

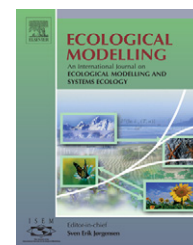


ELSEVIER

available at www.sciencedirect.com



journal homepage: www.elsevier.com/locate/ecolmodel



Where is the worm? Predictive modelling of the habitat preferences of the tube-building polychaete *Lanice conchilega*

Wouter Willems^{a,*}, Peter Goethals^b, Dries Van den Eynde^c, Gert Van Hoey^a, Vera Van Lancker^d, Els Verfaillie^d, Magda Vincx^a, Steven Degraer^a

^a Marine Biology Section, Biology Department, Ghent University, Krijgslaan 281/S8 B-9000 Gent, Belgium

^b Laboratory of Environmental Toxicology and Aquatic Ecology, Ghent University, Department of Applied Ecology and Environmental Biology, J. Plateaustraat 22 B-9000, Gent, Belgium

^c Management Unit of the North Sea Mathematical Models, Gulledele 100 B-1200 Brussels, Belgium

^d Renard Centre of Marine Geology, Department of Geology and Soil Science, Ghent University, Krijgslaan 281/S8 B-9000 Gent, Belgium

ARTICLE INFO

Keywords:

Lanice conchilega

Polychaeta

Habitat preference

Generalized linear models (GLM)

Artificial neural networks (ANN)

ABSTRACT

Grab samples to monitor the distribution of marine macrobenthic species (animals >1 mm, living in the sand) are time consuming and give only point based information. If the habitat preference of a species can be modelled, the spatial distribution can be predicted on a full coverage scale from the environmental variables. The modelling techniques Generalized Linear Models (GLM) and Artificial Neural Networks (ANN) were compared in their ability to predict the occurrence of *Lanice conchilega*, a common tube-building polychaete along the North-western European coastline. Although several types of environmental variables were in the data set (granulometric, currents, nutrients) only three granulometric variables were used in the final models (median grain-size, % mud and % coarse fraction). ANN slightly outperformed GLM for a number of performance indicators (% correct predictions, specificity and sensitivity), but the GLM were more robust in the crossvalidation procedure.

© 2007 Published by Elsevier B.V.

1. Introduction

To scientifically underpin management decisions, there is a growing need to have detailed knowledge on the distribution of marine species. Predictive modelling is a time and cost effective method to produce detailed distribution maps. Predictive modelling objectively investigates the relation between the occurrence of a species and the abiotic habitat. This habitat is quantified by a number of environmental variables, directly measured at the time of sampling, through remote sensing or derived from other models (e.g. currents).

This research will focus on *Lanice conchilega*, a common tube-building polychaete along the North-western European coastline. This species was chosen because of its role as habi-

tat engineer, increasing macrobenthic species diversity and abundance in rather low structured soft sediments through enhancement of the habitat complexity (Zühlke, 2001; Zühlke et al., 1998). *Lanice conchilega* is also an important food source for several demersal fish (Rijnsdorp and Vingerhoed, 2001) and, when occurring in high densities *Lanice* acts as a refugium against predation for many organisms (Woodin, 1978).

The aims of this research were: (1) to identify the environmental variables determining the distribution of *L. conchilega*, (2) to search for the most optimal model describing the habitat preferences of *L. conchilega* and (3) to compare the modelling performance of General Linear Models and Artificial Neural Networks when applied to a marine dataset.

* Corresponding author. Tel.: +32 92648524.

E-mail address: wouter.willems@ugent.be (W. Willems).

0304-3800/\$ – see front matter © 2007 Published by Elsevier B.V.

doi:10.1016/j.ecolmodel.2007.10.017

2. Materials and methods

2.1. Data availability

All data used were collected in the near shore part of the Belgian continental shelf (Southern North Sea) within the framework of the HABITAT-project (Degraer et al., 2002, 2003) in October 1999, March 2000 and November 2000. The major part of the samples (265) were collected in the area of the Western Coastal Banks (WCB), a small complex of sandbanks and swales covering a wide range of soft sediment habitats (Degraer et al., 1999). Outside of the WCB, 38 additional samples were collected in November 2000, along four transects perpendicular to the coastline. The samples were collected with a Van Veen grab (sampling surface area: 0.1 m²) and sieved over a 1 mm sieve. In each sample all adult *L. conchilega* individuals were counted. Since the goal was to predict presence or absence of the species, the densities were transformed to presence/absence.

