 1998). When the option of learning from sample data is applied and the learning capability of neural networks is used, the approach is usually called a neurofuzzy system (Pedrycz, Kandel, and Zhang 1998). Neurofuzzy systems represent a particular class of the broad family of fuzzy neural networks. For an extensive coverage of the different configurations of possible fuzzy neural networks, the reader is referred to Ishibuchi and Nii (1998).

Neurofuzzy systems are much more powerful than either neural networks or fuzzy logic systems alone as they can incorporate the advantages of both. It is often emphasized in the literature that fuzzy sets are focused on knowledge representation but cannot accommodate efficacies implied by the underlying data (Pedrycz, Kandel, and Zhang 1998). On the other hand, neural networks are powerful at learning from underlying data but poor in representing existing knowledge. Table 1 summarizes the strengths and weaknesses of neural networks and fuzzy logic systems in various respects. The table shows that the two technologies complement each other and are ideally geared to handling the duality of knowledge learning.

In sum, a major difference between fuzzy and neurofuzzy systems is that fuzzy systems acquire knowledge of parameters (of fuzzy membership functions and fuzzy rules) mostly from expert knowledge while neurofuzzy systems obtain it from training data (Pedrycz, Kandel, and Zhang 1998). We will use neurofuzzy inference to acquire implicitly the proximity fuzzy membership functions through existing sample cases (known as training cases).

Neurofuzzy inference for proximity modeling

An example of model structure consistent with the neurofuzzy inference approach proposed for proximity modeling is presented in Fig. 1. The structure is very similar to that of a typical neural network, except that the inputs are fuzzified and the output is defuzzified depending on the nature of the output variable. The linkages between the fuzzified layers pass through the fuzzy rules layer. The weights associated with the links between nodes are trained using the fundamental principles of neural networks. The input layer consists of input variables, which include the context variables and one type of the proximity measures (e.g., metric distance); the

Table 1 Comparison of Neural Network and Fuzzy Logic Systems (Adapted from Pedrycz, Kandel, and Zhang 1998, p. 312)

Feature	Neural network	Fuzzy logic systems
High-level knowledge	Implicit representation by weights	Explicit representation by fuzzy rules
Estimator	Trainable dynamical systems	Structured numerical systems
Knowledge acquisition	From sample data	From expert
Uncertain information	Quantitative	Quantitative and qualitative
Uncertain cognition	Perception	Decision making
Reasoning mechanism	Parallel computations	Heuristic search
Reasoning speed	High	Low
Fault-tolerance	Very high	Low
Adaptive learning	Adjusting weights	Induction
Knowledge storage	In neurons and links	In fuzzy rule base
Natural language	Implicit	Explicit

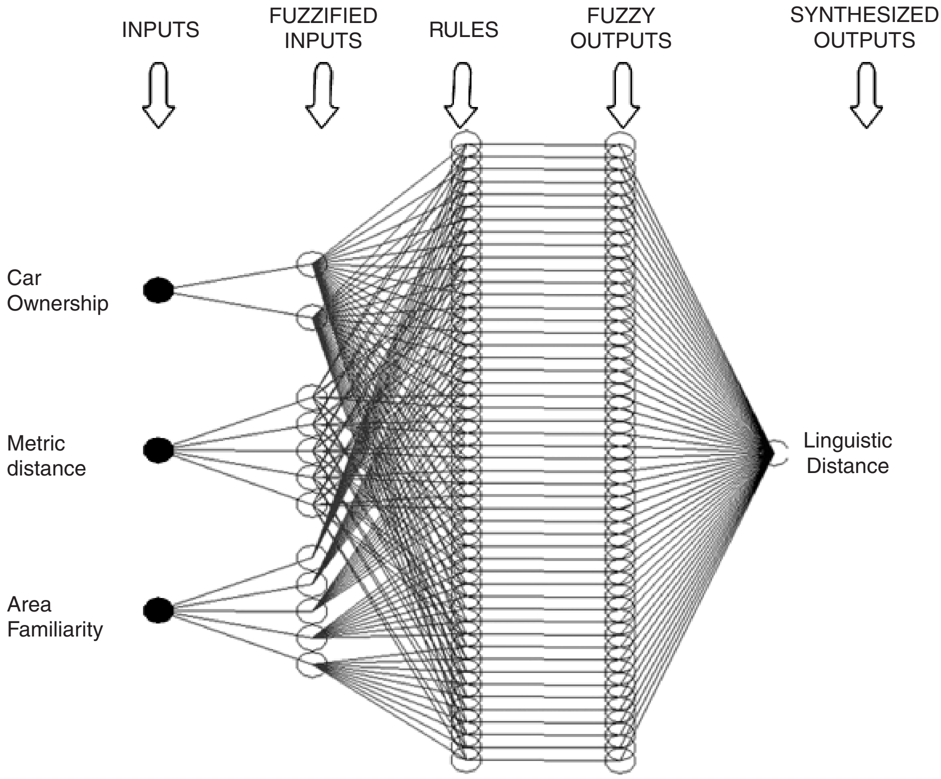

Figure 1. An example three-input model of the neurofuzzy system for proximity modeling.

output layer yields prediction of the other type of proximity measures (e.g., qualitative distance). The number of inputs is determined by the number of context variables that the modeler wants to include in the model. Fig. 1 gives an illustrative example for which three inputs variables are used to predict the perceiver's linguistic measure of the metric distance, namely metric distance, area familiarity, and car ownership. Of course, whenever necessary, we can add more context variables with each variable as an input node on the input layer. The optimal number and choices of inputs variables need to be found through a stepwise search, which will be discussed in detail in the following sections. Between the input and output layers, the system fuzzifies the input, sets up and trains the rules, obtains fuzzy outputs from the fuzzy rules, and finally generates the synthesized final output from the fuzzy outputs. We will discuss the processes in more detail below.

Fuzzification of input. As shown in Fig. 1, fuzzification takes place between the input layer and the fuzzified input layer. The fuzzification process is appropriate for proximity modeling because there are many ambiguous concepts involved in the context variables, not to mention the fuzzy nature of the proximity concept itself. This process converts crisp sets into fuzzy sets by defining membership functions for

