

Forecasting the poleward range expansion of an intertidal species driven by climate alterations

RAQUEL XAVIER^{1,2}, FERNANDO P. LIMA^{1,3} and ANTÓNIO M. SANTOS^{1,2}

¹CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal. E-mail: raq.xavier@mail.icav.up.pt

²Departamento de Zoologia-Antropologia, Faculdade de Ciências da Universidade do Porto, R. Campo Alegre, s/n, 4169-007 Porto, Portugal.

³Department of Biological Sciences, University of South Carolina, Columbia, SC, USA.

SUMMARY: Accurate distributional models can be used to reliably predict the response of organisms to climatic changes. Though such models have been extensively applied to terrestrial organisms, they have hardly ever been applied to the marine environment. Recent changes in the distribution of the marine gastropod *Patella rustica* (L.) were previously modelled with Classification and Regression Tree (CART) and the results revealed that increases in temperature were the major driver of those changes. However, the accuracy scores during the validation of the model were unsatisfactory, preventing its use for forecasting purposes. To fulfil this objective, in the present study a more robust method, Artificial Neural Network (ANN), was employed to produce a model suited to forecasting changes in the distribution of *P. rustica*. Results confirmed that the ANN model behaved better than the CART, and that it could be used for forecasting future distributional scenarios. The model forecasts that by the 2020s *P. rustica* is likely to expand its range at least 1000 km northwards. These results should be interpreted with caution considering the dispersal limitations of this species, but if such an expansion took place, major changes in the colonized ecosystems are expected due to the key role of limpets in intertidal communities.

Keywords: *Patella rustica*, climate change, range expansion, distributional model, intertidal, Artificial Neural Networks.

RESUMEN: PREDICCIÓN DE LA EXPANSIÓN HACIA EL NORTE DE LA DISTRIBUCIÓN DE UNA ESPECIE INTERMAREAL DEBIDO A ALTERACIONES CLIMÁTICAS. – Los modelos de distribución de especies son una herramienta utilizada para predecir con exactitud la respuesta de organismos a las alteraciones climáticas. Aunque estos modelos han sido frecuentemente aplicados a organismos terrestres, casi nunca han sido aplicados en ambientes marinos. Estudios sobre la distribución del gasterópodo *Patella rustica* (L.) utilizando el método de árboles de clasificación y regresión (*Classification and Regression Tree: CART*) revelaron que el aumento de temperatura sería el principal impulsor de estos cambios. Sin embargo, la precisión de este modelo no permitió la predicción de la distribución de *P. rustica* en el siglo XXI. En el presente trabajo presentamos un método más robusto, las redes neuronales artificiales (*Artificial Neural Networks: ANN*), para producir un modelo adecuado que permita predecir futuros cambios en la distribución de *P. rustica*. Los resultados confirman que las ANN pueden ser utilizadas para la previsión de futuros escenarios de distribución y muestran que en la década de 2020 esta especie podría ampliar su rango de distribución al menos 1.000 kilómetros en dirección norte. Estos resultados deben ser interpretados teniendo en cuenta las limitaciones de dispersión de este organismo, pero si esa expansión se produce, se esperan grandes cambios en los ecosistemas debido al papel fundamental de las lapas en las comunidades intermareales.

Palabras clave: *Patella rustica*, alteraciones climáticas, expansión, modelos de distribución, intermareal, Redes Neuronales Artificiales.

INTRODUCTION

Modelling species' distributions based on environmental variables has become an important field in ecology, largely sustained by the need to forecast the

effects of climate change on biodiversity (see for a review Brown *et al.*, 1997; Guisan and Thuiller, 2005). Among the vast array of techniques currently available, single-species bioclimatic envelope models have probably been the most widely used in ecology (Heikkinen

et al., 2006). These models are based on the determination of the environmental requirements and limits within which a particular species can occur (Vaughan and Ormerod, 2003), i.e. the spatial distribution of a species is correlated with several climatic variables, allowing the identification of its environmental envelope. They are frequently used to predict the distribution of species under past, current, and future climate conditions (McLaughlin *et al.*, 2002; Hijmans and Graham, 2006).

Being correlative, bioclimatic envelope models disregard factors such as species' dispersal abilities or limitations, ecological interactions and adaptation to environmental modifications (e.g. Soberón and Peterson, 2005; Parmesan, 2006). Nonetheless, this framework can be extremely useful for uncovering possible environmental constraints, and is often the first approach for assessing the potential impact of climate change on biodiversity (Pearson and Dawson, 2003). There are several statistical techniques for bioclimatic envelope modelling. Artificial Neural Networks (ANNs) are an alternative to logistic regression and to other non-linear models and have recently been applied to the study of biologically complex and/or non-linear data (e.g. Lek and Guégan, 1999; Spitz and Lek, 1999; Olden and Jackson, 2002; Park *et al.*, 2003). ANNs have the ability to cope with data even when the relationship between the presence or absence of a species and a given environmental variable is non-linear. Another important aspect of ANNs is that they generally perform well even when dealing with noisy data (Tu, 1996) and when predictor variables show a high degree of collinearity (Farrar and Glauber, 1967). However, these advantages come with a price: as ANNs are non-linear data modelling tools, the assessment of effects of individual predictors on output variables is rather difficult to determine (Tu, 1996; Brosse *et al.* 1999; Pearson *et al.*, 2004).

