Modelling spatial distribution of hard bottom benthic communities and their functional response to environmental parameters

Vinko Bandelj a, Daniele Curiel b, Sovan Lek c, Andrea Rismondo b, Cosimo Solidoro a,∗

a Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, Borgo Grotta Gigante - Brčički 42/c, 34010 Sgonico - Zgonik, Italy
b SELC, Via dell’Elettricità 3/d, 30175 Venezia-Marghera, Italy
c Laboratoire Evolution Diversité Biologique, UMR 5174, CNRS - Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cédex 4, France

ABSTRACT

In this study we analyzed and modelled spatial distribution of hard bottom benthic communities in the Lagoon of Venice, and used the model to derive functional response of these communities to changing environmental conditions.

Three communities were identified by using a fuzzy clustering algorithm to experimental observations (macrophytobenthos and macrozoobenthos data), and interpreted from a biological perspective by considering their indicator and dominant taxa. The results showed that the lagoon can be divided in areas dominated by a ‘marine’, an ‘intermediate’ and a ‘confined’ benthic community. Some differences in the extension of the three communities between the two years of monitoring were observed and related to significant differences in environmental parameters.

The relationships between benthic communities’ membership grades and a set of water quality, hydrodynamics and sediment composition parameters were then modelled by means of an artificial neural network model, using a back-propagation algorithm. Furthermore, the model was used to assess the relative contribution of predictors on membership grades of different communities, and to understand the relationships between cluster membership grades and predictors with highest relative contribution. Several responses of benthic communities to environmental parameters were non-linear and complex, which confirmed the advantage of using neural networks instead of other, more traditional statistical methods. The benthic community showed mainly ‘marine’ characteristics with low temperature, silt percentage in sediments and POC concentration in water. The ‘confined’ characteristics were pronounced in areas of low PO4 and high DOP concentrations in water, low water energy and on wooden substrates. The ‘intermediate’ community was that predicted worst and was related mainly to high water salinity and temperatures around 16 °C.

Finally, the model was used as a tool for the detection of anomalies, i.e., cases in which the predicted community was different from that observed. The anomalies were inspected and related to possible causes, such as a not optimal substrate, very low-bottom depth, complex local morphology, human disturbances, and slow adaptation of communities to changing abiotic characteristic.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Macrobenthos is a very important component of transitional environments. It lives on or near the sea bottom, and includes macrophytobenthos and macrozoobenthos organisms. Macroalgae and macrophytes are important primary producers in transitional ecosystems. Macrozoobenthos includes a variety of animals which may feed as suspension filter feeders or as detrivores, playing a very active role in recycling and remineralization. Both macrophytobenthos and macrozoobenthos are also a source of food for fishes and other organisms. Furthermore, fouling of benthic species may be a major problem for submersed structures, and the knowledge on environmental preferences of edible species (oysters, mussels) is important for aquaculture purposes. In fact, papers devoted to the analysis of distribution of macrozoobenthos (Millet and Guélorget, 1994; Sconfietti et al., 2003; Kevrekidis, 2004) or macrophytobenthos (Sfriso and Marcomini, 1996; Vincent et al., 2006) in marine systems are quite common.

However, relatively few papers deal with hard bottom benthic communities in lagoons (Guélorget and Perthisois, 1992; Saiz-Salinas and Utkiaga-Alberdi, 1998; Sconfietti et al., 2003; Benedetti-Cecchi et al., 2001; Marchini and Marchini, 2006; Marchini and Occhipinti-Ambrogi, 2007). This is probably due to
the fact that natural hard bottoms are not frequent in lagoons. Several studies can be found in literature, in which implantation and growth of hard bottom benthic organisms on purposely positioned artificial substrates has been observed (Ardisson and Bourget, 1997; Benedetti-Cecchi et al., 2001). The Lagoon of Venice is an appropriate site for studies on hard bottom benthic communities, because of its habitat heterogeneity and of the large number of submerged artificial hard substrates, which underwent a process of naturalization after a long permanence in the water.