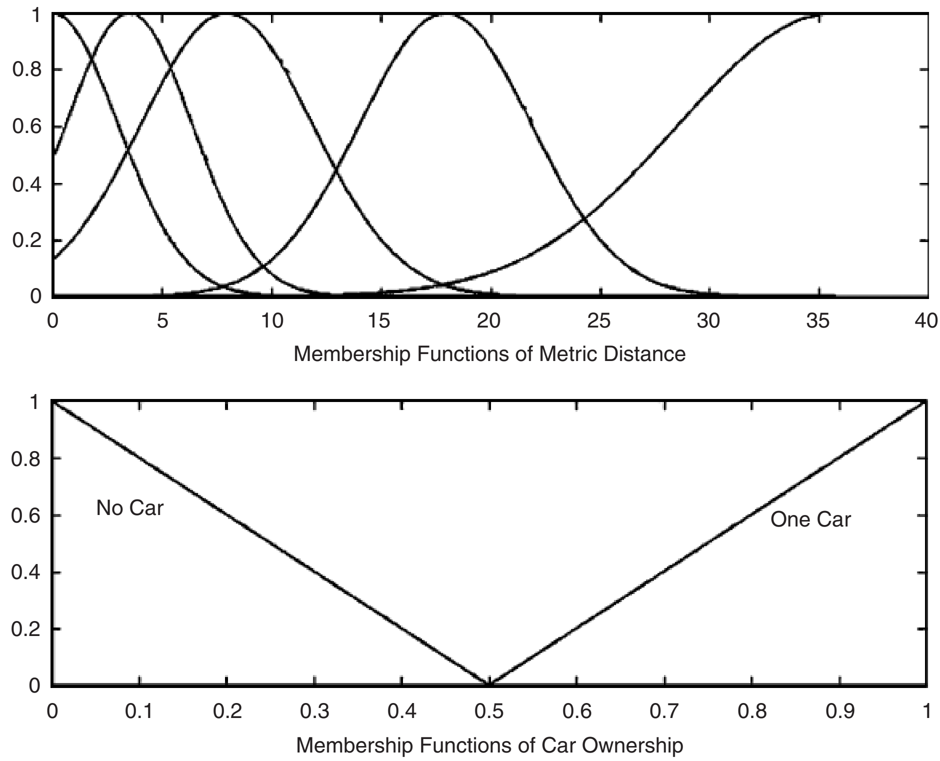

Figure 2. Examples of membership functions for input fuzzification.

each crisp input variable. The fuzzification is easy to understand if the input crisp set can be turned into fuzzy sets that express vague concepts. Fig. 2(a) illustrates such a case. The metric distances are turned into five fuzzy sets; each has a corresponding membership curve (function). We label the corresponding vague concepts from ''very short'' to ''very far'' just for illustration purpose. Labels are not necessary for the modeling process. Thus, the input variable ''metric distance'' would produce five nodes in the layer of fuzzified inputs. However, there are also some input variables that cannot be turned into vague concepts. ''Car ownership'' is such an input variable. This context variable only takes values ''0'' and ''1'' and each value has a clear semantic meaning (''no car'' and ''one car or more,'' respectively). The two crisp concepts can still be represented with fuzzy sets because a fuzzy set is a superset of a crisp set. This is shown in Fig. 2(b). Although a particular fuzzy membership function is depicted in this figure, what matter are only the membership grades at point 0 and 1. As the grades at point 0 and 1 are both 1, they essentially represent two crisp sets. The shape of the curve does not cause any distortion of the concept. This way, crisp concepts can still be used in the neurofuzzy system.

The fuzzification process is performed by the neural network through iterative updates of the link weights between the input nodes and fuzzified input nodes. It is also possible to incorporate expert knowledge into the process by setting up initial membership curves according to prior knowledge about the concepts. This will be further discussed in the case study.

Fuzzy rules. The rules layer consists of fuzzy IF–THEN rules. It serves as a rule base in the system. Fuzzy if–then rules are often used as surrogates for more complex models of reasoning that underpin humans' ability to make decisions in an environment where uncertainty is pervasive. In our approach, each rule node takes the contributing fuzzy input nodes as antecedents and yields a consequent—a node in the fuzzy output layer. An example of such a rule is as follows:

$IF\ (input1\ is\ input1mf1)\ AND\ (input2\ is\ input2mf1)\ldots AND\ (inputN\ is\ input\ Nmf1)\ THEN\ output\ is\ outputmf1$

where *input*1 through *inputN* are the input variables as nodes in the input layer, and *input*1*mf*1 through *inputNmf*1 are fuzzified concepts appearing as nodes in the fuzzified input layer. The outputmf1 is a node in the fuzzified output layer. The If-part of the fuzzy rule is called *antecedent* or *premise*, and the Then-part is called the *consequent* or *conclusion*. Fuzzy logical operations AND (fuzzy intersection or conjunction), OR (fuzzy union or disjunction), and NOT (fuzzy complement function) are applied here to generate the output fuzzy sets.

Synthesizing the final output. The rules layer is composed of multiple rules, each of which produce a fuzzy set (a node) in the fuzzy output layer. The final process in the model is to synthesize the multiple fuzzy sets in the fuzzy output layer and yield a final output node. Several strategies have been proposed, which mainly differ on whether the consequents of the fuzzy rules are crisp sets or fuzzy sets and on the procedures to carry out the aggregation and the defuzzification. Two major types of fuzzy inference are of particular interest here. With the Mamdani type of fuzzy inference (Mamdani and Assilian 1975), the output of each fuzzy rule is a fuzzy set and all fuzzy sets are ultimately aggregated and defuzzified to generate a crisp output. The second type is the Sugeno type (also known as the Takagi–Sugeno model; Takagi and Sugeno 1985; Sugeno and Kang 1988). In this type of fuzzy inference system, each fuzzy rule infers a crisp consequent and the final result is a weighted average of the crisp outputs of all fuzzy rules.

Our approach particularly concerns two issues of the existing fuzzy inference strategies. First, should we use a Mamdani or Sugeno type of fuzzy inference? We believe that both types of fuzzy inference are pertinent to our modeling approach, as each of them is suitable for different input–output relations. Accordingly, our approach should allow the consequent of each fuzzy rule to be either a fuzzy or a crisp set. As fuzzy sets are supersets of crisp sets, we generally use fuzzy sets in the system information flow as illustrated in Fig. 3. The second issue is whether the final output ought to be represented as a fuzzy set or a crisp value. In current commercial packages that have neurofuzzy inference capability, the final output is always a crisp value obtained through defuzzification (e.g., weighted average of conse-

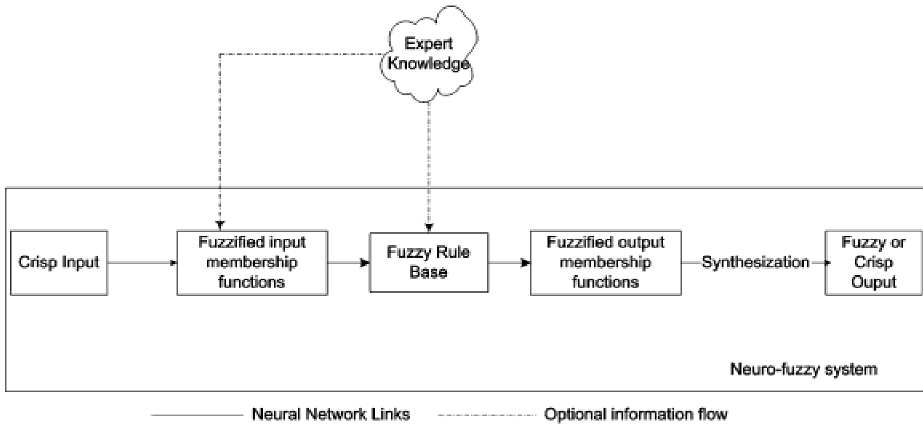

Figure 3. Process and information flow in the neurofuzzy system for proximity modeling.

quents from all fuzzy rules). However, one of two types of output may be more desirable for certain applications depending on the particular problem to be modeled. In the case that a fuzzy set is preferred, it can be generated in the synthesization process by a fuzzy operation (e.g., MIN or Product) of the fuzzy consequents from all fuzzy rules. Fig. 3 shows the sequence of processes and the flow of information in the proposed neurofuzzy system.