The main objective of the present study was to obtain a reliable bioclimatic-envelope distributional model for *Patella rustica* (L.), making it possible for the first time to forecast the evolution of the distributional range of this species in the near future. *P. rustica* occurs in the northeast Atlantic, from the Mauritanian coast (Africa) in the south (Fischer-Piètte and Gaillard, 1959) up to Capbreton (SW France) in the north (Lima *et al.*, 2007b). It also occurs throughout the entire Mediterranean basin. Significant changes in the distribution of this gastropod occurred recently on the Portuguese coast, where a historical gap of more than 200 km was colonized. These changes were attributed to anomalies related to the upwelling regime and sea surface warming of about 0.5° C in autumn and early winter (Lima *et al.*, 2006), which correspond to the reproduction and early development seasons of *P. rustica* (Ribeiro *et al.*, 2009). More recently, Lima *et al.* (2007b) modelled the distribution of *P. rustica* using a Classification and Regression Tree (CART) bioclimatic-envelope approach. Historical data from the 1950s was used for model training and data from 1990s for validation.

Although circumstantial data (recent increases in the abundance of *P. rustica* in NW Iberia and SW France and concomitant temperature increase) suggested that a range expansion along the Atlantic coast of France would be highly probable in the near future, the CART model never passed the validation phase, where it failed to hindcast some aspects of *P. rustica*'s current distribution. In other words, the model had insufficient generalization potential, which prevented any further forecasting attempt. This was a highly unsatisfactory result, given the ecological importance of *P. rustica*. Being a patellid gastropod, this species belongs to one of the most important groups of primary consumers in the intertidal and key disturbers in this environment, having a great impact on the entire species assemblage (Thompson *et al.*, 2002; Firth and Crowe, 2008).

MATERIALS AND METHODS

Biological and environmental data

Presence/absence data were used to train and validate the model. These data consisted in the historical records of *P. rustica* provided by Fisher-Piètte (1959, 1963) and Fischer-Piètte and Gaillard (1959) during the 1950s, along most of the range of this species in the northeast Atlantic. Fischer-Piètte's data set included 95 shores from Asilah (Morocco) to Trémazan (NW France), with 28 absences and 67 presences of *P. rustica*. To achieve a more balanced data set in terms of presences/absences, 27 locations in northern France (eastwards of Trémazan) were added, using data collected in the 1950s by Crisp and Southward (1958) (Fig. 1A). In total, data encompassed 122 sampling points distributed throughout Morocco, Portugal, Spain and France. The usage of historical data from the 1950s to build the ANN model was a key step in the modelling procedure, allowing contemporary data to be used independently in the validation phase of the model (see Araújo *et al.*, 2005).

Raw data for near-surface (air) temperature (NST), cloud cover (CC) and sea surface temperature (SST) along the northeastern Atlantic coast were obtained from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS, <http://icoads.noaa.gov/>, Woddruff *et al.*, 1988) and the Tyndall Centre for Climate Change Research (<http://www.cru.uea.ac.uk/>, Mitchell *et al.*, 2004) covering European land surface with a 10 minute resolution for the periods 1950-1959 and 1991-2000. Data were processed and imported into the Grass Geographic Information System (GRASS, 2006), following the methodology used by Lima *et al.* (2007b).

Exposure to wave action (Exp) was calculated following the methodology described by Ekebom *et al.* (2003). Wave climate data were downloaded from the European Environmental Agency database (<http://www.eea.europa.eu>). Data were subsequently processed as described by Lima *et al.* (2007b). A combina-

