

Assessing The Effects Of Zero Abundance Data On Habitat Preference Modelling Using A Genetic Takagi-Sugeno Fuzzy Model

Shinji Fukuda

Institute of Tropical Agriculture
Kyushu University
Fukuoka, Japan
shinji-fkd@agr.kyushu-u.ac.jp

Abstract—In this paper, the effects of zero abundance data on fish habitat modelling using a genetic Takagi-Sugeno fuzzy system were assessed with specific focus on habitat preference curves (HPCs) and model performance. Three independent data sets were prepared from a series of fish habitat surveys conducted in an agricultural canal in Japan. To quantify the effects of zero abundance data, two kinds of data (full abundance data and presence-only abundance data) were used, from which a fuzzy habitat preference model (FHPM) were developed. As a result, the HPCs obtained using presence-only abundance data resulted in similar HPCs across the different data sets used, while those obtained using full abundance data differed by the data sets. Because the model performance with regard to generalization ability was higher, the present study concluded that the use of presence-only abundance data can better capture the habitat preference of the target species.

Keywords—genetic fuzzy systems; species distribution; habitat model; model error; uncertainty

I. INTRODUCTION

Genetic fuzzy systems (GFSs) have been widely applied to various research fields because of their ability to solve different types of problems [1], [2]. The GFSs have been applied to many disciplines using such approach as linguistic fuzzy models [3] and precise fuzzy models [4], [5] (for a review, see [6]). In ecological modelling, GFSs have also been used for tuning the membership functions of Mamdani–Assilian models [7] and the consequent part of Takagi–Sugeno models [8]–[16]. They are data-driven models, which have been reported to outperform knowledge-based models in the case of habitat modelling [17]. These application studies support the strengths of GFSs, i.e., the flexibility of the model structures and their ability to tune knowledge-based systems according to the new data sets for a given problem. Because the applicability of GFSs depends on the target systems and available data, it is necessary to accumulate both theoretical and application studies to better understand the ability and limitations of GFSs and the related modelling approaches.

In practical implementation of species distribution models (SDMs), presence-only data are often used because of the availability compared to presence-absence or abundance data [18]–[21]. On the one hand, presence data indicate clear

evidence that target fish were observed at a surveyed point. On the other hand, zero abundance data (also referred to as absence data in presence-absence analysis) contain essential uncertainty. For instance, there could be a case where target species were absent due to the presence of enemy species or simply because the habitat was not yet occupied by the species. This uncertainty has been used as an explanation for the errors in SDM-related studies [17], [22], [23]. Despite the previous researches [24], [25] on the effects of absence data or prevalence of the data on the predictive performance of the models used, no study evaluated the effects of zero abundance data on habitat preference evaluation. Therefore, it is important to understand the behaviour of a target model in response to different qualities of data, from which existing models can be improved and also new models can be further developed.

The aim of this study is to assess the effects of zero abundance data on model performance and habitat preference evaluation using a genetic algorithm (GA) [26]-optimized fuzzy habitat preference model (FHPM). A series of fish habitat surveys focusing on Japanese medaka (*Oryzias latipes*) were conducted, from which three independent data sets were derived for the analyses. The effects of zero abundance data were assessed on the basis of habitat preference curves (HPCs) and predictive accuracy of spatial distribution of the fish.

II. METHODS

A. Study site and target fish

A series of field surveys was carried out in an agricultural canal located in Kurume City, Fukuoka, Japan (33°20' N, 130°42' E). The spring-fed canal runs through paddy fields and is used for both irrigation and drainage purposes. The Japanese medaka is one of the most common fish in paddy environments in Japan. This fish is considered to be one of the symbols for the conservation and restoration of rural environments, because the fish is vulnerable to changes in physical habitat conditions such as current velocity. This is partly ascribed to the body length of the Japanese medaka (2–4 cm in length for an adult).

B. Fish Habitat Surveys

The field surveys were conducted on 14 October, 5 and 9 November 2004. The surveys were conducted during a non-

This work was supported in part by the Grant-in-aid for Young Scientists B (21780225) of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

irrigation period, and thus the discharge in both reaches remained stable. The habitat use of the Japanese medaka and four physical habitat characteristics of water depth (cm, henceforth referred as depth), current velocity (cm s^{-1} , velocity), lateral cover ratio (%), cover) and percent vegetation coverage (%), vegetation) in the study reach were surveyed on sunny days. After mapping the reach, the habitat use of the fish was observed, and then the physical habitat characteristics within the reach were surveyed.