In a previous paper (Bandelj et al., 2009) we analyzed the structure and composition of hard bottom macrophytobenthos and macrozoobenthos communities in the Lagoon of Venice using data collected in 2004. The analysis of species and community distribution, and the results of a direct gradient analysis, showed that it was possible to recognize a succession of benthic communities along a gradient from inner confined areas towards areas close to the lagoon inlets influenced by exchanges with the sea. These results were partially in agreement with specific literature on the Lagoon of Venice (Occhipinti Ambrogi et al., 1998; Sconfietti et al., 2003) and with existing theory on biological zonation in transitional waters (Guélorget and Perthuisot, 1992). However, no single environmental parameter was able to adequately model the distribution of the benthic communities in the lagoon (Bandelj et al., 2009).

In the present study, additional data for 2005 were available (project funded by Magistrato alle Acque di Venezia - the Venice Water Authority), and we performed a joint analysis on the data of 2004 and 2005 in order to: (a) corroborate the previous results; (b) use neural networks to explore the existence of complex and non-linear relationships between environmental factors and benthic communities; (c) use the model as a tool for the detection of anomalies. In particular, while the number of samples in 2004 dataset permitted only a statistical analysis of community organization, the use of data from two years entailed the possibility to develop a quantitative model.

Therefore, a fuzzy k-means (Bezdek, 1981) was used to objectively group the samples in a number of clusters, which identified homogenous benthic communities. The use of a fuzzy clustering procedure was preferred to a crisp one because it may better resemble the continuum of conditions that is usually observed in natural environments (Equihua, 1990; Silvert, 2000). Then, an artificial neural network (Zupan and Gasteiger, 1999) was trained to model the functional response of benthic communities to environmental factors. Artificial neural networks were used because they are usually regarded as suitable models for high-dimensional, complex datasets, where the form of the functions is not known a priori (Lek et al., 1996; Lek and Guégan, 1999).

![Fig. 1. Map of the study area. The numbers show the sites sampled in the Lagoon of Venice in 2004 and 2005.](image-url)
The validated neural network model represents the response of the benthic communities to the environmental conditions, as they are described by the available environmental parameters. Thus, the model may be used as a tool for the identification of anomalies, i.e., cases where the observed benthic community is different from the expected. The anomalies may be due to features which possibly affect only one or a few sites, and which are not accounted for by any predictor included in the model. The anomalies thus deserve a closer inspection so as to understand what caused the model to fail in prediction.

1.1. Study site

The Lagoon of Venice is located in the Northern Adriatic Sea (Fig. 1) and it is one of the largest Mediterranean lagoons with an area of around 500 km². The present characteristics of the lagoon result from centuries of interactions and interleaveing of natural forcing, mainly the opposite influence of sea and coastal estuaries, with anthropogenic activities, such as navigation, industrial harbour activities, channel dredging, aquaculture, etc. (Ravera, 2000).

It is surrounded by heavily inhabited shores, important industrial areas, fisheries, aquaculture and tourist activities and it hosts sites of great historical and artistic value. It is a complex system, which shows strong gradients for physical and chemical (Solidoro et al., 2004a), morphological (Molinari et al., 2007) and hydrological parameters (Solidoro et al., 2004b). There are 12 major tributaries showing strong gradients for physical and chemical (Solidoro et al., 2004a), morphological (Molinari et al., 2007) and hydrological parameters (Solidoro et al., 2004b). A complex network of channels, originating from the inlets, divides the lagoon in a large number of morphological units, typically grouped in four sub-basins (Solidoro et al., 2004b). Average water depth is around 1 m in shallows, while reaches up to 10 m in navigation channels connecting the three inlets with Venice and with Marghera harbour. General circulation is almost totally tidally dominated (Gačić et al., 2005). A complex model of the circulation of the lagoon has been developed (Bandelj et al., 2009) to simulate the flow within the lagoon. The model is based on the finite element method (Umgiesser et al., 2004), forced by observed wind and wave fields and by observational data from the inlets and other sources. The model is able to reproduce the main features of the circulation of the lagoon, such as the large-scale circulation in the inner area (Cucco et al., 2009). The lagoon has high habitat heterogeneity, and different methods have been proposed in order to identify homogeneous subareas (Solidoro et al., 2004b; Marchini and Marchini, 2006).