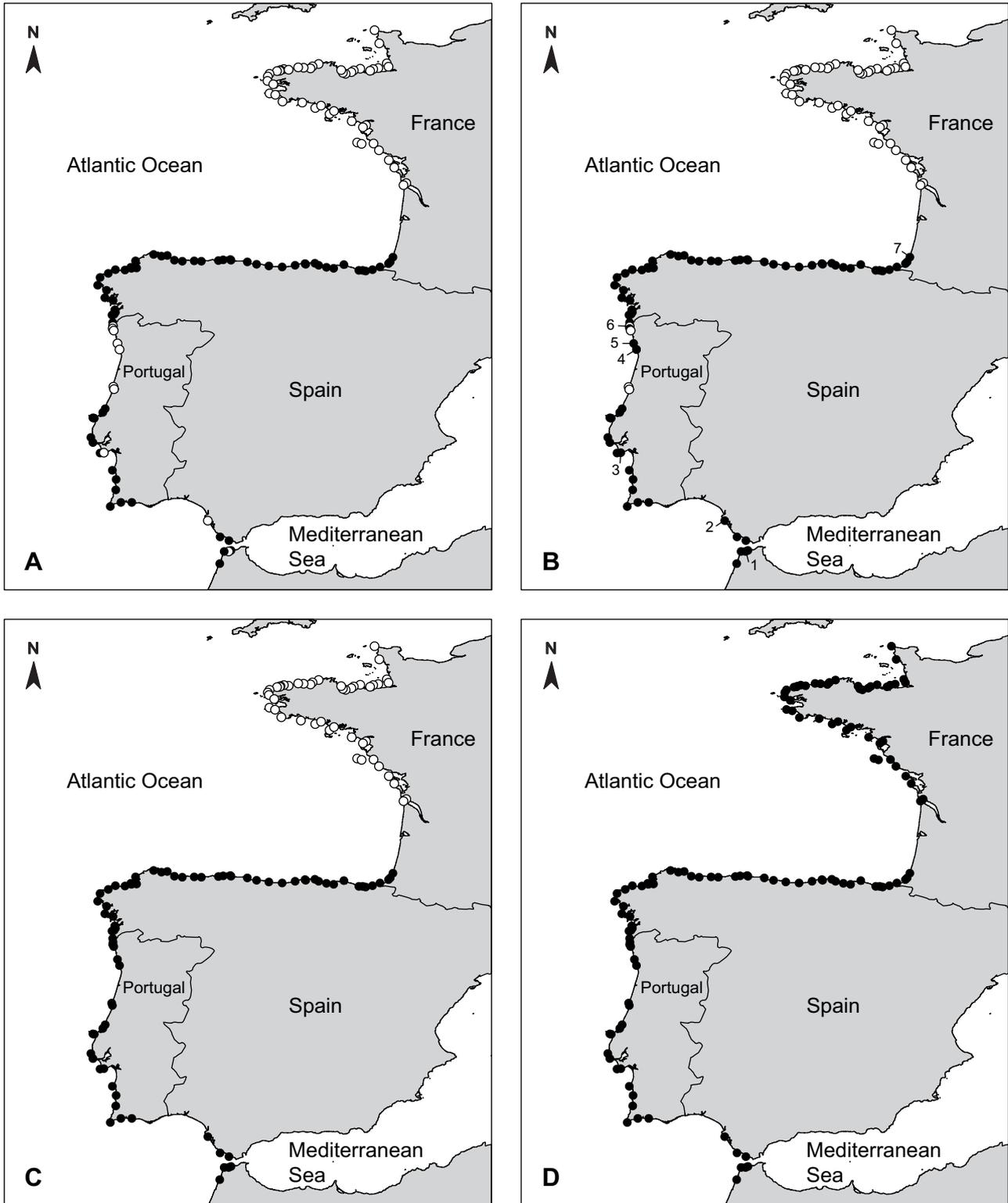


FIG. 1. – Distribution of *P. rustica*: A) during the 1950s, based on survey data from Fischer-Piètte (1959, 1963), Fischer-Piètte and Gaillard (1959) and Crisp and Southward (1958); B) hindcasted for the 1950s, resulting from ANN training. 1, Tanger; 2, Rota; 3, Sesimbra; 4, Leixões; 5, Vila do Conde; 6, Montedor; and 7, Capbreton; C) hindcasted by the ANN model for the 1990s; D) forecasted for the 2020s under a median-high emissions scenario. Presence and absence of *P. rustica* are depicted by black and white dots, respectively.

TABLE 1. – Summary of the independent variables used in the ANN model of *Patella rustica* distribution.

Environmental variables	Acronym	Units	Period	Statistics
Sea surface temperature	SST	°C	10-year	Average
	SDSST		10-year	Standard deviations
	WSST		10-year, Winters only	Average
	SSST		10-year, Summers only	Average
	RSST		10-year, Reproduction (Sep-Nov)	Average
Near surface air temperature	NST	°C	10-year	Average
	SDNST		10-year	Standard deviations
	WNST		10-year, Winters only	Average
	SNST		10-year, Summers only	Average
	RNST		10-year, Reproduction (Sep-Nov)	Average
Cloud cover	CC	%	10-year	Average
Exposure to wave action	Exp	kW	-	Average

tion of 12 variables (Table 1) was chosen based on the previous assessment of the environmental constraints of *P. rustica* distribution (Lima *et al.*, 2007b). Furthermore, the ratio between the number of sampling points and the predictors used was also taken into account. This should not be less than 10:1 to improve the analytical power of the analysis (Vaughan and Ormerod, 2003). Exposure to wave action (Exp) was considered invariant throughout time and was used equally in both the model training and testing phases. Average sea surface temperature (SST), average near-surface temperature (NST) and average cloud cover (CC) were obtained for the training and testing data sets by averaging all values over 1950-1960 and 1990-2000, respectively. Average SST and NST were also computed for the following seasons: winter (WSST and WNST), summer (SSST and SNST) and autumn (RSST and RNST), the latter referring to the reproductive period of *P. rustica*, which lasts roughly from September to November (Ribeiro *et al.*, 2009). In addition, standard deviations of SST and NST were also computed (SDSST and SDNST) and used as a measure of the average range of each of these main variables.

