



ELSEVIER

Ecological Modelling 129 (2000) 87–99

**ECOLOGICAL
MODELLING**

www.elsevier.com/locate/ecolmodel

Species density of foliage-dwelling spiders in field margins: a simple, fuzzy rule-based model

Christian Kampichler ^{a,*}, Jutta Barthel ^b, Ralf Wieland ^c

^a *GSF-Research Centre of Environment and Health, Institute of Soil Ecology, Ingolstädter Landstraße 1, D-85758 Neuherberg, Germany*

^b *Department of Nature Conservation, Faculty of Biology, University of Marburg, Lahnberge, D-35032 Marburg, Germany*

^c *ZALF Centre for Agricultural Landscape and Land Use Research, Institute of Landscape Modelling, Eberswalder Straße 84, D-15374 Müncheberg, Germany*

Received 17 August 1999; received in revised form 6 January 2000; accepted 7 January 2000

Abstract

Field margins are grassy strips or banks between arable fields or meadows that are not cultivated or ploughed and contain only single shrubs or trees. They not only support beneficial arthropods, e.g. predators of crop pests, but can also be of potential value to species of concern, making them a critical target for conservation of biodiversity in agricultural landscapes. Most ecological data and most knowledge of ecological relationships are imprecise, uncertain or ambiguous. An overemphasis on the precision of mathematical models, thus, does not necessarily translate into a greater representation of reality. In contrast, fuzzy expert systems permit the representation and processing of 'soft' ecological knowledge in terms of natural language. This is done by sets of IF–THEN rules that relate the variables (expressed in the form of fuzzy sets such as 'rather cold', 'comparably few species', etc.) among each other. Data on 96 field margins in southern Germany and their foliage-dwelling spider assemblages, which had been collected from May to September 1994, were used for fuzzy modelling of the effects of margin width, margin density, frequency of mechanical disturbance and vegetation–architectural complexity on species density of spiders. Habitat factors were ordered according to their hierarchical effectiveness: margin width and disturbance determine habitat persistence, habitat persistence and margin density determine colonisation potential, and colonisation potential and herbaceous-plant cover determine species density. Forty-five rules are necessary to relate these factors. The predictive power of the fuzzy model was surprisingly high. The mean average error between predicted and observed number of species for test data, which was not used for model development, was less than 1.4 species (observed values ranged between 3 and 25). A multiple-regression model showed a mean average error of 3.17 species. We conclude that fuzzy models are suitable for representing the soft knowledge of field-margin/spider-assemblage relationships. The model permits time- and manpower-saving validation on larger scales by qualitative rather than quantitative data collection and by the inclusion of remote sensing data. We believe that fuzzy models could provide a valuable means for formulating

* Corresponding author. Present address: Soil Zoology and Ecology Laboratory, Institute of Biology, Free University Berlin, Grunewaldstraße 34, D-12165 Berlin. Tel.: +49-30-83853948; fax: +49-30-83853886.

E-mail address: kampichl@zedat.fu-berlin.de (C. Kampichler)

measures necessary for conservation of biodiversity in agricultural landscapes. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Field margins; Spiders; Margin width; Disturbance; Herb cover; Habitat persistence; Colonisation potential; Fuzzy logic; Fuzzy model; Rule-based model

1. Introduction

During the last decades modern farming methods have led to a considerable loss of uncultivated areas, e.g. permanent fallows, field margins and hedgerows, in the agricultural landscapes of Europe (Kaule, 1991). Many species of agricultural weeds and the arthropod fauna associated with them are now rare and restricted to field edges. Recently, the importance of field margins as wildlife habitats, as refuge areas and for pest control has increasingly attracted the attention of ecologists (Morris and Webb, 1987; Frei and Manhart, 1992; Boatman, 1994; Feber et al., 1995, 1996). Field margins are grassy strips or banks between arable fields or meadows. Normally, they are not cultivated or ploughed and contain only single shrubs or trees. Field margins not only support beneficial arthropods, e.g. predators of crop pests, but can also be of potential value to species of nature-conservational importance, making them a critical target for conservation of biodiversity in agricultural landscapes (Baines et al., 1998).

In the course of ongoing research activities of the FAM Munich Research Network on Agroecosystems, Barthel and Plachter (1996) and Anderlik-Wesinger et al. (1996) investigated seven agricultural areas in southern Bavaria (Germany). They studied the influence of habitat and landscape characteristics on the occurrence of foliage-dwelling spiders in 96 field margins by visual inspection of standardised plots of 1×50 m. By means of correlation and regression analyses, they found that margin density (= length of field margins per area), margin width, percent cover of herbaceous plants and the number of mechanical treatments (e.g. mowing) were the factors that most distinctively influenced the species number of spiders. For example, species number exhibited a weak, but significant correlation with margin density (Barthel and Plachter 1996, Fig. 2). The

relationship between species number and width of the field margins, in contrast, was distinctively non-linear, with species number increasing with increasing margin width in a range from 1 to 4 m and remaining constant with further increases in margin width (Barthel and Plachter 1996, Fig. 4).

How could these more or less purely statistical relationships be translated into a model that takes the underlying causal relationships into account and that is able to predict species numbers of spiders? Although elegant in many respects mathematical equations severely limit the type of knowledge that can be represented since much ecological knowledge is qualitative and fuzzy (Rykiel, 1989). In the overwhelming majority of cases ecologists do not communicate in the form of systems of differential equations or analytical models, but they use natural language and qualitative reasoning for the description of ecological relationships. Artificial intelligence provides a means for processing knowledge that is represented in natural language, for example rule-based expert systems (Liebowitz, 1998). In this case, knowledge is represented by rules that consist of an IF (antecedent) and a THEN (consequent) part. For example, we might state the hypothetical ecological rule 'IF habitat_factor_A_is_high AND habitat_factor_B_is_intermediate (antecedents) THEN species_number_is_low (consequent)'. Another particularly useful instrument for processing vague expert knowledge and uncertain or imprecise data is the fuzzy set theory developed by Zadeh (1965). Its central idea is that members of a set may have only partial membership, which consequently may possess all possible values between 0 ('is not a member of the set') and 1 ('is a member of the set'). In classical logic there are only two possibilities: either an object is member of a set or it is not; thus, the only possible membership values are 0 and 1. Consider a hypothetical rule with the antecedent 'IF field_margin_is_high' and assume that a field margin