Training the neurofuzzy system. When no or insufficient prior expert knowledge is available on a fuzzy problem, it is useful to use a neurofuzzy system. Once the structure, membership functions, and weights are initialized either by system defaults or by a human expert, the parameters can be fine-tuned by the neural network with a training algorithm such as back-propagation. In its simplest form, the system is trained by iterating through the following four steps, while we acknowledge the fact that many variations on this basic process exist.

Step 1. Present an input data case, compute the corresponding output;
Step 2. Compute the errors between the predicted output and the actual output (from sample data);
Step 3. In light of the obtained errors in Step 2, adjust the connection weights and membership functions; and
Step 4. Repeat Step 1 through Step 3 until the change of estimates becomes minimal.

It should be noted that expert knowledge could be used on an optional basis in the modeling process. This is also reflected in Fig. 3, with dashed lines signifying it. Specifically, expert knowledge of an input variable, once available, can be used to initialize the number of shapes of fuzzy membership functions. Expert knowledge of fuzzy rules, if available, can be used to delete useless rules and/or add new rules. The neural network will then train the systems by updating parameters of the membership functions and weights on the connections among nodes of adjacent layers.

A case study

Experimental data

To test the effectiveness of this approach, we construct a set of neurofuzzy proximity models and compare their prediction power with that of an OLR proximity model reported elsewhere (Yao and Thill 2005a). The same set of data is used in this and the statistical modeling study. The data were collected in 2002 in an empirical study of proximity perception and cognition among undergraduate students at the University at Buffalo, The State University of New York. A questionnaire survey was administrated to 95 undergraduate students enrolled in three randomly selected undergraduate courses to collect qualitative and quantitative distance measures under controlled conditions, as well as contextual information. Each participant was asked to provide the above-mentioned information regarding 19 hypothetical trips. For each hypothetical trip, a trip scenario was either predefined by the survey designer or chosen by each survey participant. A participant was asked to indicate his/her psychological feeling toward each trip length, given the trip scenario—the linguistic distance. A participant could express the linguistic distance as ''very near,'' ''near,'' ''normal—not so near and not so far,'' ''far,'' or ''very far.'' The actual metric network distance (as a surrogate of trip distance) was calculated with the network analysis functionality of GIS. Overall, valid and complete data on 996 trip scenarios provided by 95 students were collected. The response database thus consists of 996 records; each record corresponds to a trip scenario evaluated by a certain participant in the survey.

Given the scope of the study and practical considerations tied to the questionnaire survey for the study (for details, refer to Yao and Thill 2005a), the following context factors are measured and tested for inclusion in the proximity model: type of activity in query (six types); transportation mode (four modes); transportation network (topology); familiarity with the area (five levels); familiarity with the route (five levels); intervening opportunities (availability and relative distance to other equivalent facilities); personal characteristics including gender (two classes), ethnicity (five classes), car ownership (two classes), years of residence in Buffalo, types of environment (four classes: urban, suburban, rural, and foreign) where the participant had lived for more than 5 years, and environment type where the participant attended high school (four classes).

Neurofuzzy modeling

The objective is to construct the fuzzy relationship between the two types of distance measures contingent upon the context variables. We want to use one of the distance variables, together with the context factors, to predict the other distance variable. Therefore, we need to construct two types of models in this research. One is to predict the linguistic proximity measure by the metric distance (equation [3]);

the other is to predict the metric distance by the linguistic proximity measure (equation [4]).

$$Linguistic\_Distance = f_1(Metric\_Distance,\ Context\_Factors) \tag{3}$$

$$Metric\_Distance = f_2(Linguistic\_Distance,\ Context\_Factors) \tag{4}$$

Software and data preprocessing

We use Matlab's add-on of Adaptive Neuro-Fuzzy Inference System (ANFIS) to carry out the modeling tasks. In this system, a back-propagation algorithm with steepest-descent optimization is used to train the fuzzy neural network (Jang 1993). When a training sample is given, the algorithm calculates the prediction error by the neural network and uses the error to adjust the weights of the neural network in the steepest descent direction (negative of the gradient). For technical details of the algorithm, refer to Mitchell (1997).

The experimental data set is split into two data sets: one for training and the other for testing. The training data set is used to adjust the initial membership function and connection weights of different layers in the system. The testing set is used to test the conformance of the prediction of the trained system to real data. To minimize the risk of overfitting and thus obtain more reliable results, the training process prefers larger data sets. We partition the 996 sample data into two data sets as follows: 896 randomly selected records for training and the remaining 100 sample data for testing and validation.

Model selection

Candidate input variables include the 12 context variables identified in "Fuzzy relationship between natural language and metric distances." One of the two distance variables will join the context factors as input variable while the other becomes the output variable. All the collected context variables are potentially influential to proximity perception to various extents. However, for the particular scenario and the subjects, we do not know a prior; how significant the influence of each context variable is. The study uses a stepwise forward training strategy to select the best combination of input variables for the fuzzy neural network. The strategy starts with a neurofuzzy system with one input variable and records its training and testing errors and then adds another input variable and records the training and testing errors of the new model. The process continues until no significant improvements in the training and testing errors are found. Table 2 details the steps of the model selection algorithm used in the research. The underlying purpose is to identify the most parsimonious model without losing significant prediction power.

Parameters of a neurofuzzy model need to be initialized before training the model. Such parameters include the coupling weights associated with links among neurons, the numbers and definitions of fuzzy membership functions for the input variables, the fuzzy rules, and the number of epochs.

Table 2 Stepwise Model Selection Algorithm for Neurofuzzy Systems

Model selection algorithm (*I, m, O, Tol, Threshold*)

I: the pool of candidate input variables

m: the total number of input candidates

n: the number of input variables of current model

O: output variable

Tol: tolerance of the difference between training error and testing error

Threshold: minimal improvement in model training error for adding an input variable to the model specification

[A]. Define membership functions for each candidate input variable and set $n = 1$.