To predict the future distribution of *P. rustica*, data from Atmosphere-Ocean General Circulation Models (AOGCMs) were coupled with the Intergovernmental Panel for Climate Change Special Report on Emission Scenarios (IPCC-SRES, Nakicenovic *et al.*, 2000). The AOGCM developed by the Hadley Centre, HadCM3 (Hadley Centre Coupled Model, version 3 Gordon *et al.*, 2000) was coupled with the A2 (medium-high) IPCC-SRES emissions scenario (Nakicenovic *et al.*, 2000) for both NST and CC at a 10 minute resolution. These data are also available at the Tyndall Centre for Climate Change Research (<http://www.cru.uea.ac.uk/>, Mitchell *et al.*, 2004).

Unfortunately, future scenarios for SST under the exact same AOGCM and IPCC-SRES scenario combination (HadCM3/A2) were not available in free data sources. SST forecasts were found under AOGCM HadGEM1, which is an improved version of HadCM3, coupled with the IPCC A1B emission scenario (Johns *et al.*, 2004) with 1° resolution. These data were down-

loaded directly from Met Office Hadley Centre website (<http://hadobs.metoffice.com>, Rayner *et al.*, 2006).

Modelling approach

The NevProp software package (version 4R1, Goodman and Rosen, Goodman Brain Computation Lab, University of Nevada, Reno) was used to build the ANN model. NevProp is freely available at <http://brain.cs.unr.edu/publications/NevProp.zip>.

ANN models are composed of multiple nodes distributed in three types of layers: input, hidden and output. The predictor variables are located at the input layer and each is connected to the “neurons” of hidden layers by a connection which bears a specific weight. In the hidden “neurons” a non-linear transformation occurs upon the weighted sum of the information arriving from the previous layer, thus allowing for the network to capture the complex and non-linear effects of the predictors (Tu, 1996). The information is then propagated to the next layer until it reaches the output neuron, which will provide the response variable value (Fig. 2).

Because ANNs are usually employed as “black-box” models (Tu, 1996), the methodology used in the present work is described in detail. There are no concrete procedures to determine the optimal number of neurons or hidden layers to be used in an ANN model, so it is necessary to try a variety of node numbers to find the optimal architecture (Yuval, 2001; Ozesmi *et al.*, 2006). Many hidden nodes correspond to a situation similar to having too many interaction terms in a logistic regression model, which may lead to an overfitted model which, in turn, has no generalization ability (Tu, 1996). Another consequence of overfitting is the tendency to model random noise as being deterministic, which will also have a strong influence on the outcome patterns (Manel *et al.*, 1999). Overtraining is another common problem in ANNs (Tu, 1996; Ozesmi *et al.*, 2006) and the choice of a method for determining when to stop the training is subjective. In this case an early stopping method based on cross-validation implemented in NevProp (Autotrain, Goodman, 1996) was employed.