2. Material and methods

2.1. Macrophytobenthos and macrozoobenthos data

Long-existing submerged artificial substrates, such as embankments and footings of buildings, concrete, wooden and iron piles, were selected as sampling sites. Their location covered evenly a wide range of lagoon habitats (shallows, tidal flats, sea inlets, channel side-shallows, etc.) with the exception of channels and industrial and urban areas. Samples were collected in 80 sampling sites, evenly distributed over the basin (Fig. 1) in 2004 and in 2005, within the frame of the MELO3 (Lagoon Ecosystem Monitoring) project, funded by Magistrato alle Acque di Venezia (Mag. Acque–SELC, 2006). Sampling was done from June to July of each year, scratching an area of 2000 cm² with a blade of 1 m in length and 0.2 m in width, from 0.2 m above the mean sea level to 0.8 m below the mean sea level.

Sampled organisms were separated during a sorting phase, determined to the microscope and the total surface covered by each species recorded in cm², according to the methodology of Boudouresque (1971). The sum of the surfaces covered by all organisms in a sampling site could exceed that of the sampled area. The majority of the organisms was identified at a species or genus level.

2.2. Water biogeochemical properties

Values of 11 physical and chemical parameters (temperature, pH, salinity, alkalinity, chlorophyll a, chlorophyll b, nitrogen, phosphorus, dissolved organic carbon, particulate organic carbon, dissolved organic nitrogen, dissolved organic phosphorus) were measured monthly during 2004 in 21 sampling sites (Mag. Acque–Cossarini et al., 2006). Values in the 80 sites where benthic organisms were sampled were computed by first interpolating the 21 data over a 250 m × 250 m grid and then averaging over the four grid points closest to each sampling site. Interpolation was performed with the inverse distance to a power method (Davis, 1986) with a searching radius of 3000 m on yearly averaged values.

2.3. Hydrodynamic and morphological properties

The interpolation procedures applied for water biogeochemical properties were used also to compute sediment composition as percentage of sand, clay and silt for each sampling site. Original data on the first 5 cm of sediment were measured in 103 locations in the Lagoon of Venice in 2004 (Mag. Acque–SELC, 2004).

The meadow residence time (TRES_M) and the root mean square velocity (RMS_V) for each sampling site were computed using a finite element model (Umgiesser et al., 2004), forced by observed values of wind stress and sea level in 2004–2005, following the procedures described in Bandelj et al. (2009). Values for the two parameters were computed for each sampling site every 3 days and then averaged.

Water depth and type of substrate for each sampling site were recorded during sampling activities, and the type of substrate was coded as 3 dummy variables: concrete, wood, iron.

3. Statistical methods

Fig. 2 summarizes the statistical procedure and methods used in this study. More details are given below.

3.1. Benthic communities

The so-called ‘double zero’ problem (Legendre and Legendre, 1998) may arise when methods relying on Euclidean distance, such as fuzzy k-means, are applied on data with extremely skewed distribution. A possible solution is to transform the data with adequate transformations, such as those presented in Legendre and Gallagher (2001). Since the distribution of the cover data was extremely skewed (95 out of 139 identified taxa were found in less than 10 sampling sites), we transformed them with the Hellinger transformation (Legendre and Gallagher, 2001). In this way fuzzy k-means preserved the Hellinger distance instead of the Euclidean distance between samples.