[B]. Find the best configuration of n-input variables:

1. Choose the first combination of n variables from the m input variable candidates in I.
2. Train the model that uses chosen n variables as input, O as output. Record the square root of the mean of errors $E_t[n]$.
3. Evaluate the model with testing data. Record the square root of the mean of testing errors $E_c[n]$.
4. If $|E_t[n] - E_c[n]| > Tol$, give this model no further consideration.
5. Repeat steps 1–4 for all the possible combination of n variables from the m candidate input variables.
6. Choose the configuration that yields the lowest $E_t[n]$ and for which $E_c[n]$ is not larger than $E_t[n] + Tol$.

[C]. If $(n > 1)$ compare $\min(E_t[n])$ with $E_t[n-1]$. If the difference is smaller than *threshold*, go to [E].

[D]. If $(n < m)$ $\{n = n+1;$ Go to [B]$\}$ else go to [E].

[E]. Stop: the selected model has a configuration of $n-1$ input variables and one specified output variable O. The $n-1$ input solution is the combination which yields $\min(E_t[n])$.

Because the model selection procedure detailed in Table 2 entails the training of a large number of model structures, it can be computationally very intensive. To decrease the computational burden, model selection is performed in a sequence of crude selection and fine-tuning stages. In the crude selection stage, we preset parameters by defaults for all the network configurations. The default maximum number of epochs is set to ten. Each input variable is assigned three Gaussian membership functions. We choose Gaussian distributions to fuzzify the inputs because they are known to be a good approximation for random effects according to the central limit theorem. When these parameters are defined, the model selection procedure described in Table 1 can be carried out by ANFIS without human intervention. ANFIS will adjust all parameters through the training process. This procedure gives us preliminary results on the performance of candidate model configurations.

After the crude selection stage, a small set of the best performing models is selected for fine tuning so that we can work more closely on the parametric specification with consideration for expert knowledge. In the fine-tuning stage, some parameters are chosen according to prior knowledge through human intervention. Expert knowledge includes one's understanding of the nature of each variable, data histograms, and so forth. The number and definitions of the membership functions for each variable are initialized according to prior knowledge of the data. Fig. 2 depicts two examples of membership curve initialization through human intervention. The upper part of this figure shows the membership functions for the variable of metric distance. Five membership functions are initialized, which correspond to the concepts of *very short*, *short*, *normal*, *long*, and *very long* distances, respectively. The figure shows that the variable of metric distance is assigned unevenly distributed membership curves. Indeed, examination of descriptive statistics of the experimental data collected suggests that shorter distance ranges should be applied to shorter proximity types. To reflect this observation, the membership function curves are given higher kurtosis and are closer to each other at the lower end of the metric distance scale. Parameters of the membership functions are subsequently adjusted by ANFIS through data training.

In the fine-tuning stage, epoch size is endogenously determined by the convergence of the training process, which is achieved when the difference between training and testing errors has reached an acceptably small value. At convergence, goodness of fit is the best and overfitting is avoided. When the results of the models constructed at the two stages (crude selection stage and the fine-tuning stage) are compared, it is found that the fine-tuned models exhibit better or at least as good prediction accuracies as their original counterparts. The final best models can be selected based on results at both stages.

Modeling results

Modeling to predict linguistic distance. The modeling procedure for predicting the linguistic distance by the metric distance and context factors starts with the algorithm described in Table 1. The output variable is the linguistic distance. For the input variables, we start with one-input variables up to five-input variables, as no improvement in prediction accuracy is seen from five-input variables models compared with their four-input counterparts. Fig. 4 shows the training and testing errors of the nine best-performing models (ranked in ascending order of training errors) for the same configuration (same number of input variables). Here, we report only the nine best models for each configuration to ensure the readability of the figure. The line pattern indicates the number of input variables. The horizontal axis shows the sequence number of each model. For example, Model 2-2 means that it is the second of all two-input models tested. We choose the best models of each configuration (except the one-input configuration due to its significantly larger errors) for further fine tuning. The selection is based on the criteria of low training and testing errors. The chosen models include a two-input model (Model 2-4), two three-input models (Model 3-1 and Model 3-4), and a four-input model (Model 4-1). In the fine-tuning stage, all four retained models are trained and tested once again. This time, we choose the number and shapes of the membership functions according to expert

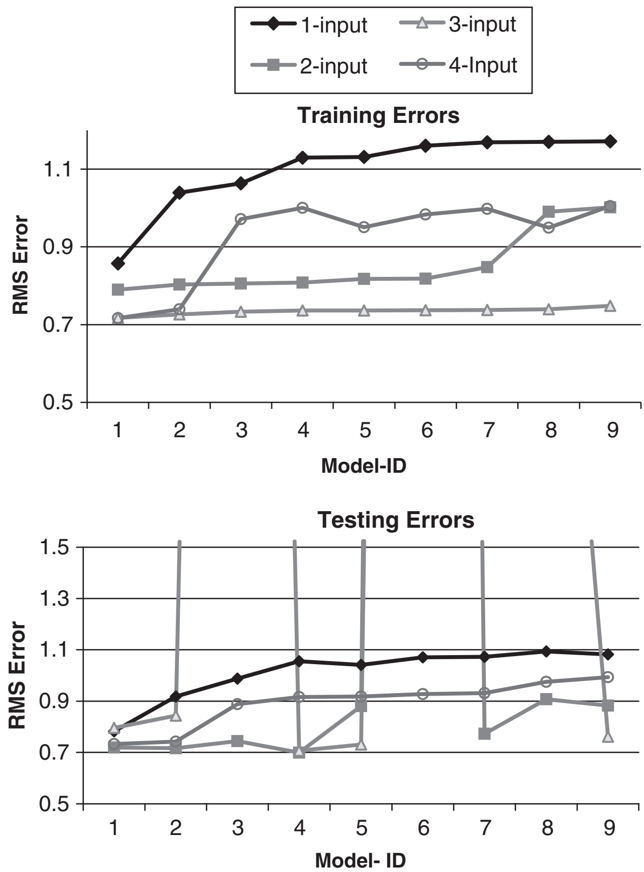

Figure 4. Comparing the training and testing errors of models to predict linguistic distances.

knowledge on the data as explained in ''Model selection.'' For example, five (instead of the default 3) unevenly distributed membership curves are associated with the metric distance variable as suggested by the scatter plot of linguistic distances versus metric distances in the experimental data.

The modeling results of the two-stage training process are reported and compared in Table 3. The columns labeled ''Original error'' report model errors at the crude-selection stage, while the columns labeled ''New error'' report errors by models produced at the fine-tuning stage. Comparing the original and the new sets of errors, one can find that most new models perform better or at least as well as their original counterparts. Although our use of expert knowledge improves the model performance to a limited extent, the improvements are so insignificant that it proves that the initialization of the membership functions does not dramatically affect the construction of a model. In other words, the modeling is very resilient to the exact specification of membership functions. In spite of the improvement in

Table 3 Comparing the Models of Linguistic Distance Before and After Adding Prior Knowledge

Model ID	Input variables	Original training error	New training error	Original testing error	New testing error
2-4	*Metric_Distance*, *Car_Ownership*	0.8082	0.8059	0.6994	0.6993
3-1	*Metric_Distance*, *Area_Familiarity*, *Ethnicity*	0.7172	0.7008	0.7956	0.7490
3-4	*Metric_Distance*, *Area_Familiarity*, *Car_Ownership*	0.7362	0.7256	0.7075	0.6799
4-1	*Metric_Distance*, *Area_Familiarity*, *Gender*, *Car_Ownership*	0.7165	0.6858	0.7331	0.7352

goodness of fit of all four models, their relative performances have remained unchanged after adding prior knowledge.