The habitat use of Japanese medaka was observed visually from the bank; the observer moved slowly and carefully to avoid disturbing the fish. The number of the fish was counted in units of five to consider the habit of school formation; fish in a small school (less than five) were not counted. The fish observations were repeated eight times and the results were averaged in order to reduce observational errors.

Immediately after completing the fish observation, the four physical habitat variables of depth, velocity, cover and vegetation were surveyed to establish a relationship between the habitat variables and the habitat preferences of Japanese medaka. First, depth and velocity were measured, in which depth was measured with a stainless steel ruler and velocity was measured with a portable propeller current meter (KENEK, V-303) at three lateral points comprising a midpoint and two near-shore points at longitudinal intervals of 1 m. Next, the study reach were divided into small water units having a similar condition with regard to depth and velocity. The other two factors of cover and vegetation were then calculated at each water unit. The lateral cover ratio is defined as a function of the presence of a lateral cover, which comprises the water's edge, a dike, or anything that emerges from the water surface and surrounds the water body. The cover thus consists of four components (four lateral sides). The maximum cover ratio is 100%, and each of the cover

components is assigned a score of 25%. In the definition of the cover, objects attached to more than 90% of the boundary between water units were regarded as cover components. That is, only instream and undersurface cover structures that may have the same effects as the margin of the stream were considered as cover components. Percent vegetation coverage is defined as the percentage of the area covered with aquatic vegetation in each water unit. Both submerged and emergent vegetation were pooled because of their same roles in providing foods and shelters from predators and fast-flowing currents.

The fish distribution data used in the following analyses were the log-transformed fish population density observed at the i^{th} water unit $\text{FPD}_{0,i}$ (individuals per square metre), where i ($= 1, 2, \dots, N$) denotes the index of the water units and N is the total number of water units (Fig. 1).

C. Fuzzy Habitat Preference Model

In the present study, a fuzzy habitat preference model (FHPM; [8]) was employed. The FHPM is a 0-order Takagi-Sugeno model [27] which relates habitat variables to habitat preference by considering uncertainties such as fish behaviour and measurement errors of the physical habitat variables. FHPMs have been applied to modelling the habitat use of Japanese medaka in a laboratory experiment [7] and at a microhabitat scale [9–15]. The FHPMs has also been applied to habitat evaluations at a larger spatial scale [16]. A FHPM can be viewed as a fuzzified version of a GA-optimized piecewise constant model [12], in which simultaneous optimization of all model parameters enables an FHPM to evaluate habitat preference, even under the nonlinear, complex interactions between habitat variables and habitat preference [11]. The brief description of an FHPM follows.

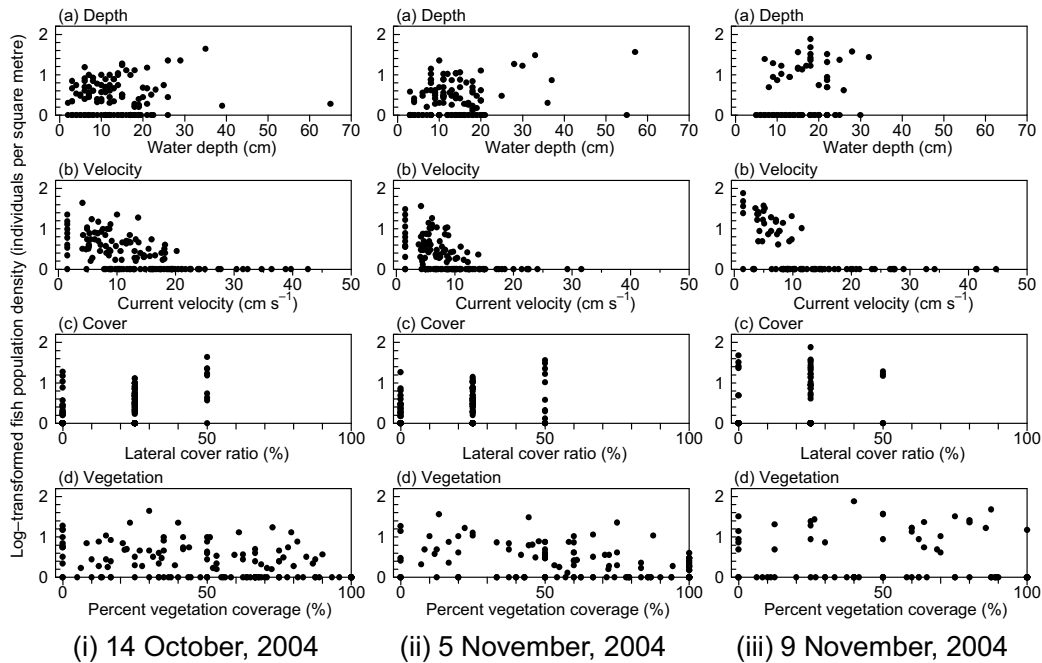


Figure 1. Results of fish habitat surveys.