is regarded to be high at a value of 5 m and more. According to classic set theory, for a margin which is 5.01 m wide, the antecedent is true and the rule will be activated; for a margin which is 4.99 m wide, however, the antecedent is false (Fig. 1a). Biologically, this does not make much sense. In fuzzy set theory the set 'high' may have the form which is shown in Fig. 1b: with increasing margin width, the membership in the set 'high' gradually rises from 0 to 1. For example, a field margin which is 5 m wide belongs to 'high' with the membership value 0.5. Briefly, the closer the membership of an element is to 1, the more it belongs to the set; the closer the membership of an element is to 0, the less it belongs to the set.

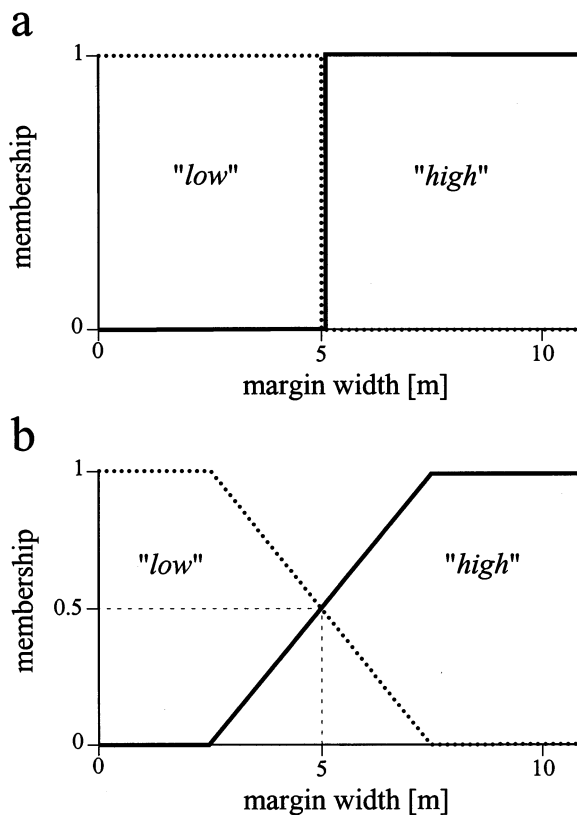


Fig. 1. Example for crisp sets (a) and fuzzy sets (b). Membership of the crisp sets 'low' (dotted line) and 'high' (unbroken line) have a sharp upper and lower border, respectively, while membership of the fuzzy sets 'low' (dotted line) and 'high' (unbroken line) increase and decrease gradually, respectively (b).

Thus, wherever we have been forced to draw artificially sharp distinctions in ecology, we can now draw more realistic boundaries by means of fuzzy sets. Fuzzy logic is a generalisation of Boolean logic; it provides means for a number of operations and can manage truth values between 'completely true' and 'completely false'. The combination of expert systems, fuzzy set and fuzzy logic to fuzzy rule-based models has proven to be a promising approach to environmental modelling, for example in water management (Pesti et al., 1996; Lee et al., 1997), impact assessment (van der Werf and Zimmer, 1998), population ecology (Daunicht et al., 1996; Schröder, 1997; Bock and Salski, 1998) or ecosystem-behaviour analysis (Uhrmacher et al., 1997). An excellent overview over the various applications of fuzzy logic can be found in two special issues of Ecological Modelling (volume 85, issue 1, Fuzzy Logic in Ecological Modelling; volume 90, issue 2, Fuzzy Modelling in Ecology) (Li and Rykiel, 1996; Salski, 1996).

The aim of this work was to represent the statistical results of Barthel and Plachter (1996) and Anderlik-Wesinger et al. (1996) and the ecological knowledge about the relationships between characteristics of field margins and species number of foliage-dwelling spiders by a fuzzy rule-based model and to test its predictive power.

2. Material and methods

2.1. Data origin

The investigations by Barthel and Plachter (1996) and Anderlik-Wesinger et al. (1996) were carried out in seven agricultural areas in the same geographic region, a hilly landscape between the rivers Isar and Danube in southern Germany (see Barthel and Plachter (1996), Fig. 1). The number of 1×50 m field-margin study plots ranged between 12 and 17 per study area, with the total number of study plots being 96. Data were collected by standardised visual search in the herbaceous vegetation. Each plot was sampled once per month from May to September 1994. For a more detailed description of the data collection, see Barthel and Plachter (1996) and Barthel (1997).

2.2. Model structure

The statistical relationships between field-margin characteristics and species density of spiders had to be translated into a framework of ecological interpretations. This included:

1. the division of each variable into fuzzy sets

For every factor, the adequate number and shape of the fuzzy sets had to be chosen. Although any number and any shape of fuzzy sets are possible, we tried to keep the model as simple as possible. We started with three sets per variable, thus defining a ‘low’, a ‘medium’ and a ‘high’ range of values. The shape of the fuzzy sets was chosen to be triangular or trapezoid, with memberships of the single fuzzy sets adding to 1 (cf. Fig. 2).
2. the construction of a rule base

The rule base must cover the entire variable space, meaning that, for any possible combination of habitat characteristics, a rule must be provided. (This is a general drawback of all rule-based systems and can cause considerable difficulties for larger models.) The structure of the rule base is up to the modeller. For example, if three habitat factors are included in the model and each variable is divided into three fuzzy sets, then the rules can be organised:

 - (a) in a single rule-node

In this case, all rules are of the form ‘IF habitat_factor_1_is_(attribute) AND habitat_factor_2_is_(attribute) and habitat_factor_3_is_(attribute) THEN species_number_is_(attribute)’, giving a total of $3^3 = 27$ rules.
 - (b) in two separate rule-nodes

In this case, two variables are combined to an intermediate variable by rules of the form ‘IF habitat_factor_1_is_(attribute) AND habitat_factor_2_is_(attribute) THEN intermediate_variable_is_(attribute)’. The intermediate variable is then combined with the third habitat factor with rules of the form ‘IF intermediate_variable_is_(attribute) AND habitat_factor_3_

is_(attribute) THEN species_number_is_(attribute)’. By structuring the model into two rule-nodes, the total number of rules can be kept smaller than with a single node at $3^2 + 3^2 = 18$ rules.