A sediment subsample was taken with a 3.6 cm diameter core to measure nutrient concentrations (NO₃ + NO₂, NH₄, PO₄ and Si) in the interstitial water. Sediment granulometry was determined: the sediment fraction <1 mm was analysed with a LS Coulter laser counter (vol.%), while the sediment fraction >1 mm was weighted (mass%). The following variables were calculated: median grain-size, mean grain-size, mean/median grain-size ratio, mode, variance, skewness, kurtosis, the volume percentages of the 0–63 μm (hereafter: % mud), 63–125 μm, 125–250 μm, 250–500 μm and 500–800 μm fractions, as well as the mass percentage of the >1 mm fraction (hereafter % coarse fraction).

Bottom current speed and bottom shear stress were obtained from the 3D baroclinic hydrodynamic COHERENS model (Luyten et al., 1999). This model has a horizontal resolution of about 250 × 250 m and a vertical resolution of ten layers. *U* and *U*_{max} are the maximum and median bottom current, and BSRTM and BSRTX are the median and maximum bottom shear stress. Median and maximum chlorophyll-*a* concentration in the surface water were obtained from MERIS satellite images of 2003 from the REVAMP-project (Peters et al., 2005).

3. Modelling techniques

3.1. Variable selection

Since related variables (i.e. granulometry, nutrients) were expected to be highly correlated and thus redundant, Principal Component Analysis (PCA) was used to analyse the relationships between the variables for inclusion the models. A varimax rotation was performed to maximise the independence of the Principal Components (PCs). The non-parametric correlation coefficient Kendall's τ was used to explore the correlation between the potential environmental variables for the modelling, because it can deal better with outliers and extreme distributions of the variables (Arndt et al., 1999). Based on the PCA and the correlation analysis different sets of variables were offered to the forward selection algorithm

of the GLM (see further). It was avoided to enter highly correlated variables in such a set or too much variables which were highly associated with one PC.

3.2. GLM: logistic regression

To predict the absence or presence of *L. conchilega* multiple logistic regression (Trexler and Travis, 1993), a type of GLM, was used. Logistic regression has been widely used in ecology (Paruelo and Tomasel, 1997; Ysebaert et al., 2002) and predicts the probability (between 0 and 1) that a species will occur, based on the environmental variables. Since the sample distribution was binary (absent or present), the logit link was used. The forward stepwise likelihood ratio method was used to select the best set of variables. Interaction terms and non-linear terms (i.e. quadratic) of each variable were also included in the set of variables. A cut-off value for species presence was based on the percentage of the samples in which *L. conchilega* was present in the data set (26% of the samples, cut-off of 0.26) (Ysebaert et al., 2002). The analysis was performed with SPSS version 12.0 (SPSS, Inc., Chicago, IL).

Threefold crossvalidation was used to test the robustness of the models. The data set was randomly split in three parts and two parts were iteratively used to construct a model, and the third part to test the model. If the predictive performance of the model for each fold was similar, a final model was constructed with all the data.

3.3. Artificial neural networks

Artificial Neural Networks (ANN) are a technique from the field of artificial intelligence (Lek and Guégan, 1999). They have a similar structure as the human brain: a network of connected neurons. The neurons are the building blocks of the ANN. Data enters a neuron from several other neurons, is summed and then fed into an activation function, which generates the output of the neuron. Neurons can pass on information because they are connected. The importance of a connection is expressed as an interconnection weight. The adjustment of these weights will influence the model output (Lek and Guégan, 1999). Through a learning algorithm, the weights will be adjusted iteratively, increasing the agreement between the observed and predicted presence of the species (Lek and Guégan, 1999). The ANN's in this research have their neurons organised in three layers: environmental variables are presented at the input layer, are passed on to the hidden layer which processes the information and finally the output layer generates the prediction of the probability of presence of *L. conchilega*. As with GLM, threefold crossvalidation and output visualisation were used (Fig. 1b and d) was used to test the robustness of the models. For the ANN the species was predicted to be present if the model output was larger than 0.5. The ANN's were constructed in MATLAB 6.1 using the neural networks toolbox.