Now we will choose the best model(s) from the four models listed in Table 3. Three criteria are used for this purpose: a low training error, a low testing error, and finally, a small difference between training and testing errors. If two models are equally good according to the above three criteria, the most parsimonious model is selected. The first two criteria have obvious reasons. The third criterion can be justified on the following grounds. Two situations produce a large difference between the two types of errors. In the first situation, training error is low but testing error is much higher. It is evidence of overfitting and implies that the model has low prediction accuracy and therefore is not a good model. The second situation is characterized by a high training error but a much lower testing error. Although a low testing error may imply high prediction accuracy, it is not guaranteed. In principle, a good model fits its training data set better than its testing set because that is precisely the purpose of training samples. If the training error is high, the model is not good even though it yields a low testing error. In fact, the low testing error might have occured by chance, considering that our testing data set (100 cases) is much smaller than the training data set (896 cases). In short, the second situation does not apply to good models either.

Model 2-4 still has a significantly larger training error than the other models in spite of its low testing errors. Therefore, it will not be further considered. Model 3-1 is inferior to Model 4-1 in both training error and testing error. The performances of Models 3-4 and 4-1 are not significantly different from one another: Model 4-1 has the smallest training error and Model 3-4 has the smallest testing error. Model 3-4 relies on *Metric_Distance*, *Area_Familiarity*, and *Car_Ownership* to predict linguistic distance. We can see that Model 4-1 has the *Gender* variable in addition to all the input variables in Model 3-4. Model 4-1 is a more complex model while yielding similar performance. Thus, for parsimony reasons, Model 3-4 is chosen to be the best model.

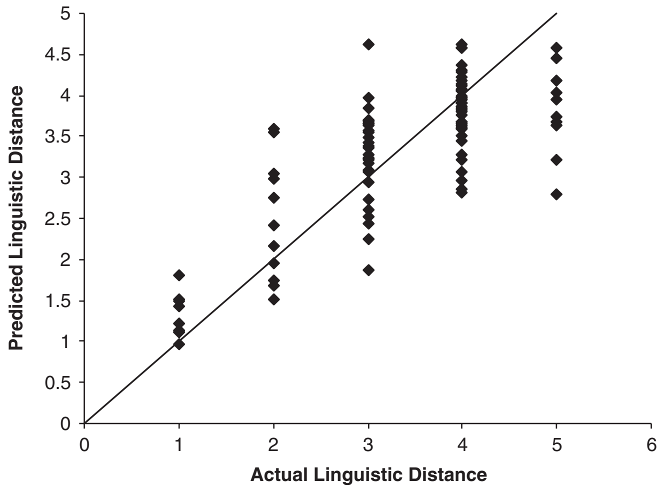

Figure 5. Comparison of actual and predicted linguistic distances by Model 3-4.

Fig. 5 plots the linguistic distances predicted by Model 3-4 against the actual linguistic distances in the testing data set. The 45° line shows the locations where predicted distances equal actual distances. The figure shows that Model 3-4 tends to estimate linguistic distances longer than that perceived by the subjects for a given metric distance that is short; conversely, Model 3-4 tends to predict linguistic distances to destinations that are metrically known to be very far as shorter than perceived. This systematic bias of the model is probably due to the small number of cases in these two categories of linguistic distances. Because the model has been trained to give the smallest overall prediction error, it gives the best prediction of linguistic distances for cases around the mean, where most cases are distributed. It should be pointed out that the expert knowledge brought to bear on the model construction is of little use in this situation because it is about structures embedded in the data, not about the relationships (or model) between inputs and outputs themselves. As ANFIS trains the model to best fit the data, the prediction power cannot exceed what the sample data allow for. Therefore, the quality of the model is largely determined by the characteristics of the sample data.

Modeling to predict metric distance. The modeling procedure is the same as that used to predict linguistic distances. However, here the output is metric distance. Fig. 6 shows the training and testing errors of the nine best-performing models (ranked in ascending order of training errors) for each configuration of input variables. The two-input models surpass the one-input model in both training and testing performance. Three-input models improve the performance further on both data sets. The four-input and five-input models, however, start to exhibit higher testing errors, although their training errors continue to decrease. This is a sign of model overfitting. It means that the predictive power (accuracy) of these models starts to

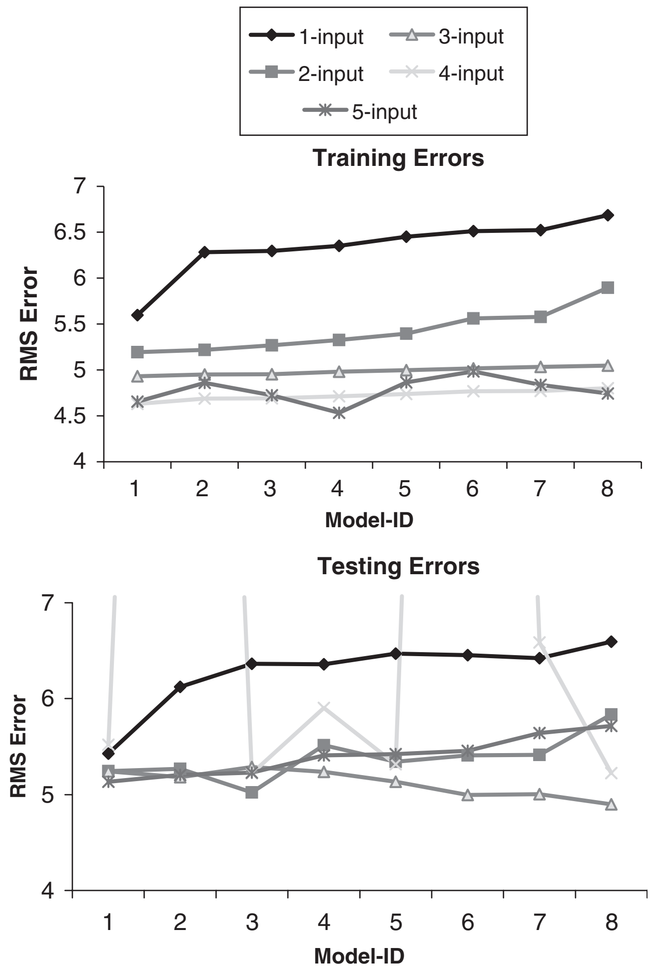

Figure 6. Comparing the training and testing errors of models for predicting metric distances.

decrease in spite of the fact that more variables are used. Therefore, adding more variables is not going to improve the model any more. Thus, there is no need to test models with more complex configurations and we can look for the best among the current models. Table 4 reports the three best-performing three-input models at the end of the crude selection and after fine tuning.