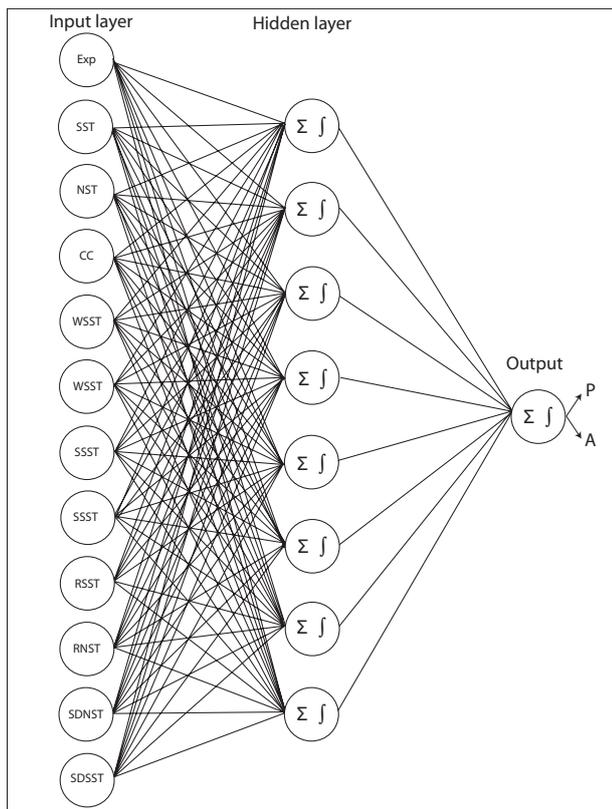


FIG. 2. – Simplified schematic outline of the ANN model employed. The input layer is composed of 12 environmental variables: Exp (Exposure), SST (average sea surface temperature), NST (average near-surface temperature), CC (average cloud cover), WSST (average winter sea surface temperature), WNST (average winter near-surface temperature), SSST (average summer sea surface temperature), SNST (average summer near-surface temperature), RSST (average sea surface temperature during the reproductive period), RNST (average near-surface temperature during the reproductive period), SDSST (sea surface temperature standard deviation) and SDNST (near surface temperature standard deviation). Each input variable is connected to 8 neurons located in a single hidden layer. The single neuron in the output layer translates the available information into presence or absence of *P. rustica* at a given site. Lines connecting neurons bear specific weights computed during the training phase. Σf represents the non-linear transformation that occurs upon information arrival from previous layers conditioning the propagation of the signal to the next layer. The occurrence of *P. rustica* is represented in ANN output by either the letter P (Presence) or the letter A (Absence).

Two temporary independent data sets (training and test) were used to build the model, both containing pairs of input and output vectors taken from real data. The first set (data from the 1950s) was used for training and internal testing, and the second set (data from the 1990s) was used for model validation.

Training was performed with a standard feed-forward neural network (FNN), with an error back-propagation algorithm. Learning rate was of 0.02 and weight decay was set to -0.002. In Nevprop, the Concordance Index (*c-index*) provides a measure of model accuracy. The *c-index* is a nonparametric measure of discrimination (the ability to separate output categories) and is

approximately equal to the area under curve (AUC) obtained from receiver operator characteristic (ROC) curves (McPherson *et al.*, 2004; Ozesmi *et al.*, 2006). To test for the best FNN architecture, the *c-index* was calculated for each produced model as the average value resulting from 10 model runs. After testing several combinations of number of neurons, the network that provided the best results was composed of 8 neurons, displayed in a single hidden layer, connected to one output that gave the probability of presence/absence of *P. rustica*. The input layer consisted of 12 environmental variables, as described above (Fig. 2).

Data were also analyzed through an FNN with exactly the same architecture, running parameters and replicas, but with no hidden layers, mimicking a general linear model (GLM). Comparison of the FNN with the GLM is recommended as a standard procedure, since if a GLM performs as well as or better than the FNN during training, it means that interactions or non-linear terms do not need to be modelled (Goodman, 1996; Ozesmi *et al.*, 2006). The final FNN model was used to forecast the distribution of *P. rustica* within the study area from the 2020s and onwards, using decadal averages up to the 2090s. Environmental variables for these periods were used along with the FNN architecture parameters and the weights obtained during the training phase. Full details of the FNN model (weights and activation functions) are available from the authors upon request.

RESULTS

During training, GLM achieved a *c-index* of 0.939, which means that only a few classification errors occurred. However, the GLM was not able to correctly model the absences in the distributional gap observed in the 1950s, filling it with false presences. Moreover, all model misclassifications (a total of 13) in the training phase corresponded to false presences.

The FNN model performed generally better than the GLM, with an average *c-index* in the training set of 0.986. The FNN correctly modelled the absence of *P. rustica* on some shores within the distributional gap in northern Portugal during the 1950s. Moreover, it correctly identified the northern boundary at Capbreton (location 7 in Fig. 1B). There were only 6 misclassification errors, all false presences, located in Portugal and southern Spain. From these, three were in Tanger, Rota and Sesimbra (locations 1, 2 and 3 in Fig. 1B), with probability values of 0.82, 0.51 and 0.69, respectively. These are the only three shores within the 'normal' range of *P. rustica* (gap excluded) where this limpet was not observed during the 1950s surveys. The other misclassifications occurred in the localities of Leixões, Vila do Conde and Montedor (locations 4, 5 and 6 in Fig. 1B), within the gap. The presence probabilities in these shores were very close to the score threshold (0.55, 0.54 and 0.54, respectively). It is worth noticing that during the Fischer-Piètte and Gaillard (1959) sur-

veys a single *P. rustica* specimen was observed at Vila do Conde, but in Lima *et al.* (2007b) and in the present study that record was considered as an absence. When the FNN model was fed with the environmental and biological data for the 1990s, the *c-index* was 0.996, reflecting a high accuracy and no misclassification errors (Fig. 1C).

The model's forecast for the potential distribution of *P. rustica* in the 2020s resulted in occurrence probabilities close to one, varying between 0.91 and 0.99, throughout the study area. These results indicate that the climatic changes occurring during the next decade will probably lead to increases in habitat suitability northwards from Capbreton (Fig. 1D). It was also predicted that the ongoing climatic changes will continue to favour the presence of *P. rustica* in northern Portugal, indicating the permanent closure of the historical distributional gap. Since the forecasts for the 2020s revealed that *P. rustica* can potentially exist throughout the entire study area, forecasts for 2030s onwards were not included in the present study.