3.4. Model performance and variable importance

In order to assess and compare the predictive power of GLM and ANN several performance indicators were cal-

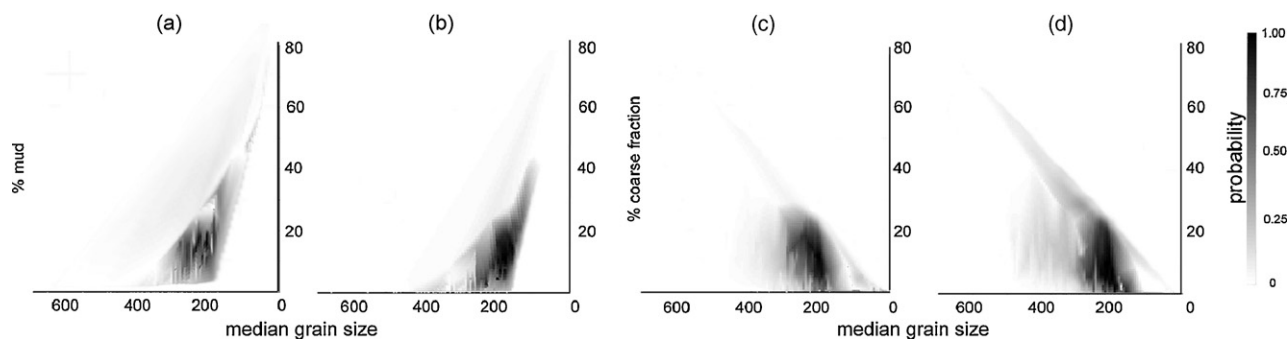


Fig. 1 – Predicted probability of presence of *L. conchilega* for the GLM (a and c) and ANN-models (b and d) with all samples.

134 culated. Most indicators were based on a two by two
 135 contingency table containing the number of True Positive
 136 (TP), False Negative (FN), True Negative (TN) and False
 137 Positive (FP) predictions. The overall percent of correct
 138 predictions is expressed as Correctly Classified Instances
 139 (CCI) = (TP + TN)/(total # samples). The number of correctly
 140 classified species absences on the total number of predicted
 141 species absences was calculated as the Negative Predictive
 142 Value: NPV = TN/(TN + FN). Similarly, the Positive Predictive
 143 Value was calculated: PPV = TP/(TP + FP). The model speci-
 144 ficity and sensitivity were calculated as TN/(FP + TN) and
 145 TP/(TP + FN), respectively. Cohen's κ (Dedecker et al., 2004)
 146 assesses the correctness of the prediction independent of the
 147 prevalence of a species in the original data set. An alternative
 148 statistical parameter to express the performance of a model
 149 is the Area Under the Curve (AUC) (Fielding and Bell, 1997).
 150 The AUC is not dependent on a cut-off value at which the
 151 species is present. In this way it evaluates the model output
 152 in a continuous, instead of discrete manner. Finally, the Pear-
 153 son correlation between the output of the ANN and GLM final
 154 models was calculated to assess the similarity of the predic-
 155 tions.

156 The relative importance of an environmental variable
 157 in the prediction of the occurrence of *L. conchilega* was
 158 assessed. For the GLM, the Wald statistic for each variable
 159 in the GLM was compared (Ysebaert et al., 2002). Equiv-
 160 alently, the partial derivatives (PaD) method was used for
 161 assessing variable importance for the ANN (Dedecker et al.,
 2004).

4. Results

4.1. Variable selection

162 The first five PCs explain 78% of the variance in the data and
 163 were clearly associated with groups of related environmental
 164 variables (Table 1). PC1 was most correlated with the sorting
 165 of the sediment and the fine sediment fraction, PC2 with the
 166 coarser fraction and with general sediment variables, PC3 only
 167 with current characteristics, PC4 with the shape of the grain-
 168 size distribution and PC5 with the nutrient concentrations in
 169 the interstitial water.
 170

171 The results of the correlation analysis agree with those of
 172 the PCA. The sediment variables were highly correlated with
 173 each other, but % mud and % coarse fraction have a low corre-
 174 lation with the rest of the sediment variables and each other.
 175 The bottom current variables were moderately to highly corre-
 176 lated ($r=0.4-0.7$) with each other, while there was a very
 177 low correlation between the nutrient concentrations ($r < 0.3$)
 178 and very little ($r = -0.2$) correlation between the maximum and
 179 median chlorophyll-*a* concentration. The correlation between
 180 the variables of the final models is low (median grain-size-
 181 % mud: $r = -0.45$; median-% coarse: $r = 0.31$; % mud-% coarse:
 182 $r = 0.0$).