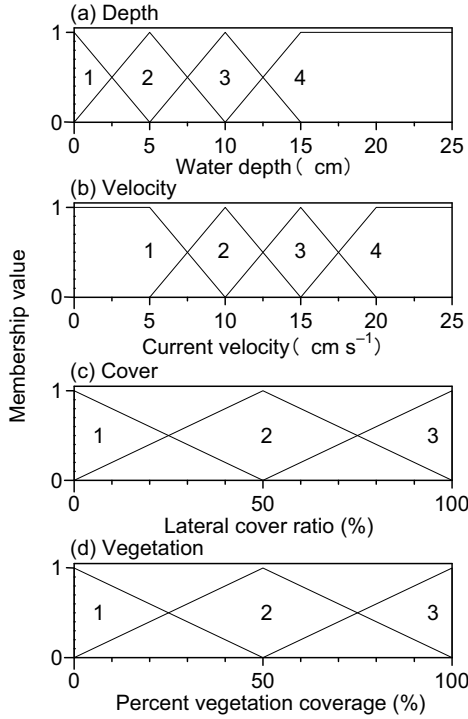


Figure 2. Membership functions.

The present FHPM is a single-input single-output fuzzy system and has 4 sets of if-then rules, comprised of 4 rules for depth and velocity, and 3 rules for cover and vegetation (14 rules in total). The input values (four habitat variables) were expressed by symmetric triangular fuzzy numbers with centre α_m and spread σ_m , where m denotes the habitat conditions of depth, velocity, cover and vegetation, to consider the measurement errors and the spatial variances of the habitat variables. The centre α_m is the observed value and the spread σ_m is determined from the allowable variances when dividing the reach into water units: 1 cm was given for the spread of depth σ_d , 1 cm s⁻¹ for velocity σ_v and 10% each for cover σ_c and vegetation σ_{veg} . The abbreviations d, v, c and veg denote depth, velocity, cover and vegetation, respectively. The antecedent parts of the FHPM (the membership functions) were defined to extract habitat requirements of the target fish (Fig. 2). Each membership function has a corresponding singleton value in the consequent part. The habitat preference was calculated by taking a weighted mean of the singleton values in the consequent parts, with the membership degrees as weights. The singleton values in the consequent part are in the range [0, 1] and were determined by a GA so as to minimize the mean squared error (MSE) between the modelled and observed fish population density. That is, the fitness function of the GA is a reciprocal of the MSE. The present GA consisted of a population of 100 individuals, each of which has a 56-bit string (4 bits $\times$ 14 singletons). The shape of the habitat preference curves (HPCs), i.e., a set of singletons for a habitat variable, was constrained to have a unimodal or monotone form with maximum preference of one. Specifically, the model was penalized by giving large MSE values when the above regulation was violated. In the estimation of the fish population

density, a composite habitat preference according to the four habitat variables was calculated as

$$HP_i = P_{d,i} \times P_{v,i} \times P_{c,i} \times P_{veg,i} \quad (2)$$

where HP_i denotes the habitat preference in the i^{th} water unit, and $P_{d,i}$, $P_{v,i}$, $P_{c,i}$, and $P_{veg,i}$ are the habitat preference to individual habitat variables of depth, velocity, cover and vegetation, respectively. The product operation is reported to produce a FHPM with higher performance with consistent habitat preference curves [16]. The habitat use of Japanese medaka was then estimated using the equation:

$$FPD_{c,i} = \left(HP_i / \sum_{i=1}^N HP_i \right) \cdot \sum_{i=1}^N FPD_{o,i} \quad (3)$$

where $FPD_{c,i}$ is the calculated fish population density in the i^{th} water unit, and $FPD_{o,i}$ is the observed fish population density. The MSE was calculated accordingly and the GA repeatedly modified the model structure (the singletons in the consequent part) using three genetic operations: roulette wheel selection with an elitist strategy, uniform crossover and mutation at a probability of 5%. After 2,000 iterations, the optimal model was obtained. The GA optimization was repeated using 20 different sets of initial conditions.