DISCUSSION

Although the simple study of past or recent distributional changes can be used to retrospectively describe the events and the environmental framework correlated with such changes, it does not permit hypothesis testing and therefore has limited scientific value. On the other hand, distributional models can be used to generate predictions and forecasts that can be subsequently tested. Thus, it is unfortunate that although descriptions of changes in the distribution of marine species are abundant in the recent scientific literature (Lima *et al.*, 2007a), distributional models have seldom been used to forecast future changes. There are several apparent reasons for this. First, the inaccessible nature of marine habitats, coupled with the high mobility of many of their inhabiting species, make it difficult to establish their distributional boundaries with precision (MacLeod *et al.*, 2008). Second, although historical datasets are fundamental for model training and independent validation (Araújo *et al.*, 2005), the scarcity and constraints associated with historical data frequently limit both the number of species likely to be modelled and the geographical/temporal domain of the models. In many cases, historical inventories of species are published in non-digitized regional journals, herbarium records, museum reprints, technical reports and unpublished academic theses. Sometimes the usable information is limited because most of the earlier studies did not cover the entire distributional range of target species. In many cases, abundance data are too vague and species lists are incomplete, often lacking exact sampling dates. Additionally, ambiguous taxonomic classifications add another uncertainty level. Considering the aforementioned difficulties, *P. rustica* offers a unique opportunity in this research field. Its taxonomy has been stable since the 1800s, and its historical and

current distributional range are well documented (e.g. Fischer-Piètte, 1955, 1959; Lima *et al.*, 2006; Lima *et al.*, 2007b). Similarly to other studies (Muñoz and Felicísimo, 2004; Ozesmi *et al.*, 2006), results from the present work indicate that the relations between the environment and organisms are often very complex. These may partly explain why the GLM failed to model the distribution of *P. rustica*, especially within the historical distributional gap observed in the 1950s. It is also possible that some of the relations between the predictor variables and the presence/absence of *P. rustica* are highly non-linear, a situation that GLMs cannot easily handle. Furthermore, despite the efforts to use climatic variables that are as independent as possible, they are still correlated (e.g. air temperature and sea surface temperature), which may have also decreased the performance of the GLM (Graham, 2003).

Several studies have compared the performance of various modelling techniques for predicting the future distribution of species (Guisan and Zimmermann, 2000; Heikkinen *et al.*, 2006). When weighted against other linear and non-linear modelling techniques (such as GLM and CART analyses), most studies illustrate the greater robustness of ANNs and their remarkable generalization capacity (e.g. Olden and Jackson, 2002; Araújo *et al.*, 2005). Similarly, the approach adopted in this work performed slightly better during the training phase than the CART analysis carried out by Lima *et al.* (2007b), and more importantly, passed the validation step after correctly predicting both the bridging of the historical gap in northern Portugal and the northern distributional limit observed at Capbreton (SW France) in the late 1990s. Also, unlike previous studies, the ANN model suggested that the interplay between the environmental factors considered in the analysis was sufficient to fully explain the current distribution of *P. rustica*. The success of this validation meant that the model could then be used for producing forecasts of the future distribution of this species.

When fed with a median-high emission scenario, the ANN model predicted a 1000-km range expansion northwards from Capbreton, encompassing the entire study area. The model's forecasts also indicate that the presence of *P. rustica* in northern Portugal will continue in the future, and what could be initially argued as a temporary change in the distribution of the species will probably become permanent. It has recently been argued that climate change is likely to have more complex outcomes than simple shifts in the abundances or distribution of isolated species (Poloczanska *et al.*, 2008). Changes in the composition or in the abundance of key species in response to environmental changes have been shown to cause cascading effects on ecosystems functioning (see for example Chapin *et al.*, 2000). Species of *Patella* are considered to be key ecosystem disturbers, due to their role as primary consumers and regulators of macro-algal biomass and composition (Thompson *et al.*, 2002; Firth and Crowe, 2008). Within the group of NE Atlantic patellid species, *Patella*

rustica is highly selective in relation to microhabitat occupation, living on steep or vertical rocky walls directly exposed to wave action (Lima *et al.*, 2006). Since other grazers are absent in this particular microhabitat (with the exception of the small littorinid *Melaraphe neritoides*), a northwards range expansion of *P. rustica* may impact local communities as it is highly likely that they have not experienced such a grazing pressure in their recent evolutionary history. Also, the absence of strong competitors will probably facilitate the geographical expansion of *P. rustica* into an unoccupied ecological niche. Therefore, although dispersal barriers are likely to occur north of Capbreton (Lima *et al.*, 2007b), it is reasonable to expect a northward range expansion of *P. rustica* once the habitat becomes favourable in NW France. Considering the absence of competition, even a stochastic event such as an extraordinary storm carrying a small number of larvae or a human-mediated introduction may be sufficient to allow for successful colonization of NW France. Judging from the recent colonization of northern Portugal by *P. rustica*, a suitable habitat disjunct from its current distributional range may be occupied in timeframes as short as a few years (Lima *et al.*, 2006).