4.2. Comparison techniques and variable importance

183 The forward selection algorithm allowed to select the most
 184 optimal set of environmental variables for the models. The
 185

Table 1 – Principal component analysis: rotated scores (Varimax rotation) of the variables for the first five principal components (PC)

PC1	38%	PC2	16%	PC3	10%	PC4	9%	PC5	5%
S.D.	0.96	500-800 μm	0.94	U	0.93	Skewness	0.96	NH ₄	0.73
38–63 μm	0.92	Mode	0.91	U _{max}	0.93	Kurtosis	-0.93	PO ₄	0.65
% mud	0.87	d90	0.89	BSTRM	0.92			NO ₃ + NO ₂	-0.56
63–125 μm	0.78	Median	0.84	BSTRX	0.76				
M/M ratio	-0.78	125–250 μm	-0.79						
d10	-0.77	Mean	0.75						
Mean	-0.62	% coarse fr.	0.71						

The percentage of variance explained is shown in the upper right for each PC. Only scores with an absolute value above 0.50 are shown.

Table 2 – Model performance indicators for the threefold and the models with all data

	GLM				ANN			
	Fold 1	Fold 2	Fold 3	All data	Fold 1	Fold 2	Fold 3	All data
Learning set								
CCI	78.1	74.5	73.5	78.0	82.1	85.7	82.0	80.6
NPV	90.4	89.8	91.1	92.3	88.1	89.7	85.6	89.4
PPV	56.3	51.3	50.0	54.6	66.0	74.0	69.2	56.4
Spec.	78.5	73.6	70.8	77.0	87.5	91.0	90.7	85.0
Sens.	76.9	76.9	80.8	81.1	67.3	71.2	57.7	65.7
Test set								
CCI	74.5	71.4	73.5	/	78.6	72.5	81.6	/
NPV	51.2	47.4	50.0	/	84.7	80.8	91.7	/
PPV	91.2	86.7	88.3	/	61.5	50.0	53.9	/
Spec.	72.2	72.2	73.6	/	85.9	82.0	84.6	/
Sens.	80.8	69.2	73.1	/	59.3	48.2	70.0	/
Cohen's κ	0.45	0.36	0.41	(0.49)	0.46	0.30	0.49	(0.48)
All								
AUC	82.9	80.80	82.8	82.5	84.6	85.0	88.3	85.2

CCI, Correctly Classified Instances; NPV, Negative Predictive Value; PPV, Positive Predictive Value; spec., specificity; sens., sensitivity; AUC, Area Under the Curve.
The Cohen's κ for the models with all data was calculated from the performance on the training set.

selected set contained only three granulometric variables: median grain-size, % mud and % coarse fraction, along with the non-linear terms % mud² and % coarse fraction² (all interaction terms were rejected). ANN with one hidden layer containing three neurons were trained. The same variables were used as with the GLM. Only sigmoid activation functions were used in the neurons (Table 2).

In the visualization (Fig. 1), the similarity between the GLM and ANN predictions was higher within the range of the original data set, but lower outside of this range. After conversion of the predicted outcomes to a binary coding (using the cut-off values), a number of performance indicators were calculated (Table 2). In a paired t-test the CCI, specificity and sensitivity of the ANN-models was significantly higher than their GLM counterpart (6 df; $p < 0.05$). The NPV and PPV and the Cohen's κ were not significantly higher for the neural network models (NPV and PPV: 6 df, $p > 0.05$; Cohen's κ : 3 df; $p > 0.05$). The AUC was significantly larger for the ANN models (paired t-test; 3 df; $p < 0.05$). The Pearson correlation between the ANN and GLM final model output was high: $r = 0.89$ ($p < 0.01$).

For the GLM, the relative importance of the variables in predicting the presence of *L. conchilega* was similar between the threefold and the final model with all data (Table 3). Overall the median grain-size was most important in GLM, the other two variables and their quadratic terms had similar importances, the constant was less important (Table 3). The ANN models showed a high variability in variable importance between the folds and the final model. The order of importance in the final model was reversed compared to the GLM: from % coarse fraction over % mud to median grain-size.