Comparing the original and the new sets of errors, one can find that fine tuning always improves model performance. It again confirms that the use of prior knowledge in specifying model initial parameters does help to improve the goodness of fit, although the improvement is not significant. In other words, the neurofuzzy system of proximity modeling is not very sensitive to the initial specifications of membership functions.

Table 4 Comparing the Models of Metric Distance Before and After Using Prior Knowledge

Model ID	Input variables	Original training error	New training error	Original testing error	New testing error
3-6	*Activity, Linguistic_Distance, Car_Ownership*	5.0162	4.8765	4.9966	4.8715
3-7	*Ethnicity, Linguistic_Distance, Car_Ownership*	5.0327	4.9912	5.0056	4.9016
3-8	*Area_Familiarity, Linguistic_Distance, Car_Ownership*	5.0465	4.9848	4.8984	4.8433

Model 3-6 only marginally outperforms Mode 3-7 and Model 3-8, which may be due to random factors. The training errors and testing errors of the three models are indeed very close to each other. Therefore, it can be concluded that Models 3-6, 3-7, and 3-8 are the three best models for predicting metric distance. Because the performances of the three models are similar, the following two figures show the prediction results of only one model, Model 3-8, for demonstration purposes. Fig. 7 plots the predicted metric distances against the actual metric distances. We can see that the model predicts better for short- and middle-range distances than it does for the long distances. The prediction of metric distances tends to be smaller than the actual metric data for destinations that are far away. The observation is confirmed by the histogram of the prediction residuals of Model 3-8 (Fig. 8a). The mean error is 0.3 with a standard deviation of 4.86, but the error distribution is clearly skewed to the left. It means that larger errors are seen when the model underestimates the distance. This is probably due to the small number of cases in the range of distances

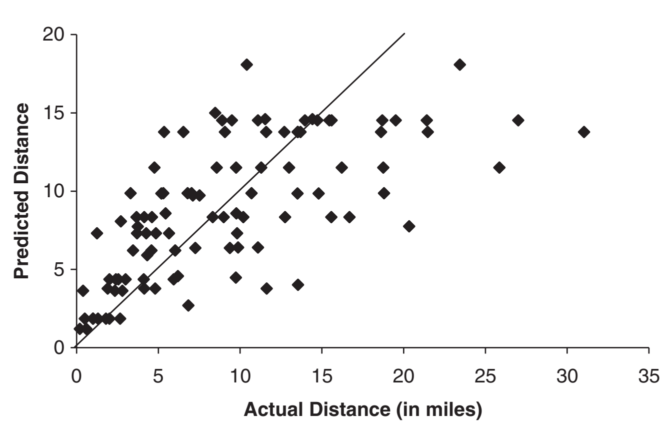

Figure 7. Comparison of actual metric distance and metric distance predicted by Model 3-8 on the testing data set.

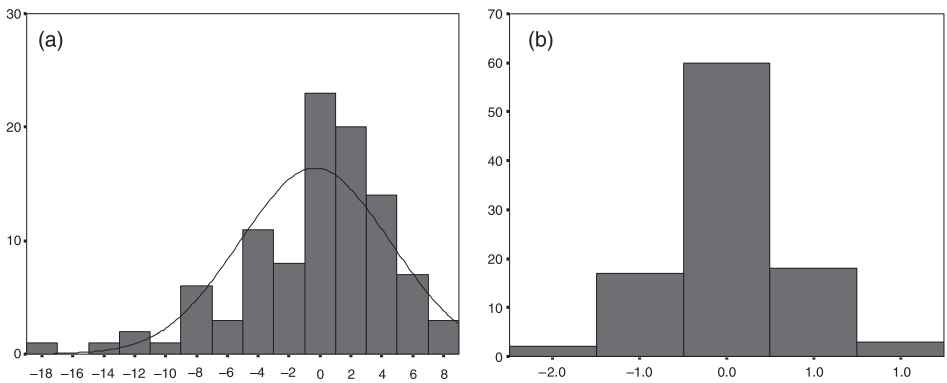

Figure 8. Histograms of prediction errors: (a) Model 3-8 of metric distance; (b) Model 3-4 of linguistic distance.

over 15 miles and, consequently, the lower predictive ability of the neurofuzzy system in this distance range.

Comparing the neurofuzzy modeling results with a statistical model

To evaluate the prediction power of the neurofuzzy systems approach for proximity modeling, we compare it with the results of an OLR model estimated on the same data and reported in Yao and Thill (2005a). Estimates of the neurofuzzy model are defined on a continuous scale (real numbers ranging from 1 to 5 representing the gradual change from very near to very far), while that of the OLR model is defined on an ordinal scale (integers 1–5 representing five linguistic measures very near to very far). To make the results of the two models comparable, we convert the neurofuzzy predictions into discrete linguistic distances by rounding to the nearest integer. As the OLR model cannot be used to estimate metric distance measures because they are defined on an interval scale, we will compare the OLR model of linguistic distances with its counterpart of the neurofuzzy systems model only.

Fig. 8b shows the histogram of residuals of the best neurofuzzy model (Model 3-4) to predict linguistic distance measures. The mean error is 0.0 with a standard deviation of 0.74. The OLR model (Yao and Thill 2005a) gives a distribution of errors that is normally distributed with a mean of 0.056 and a standard deviation of 0.84. To further compare the prediction accuracy, we can look at the root mean squared errors (RMSE) of the two models. The RMSE of the OLR model is 0.8654, while the RMSE of neurofuzzy Model 3-4 is 0.7256 on the training data and 0.7416 on the testing data (Fig. 8b). Following the inferential approach proposed by Granger and Newbold (1977, p. 281), the difference in RMSE between the neurofuzzy and OLR models on the testing data is found to be statistically significant at the 0.05 level. Hence, a comparison based on these statistics indicates that the goodness of fit of neurofuzzy models of linguistic proximity is superior to that of the statistical approach of OLR. Furthermore, the neurofuzzy model turns out to use fewer context variables than the statistical model to achieve better prediction performance. Not only does the neurofuzzy model have a higher predictive power, it is also more parsimonious.

In addition to providing better prediction accuracy, more information can be obtained from the neurofuzzy approach. Through the stepwise process of neurofuzzy model building, the order of prediction power of input variables (to linguistic distance) is elicited. Some variables consistently appear in all the well-performing models. Of all one-input models, the one with the variable of *Metric_Distance* gives the best results. All the highly ranked well-performing two-input models include *Metric_Distance* as a predictor. It suggests that *Metric_Distance* is the most important input variable, which conforms to our intuitive understanding of the mapped relationship. Similarly, the majority of well-performing three-input models consist of the variables of *Metric_Distance* and *Area_Familiarity*, which suggests that *Area_Familiarity* is the next most important input variable. This kind of observations helps us determine the order of importance of variables that account for the prediction power of models. The order in Model 3-4, for example, is *Metric_Distance*, *Area_Familiarity*, and then *Car_Ownership*.