In analogy, for a larger number of input variables, the number of rules can be kept small by choosing an optimum model structure.

2.3. Fuzzy control

The main field of application of fuzzy logic is control; thus, the processing of rules is often called fuzzy control. Fuzzy control is organised in three steps: fuzzification, inference and defuzzification. For each step a number of methods is possible (Zimmermann, 1991). Here, only the methods used in this paper are presented, following the introductory texts by Traeger (1994) and Bothe (1995).

2.3.1. Fuzzification

Let A and B be fuzzy sets of a variable Var, e.g. the sets ‘low’ and ‘high’ of the variable ‘disturbance’ (Fig. 2). Now determine the membership functions, $\mu_A(x)$ and $\mu_B(x)$, that determine the membership of the measurement x in the sets A and B . For example, in Fig. 2 the measurement x ‘disturbance_events_per_year = 2’ would have the membership values $\mu_{\text{low disturbance}}(x) = 0.25$ and $\mu_{\text{high disturbance}}(x) = 0.75$.

2.3.2. Inference

Assume a case x, y with the measurement x for a first input variable $\text{Var1}_{\text{input}}$ and the measurement y for a second input variable $\text{Var2}_{\text{input}}$. Further, assume a rule of the form ‘IF $x = A$ AND $y = B$ THEN $x, y = C$ ’ that infers the membership function of the case x, y in the fuzzy set C of an output variable $\text{Var}_{\text{output}}$ when x is an element of fuzzy set A and when y is an element of a fuzzy set B . Combine the membership values $\mu_A(x)$ and $\mu_B(y)$ by the operation:

$$\mu_C(x, y) = \min \{ \mu_A(x); \mu_B(y) \} \quad (1)$$

and get the membership function of x, y in the fuzzy set C (min is the simplest operator for AND

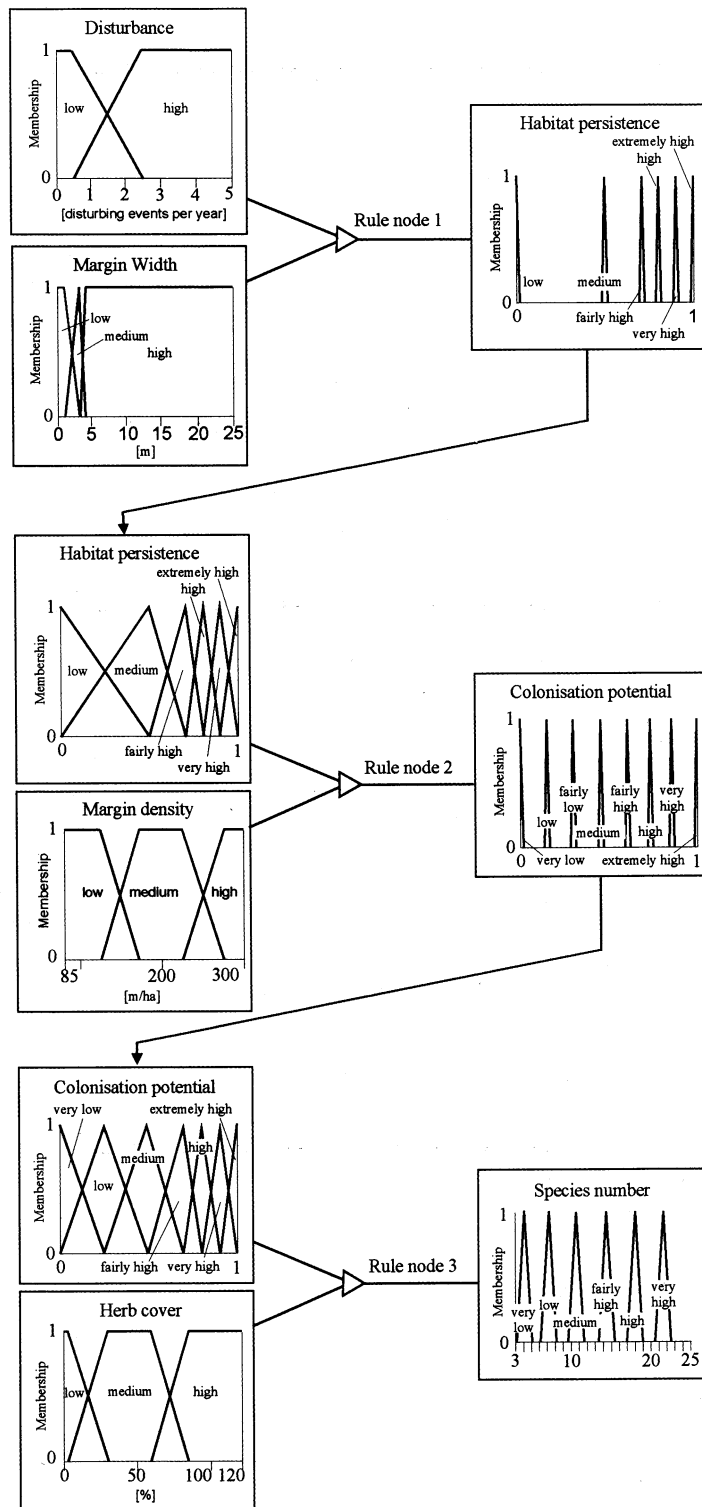


Fig. 2. Model structure of the fuzzy rule-based model for the prediction of species number of foliage-dwelling spiders in field margins. See the text for the explanation.

in fuzzy control). For example, take rule 1 (Appendix A) and let the measurements x and y be ‘disturbing_events_per_year = 2’ and ‘margin_width = 2 m’. The respective membership values are $\mu_{\text{low disturbance}}(x) = 0.25$ and $\mu_{\text{low margin width}}(y) = 0.5$ (Fig. 2). Then, the case x, y has the membership value $\mu_{\text{high habitat persistence}}(x, y) = \min \{0.25; 0.5\} = 0.25$.