D. Model Development and Analysis

To assess the effects of zero abundance data (i.e., data at which no fish was present) on the habitat evaluation of Japanese medaka, two kinds of data were used: full abundance data (henceforth referred to as full data) and the data without zero abundance (henceforth referred to as presence-only abundance data). The full data consist of both abundance data and zero abundance data, whereas presence-only abundance data consist of only abundance data (i.e., full data without zero abundance data). FHPMs were developed using three independent data sets with two kinds (full and presence-only abundance data), in which a data set was used for model calibration and each of the remaining two data sets were used for model validation. HPC for a habitat variable was obtained by providing consecutive values, which were in the range of the corresponding membership functions (in steps of 0.1), to the FHPM of which the output values were plotted against the habitat variable. HPCs were evaluated according to the shapes and variances of all HPCs from 20 different initial conditions. Consistent HPCs indicate that consistent ecological knowledge was extracted by FHPMs. A variance in HPCs reflects the variance in model structures optimized using 20 different initial conditions in the GA optimization, and quantifies the convergence of model parameters. Consistency or a smaller variance in HPCs would thus demonstrate the appropriateness of the model developed.

The accuracy of the model was evaluated based on the MSE between predicted and observed fish population density. The MSE was calculated for the three cases of species distribution modelling of the fish: prediction using full data by FHPMs developed from full data, prediction using full data by FHPMs developed from presence-only abundance data, and prediction using presence-only abundance data by FHPMs

developed from presence-only abundance data. Mean and standard deviation of the MSE values across 20 different initial conditions were used as performance measures. The smaller MSE values in validation data sets indicate the higher generalization ability of the models.

III. RESULTS

The full data and presence-only abundance data result in slightly different HPCs (Figs. 3–4), in which differences are also found across data sets used. The difference can be seen from the shape or trend of HPCs and the degree of preference. For instance, it is observed from the HPCs for depth derived using the full data (Figs. 3a(i)–3a(iii)) that Japanese medaka

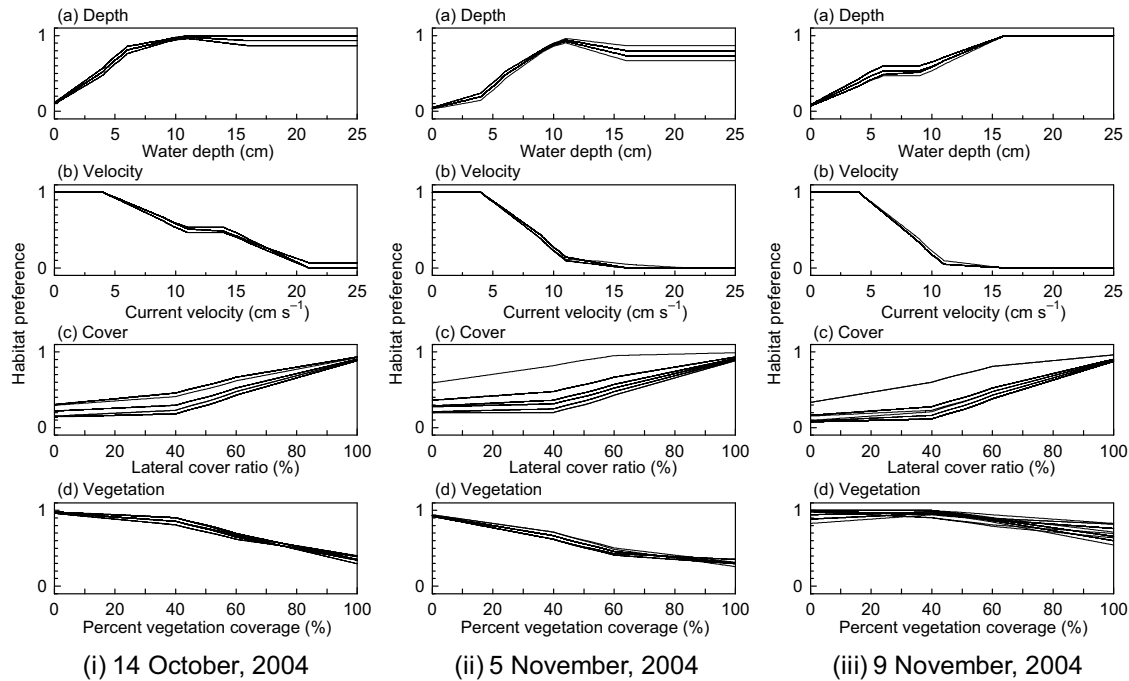


Figure 3. Habitat preference curves of fuzzy habitat preference models developed using full abundance data