Other variables, including *Transportation_Mode*, *Activity_Type*, *Perceiver's Ethnicity*, and *Gender*, are statistically significant predictors in the OLR model (Yao and Thill 2005a) but do not appear to enhance goodness of fit in the neurofuzzy model significantly. Two comments are in order on this matter. First, as discussed above, there is an implicit order of prediction power associated with each factor to predict linguistic distance with the neurofuzzy model. The absence of these variables in our best neurofuzzy models suggests that their prediction power is lower than that of the variables retained in the models, and yet not null. Furthermore, because the neural network approach offers greater flexibility to capture complex relationships in the training data that do not necessarily conform to simple functional relationships as those assumed in OLR models, the explanatory power of the most fundamental predictors may be strengthened, while other potential predictors may become extraneous.

Conclusions

This article explores a neurofuzzy approach to context–contingent proximity modeling. We conclude that the neurofuzzy approach is superior to the previous statistical (OLR) approach in several respects. First, the neurofuzzy approach is grounded in sounder theoretical considerations on two counts. The first is that fuzzy logic is better suited to represent the fuzziness inherent to the relationship between natural language and metric proximity statements. The other is related to the definition of fuzzy membership functions. It is our contention that no particular relationship should be presumed between the proximity measures and context variables. In our approach, we use a neural network to figure out the membership functions from sample data. This is more convincing than the previous statistical approach and other previous fuzzy logic studies, which all assume the form and/or parameterization of the relationship. Second, the modeling results show that the neurofuzzy approach gives higher prediction accuracy on both training and testing data, while being more parsimonious. Third, the neurofuzzy approach provides a two-way prediction, meaning that either metric distance or linguistic distance can be the predicted variable. The previous statistical model can only predict linguistic distance.

Although the theoretical construction of the approach is well justified and the preliminary study shows the obvious promise of this approach, our experiment is still constrained by the capability of current commercial neurofuzzy network packages. As the fuzzy inference system was originally proposed for industry control, crisp outputs had to be generated for decision making (e.g., Mamdani and Assilian 1975; Takagi and Sugeno 1985; Jang 1993). This legacy is still prevalent in current neurofuzzy inference systems. For example, Matlab's ANFIS (as well as other similar commercial packages) uses Sugeno-type inference, which defuzzifies the fuzzy outputs to produce a single output value. While the type of inference is appropriate for systems where a single number is required, it is not well suited for the prediction of metric distances from a linguistic distance measure. The conceptual structure of our approach as shown in Fig. 3 indicates that both fuzzy and crisp output should be possible, depending on the particular requirement of the modeling problem. Future research will necessitate the implementation of neurofuzzy inference systems that can produce fuzzy output as well. Another future research avenue is the integration of this inference capability in GIS. As suggested in the previous sections, potential applications of the context–contingent proximity models are personalized GIS applications for the general public. The model needs to receive data from, and send results to, the GIS applications. As the potential users are the general public at large, information transfer between the model and GIS should be as automatic as possible. A tight integration between the proximity models and GIS thus expects to receive its due research endeavors in the days to come.

References

Bera, R., and C. Claramunt. (2003). ''Topology-Based Proximities in Spatial Systems.'' *Journal of Geographical Systems* 5(4), 353–80.

Brown, W., D. Groves, and T. Gedeon. (2003). ''Use of Fuzzy Membership Input Layers to Combine Subjective Geological Knowledge and Empirical Data in a Neural Network Method for Mineral-Potential Mapping.'' *Natural Resources Research* 12(3), 183–200.

Davis, T. J., and C. P. Keller. (1997). ''Modeling and Visualizing Multiple Spatial Uncertainties.'' *Computers and Geosciences* 23(4), 397–408.

Egenhofer, M. J., and D. M. Mark. (1995). ''Naïve Geography.'' In *COSIT '95 Proceedings: Spatial Information Theory: A Theoretical Basis for GIS*, 1–15, edited by A. U. Frank and W. Kuhn. Berlin, Germany: Springer-Verlag.

Ferrier, G., and G. Wadge. (1997). ''An Integrated GIS and Knowledge-Based System as an Aid for the Geological Analysis of Sedimentary Basins.'' *International Journal of Geographical Information Science* 11(3), 281–97.

Fischer, M. M., and R. J. Abrahart. (2000). ''Neurocomputing—Tools for Geographers.'' In *Geocomputation*, 187–218, edited by S. Openshaw and R. J. Abrahart. New York: Taylor & Francis.

Fischer, P. (2000). ''Fuzzy Modelling.'' In *Geocomputation*, 161–86, edited by S. Openshaw and R. J. Abrahart. New York: Taylor & Francis.

Fiset, R., F. Cavayas, M. C. Mouchot, B. Solaiman, and R. Desjardins. (1998). ''Map-Image Matching Using a Multi-Layer Perceptron: The Case of the Road Network.'' *Journal of Photogrammetry and Remote Sensing* 53(2), 76–84.

Frank, A. U. (1992). ''Qualitative Spatial Reasoning About Distances and Directions in Geographic Space.'' *Journal of Visual Languages and Computing* 3, 343–71.

Freeman, J. (1973). ''The Modelling of Spatial Relations.'' *Computer Graphics and Image Processing* 4, 156–71.

Gahegan, M. (1995). ''Proximity Operators for Qualitative Spatial Reasoning.'' In *COSIT '95 Proceedings: Spatial Information Theory: A Theoretical Basis for GIS*, 31–44, edited by A. U. Frank and W. Kuhn. Berlin, Germany: Springer-Verlag.

Goodchild, M. (2000). ''GIS and Transportation: Status and Challenges.'' *Geoinformatica* 4, 127–40.

Gopal, S., W. Liu, and C. Woodcock. (2001). ''Visualization Based on the Fuzzy ARTMAP Neural Network for Mining Remotely Sensed Data.'' In *Geographical Knowledge Discovery and Data Mining*, 315–36, edited by H. Miller and J. Han. New York: Taylor & Francis.

Granger, C. W. J., and P. Newbold. (1977). *Forecasting Economic Time Series*. New York: Academic Press.

Guesgen, H. W. (2002a). ''Reasoning About Distance Based on Fuzzy Sets.'' *Applied Intelligence* 17, 265–70.

Guesgen, H. W. (2002b). ''Fuzzyfying Spatial Relations.'' In *Applying Soft Computing in Defining Spatial Relations*, 1–16, edited by P. Matsakis and L. M. Sztandera. Berlin, Germany: Springer-Verlag.