The case above also activates the rules 2, 4 and 5 and, by analogy, leads to membership values of their consequents of $\mu_{\text{very high habitat persistence}}(x) = \min \{0.25; 0.5\} = 0.25$, $\mu_{\text{low habitat persistence}}(x) = \min \{0.75; 0.5\} = 0.5$ and $\mu_{\text{medium habitat persistence}}(x) = \min \{0.75; 0.5\} = 0.5$, respectively.

For a case where several activated rules have the same consequent, their membership values are combined by the operation:

$$\mu_{A \text{ or } B}(x, y) = \max \{ \mu_A(x, y); \mu_B(x, y) \} \quad (2)$$

For example, if a case x, y activates rule 8, leading to $\mu_{\text{low colonization potential}}(x, y) = 0.4$, and rule 10, leading to $\mu_{\text{low colonization potential}}(x, y) = 0.6$, then this operation would yield $\mu_{\text{low colonization potential}}(x) = \max \{0.4; 0.6\} = 0.6$.

Each fuzzy-set of the output variable is cut off at the height of its respective membership-value. For example, if a case x, y leads to $\mu_{\text{low colonization potential}}(x) = \max \{0.4; 0.6\} = 0.6$, then the fuzzy set ‘low’ is cut off at height 0.6. Fuzzy-sets that do not appear in the consequents of the activated rules are cut off at height 0. The result of the fuzzy inference process is the polygon that is eventually formed by all cut-off fuzzy-sets of the output variable.

2.3.3. Defuzzification

The result, in terms of membership values of fuzzy output sets, is back-translated into a ‘crisp’ output value. The centre of gravity of the polygon resulting from the inference process is projected onto the x -axis, and its x -coordinate x_{cg} can be calculated according to:

$$x_{\text{cg}} = \frac{\int_{x_1}^{x_r} x \cdot f(x) \, dx}{\int_{x_1}^{x_r} f(x) \, dx} \quad (3)$$

where x_l and x_r are the left and the right end of the polygon, and $f(x)$ is the border-line function of the polygon.

The defuzzification is least biased by the shape of the polygon if the fuzzy sets of the output variable are chosen as narrowly as possible; thus, not yielding a contiguous polygon, but several disjunctive narrow isosceles triangles (Fig. 2) or even single lines, so-called singletons. Eq. (3) then simplifies to:

$$x_{\text{cg}} = \frac{\sum_{i=1}^n x_i \mu_i}{\sum_{i=1}^n \mu_i} \quad (4)$$

where n is the number of narrow triangles (or singletons) of the output variable, x_i is the x -coordinate of the midpoint of the i -th triangle (or of the i -th singleton), and μ_i is the membership value of the i -th triangle (or of the i -th singleton). There might arise the impression that values in-between the non-overlapping singletons do not exist. The following example, however, shows that by computing the centre of gravity of the singletons and projecting it onto the x -axis the crisp model output may take any possible value: if a case x activates rules 43, 44 and 45 and yields the fuzzy output $\mu_{\text{fairly high species number}}(x) = 0.5$, $\mu_{\text{high species number}}(x) = 0.5$ and $\mu_{\text{very high species number}}(x) = 0.25$, then the crisp model output for the variable ‘species number’ would be $(14.5 * 0.5 + 18 * 0.5 + 21.5 * 0.25) / (0.5 + 0.5 + 0.25) = 17.3$.

As pointed out above, the shape of the fuzzy sets may be arbitrarily chosen (as long as ecologically meaningful). Since singletons are not suitable as input variables, the intermediate variables ‘habitat persistence’ and ‘colonisation potential’ are reshaped from singletons to triangular fuzzy-sets when used as input variables for the next rule-node (Fig. 2).

2.4. Model tuning and testing of predictive power

The 96 study plots were randomly split into a development set with 87 plots and a test set with nine plots. The development set was used for the tuning of the model. An initial model was run and the output of the model compared with the de-

sired output. By slightly altering the shapes (e.g. by shifting the feet of the triangular or trapezoid fuzzy-sets to lower or higher values) and number of the fuzzy sets of a variable (e.g. by dividing a variable into $n + 1$ instead of n fuzzy-sets) and/or by altering the rules, a number of model generations were developed and the desired output was progressively approximated.

When the model output eventually fitted the desired output, the model was tested with the test set, since the correspondence of predicted with observed values of unseen cases must be the criterion used for deciding whether there is a generalisable relationship between the independent and dependent variables (= predictive validation sensu Rykiel (1996)). In case that the model was overfitted, i.e. it was not able to generalise for unseen cases, the model generations were stepped backwards until the model generation was chosen that yielded the best compromise between two fundamental, but conflicting requirements of ecological models: precision (here: correspondence between model output and desired output according to the generation data) and generalisation (here: correspondence between model output and observed values of unseen cases) (Wissel, 1989). As the criterion for measuring correspondence, we used the mean absolute error (MAE), i.e. the mean absolute difference between prediction and observation, a deviance measure recommended by Mayer and Butler (1993) for model validation. We preferred MAE to the often-used coefficient of determination, R^2 , since, for a comparison of models, R^2 has to be adjusted for the respective degrees of freedom (Kvålseth, 1985). Degrees of freedom, however, cannot be unequivocally determined for a fuzzy model (B. Schröder, personal communication). The performance of the fuzzy rule-based model was compared with the predictive power of a multiple-regression model, with habitat factors as the independent and species number as the dependent variable.