Most clustering methods can only give a crisp partition of samples between groups, thus introducing discontinuities which may not exist in reality, where gradual transitions from one habitat or community to another are observed. The fuzzy classification procedures may be more appropriate for clustering of natural environments. Fuzzy k-means (Bezdek, 1981; Odeh et al., 1992), a general case of the k-means, gives to all samples a membership grade in respect to a number of classes. Thus, each sample is not associated to a single class, as in crisp clustering, but with different membership grades to all classes. A confusion index (CI) is also associated to each sample and it is a measure of the degree of class overlap in the space of attributes. A low CI indicates a sample which
is clearly associated to one class, a high CI indicates a sample associated with similar membership value to more clusters. Prior to clustering, the analyst has to choose the number of classes \( N \) and the fuzzy exponent \( \Phi \). The fuzzy \( k \)-means was run on all 160 samples for 2–8 classes with \( \Phi \) increasing by steps of 0.1 from 1.1 to 2. The best combination of \( N \) and \( \Phi \) was obtained by minimizing the Normalized Classification Entropy and the Fuzziness Performance Index (Roubens, 1982; Odeh et al., 1992).

The biological description of the clusters was based on their indicator and dominant species. The IndVal method (Dufrêne and Legendre, 1997) was used to identify indicator species of each cluster. A species is an indicator (high IndVal) if in a certain class shows high specificity (i.e., it was observed only in that class) and high fidelity (i.e., it was observed in all samples of that class). Dufrêne and Legendre (1997) proposed also two tests which give an indication on the strength of relation between indicator taxa and classes. Since IndVal can be applied only to discrete clusters, a defuzzification of fuzzy \( k \)-means results was obtained assigning each sample to the class with highest membership grade. Indicators were considered taxa with a value of IndVal > 25%, significant after 499 permutations between samples and cluster memberships (Dufrêne and Legendre, 1997). As for dominant species, a species was defined as dominant if it contributed to 90% of the total sum of covers in each discrete cluster.

### 3.2. Artificial neural network model

We used the multilayer perceptron with back-propagation algorithm (back-propagation neural network-BPNN) (Zupan and Gasteiger, 1999) to model the relationships between the membership grades of the samples and the environmental variables. All three membership grades for each sample were predicted together, using all 20 environmental variables as predictors. 8 BPNNs were computed with 1000 training epochs and with increasing number of hidden neurons from 3 to 10. Because the number of samples was not enough to split the dataset in a training and testing set, the leave-one-out (LOO) procedure was used to validate the model. In the LOO the target value of each sample is in turn predicted using the predictors of all other samples. The best model was chosen according to the correlation coefficient between the target and predicted values in LOO. The final model was then built using all available samples: 10 runs of the BPNN were performed and the mean was considered as the best prediction for each sample.

With the partial derivatives method (Dimopoulos et al., 1995) and the profile method (Lek et al., 1995) we assessed the importance of environmental parameters in predicting the communities, and estimated the functional form of the responses of benthic communities to the environmental parameters.

The relative importance, i.e., variable contributions of each predictor on membership grades of each cluster, was assessed with the partial derivatives method. For each target variable the partial derivative in respect to all values of each predictor is computed. The sum of squares of partial derivatives for each predictor gives an estimate of the relative importance of that predictor on each target variable.

With the profile method we inspected how the target values change in response to variations in predictors. The range of variation of each predictor is divided in 12 bins and for each value a BPNN is run, while all other predictors are simultaneously set at their minimum, then at their first quartile, median, third quartile and maximum value. The results of the 5 runs are averaged and assumed to be the response of the target variable for each of the 12 fixed values of the predictor. The procedure is then repeated for each predictor.
To assure a more robust assessment, we searched for anomalies only after defuzzification of model results, i.e., we considered as possibly anomalous only samples for which the cluster with highest membership grade predicted by the BPNN was different from the cluster with highest membership grade as resulting from fuzzy classification.