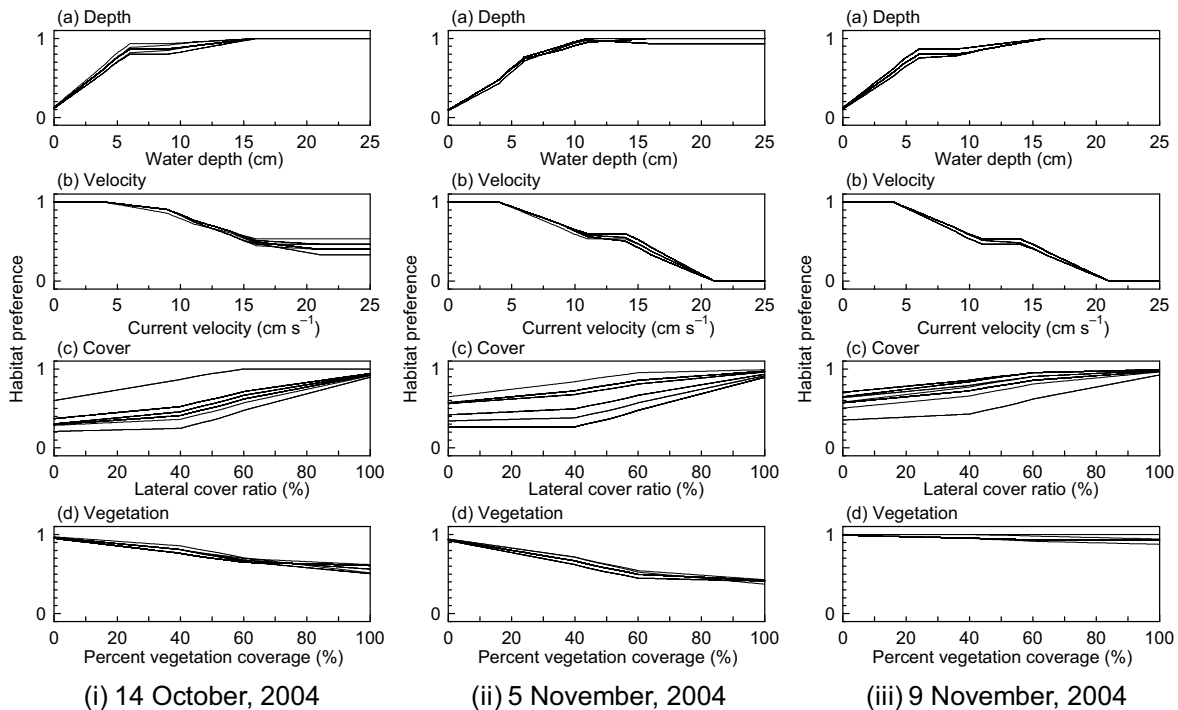


Figure 4. Habitat preference curves of fuzzy habitat preference models developed using presence-only abundance data.

TABLE I. MEAN SQUARED ERRORS BETWEEN PREDICTED AND OBSERVED FISH POPULATION DENSITY, IN WHICH RESULTS FROM DIFFERENT SETTING ARE SHOWN: (A) FUZZY HABITAT PREFERENCE MODELS (FHPMs) CALIBRATED AND EVALUATED USING FULL ABUNDANCE DATA, (B) FHPMs CALIBRATED USING PRESENCE-ONLY ABUNDANCE DATA AND EVALUATED USING FULL ABUNDANCE DATA, AND (C) FHPMs CALIBRATED AND EVALUATED USING PRESENCE-ONLY ABUNDANCE DATA. EACH OF THREE DATA SETS WAS USED FOR MODEL CALIBRATION AND EVALUATION.