5. Discussion

5.1. Selection of environmental variables

A good prior selection of environmental variables for the final model was found to be necessary. The inclusion of redun-

dant variables will complicate the assessment of the relative variable importance, which is very helpful to derive ecological insights from a model. The relative importance of two highly redundant variables will be “shared” between them, underestimating their importance. PCA and correlation allowed only to assess the mutual relations of the variables, but did not allow to distinguish if the variables were important in the prediction of *L. conchilega*. A reason for this is that these techniques hypothesise a linear relation between variable and presence of the species, while species response curves most likely have a bell-shape form. The forward selection algorithm was very important in selecting sets of appropriate environmental variables.

5.2. Modelled habitat preference

Although a data set with various types of environmental variables was available, only three granulometric variables were

Table 3 – The relative importance (%) of the variables in the prediction of the presence of *L. conchilega*

Fold	1	2	3	All
GLM				
Median	27.5	25.5	37.8	26.2
% mud	20.3	14.0	10.8	15.9
% mud ²	16.3	13.0	14.9	15.8
% coarse	19.2	23.6	15.9	19.9
% coarse ²	11.6	17.0	5.6	15.2
Constant	5.0	6.9	15.0	7.0
ANN				
Median	21.2	5.1	2.6	18.7
% mud	49.8	6.3	5.9	37.2
% coarse	29.0	88.6	91.6	44.1

For the GLM the percentages are based on the Wald statistic, for the ANN on the partial derivatives method. “all” indicates the model created with all samples.

selected in the models. The relative importance of these variables in the GLM-models was median grain-size, % coarse (sum of relative importance of linear and quadratic term) and % mud (idem % coarse). For the ANN-models the order of importance was not consistent between the folds and the final model with all samples. Previous studies did point out the importance of granulometric variables (Gray, 1981; Snelgrove, 1994), but Buhr and Winter (1977) indicated that currents also have an important effect on a much smaller scale for *L. conchilega*. In the models presented here, currents were found to be not important. However this could be due to the resolution of the current model which was only 250 × 250 m. Chlorophyll-a was expected to be important in predicting the occurrence of *L. conchilega*, as it is a proxy for primary production. However it was not selected for the models. Probably because it showed little variance on such a small scale and the fact that only data from the year 2003 were available. The reason that nutrients were not used in the final models could be due to a seasonal effect. Because not all of the samples were collected simultaneously, no relation with was found between species occurrence and nutrient levels, as these fluctuate seasonally.

The modelled habitat preferences are a result of the active habitat selection of *L. conchilega*: pelagic larvae settle on the bottom and can return to the water column again if the habitat is not suitable (Bhaud and Cazaux, 1990). The models did not perform equally well through the whole range of the variables. This is due to the fact that there were only very little samples in some parts of the range of a variable. The model should therefore only be used with in the variable range of original dataset used to construct the models (Fig. 1, dashed areas). The maximum values of the CCI are 80–85%. This could be explained by the effect variables which were not measured and the patchy distribution of *L. conchilega* (Heuers et al., 1998). This distribution, due to biological interactions and recruitment fluctuations, introduces noise in the data set: samples in suitable environments would have no *L. conchilega*.

6. GLM versus ANN

Although there was a very strong correlation of the model outputs, ANN outperformed GLM for a number of performance indicators (CCI, spec. and sens.). Also the Area Under the Curve was significantly higher for the ANN, indicating that the difference in performance is not due to the difference in the cut-off for presence of a species. For the GLM each non-linear relation (i.e. quadratic) and each interaction term had to be explicitly presented to the selection algorithm. The superior performance of ANN could be explained by the fact that non-linear functions and variable interaction are inherent to the architecture of the ANN, because of the connections between the neurons. The effect of only two variables could be visualised simultaneously, the graphs were only a simplification of the model predictions, which are in a multivariate space. However, performance and relative variable importance showed a higher dissimilarity between the folds for the ANN-models. This could be due to the higher internal complexity of the ANN or a high dependence on the initial conditions during the training of the network. From a parsimonious point of view the

GLM were superior, as the models were simpler and ANN only performed slightly better.