Model generation and manipulation was performed with the Fuzzy Tool of MATLAB/SIMULINK.

3. Results

3.1. Model structure

The statistical analysis by Anderlik-Wesinger et al. (1996) had shown that four habitat factors had the strongest impact on the species density of the foliage-dwelling spider community, namely margin width, margin density (total length of margins per hectare), disturbance (number of mechanical disturbance events per year, such as mowing and ploughing) and architectural complexity of the foliage (percent cover of herbaceous plants in the margin vegetation). Thus, only these factors were taken into account for the development of the fuzzy model. The factors were arranged according to ecological plausibility in order to form a model structure that minimised the number of rules necessary.

Species density increased with increasing margin width, but levelled off at a margin width of approximately 3–4 m (Anderlik-Wesinger et al., 1996). We interpreted this observation as a demonstration of increasing insensitivity against mechanical disturbances: narrow field margins — and thus the spider community it accommodated — may be completely destroyed by even a single incidental mowing or ploughing event. Wider field margins had a higher chance that at least a narrow strip survived mechanical manipulation. Field margins of 3 m width and more faced nearly no risk of being erased even by frequent and regular disturbance. Moreover, Barthel and Plachter (1996) had observed that farmers were more disposed to mow or plough narrow margins than wider margins, thus leading to a negative correlation between margin width and disturbance. We described the relationship between margin width and disturbance by a set of rules and defined an intermediate variable, habitat persistence, relevant for the survival of the spider community (Fig. 2 and Fig. 3a).

Field margins that faced the risk of being frequently destroyed needed a high rate of colonising spiders from nearby habitats to accommodate a high species number. This is warranted only if a sufficient amount of suitable habitat is provided in the vicinity, characterised by

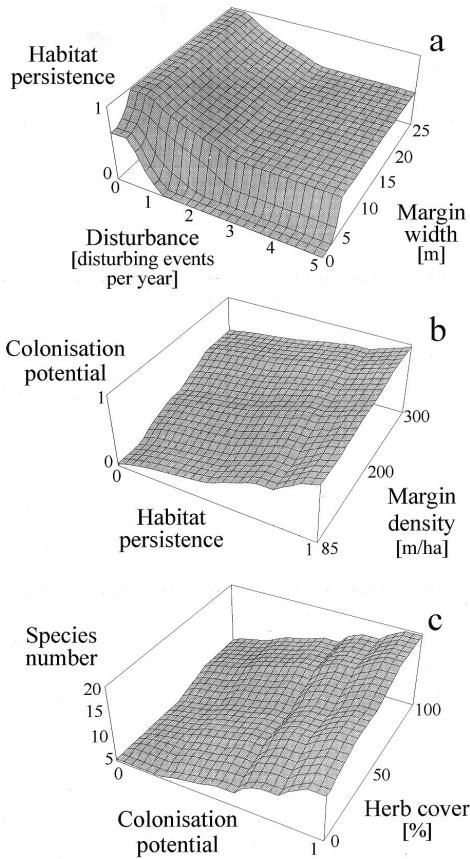


Fig. 3. Response surfaces of rule-node 1 relating margin width and disturbance to habitat persistence (a), rule-node 2 relating margin density and habitat persistence to colonisation potential (b) and rule-node 3 relating cover of herbaceous plants and colonisation potential to species number of foliage-dwelling spiders in field margins (c)

the margin density in a given area. Margins facing a negligible extinction risk may harbour large numbers of species even when the area shows only small margin density. The habitat persistence/margin-density relationships were described by a set of rules, defining another intermediate variable, colonisation potential (Fig. 2 and Fig. 3b).

Field margins with a high colonisation potential may differ in their suitability for harbouring large numbers of foliage-dwelling species by the percent

of herbaceous plants in the vegetation. Even with an optimum supply of an architecturally complex foliage, field margins may not harbour a rich spider community if the colonisation potential is low. We described the relationships between colonisation potential and herbaceous-plant cover by a third set of rules, yielding an estimate for the species number of spiders (Fig. 2 and Fig. 3c).

3.2. Model tuning and testing

The initial model structure outlined above was altered successively by modifying the number and shapes of the different fuzzy sets as well as the corresponding rules until a maximum correspondence between model prediction and desired output for the 87 cases of the development data was reached (MAE = 2.14). This model, however, showed a larger error for the test data (MAE = 5.30), demonstrating over-fitting of the development data. One step back in the series of model generations yielded less precision regarding the modelling of the development data (MAE = 2.96), but a much better predictive power for the test data (MAE = 1.38) (Fig. 4). The model is shown

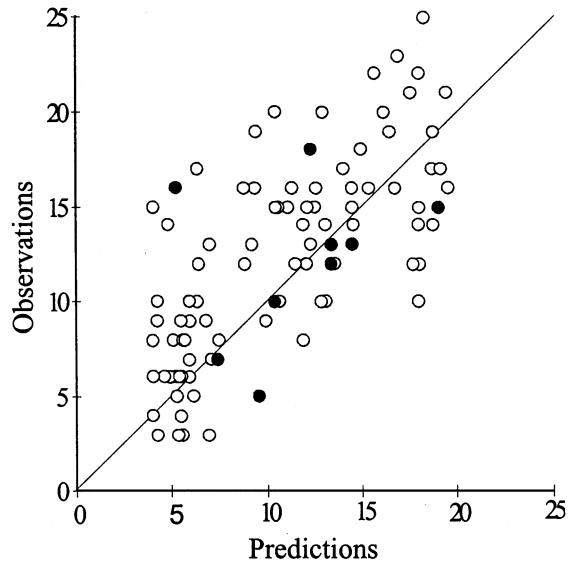


Fig. 4. Fuzzy rule-based model predictions for the species number of foliage-dwelling species in field margins (species per 50 m²) plotted against the observed species number. Development data (○): n = 87; test data (●): n = 9.

in Fig. 2, the complete set of 45 rules necessary for the chosen model structure is presented in Appendix A. Fig. 3 shows the response surfaces of the final model. They illustrate the quantitative relationships between input and output variables for each rule node. The non-linear response of habitat persistence to an increase in margin width and a decrease in disturbance is clearly visible (Fig. 3a). Margin density and cover of herbaceous plants responded in a more linear way to increases in margin density and habitat persistence (Fig. 3b), and colonisation potential and cover of herbaceous plants (Fig. 3c), respectively. Fig. 3b and c illustrate the assumed compensatory effects of the input variables (for example, high margin density can compensate for low habitat persistence and vice versa). They also illustrate the assumption that one of the input variables exerts a stronger influence on the output variable than another (for example, high colonisation potential completely compensates for low cover of herbaceous plants, whereas high cover of herbaceous plants has only a low compensative power).