(a) FHPM	Data Used For Model Evaluation		
	14 Oct.	5 Nov.	9 Nov.
Model 1	0.0691 ±0.0005	0.0842 ±0.0017	0.2163 ±0.0057
Model 2	0.1457 ±0.0071	0.0665 ±0.0004	0.2056 ±0.0088
Model 3	0.2399 ±0.0418	0.1148 ±0.0162	0.1431 ±0.0013
(b) FHPM ₀₋₁			
Model 1	0.0933 ±0.0017	0.1091 ±0.0012	0.2613 ±0.0036
Model 2	0.0768 ±0.0011	0.0927 ±0.0015	0.2361 ±0.0051
Model 3	0.0846 ±0.0018	0.1064 ±0.0025	0.2127 ±0.0049
(c) FHPM ₀₋₂			
Model 1	0.0535 ±0.0004	0.0648 ±0.0011	0.1368 ±0.0072
Model 2	0.0698 ±0.0025	0.0566 ±0.0005	0.1862 ±0.0128
Model 3	0.0754 ±0.0028	0.0820 ±0.0020	0.0697 ±0.0005

Mean ± standard deviation are shown.

Model 1 was developed using the data set on 14 October.

Model 2 was developed using the data set on 5 November.

Model 3 was developed using the data set on 9 November.

prefers a larger water depth as a trend, while different degrees of preference around the depth over 10 cm are observed between the three data sets. For velocity, the degrees of preference at the velocity of 10 to 15 cm s⁻¹ are found to be different between the data sets of 14 October and the other data sets (Figs. 3b(i)–3b(iii)). The difference of HPCs are not clear for cover (Figs. 3c(i)–3c(iii)), of which the convergence are slightly better in the data sets on 14 October. For vegetation, the degree of preference to the larger vegetation is higher in the data sets on 9 November (Figs. 3d(i)–3d(iii)). These differences between data sets can be the source of errors in validation data sets (Table Ia). Large errors are observed at model 2 for the data set of 14 October and model 3 for the data sets of 14 October and 5 November, which is partly ascribed to the overestimations of fish population density.

In the case where presence-only abundance data were used, the differences in HPCs between data sets are found to be small (Fig. 4). While the degrees of preference are slightly different between the HPCs from different data sets, the trends of HPCs are the same across the data sets. The slight differences in the degree of preference are observed at the HPCs for velocity over 20 cm s⁻¹ and those for vegetation (Figs. 4b and 4d). The HPCs for vegetation developed using the presence-only abundance data show weak trend over the variable range (Fig. 4d(iii)), which would indicate a small contribution of the variable in modelling spatial distributions of the fish using the data set. The HPCs obtained from presence-only abundance data result in worse accuracy when testing the FHPMs by the full data

(Table Ib) but better accuracy when testing the FHPMs by the presence-only abundance data (Table Ic). Despite the different accuracy between the FHPMs developed using the presence-only abundance data (Table Ib–Ic), the accuracies are better than that of FHPMs developed using the full data (Table Ia).

IV. DISCUSSION

The present results illustrated that the zero abundance data affect model performance of the FHPMs (Figs. 3–4, Table I), which was observed through the HPCs and resulting MSE values. The reduced standard deviation values of MSEs indicate that the zero abundance data can be a source of uncertainties in species distribution modelling (Table I). This is clear, to some extent, because zero abundance contains two possible cases of true and chance absences. The former means that the fish was absent because the fish avoided it, while the latter means that the fish was absent because the habitat was not yet occupied or because of presence of predator or competitive species, even the habitat was preferable for the fish. As a result, model predictions in the former case are evaluated to be perfect fits, while model predictions in the latter case are to be errors. This type of errors is the same as the errors that are reported as false positives in presence-absence modelling [17], [22], [23]. It is also reported that the proportion of presence data to total data, often termed as prevalence, affects the performance of SDMs (see [24], [25] for reviews). Accordingly, it could be concluded that the model performance is affected by not only the prevalence of data but also the proportion of true and chance absences in absence data or zero abundance data. However, it is impossible to distinguish true and chance absences at the moment. Novel methods are, therefore, needed to be developed for better understanding the model behaviour to uncertainties associated with the data used.