4. Results

We identified a total number of 168 taxa in year 2004 and 164 taxa in year 2005. Only the 139 taxa sampled both years were retained for the analysis (Supplementary data Table 1). Of these, there were 72 macrophytobenthos taxa (39 rhodophytes, 24 chlorophytes, and 9 phaeophytes) and 67 macrozoobenthos (11 molluscs, 10 tunicates, 9 bryozoans, 9 polychaetes, 8 crustaceans, 7 poriferans, 6 hydrozoans, 5 echinoderms, 1 anthozoans, 1 polyplacophors). The total cover of macrozoobenthos (211290 cm²) was almost four times that of macrophytobenthos (58629 cm²).

4.1. Fuzzy clustering

The best combination of fuzzy $k$-means parameters was $\Phi = 1.5$ and $N = 3$. The median CI was 0.797, minimum was 0.377 (site 43 in 2004) and maximum 0.999 (site 6 in 2005). In Supplementary data Table 2 the results of the fuzzy clustering are reported for the two years: it should be kept in mind that all 160 samples were clustered together. Several sampling sites were clearly associated to one of the 3 clusters (low CIs): e.g., sites 68, 70 and 73 in the inner southern lagoon, or the sites close to the lagoon inlets. Other sites had membership grades almost equal for 2 (e.g., sites 26, 27, 34 in 2004, and 50, 62 in 2005), or even all 3 clusters (e.g., sites 10, 38 in 2004, and 3, 7 in 2005) and showed high CI values.

The three clusters are represented with white, grey and black markers in Fig. 3. Sites showing highest membership for the ‘white’ cluster had a mean cover of 2361 cm² and a mean number of 42 taxa per site. For the ‘grey’ cluster the mean cover was 1651 cm² and the mean number of taxa per site was 36. For the ‘black’ cluster 1004 cm² was the mean cover per site and 25 was the mean number of taxa per site.

4.2. Interannual variability

We observed some differences in the clusters of maximum membership grades for single sampling sites in the two years under study. In particular, 21 sites were assigned with maximum membership grade to different clusters in 2005 compared to 2004. In 2004 20 sampling sites showed the highest membership grade for the ‘white’ cluster, 30 for the ‘grey’ and 26 for the ‘black’. In 2005 30 sampling sites showed the highest membership grade for the ‘white’ cluster, 30 for the ‘grey’ and 20 for the ‘black’. This can be seen better in Fig. 3, where each sampling site is associated to the cluster of maximum membership grade. 6 sites changed from ‘black’ to ‘grey’ cluster and 3 sites from ‘black’ to ‘white’ cluster, 8 sites changed from ‘grey’ to ‘white’ cluster and 3 from ‘grey’ to ‘black’ cluster, and only one site changed from ‘white’ to ‘grey’. The CI for these sites was rather high in both years (median 0.883 in 2004 and 0.859 in 2005), but for many of them decreased from 2004 to 2005.

An ANOVA revealed significant differences ($p < 0.05$) in the mean values between 2004 and 2005 for 7 variables (Fig. 4): temperature, chlorophyll, NO$_3$, and DOP values decreased, while salinity, DOC and DON values increased.

4.3. Biological characterization of clusters

There were 18 indicator taxa (11 zoobenthos, 7 phytobenthos) in the ‘white’ cluster (Table 1), 6 (3 zoobenthos, 3 phytobenthos) in the ‘grey’ cluster and 5 (4 zoobenthos, 1 phytobenthos) in the ‘black’ cluster. The maximum IndVal value amounted at more than 80% in the ‘black’ (Balanus eburneus 84.76%) and in the ‘white’ cluster (Mytilus galloprovincialis 83.78%, Hiatella arctica 80.13%), while it was around 50% for the ‘grey’ cluster (Crassostrea gigas 53.46%).

There were 14 dominant taxa in the ‘white’ cluster (Table 2), and 16 in both, the ‘grey’ and the ‘black’ cluster. Of these, 9 taxa were dominant in all 3 clusters, and 4 taxa were dominant in 2 clusters.

Fig. 3. Discretization of the results of fuzzy k-means. Each site was associated to the cluster of maximum membership grade for years 2004 (left) and 2005 (right).
thus, only 24 taxa of phytobenthos and zoobenthos built up more than 90% of cover of all 3 clusters.