Further research will deal with the testing of these models on a larger geographical scale and the construction of models that try to predict the abundance (individuals/m²) of macrobenthic species.

Uncited references

Grimm (1999), Guisan and Zimmerman (2000), Haelters et al. (2004), Manel et al. (1999), Ropert and Dauvin (2000), Van Hoey et al. (2004) and Ysebaert et al. (2000).

Acknowledgements

The data was obtained from the project “Intensive evaluation of the evolution of a protected benthic habitat (HABITAT)” funded by OSTC (project number: MN/02/89) and AWK (Coastal Waterways Division; file numbers: 99380 and 200.455). Chlorophyll data were obtained from the REVAMP-project: (Peters et al., 2005). The first author is grant holder of the Institute for the promotion of Innovation through Science and Technology in Flanders (IWT).

REFERENCES

- Arndt, S., Turvey, C., Andreasen, N.C., 1999. Correlating and predicting psychiatric symptom ratings: Spearman's r versus Kendall's Tau correlation. *J. Psychiat. Res.* 33, 97–104.
- Bhaud, M.R., Cazaux, C.P., 1990. Buoyancy characteristics of *Lanice conchilega* (Pallas) larvae (Terebellidae). Implications for settlement. *J. Exp. Mar. Biol. Ecol.* 141 (1), 31–45.
- Buhr, K., Winter F.E., 1977. Distribution and maintenance of a *Lanice conchilega* association in the Weser estuary, with special reference to the suspension-feeding behaviour of *Lanice conchilega*. In: Keegan, B.F., Ceidigh, Boaden, P.J.S. (Eds.), *Biology of Benthic Organisms*. 11th European Symposium on Marine Biology, Gallaway, October 1976. Pergamon Press, Oxford and New York, pp. 101–113.
- Dedecker, A., Goethals, P.L.M., Gabriels, W., De Pauw, N., 2004. Optimisation of Artificial Neural Network (ANN) model design for prediction of macroinvertebrate communities in the Zwalm river basin (Flanders Belgium). *Ecol. Model.* 174 (1/2), 161–173.
- Degraer, S., Van Lancker, V., Moerkerke, G., Van Hoey, G., Vincx, M., Jacobs, P., Henriët, J.-P., 2002. *Intensive evaluation of the evolution of a protected benthic habitat: HABITAT*. Final report. Federal Office for Scientific, Technical and Cultural Affairs (OSTC) – Ministry of the Flemish Community, Environment and Infrastructure Department Waterways and Marine Affairs Administration, Coastal Waterways, 124 pp.
- Degraer, S., Van Lancker, V., Moerkerke, G., Van Hoey, G., Vanstaen, K., Vincx, M., Henriët, J.-P., 2003. Evaluation of the ecological value of the foreshore: habitatmodel and macrobenthic side-scan sonar interpretation: extension along the Belgian Coastal Zone. Final report. Ministry of the Flemish Community, Environment and Infrastructure Department Waterways and Marine Affairs Administration, Coastal Waterways, 63 pp.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24 (1), 38–49.

Q3

300
301
302
303
304
305
306
307
308
309310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345