To achieve linearity, margin width was subjected to a reciprocal transformation prior to calculating the multiple regression (Anderlik-Wesinger et al., 1996). The resulting regression model was $\text{species_number} = 8.27 - 2.13(1/\text{margin_width}) + 0.02 \text{ margin_density} - 1.53 \text{ disturbance} + 0.05 \text{ herb_cover}$. All factors were statistically significant ($P < 0.001$) except $1/\text{margin_width}$ ($P = 0.14$). The regression was slightly more precise than the fuzzy model (MAE = 2.73 for development data), but the MAE for the test data (3.17) was 130% higher than the MAE of the fuzzy model.

4. Discussion

4.1. Plausibility of model structure

The negative effect of mechanical disturbances on species numbers of spiders has been shown repeatedly: Baines et al. (1998) reported that cutting of field-margins reduced both abundance and species richness; Kajak (1971) and Barthel (1997)

(p. 88f) demonstrated a similar effect on the spider community of meadows. The sensitivity to destruction of the vegetation has also been shown for a number of individual foliage-dwelling species (Schaefer, 1978; Oxford, 1993). Since margin width has a considerable influence on species numbers of plants and animals (Molthan, 1990; Kaule, 1991) and since margin width and the frequency of mechanical disturbance in the study areas were significantly negatively correlated (Barthel and Plachter, 1996), the relationships described by rule-node 1 appear to be plausible.

Spiders have pronounced dispersal abilities and exhibit a high potential for colonising newly established habitats. For example, Thomas et al. (1992) reported rapid colonisation of newly sown grass strips from adjacent field margins; Asselin (1988) observed considerable immigration into margins that had been adversely affected by agricultural-management activities. According to metapopulation theory, we assume the colonisation rate to be distance-dependent (Hanski 1998) and relate it to margin density, which characterises both the distance and size of other populations. Also, Baines et al. (1998) described rapid colonisation of new margins, but species numbers remained poorer than in old margins throughout a 5-year experiment. Their observation supports the model-structure assumption that margin density can compensate for low habitat persistence to only a certain degree (Fig. 3b).

Foliage-dwelling spiders need a complex vegetation structure for satisfying their various foraging and reproduction (particularly supply of web and cocoon sites) demands. Grasses alone do not supply sufficient architectural complexity; thus, the richness of the spider community depends, in particular, on the diversity of the vegetation and on the presence of herbaceous plants (Gibson et al., 1992). Baines et al. (1998) attribute the beneficial effect of sowing a wild grass and forb mixture on spiders (as compared with natural regeneration) to the dominance of robust, branching species. Also, Anderlik-Wesinger et al. (1996) found a moderate, but highly significant correlation between the number of herb species and the species

density of spiders ($r = 0.47$, $P < 0.001$). Tuning the model led to a very narrow fuzzy set ‘low’ of the variable *cover_of_herbs*, suggesting that already a cover of $> 15\%$ supplied the community with sufficient spatial structure. Only very high herbaceous-plant cover ($> 70\%$) provided even better conditions. However, if *cover_of_herbs* was given too much weight in the rules relating it to *colonisation_potential*, this led to a considerable under- or overestimation of species density at low or high herbaceous-plant cover, respectively. The combined effects of disturbance, margin width and margin density, thus, were only slightly modulated by herbaceous-plant cover, illustrating a hierarchical influence of habitat factors.

4.2. Predictive power of the fuzzy model

The correspondence between prediction and observation for the test data is surprisingly good; the MAE of 1.38 is far better than was expected beforehand. Significant deviations from the predicted values, however, may be observed in the left upper corner of the scatter-plot of prediction versus observation (Fig. 4): some field margins that were expected to harbour only low numbers of species actually showed moderate to considerable species density. An underestimation of species numbers, in contrast, was rarely observed. We interpret this as a consequence of the data-collection method: mechanical disturbance by agricultural management measures could not be directly observed, but had to be identified afterwards by interpreting its consequences. Therefore, different types of treatment, e.g. mowing, cutting, ploughing or trampling, could not be discriminated and their intensity could not be estimated (Barthel and Plachter, 1996). The classification of different treatments into a single habitat variable, disturbance, hindered a closer differentiation of their impact on species number of spiders. Probably, a certain number of margins were exposed to low-intensity mechanical disturbance. The respective local spider assemblages were thus affected only to a minor extent, leading to the observed deviations from the model predictions. No modification of the model structure could correct for this problem.

4.3. Advantage of the fuzzy model

The majority of ecological knowledge is fuzzy. At first sight, uncertainty, imprecision and ambiguity might be considered removable artefacts that can be eliminated by increasing our knowledge. In reality, however, they are inevitable or inherent parts of natural systems. Overemphasised precision of models, thus, does not always translate into greater representation of reality (McBratney and Odeh, 1997). For example, how meaningful are the decimal places in the multiple regression model of spider species density given above? Fuzzy models do not pretend to offer an unrealistic degree of precision, but they can handle knowledge in the form in which it is normally communicated among ecologists, in natural language. This is a major advantage since ecologists, until recently, had no effective technology for using this knowledge in a meaningful way (Rykiel, 1989). By using natural language the data-collecting researcher himself (field or laboratory ecologist) is enabled to process his knowledge without being forced to translate it into mathematical formalisms (or to pass it on to a mathematically skilled modeller). The basic principles of fuzzy control are comparably easy to grasp and easy to apply. Easy-to-use tools for fuzzy modelling are available (cf. the URL <http://www.cs.cmu.edu/Groups/AI/html/faqs/ai/fuzzy/part1/faq.html>).