4.4. Modelling relationships between communities and habitat characteristics

Table 3 shows the correlation coefficients between target and predicted values for BPNNs with a number of hidden neurons from 3 to 10 and 1000 learning epochs, computed using leave-one-out procedure. The best overall prediction was obtained with the simplest architecture (3 hidden neurons, \( r = 0.76 \)) and the prediction quality was different for the three clusters. Membership grades of the ‘black’ cluster were predicted best \( (r = 0.83) \), followed by those of the ‘white’ cluster \( (r = 0.79) \) and of the ‘grey’ cluster \( (r = 0.62) \). Thus, the BPNN with 3 hidden neurons was retained as the best model.

The relative contribution of the predictors on membership grades of the 3 clusters was computed with the partial derivative method averaging the results of 10 BPNNs built with the entire dataset (Fig. 5). Predictors with highest relative effect for the ‘white’ cluster were temperature, POC and percentage of silt in the sediment. Salinity and temperature had the highest effect on the prediction of the ‘grey’ cluster membership grade. The ‘black’ cluster was best predicted by RMS \(_V\), wooden substrate, PO\(_4\) and DOP.

We used the profile method in order to understand the relationships between cluster membership grades and the predictors.
whose effect, according to the partial derivative method, exceeded 10%. Membership grade of the ‘white’ cluster decreased with increasing temperature, POC and silt percentage (Fig. 6). The shape of the curves was very similar for POC and silt, while the decrease as function of temperature was sharp for low and high temperature, but it was slow for intermediate temperature. The membership grade of the ‘grey’ cluster increased with increasing salinity, first sharply, than only slowly at higher salinity (Fig. 7). As function of temperature the membership grade of the ‘grey’ cluster showed first an increase with increasing temperature (Fig. 7). After a maximum, reached between 15.28 and 16.07 °C, a sharp decrease of membership grade could be observed for higher temperature. The relationship between membership grade of the ‘black’ cluster with PO4 and DOP seems linear (Fig. 8), but with inverse behaviour: the membership grade increased with decreasing PO4 and increasing DOP. A sharp decrease was observed for the ‘black’ cluster with increasing RMSV until a value of around 0.2 m s⁻¹ (Fig. 8). After that, increasing RMSV seems not to have had effect on membership grade of the ‘black’ cluster. Finally, the ‘black’ cluster membership was maximum for wooden substrate and minimum when the substrate was not wooden (Fig. 8).

4.5. Identification of anomalies

After defuzzification of model results, when each sample was assigned to the cluster with highest membership grade, we identified 22 cases of anomalies. For 7 sites (12, 26, 38, 46, 49, 63, 65) the samples were wrongly predicted only in 2004, for 10 (4, 10, 11, 24, 28, 45, 55, 59, 66, and 67) only in 2005 and for 2 sites (14, 64) the prediction was wrong for both years. Prediction for site 6 in 2005 was not considered as wrong because of the small difference as those for each of the three clusters. All reported correlation coefficients were statistically significant (two-tailed t-test, p-level = 0.05).

5. Discussion

5.1. Spatial distribution and biological characterization of benthic communities

The spatial distribution of the clusters and their biological description confirmed the results obtained analyzing only the 2004
Fig. 5. Relative importance of environmental variables as predictors for the 'marine' cluster membership (above), for the 'intermediate' cluster membership (middle) and for the 'confined' cluster membership (below), as computed with the partial derivatives method.

Data set (Bandelj et al., 2009). In particular, it is clear that the biological zonation in the Lagoon of Venice is influenced by the lagoon morphology and by hydrographical conditions that prevail in different lagoon areas: the 'white' cluster was found mainly close to the inlets and along the main channels; the 'grey' cluster had high membership grades for sites in low-depth areas of the central and southern basins; sites which showed preference for the 'black' cluster were those positioned landward, away from the main channels and near freshwater inputs. Thus, and according to Bandelj et al. (2009), we labelled the 'white' cluster as 'marine', the 'grey' cluster as 'intermediate' and the 'black' cluster as 'confined'.