ACKNOWLEDGEMENTS

R. Xavier was supported by an FCT grant (ref SFRH/BD/29370/2006). F. P. Lima was supported by FCT and NOAA Ecological Forecasting grants (ref SFRH/BPD/34932/2007 and NA04NOS4780264, respectively). The authors would like to thank Pedro Tarroso, Pedro Ribeiro and Nuno Queiroz for their help and support during the several stages of this work.

REFERENCES

- Araújo, M.B., R.G. Pearson, W. Thuiller and M. Erhard. – 2005. Validation of species–climate impact models under climate change. *Glob. Change Biol.*, 11(9): 1504–1513.
- Brosse, S., S. Lek and F. Dauba. – 1999. Predicting fish distribution in a mesotrophic lake by hydroacoustic survey and artificial neural networks. *Limnol. Oceanogr.*, 44(5): 1293–1303.
- Brown, J.H., T.J. Valone and C.G. Curtin. – 1997. Reorganization of an arid ecosystem in response to recent climate change. *Proc. Natl. Acad. Sci. U.S.A.*, 94: 9729–9733.
- Chapin, F.S., E.S. Zavaleta, V.T. Eviner, R.L. Naylor, P.M. Vitousek, H.L. Reynolds, D.U. Hooper, S. Lavorel, O.E. Sala, S.E. Hobbie, M.C. Mack and S. Díaz. – 2000. Consequences of changing biodiversity. *Nature*, 405: 234–242.
- Crisp, D.J. and A.J. Southward. – 1958. The distribution of intertidal organisms along the coasts of English Channel. *J. Mar. Biol. Ass. U. K.*, 37: 1031–1048.
- Ekeboom, J., P. Laiho and T. Suominen. – 2003. A GIS-based step-wise procedure for assessing physical exposure in fragmented archipelagos. *Estuar. Coast. Shelf Sci.*, 57(5): 887–898.
- Farrar, D.E. and R.R. Glauber. – 1967. Multicollinearity in regression analysis: the problem revisited. *Rev. Econ. Statist.*, 49: 92–107.
- Firth, L.B. and T.P. Crowe. – 2008. Large-scale coexistence and small-scale segregation of key species on rocky shores. *Hydrobiologia*, 614: 233–241.
- Fischer-Piètte, E. – 1955. Répartition, le long des côtes septentrionales de l'Espagne, des principales espèces peuplant les roches intercotidales. *Ann. Inst. Océanogr. Paris*, 31: 37–124.
- Fischer-Piètte, E. – 1959. Répartition des principales espèces intercotidales de la côte atlantique Française en 1954–1955. *Ann. Inst. Océanogr. Paris*, 36: 275–388.
- Fischer-Piètte, E. – 1963. La distribution des principaux organismes intercotidaux Nord-Ibériques en 1954–1955. *Ann. Inst. Océanogr. Paris*, 40: 165–312.
- Fischer-Piètte, E. and J.-M. Gaillard. – 1959. Les patelles au long des côtes atlantiques Ibériques et Nord-Marocaines. *J. Conchol.*, 99: 135–200.
- Goodman, P.H. – 1996. NevProp software, version 3. In: URL:ftp://ftp.scs.unr.edu/pub/cbmr/nevpropdir/ (ed.), University of Nevada, Reno, NV.
- Gordon, C., C. Cooper, C.A. Senior, H.T. Banks, J.M. Gregory, T.C. Johns, J.F.B. Mitchell and R.A. Wood. – 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dynam.*, 16(2/3): 147–168.
- Graham, M.H. – 2003. Confronting multicollinearity in ecological multiple regression. *Ecology*, 84(11): 2809–2815.
- GRASS. – 2006. Geographic Resources Analysis Support System (GRASS). Trento, Italy.
- Guisan, A. and W. Thuiller. – 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.*, 8(9): 993–1009.
- Guisan, A. and N.E. Zimmermann. – 2000. Predictive habitat distribution models in ecology. *Ecol. Model.*, 135(2): 147–186.
- Heikkinen, R.K., M. Luoto, M.B. Araújo, R. Virkkala, W. Thuiller and M.T. Sykes. – 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Prog. Phys. Geogr.*, 30: 751–777.
- Hijmans, R.J. and C.H. Graham. – 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Glob. Change Biol.*, 12: 2272–2281.
- Johns, T., C. Durman, H. Banks, M. Roberts, A. McLaren, J. Ridley, C. Senior, K. Williams, A. Jones, A. Keen, G. Rickard, S. Cusack, M. Joshi, M. Ringer, B. Dong, H. Spencer, R. Hill, J. Gregory, A. Pardaen, J. Lowe, A. Bodas-Salcedo, S. Stark and Y. Searl. – 2004. HadGEM1 – Model description and analysis of preliminary experiments for the IPCC Fourth Assessment Report. *Hadley Centre technical note 55*.
- Lek, S. and J.F. Guégan. – 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.*, 120(2/3): 65–73.
- Lima, F.P., N. Queiroz, P.A. Ribeiro, S.J. Hawkins and A.M. Santos. – 2006. Recent changes in the distribution of a marine gastropod, *Patella rustica* Linnaeus, 1758, and their relationship to unusual climatic events. *J. Biogeogr.*, 33: 812–822.
- Lima, F.P., P.A. Ribeiro, N. Queiroz, S.J. Hawkins and A.M. Santos. – 2007a. Do distributional shifts of northern and southern species of algae match the warming pattern? *Glob. Change Biol.*, 13: 2592–2604.
- Lima, F.P., P.A. Ribeiro, N. Queiroz, R. Xavier, P. Tarroso, S.J. Hawkins and A.M. Santos. – 2007b. Modelling past and present geographical distribution of the marine gastropod *Patella rustica* as a tool for exploring responses to environmental change. *Glob. Change Biol.*, 13: 2065–2077.
- MacLeod, C.D., L. Mandleberg, C. Schweder, S.M. Bannon and G.J. Pierce. – 2008. A comparison of approaches for modelling the occurrence of marine animals. *Hydrobiologia*, 612: 21–32.
- Manel, S., J.M. Dias, S.T. Buckton and S.J. Ormerod. – 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *J. Appl. Ecol.*, 36(5): 734–747.
- McLaughlin, J.F., J.J. Hellmann, C.L. Boggs and P.R. Ehrlich. – 2002. Climate change hastens population extinctions. *Proc. Natl. Acad. Sci. USA*, 99(9): 6070–6074.
- McPherson, J.M., W. Jetz and D.J. Rogers. – 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.*, 41(5): 811–823.
- Mitchell, T.D., T.R. Carter, P.D. Hulme and M. New. – 2004. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). *Tyndal Working Paper*, 55: 1–30.
- Muñoz, J. and Á.M. Felicísimo. – 2004. Comparison of statistical methods commonly used in predictive modelling. *J. Veg. Sci.*, 15: 285–292.
- Nakicenovic, N., O. Davidson, G. Davis, A. Grübler, T. Kram, E.L.L. Rovere, Bert Metz, T. Moritz, W. Pepper, H. Pitcher, A.