The present results demonstrated the higher generalization ability of FHPMs developed using presence-only abundance data, which can be mainly because of the consistent HPCs obtained (Figs 3–4). For instance, the present HPCs show a consistent trend of habitat preference that the fish prefer a larger depth, a slower velocity, a larger cover and a smaller vegetation with a small difference of the degree of preference. This supports the appropriateness of presence-only abundance data in species distribution modelling, especially where HPCs are necessary or important. The differences in HPCs have been reported as a ‘transferability’ of a habitat model [10], [14], [28]–[31]. References [32], [33] suggested the necessity of a response curve of a model, i.e., a HPC in the present study, for a better understanding towards an improvement of the existing models or development of a new model. Since the response curve of a model illustrate how target species response to a given habitat condition, SDMs should be compared on the basis of both accuracy measures and response curves derived from the models.

V. CONCLUSION

This case study of Japanese medaka demonstrated the effect of zero abundance data in modelling fish habitat preference using a 0-order Takagi-Sugeno fuzzy system. The dual nature of zero abundance (i.e., true and chance absences) would be one of the sources of performance deterioration of the model.

Further studies are needed to develop an approach to distinguish true and chance absences, which contributes to the better understandings of target species and the improvement of SDMs

ACKNOWLEDGMENT

The author thanks Prof. K. Hiramatsu at Faculty of Agriculture, Kyushu University for his sincere assistance in the early phase of this study.