The non-linearities of ecological relationships are easily preserved in rule-based models. For example, for a widespread standard modelling approach, multiple regression, data must be linearized prior to calculation. This potentially causes difficulties in interpreting the resulting model when independent variables appear in root-transformed, logarithmic and/or reciprocal form, especially at high data dimensionality. By choosing an adequate number of fuzzy sets for each variable and a set of simple IF-THEN-rules, any non-linear relationship (e.g. smooth gradients, sharp discontinuities) known (or believed) to exist between two independent and a dependent variable can be represented (Fig. 3a). This characteristic of fuzzy models together with the reduced risk of overfitting — a least-square regression model by definition is fitted as tightly as possible to the

available data — make the approach suitable for predictive purposes. The fuzzy model on species density of foliage-dwelling spiders has an amazingly high predictive power; by processing information on only four habitat characteristics (margin width, margin density, disturbance, herb cover) predicted species density deviated, on average, less than 1.4 species from the observed value, a considerably better result than achieved by the linear regression model.

The fuzzy model also offers an advantage for further model validation. Since the model does not depend on crisp input data, future data collection may be carried out in a time-saving qualitative manner. Instead of precise metric measurements of the habitat variables, a fuzzy estimation is sufficient (e.g. ‘margin_width = rather narrow’, ‘herb_cover = somewhere between 20 and 40%’), because fuzzy logic provides a number of features (e.g. modifiers, fuzzy numbers) for relating qualitative data of that form to the fuzzy sets of the model (Traeger, 1994; Bothe, 1995). Also, data collected by remote sensing (margin width, margin density) could be used for further model validation, thus, reducing expenditure in time and man-power even more and allowing for testing the model’s generalisation potential for entire landscapes. If, at smaller or larger geographic scales, additional factors (e.g. the vicinity of certain agricultural managing systems; altitude, climatic conditions) turn out to be effective, the model can be altered by inserting supplementary rule-nodes at the proper hierarchical level within the model structure. We thus believe that a validated fuzzy model of the species density of foliage-dwelling spiders in field margins could provide a valuable means for elucidation measures necessary for conservation of biodiversity in agricultural landscapes.

Acknowledgements

The scientific activities of the FAM Munich Research Network on Agroecosystems (Forschungsverbund Agrarökosysteme München) are financially supported by the Federal Ministry of Culture and Science, Research, and Technol-

ogy. Rent and operating expenses are paid for by the Bavarian State Ministry for Education and Culture, Science and Arts.

Appendix A

Rules of the fuzzy rule-based model for species number of foliage-dwelling spiders in field margins. Read every rule as: IF antecedent_1 AND antecedent_2 THEN consequent (e.g. rule 1: IF disturbance_is_low AND margin_width_is_low THEN habitat_persistence_is_high)

Rule node 1

Rule	Disturbance	Margin width	Habitat persistence
1	Low	Low	High
2	Low	Medium	Very high
3	Low	High	Extremely high
4	High	Low	Low
5	High	Medium	Medium
6	High	High	Fairly-high

Rule node 2

Rule	Habitat persistence	Margin density	Colonisation potential
7	Low	Low	Very low
8	Low	Medium	Low
9	Low	High	Fairly low
10	Medium	Low	Low
11	Medium	Medium	Fairly low
12	Medium	High	Medium
13	Fairly high	Low	Low
14	Fairly high	Medium	Medium
15	Fairly high	High	Fairly high
16	High	Low	Medium
17	High	Medium	Fairly high
18	High	High	High
19	Very high	Low	Fairly high
20	Very high	Medium	High
21	Very high	High	Very high

22	Extremely high	Low	High
23	Extremely high	Medium	Very high
24	Extremely high	High	Extremely high

Rule node 3

Rule	Colonisation potential	Herb cover	Species number
25	Very low	Low	Very low
26	Very low	Medium	Very low
27	Very low	High	Low
28	Low	Low	Very low
29	Low	Medium	Low
30	Low	High	Medium
31	Medium	Low	Low
32	Medium	Medium	Medium
33	Medium	High	Fairly high
34	Fairly high	Low	Medium
35	Fairly high	Medium	Fairly high
36	Fairly high	High	High
37	High	Low	Medium
38	High	Medium	Fairly high
39	High	High	High
40	Very high	Low	Fairly high
41	Very high	Medium	High
42	Very high	High	Very high
43	Extremely high	Low	Fairly high
44	Extremely high	Medium	High
45	Extremely high	High	Very high

References

Anderlik-Wesinger, G., Barthel, J., Pfadenhauer, J., Plachter, H., 1996. Einfluß struktureller und floristischer Ausprägungen von Rainen in der Agrarlandschaft auf die Spinnen