Guesgen, H. W., and J. Albrecht. (2000). ''Imprecise Reasoning in Geographic Information Systems.'' *Fuzzy Sets and Systems* 113, 121–31.

Hernandez, D., E. Clementini, and P. D. Felice. (1995). ''Qualitative Distances.'' In *COSIT '95 Proceedings: Spatial Information Theory: A Theoretical Basis for GIS*, 45–57, edited by A. U. Frank and W. Kuhn. Berlin; Germany: Springer-Verlag.

Ishibuchi, H., and M. Nii. (1998). ''Fuzzy Neural Networks Techniques and Their Applications.'' In *Fuzzy Logic and Expert Systems Applications*, 1–56, edited by C. T. Leondes. San Diego, CA: Academic Press.

Jang, J. S. R. (1993). ''ANFIS: Adaptive-Network-Based Fuzzy Inference System.'' *IEEE Transactions on Systems, Man and Cybernetics* 23, 665–85.

Karimi, H. A., and A. Hammad. (2004). *Telegeoinformatics: Location-Based Computing and Services*. Boca Raton, FL: CRC Press.

Kathmann, R. M. (1993). ''Neural Networks for the Mass Appraisal of Real-Estate.'' *Computers, Environment and Urban Systems* 17(4), 373–84.

Kaufmann, A. (1975). *Introduction to the Theory of Fuzzy Subsets*. New York: Academic Press.

Kuipers, B. (2004). ''Making Sense of Common Sense Knowledge.'' *Ubiquity* 4(45), www.acm.org/ubiquity.

Li, X., and G. Yeh. (2002). ''Neural-Network-Based Cellular Automata for Simulating Multiple Land Use Changes Using GIS.'' *International Journal of Geographical Information Science* 16(4), 323–43.

Lloyd, R. (1994). ''Learning Spatial Prototypes.'' *Annals of the Association of American Geographers* 84(3), 418–40.

Mamdani, E. H., and S. Assilian. (1975). ''An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller.'' *International Journal of Man–Machine Studies* 7(1), 1–13.

Miller, H., and J. Han (eds). (2001). *Geographical Knowledge Discovery and Data Mining*. New York: Taylor & Francis.

Mitchell, T. M. (1997). *Machine Learning*. Boston: McGraw-Hill Higher Education.

Montello, D. R. (1997). ''The Perception and Cognition of Environmental Distance: Direct Sources of Information.'' In *Spatial Information Theory: A theoretical Basis for GIS*, 297–311, edited by S. C. Hirtle and A. U. Frank. Berlin, Germany: Springer-Verlag.

Nauck, D., and R. Kruse. (1998). ''A Neuro-Fuzzy Approach to Obtain Interpretable Fuzzy Systems for Function Approximation.'' In *Proceedings of IEEE International Conference on Fuzzy Systems 1998 (FUZZ-IEEE'98)*, 1106–11. Piscataway, NJ: IEEE Service Centre.

Nguyen, H. T., and M. Sugeno (eds). (1998). *Fuzzy Systems Modeling and Control*. Boston: Kluwer Academic Publishers.

Odhiambo, L. O., R. S. Freeland, R. E. Yoder, and J. W. Hines. (2004). ''Investigation of a Fuzzy-Neural Network Application in Classification of Soils Using Ground-Penetrating Radar Imagery.'' *Applied Engineering in Agriculture* 20, 109–17.

Pedrycz, W., A. Kandel, and Y. Q. Zhang. (1998). ''Neurofuzzy Systems.'' In *Fuzzy Systems: Modeling and Control*, 311–80, edited by H. T. Nguyen and M. Sugeno. Boston: Kluwer Academic Publishers.

Robinson, V. B. (1990). ''Interactive Machine Acquisition of a Fuzzy Spatial Relation.'' *Computers and Geosciences* 16(6), 857–72.

Robinson, V. B. (2000). ''Individual and Multipersonal Fuzzy Spatial Relations Acquired Using Human–Machine Interaction.'' *Fuzzy Sets and Systems* 113, 133–45.

Robinson, V. B., M. Blaze, and D. Thoongs. (1986). ''Representation and Acquisition of a Natural Language Relation for Spatial Information Retrieval.'' In *Proceedings of the Second International Symposium of Spatial Data Handling*, 472–87, edited by D. F. Marble. Williamsville, NY: International Geographical Union.

Robinson, V. B., A. U. Frank, and H. Karimi. (1987). ''Expert Systems for Geographic Information Systems in Resource Management.'' *AI Applications in Natural Resources Management* 1, 47–57.

Sharma, J., D. M. Flewelling, and M. J. Egenhofer. (1994). ''A Qualitative Spatial Reasoner.'' In *Proceedings of Spatial Data Handling '94*, 665–81, edited by T. C. Waugh and R. G. Healey. London: Taylor & Francis.

Sugeno, M., and G. T. Kang. (1988). ''Structure Identification of Fuzzy Model.'' *Fuzzy Sets and Systems* 28, 15–33.

Takagi, T., and M. Sugeno. (1985). ''Fuzzy Identification of Systems and its Applications to Modeling and Control.'' *IEEE Transactions on Systems, Man, and Cybernetics* 15, 116–32.

Thill, J. C. (2000). ''Geographic Information Systems for Transportation in Perspective.'' *Transportation Research Part C* 8, 3–12.

Thill, J. C., and M. Mozolin. (2000). ''Feed-Forward Neural Networks for Spatial Interaction: Are They Trustworthy Forecasting Tools?'' In *Spatial Economic Science: New Frontiers in Theory and Methodology*, 355–81, edited by A. Reggiani. Heidelberg, Germany: Springer.

Thill, J. C., and D. Z. Sui. (1993). ''Mental Maps and Fuzzy Preferences.'' *The Professional Geographer* 45, 264–76.

Wang, F. (2003). ''Handling Grammatical Errors, Ambiguity and Impreciseness in GIS Natural Language Queries.'' *Transactions in GIS* 7, 103–21.

Worboys, M. F. (1996). ''Metrics and Topologies for Geographic Space.'' In *Advances in Geographic Information Systems Research II: Proceedings of the International Symposium on Spatial Data Handling*, 365–76, edited by M. J. Kraak and M. Molenaar. London: Taylor & Francis.

Worboys, M. F. (2001). ''Nearness Relations in Environmental Space.'' *International Journal of Geographical Information Science* 15(7), 633–51.

Yao, X., and J. C. Thill. (2005a). ''How Far is Too Far?—A Statistical Approach to Context–Contingent Proximity Modeling.'' *Transactions in GIS* 9(2), 157–78.

Yao, X., and J. C. Thill. (2005b). ''Spatial Queries with Qualitative Locations in Spatial Information Systems.'' *Computers, Environment and Urban Systems* 30(4), 485–502.

Zadeh, L. (1965). ''Fuzzy Sets.'' *Information and Control* 8, 339–353.