REFERENCES

- [1] O. Cordon, F. Herrera, F. Hoffmann, and L. Magdalena, "Genetic fuzzy systems. Evolutionary tuning and learning of fuzzy knowledge bases," World Scientific, Singapore, 2001.
- [2] F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects," *Evol. Intel.*, vol. 1, pp. 27–46, 2008.
- [3] H. Ishibuchi, T. Nakashima and M. Nii, "Classification and modeling with linguistic information granules: Advanced approaches to linguistic data mining," Springer Verlag, 2004.
- [4] N. Xiong, "Evolutionary learning of rule premises for fuzzy modelling," *Int. J. Syst. Sci.*, vol. 32, pp. 1109–1118, 2001.
- [5] O. Cordon and F. Herrera, "A two-stage evolutionary process for designing TSK fuzzy rule-based systems," *IEEE Trans. Systems Man Cybernet. B*, vol. 29, pp. 703–715, 1999.
- [6] M.J. Gacto, R. Alcalá and F. Herrera, "Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures," *Inform. Sci.*, in press.
- [7] E. Van Broekhoven, V. Adriaenssens, and B. De Baets, "Interpretability-preserving genetic optimization of linguistic terms in fuzzy models for fuzzy ordered classification: An ecological case study," *Int. J. Approx. Reasoning.*, vol. 44, pp. 65–90, 2007.
- [8] K. Hiramatsu, S. Fukuda, and S. Shikasho, "Mathematical modelling of habitat preference of Japanese medaka for instream water environment using fuzzy inference," *Trans. JSIDRE*, vol. 228, pp. 65–72, 2003 (in Japanese with English abstract).
- [9] S. Fukuda, K. Hiramatsu, M. Mori, and S. Shikasho, "Mathematical Characterization of fuzziness in fish habitat preference of Japanese medaka (*Oryzias latipes*) in agricultural canal," *Trans. JSIDRE*, vol. 239, pp. 43–49, 2005 (in Japanese with English abstract).
- [10] S. Fukuda and K. Hiramatsu, "Prediction ability and sensitivity of artificial intelligence-based habitat preference models for predicting spatial distribution of Japanese medaka (*Oryzias latipes*)," *Ecol. Model.*, vol. 215, pp. 301–313, 2008.
- [11] S. Fukuda and S. Okushima, "Assessing nonlinearity in fish habitat preference of Japanese medaka (*Oryzias latipes*) using genetic algorithm-optimized habitat prediction models," *JARQ-Jpn. Agric. Res. Quart.*, vol. 42 (2), pp. 97–107, 2008.
- [12] S. Fukuda, "Consideration of fuzziness: Is it necessary in modelling fish habitat preference of Japanese medaka (*Oryzias latipes*)?" *Ecol. Model.*, vol. 220, pp. 2877–2884, 2009.
- [13] S. Fukuda, "A preliminary analysis for improving model structure of fuzzy habitat preference model for Japanese medaka (*Oryzias latipes*)," *Proc. of 13th IFSA World Congress and 6th Conference of EUSFLAT*, pp. 1258–1263, 2009.
- [14] S. Fukuda, "Assessing transferability of genetic algorithm-optimized fuzzy habitat preference models for Japanese medaka (*Oryzias latipes*)," *Proc. of GEFS2010*, pp. 57–62, 2010.
- [15] S. Fukuda, "Effect of data quality on habitat preference evaluation for Japanese medaka (*Oryzias latipes*) using a simple genetic fuzzy system," *Proc. of FUZZ-IEEE2010*, pp. 1231–1237, 2010.
- [16] S. Fukuda, B. De Baets, A.M. Mouton, W. Waegeman, J. Nakajima, T. Mukai, K. Hiramatsu, and N. Onikura, "Effect of model formulation on the optimization of a genetic Takagi-Sugeno fuzzy system for fish habitat suitability evaluation," *Ecol. Model.*, vol. 222, pp. 1401–1413, 2011.
- [17] A.M. Mouton, B. De Baets, and P.L.M. Goethals, "Knowledge-based versus data-driven fuzzy habitat suitability models for river management," *Environ. Modell. Softw.*, vol. 24, pp. 982–993, 2009.
- [18] A. Guisan, and N.E. Zimmermann, "Predictive habitat distribution models in ecology," *Ecol. Model.*, vol. 135, pp. 147–186, 2000.
- [19] A. Guisan, and W. Thuiller, "Predicting species distribution: offering more than simple habitat models," *Ecol. Lett.*, vol. 8, pp. 993–1009, 2005.
- [20] J. Elith, C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, and N.E. Zimmermann, "Novel methods improve prediction of species' distributions from occurrence data," *Ecography*, vol. 29, pp. 129–151, 2006.
- [21] R. Pino-Mejías, M.D. Cubiles-de-la-Vega, M. Anaya-Romero, A. Pascual-Acosta, A. Jordán-López, and N. Bellinfante-Crocci, "Predicting the potential habitat of oaks with data mining models and the R system," *Environ. Modell. Softw.*, vol. 25, pp. 826–836, 2010.
- [22] A.M. Mouton, B. De Baets, E. Van Broekhoven, and P.L.M. Goethals, "Prevalence-adjusted optimisation of fuzzy models for species distribution," *Ecol. Model.*, vol. 220, pp. 1776–1786, 2009.
- [23] A.M. Mouton, A.P. Dedecker, S. Lek, P.L.M. Goethals, "Selecting Variables for Habitat Suitability of Asellus (Crustacea, Isopoda) by Applying Input Variable Contribution Methods to Artificial Neural Network Models," *Environ. Model. Assess.*, vol. 15, pp. 65–79, 2009.
- [24] A.H. Fielding, and J.F. Bell, "A review of methods for the assessment of prediction errors in conservation presence/absence models," *Environ. Conserv.*, vol. 24(1), pp. 38–39, 1997.
- [25] A.M. Mouton, B. De Baets, P.L.M. Goethals, "Ecological relevance of performance criteria for species distribution models," *Ecol. Model.*, vol. 221, pp. 1995–2002, 2010.
- [26] D. Goldberg, "Genetic algorithms in search, optimization, and machine learning," Addison-Wesley, 1989.
- [27] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modelling and control," *IEEE Trans. Systems Man Cybernet. B*, vol. 15, pp. 116–132, 1985.
- [28] A. Mäki-Petäy, A. Huusko, J. Erkinaro and T. Muotka, "Transferability of habitat suitability criteria of juvenile Atlantic salmon (*Salmo salar*)," *Can. J. Fish. Aquat. Sci.* vol. 59, pp. 218–228, 2002.
- [29] J.C. Guay, D. Boisclair, M. Leclerc and M. Lapointe, "Assessment of the transferability of biological habitat models for Atlantic salmon parr (*Salmo salar*)," *Can. J. Fish. Aquat. Sci.* vol. 60, pp. 1398–1408, 2003.
- [30] M. Nykänen and A. Huusko, "Transferability of habitat preference criteria for larval European grayling (*Thymallus thymallus*)," *Can. J. Fish. Aquat. Sci.* vol. 61, pp. 185–192, 2004.
- [31] W. Vanreusel, D. Maes and H. Van Dyck, "Transferability of species distribution models: a functional habitat approach for two regionally threatened butterflies," *Conserv. Biol.* vol. 21(1), pp. 201–212, 2007.
- [32] J. Elith, S. Ferrier, F. Huettmann, and J. Leathwick, "The evaluation strip: A new and robust method for plotting predicted responses from species distribution models," *Ecol. Model.*, vol. 186, pp. 280–289, 2005.
- [33] J. Elith and C.H. Graham, "Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models," *Ecography*, vol. 32, pp. 66–77, 2009.