- (Araneae) der Krautschicht. Verh. Ges. Ökol. 26, 711–720.
- Asselin, A., 1988. Changes in grassland use: consequences on landscape patterns and spider distribution. Münst. Geogr. Arb. 29, 85–88.
- Baines, M., Hambler, C., Johnson, P.J., Macdonald, D.W., Smith, H., 1998. The effects of arable field margin management on the abundance and species richness of Araneae (spiders). Ecography 21, 74–86.
- Barthel, J., 1997. Einfluß von Nutzungsmuster und Habitatkonfiguration auf die Spinnenfauna der Krautschicht (Araneae) in einer süddeutschen Agrarlandschaft. Agrarökologie Vol. 25. Verlag Agrarökologie, Bern.
- Barthel, J., Plachter, H., 1996. Significance of field margins for foliage-dwelling spiders (Arachnida, Araneae) in an agricultural landscape of Germany. Rev. Suisse Zool., vol. hors série, 45–59.
- Boatman, N., 1994. Field margins: integrating agriculture and conservation. British Crop Protection Council Monograph, Vol. 58, BCPC, Farnham.
- Bock, W., Salski, A., 1998. A fuzzy knowledge-based model of population dynamics of the Yellow-necked mouse (*Apodemus flavicollis*) in a beech forest. Ecol. Model. 108, 155–161.
- Bothe, H.-H., 1995. Fuzzy Logic, 2nd edn. Springer, Berlin.
- Daunicht, W., Salski, A., Nöhr, P., Neubert, C., 1996. A fuzzy knowledge-based model of annual production of skylarks. Ecol. Model. 85, 67–73.
- Feber, R.E., Johnson, P.J., Smith, H., Baines, M., Macdonald, D.W., 1995. The effects of arable field margin management on the abundance of beneficial arthropods. In: McKinlay, R.G., Atkinson, D. (Eds.), Integrated Crop Protection: Towards Sustainability? British Crop Protection Council Symposium Proceedings, vol. 63. BCPC, Farnham, pp. 163–170.
- Feber, R.E., Smith, H., Macdonald, D.W., 1996. The effects of arable field margin management on butterfly abundance. J. Appl. Ecol. 33, 1191–1205.
- Frei, G., Manhart, C., 1992. Nützlinge und Schädlinge an künstlich angelegten Ackerkrautstreifen in Getreidefeldern. Agrarökologie Vol. 4. Verlag Agrarökologie, Bern.
- Gibson, C.W., Hambler, C., Brown, V.K., 1992. Changes in spider (Araneae) assemblages in relation to succession and grazing management. J. Appl. Ecol. 29, 132–142.
- Hanski, I., 1998. Metapopulation dynamics. Nature 396, 41–49.
- Kajak, A., 1971. Productivity investigations of two types of meadows in the Vistula Valley. IX. Production and consumption of field layer spiders. Ecol. Pol. 19, 197–211.
- Kaule, G., 1991. Arten- und Biotopschutz, 2nd edn. Ulmer, Stuttgart.
- Kvålseth, T.O., 1985. Cautionary note about R^2 . Am. Stat. 39, 279–285.
- Lee, H.K., Oh, K.D., Park, D.H., Jung, J.H., Yoon, S.J., 1997. Fuzzy expert system to determine stream water quality classification from ecological information. Water Sci. Technol. 36, 199–206.

- Li, B.-L., Rykiel, E.J. Jr, 1996. Introduction. *Ecol. Model.* 90, 109–110.
- Liebowitz, J., 1998. *Handbook of Applied Expert Systems*. CRC Press, Boca Raton, FL.
- Mayer, D.G., Butler, D.G., 1993. Statistical validation. *Ecol. Model.* 68, 21–32.
- McBratney, A.B., Odeh, I.O.A., 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma*, 77, 85–113.
- Molthan, J., 1990. Artenspektren, Dominanzverhältnisse und Abundanzdynamik von Schwebfliegen (Diptera, Syrphidae) in Feldrandbiotopen im Hessischen Ried. *Mitteilung. Ges. Allg. Angew. Entomol.* 7, 368–379.
- Morris, M.G., Webb, N.R., 1987. The importance of field margins for the conservation of insects. In: Way, J.M., Grieg-Smith, P.W. (Eds.), *Field Margins*. British Crop Protection Council Monograph, Vol. 35. BCPC, Thornton Heath.
- Oxford, G.S., 1993. Components of variation in population size in the spider, *Enoplognatha ovata* (Clerck) sensu stricto (Araneae: Theridiidae). *Bull. Br. Arachnol. Soc.* 9, 193–202.
- Pesti, G., Shrestha, B.P., Duckstein, L., Bogardi, I., 1996. A fuzzy rule-based approach to drought assessment. *Water Resour. Res.* 32, 1741–1747.
- Rykiel, E.J. Jr, 1989. Artificial intelligence and expert systems in ecology and natural resource management. *Ecol. Model.* 46, 3–8.
- Rykiel, E.J. Jr, 1996. Testing ecological models: the meaning of validation. *Ecol. Model.* 90, 229–244.
- Salski, A., 1996. Introduction. *Ecol. Model.* 85, 1–2.
- Schaefer, M., 1978. Some experiments on the regulation of population density in the spider *Floronia bucculenta* (Araneidae: Linyphiidae). *Symp. Zool. Soc. Lond.* 42, 203–210.
- Schröder, B., 1997. Fuzzy Logik und klassische Statistik-ein kombiniertes Habitatsignungsmodell für *Conocephalus dorsalis* (LATREILLE, 1804) (Orthoptera: Tettigoniidae). *Verh. Ges. Ökol.* 27, 219–226.
- Thomas, M.B., Wratten, S.D., Sotherton, N.W., 1992. Creation of 'island' habitats in farmland to manipulate populations of beneficial arthropods: predator densities and species composition. *J. Appl. Ecol.* 29, 524–531.
- Traeger, D.H., 1994. *Einführung in die Fuzzy-Logik*, 2nd edn. B.G.Teubner, Stuttgart.
- Uhrmacher, A.M., Cellier, F.E., Frye, R.J., 1997. Applying fuzzy-based inductive reasoning to analyse qualitatively the dynamic behaviour of an ecological system. *AI Appl.* 11, 1–10.
- van der Werf, H.M.G., Zimmer, C., 1998. An indicator of pesticide environmental impact based on a fuzzy expert system. *Chemosphere* 36, 2225–2249.
- Wissel, C., 1989. *Theoretische Ökologie*. Springer, Berlin.
- Zadeh, L., 1965. Fuzzy sets. *Inf. Control* 8, 338–353.
- Zimmermann, H.-J., 1991. *Fuzzy Set Theory and its Applications*. Kluwer, Dordrecht.