- Sankovski, P., Shukla, R., Swart, R., Watson and Z. Dadi – 2000. IPCC Special Report on Emission Scenarios. *IPCC WGIII. Intergovernmental Panel on Climate Change*.
- Olden, J.D. and D.A. Jackson. – 2002. A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biol.*, 47(10): 1976-1995.
- Ozesmi, S.L., C.O. Tan and U. Ozesmi. – 2006. Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecol. Model.*, 195: 83-93.
- Park, Y.-S., R. Céréghino, A. Compin and S. Lek. – 2003. Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters. *Ecol. Model.*, 160(3): 265-280.
- Parmesan, C. – 2006. Ecological and Evolutionary Responses to Recent Climate Change. *Annu. Rev. Ecol. Evol. Syst.*, 37: 637-669.
- Pearson, R.G. and T.P. Dawson. – 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecol. Biogeogr.*, 12(5): 361-371.
- Pearson, R.G., Dawson, T.P. and C. Liu. – 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, 27: 285-298.
- Poloczanska, E.S., S.J. Hawkins, A.J. Southward and M.T. Burrows. – 2008. Modeling the response of populations of competing species to climate change. *Ecology*, 89(11): 3138-3149.
- Rayner, N.A., P. Brohan, D.E. Parker, C.K. Folland, J.J. Kennedy, M. Vanicek, T. Ansell and S.F.B. Tett. – 2006. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the new HadSST2 data set. *J. Clim.*, 19(3): 446-469.
- Ribeiro, P.A., R. Xavier, A.M. Santos and S.J. Hawkins. – 2009. Reproductive cycles of four species of *Patella* (Mollusca: Gastropoda) on the northern and central Portuguese coast. *J. Mar. Biol. Ass. U. K.*, 89(6): 1215-1221.
- Soberón, J. and A.T. Peterson. – 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Inform.*, 2: 1-10.
- Spitz, F. and S. Lek. – 1999. Environmental impact prediction using neural network modelling. An example in wildlife damage. *J. Appl. Ecol.*, 36(2): 317-326.
- Thompson, R.C., T.P. Crowe and S.J. Hawkins. – 2002. Rocky intertidal communities: past environmental changes, present status and predictions for the next 25 years. *Environ. Conserv.*, 29(2): 168-191.
- Tu, J.V. – 1996. Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes. *J. Clin. Epidem.*, 49(11): 1225-1231.
- Vaughan, I.P. and S.J. Ormerod. – 2003. Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conserv. Biol.*, 17(6): 1601-1611.
- Woddruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley. – 1988. ICOADS release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, 23: 517-526.
- Yuval, N. – 2001. Enhancement and Error Estimation of Neural Network Prediction of Niño-3.4 SST Anomalies. *J. Clim.*, 14: 2150-2163.

Scient. ed.: F. Maynou.

Received September 1, 2009. Accepted January 27, 2010.

Published online June 29, 2010.