- 346 Gray, J.S., 1981. The Ecology of Marine Sediments. Cambridge
347 Studies in Modern Biology, vol. 2. Cambridge University Press,
348 Cambridge, p. 185.
- 349 Grimm, V., 1999. Modelling the spatial and temporal distribution
350 of *Lanice conchilega*. In: The Wadden Sea ecosystem—Stability
351 Properties and Mechanisms. Springer, Berlin Heidelberg New
352 York, pp. 147–152.
- 353 Guisan, A., Zimmerman, N.E., 2000. Predictive habitat
354 distribution models in ecology. *Ecol. Model.* 135, 147–186.
- 355 Haelters, J., Vigin, L., Stienen, E.W.M., Scory, S., Kuijken, E.,
356 Jacques, T.G., 2004. Ornithologisch belang van de Belgische
357 zeegebieden. Identificatie van mariene gebieden die in
358 aanmerking komen als Speciale Beschermingszones in
359 uitvoering van de Europese Vogelrichtlijn. Rapport van de
360 Beheerseenheid Mathematisch Model Noordzee en het
361 Instituut voor Natuurbehoud. Bulletin van het Koninklijk
362 Belgisch Instituut voor Natuurwetenschappen, Biologie 74
363 (Suppl.), 91 pp. (in Dutch).
- 364 Heuers, J., Jaklin, S., Zühlke, R., Dittman, S., Günther, C.P.,
365 Hildebrandt, H., Grimm, V., 1998. A model on the distribution
366 and abundance of the tube-building polychaete *Lanice*
367 *conchilega* (Pallas 1766) in the intertidal of the Wadden Sea.
368 Verhandlungen Gesellschaft Ökologie 28, 207–215.
- 369 Lek, S., Guégan, J.F., 1999. Artificial neural networks as a tool in
370 ecological modelling, an introduction. *Ecol. Model.* 120, 65–73.
- 371 Luyten, P.J., J.E. Jones, R. Proctor, A. Tabor, P. Tett and K.
372 Wild-Allen, 1999. COHERENS—A Coupled
373 Hydrodynamical-Ecological Model for Regional and Shelf
374 Seas: User Documentation, MUMM Report, Management Unit
375 of the Mathematical Models of the North Sea, 914 pp.
- 376 Manel, S., Dias, J.-M., Ormerod, S.J., 1999. Comparing discriminant
377 analysis, neural networks and logistic regression for
378 predicting species distributions: a case study with a
379 Himalayan river bird. *Ecol. Model.* 120, 337–347.
- 380 Paruelo, J.M., Tomasel, F., 1997. Prediction of functional
381 characteristics of ecosystems: a comparison of artificial neural
382 networks and regression models. *Ecol. Model.* 98, 173–186.
- Peters, S.V.M., Eleveld, M., Pasterkamp, H., vann der Woerd, H.,
DeVolder, M., Jans, S., Park, Y., Ruddick, K., Block, T.,
Brockmann, C., Doerffer, R., Krassemann, H., Schoenfeld, W.,
Jørgenson, P.V., Tislstone, G., Moore, G., Sørensen, K., Hokedal,
J., Aas, E., 2005. Atlas of the chlorophyll-*a* concentration in the
North Sea based on MERIS imagery of 2003, Edition 1.0. Vrije
Universiteit, Amsterdam.
- Rijnsdorp, A.D., Vingerhoed, B., 2001. Feeding of plaice
Pleuronectes platessa L. and sole *Solea solea* L. in relation to the
effect of bottom trawling. *J. Sea Res.* 45, 219–229.
- Ropert, M., Dauvin, J.-C., 2000. Renewal and accumulation of a
Lanice conchilega (Pallas) population in the baie des Veys,
western Bay of Seine. *Oceanol. Acta* 23 (4), 529–546.
- Snelgrove, P.V.R., 1994. Animal–sediment relationships revisited:
cause versus effect. *Oceanogr. Mar. Biol.: Annu. Rev.* 32,
111–177.
- Trexler, J.C., Travis, J., 1993. Non-traditional regression analyses.
Ecology 74 (6), 1629–1637.
- Van Hoey, G., Degraer, S., Vincx, M., 2004. Macrobenthic
community structure of soft-bottom sediments at the Belgian
Continental Shelf. *Estuar. Coast. Shelf Sci.* 59 (4), 599–613.
- Woodin, S.A., 1978. Refuges, disturbance, and community
structure: a marine soft bottom example. *Ecology* 59, 274–284.
- Ysebaert, T., de Neve, L., Meire, P., 2000. The subtidal
macrobenthos in the mesohaline part of the Schelde estuary
(Belgium): influenced by man? *J. Mar. Biol. Assoc.* 80, 587–597.
- Ysebaert, T., Meire, P., Herman, P.M.J., Verbeek, H., 2002.
Macrobenthic species response surfaces along estuarine
gradients: prediction by logistic regression. *Mar. Ecol. Prog.*
Ser. 225, 79–95.
- Zühlke, R., Blome, D., van Bernem, D., Dittman, S., 1998. Effects of
the tube-building polychaete *Lanice conchilega* (Pallas) on the
benthic macrofauna and nematodes in an intertidal sandflat.
Senckenbergiana Maritima 29, 131–138.
- Zühlke, R., 2001. Polychaete tubes create ephemeral community
patterns: *Lanice conchilega* (Pallas, 1766) associations studied
over six years. *J. Sea Res.* 46, 261–272.