Fig. 6. Relationship between the membership grade of the 'marine' cluster and the three most important predictors (relative importance >10%): temperature (above), POC (middle) and silt percentage in the sediment (below).
Many sampling sites were not associated clearly to only one cluster, but rather to 2 or even all 3. This confirmed that the spatial succession between different benthic communities is gradual and many sites show a benthic community with mixed characteristics. From Fig. 3 is evident that in some areas of the lagoon, most notably in the northern basin, the ‘marine’ and the ‘confined’ benthic community were found in sites close one to another, apparently without the interposition of the ‘intermediate’ cluster. The incomplete zonation was even more pronounced in 2005. Thus, from our results emerged that the complexity of the Lagoon of Venice reflects sharp biological gradients.

The biological description of the clusters is given in Tables 2 and 3. Dominant and indicator taxa bring different information about benthic communities. Indicator taxa show high specificity and fidelity to corresponding clusters (Dufrêne and Legendre, 1997), indicating particular environmental conditions. They can also be rare organisms and contribute little to the total cover of clusters. By definition, dominant taxa show highest cover values in each cluster, so that they are important also for biomass budgets or in marine fouling studies.

The 3 clusters showed a clear gradient in cover and biodiversity, as already observed in the Lagoon of Venice in 2004 (Bandelj et al., 2009), and in other systems (Guélorget and Perthuisot, 1992). The mean number of taxa per site and the mean cover per site were highest for the ‘marine’ cluster and lowest for the ‘confined’ cluster. The ‘marine’ cluster was characterized by species characteristic of high salinity, low turbidity and high energy waters, such as the bivalve *Mytilus galloprovincialis*, the red alga *Polysiphonia denudata*, the porifera *Hymeniacidon sanguinea* and the bryozoan *Tricellaria inopinata* (Ceccherelli and Rossi, 1984; Carballo et al., 1996; Occhipinti Ambrogi, 2000; Corriero et al., 2007). In the ‘intermediate’ cluster the most important species were molluscs *Crassostrea gigas* and green algae of the *Ulva* genus. Stronger penetration of *Crassostrea gigas* into the lagoon than *Mytilus galloprovincialis* is partially due to the high eurivalent characteristic of the first species (Relini et al., 1972). The low-bottom areas of the central and southern basin are the most suitable habitats for Ulvaceae (Sfriso et al., 2002; Curiel et al., 2001). In the ‘confined’ cluster crustaceans *Balanus balanoides* and *B. amphitrite*, and bryozoan *Bowerbankia gracilis* were the most important species. They are usually found in areas where conditions are critical for other species (Relini, 1980; Candela et al., 1983; Sconfietti, 1998).

5.2. Interannual variability

In the joint analysis on two years of samples we noted several differences between 2004 and 2005 in membership grades of the sites. This can be clearly seen when considering clusters’ maximum membership grades (Fig. 3). All sites with pronounced differences between the two years showed a rather high CI for both years, but in most cases it was lower in 2005 than in 2004. Also the median CI for all sites was lower in 2005 than in 2004, indicating that possibly the zonation in 2005 was sharper. The changes from 2004 to 2005 resulted in a net increase in memberships of the ‘marine’ cluster and in a decrease in memberships of the ‘confined’ and of the ‘intermediate’ cluster. The ‘marine’ cluster membership increased especially in proximity of the main channels and in the northern basin, where the ‘confined’ cluster retreated. In the southern basin the ‘confined’ cluster lose positions to the ‘intermediate’ cluster in 2005.

Analysis on biogeochemical properties evidenced an analogous interannual variability. Statistically significant differences were observed between 2004 and 2005 in values of 7 environmental variables: temperature, salinity, chlorophyll, NO$_x$, DOC, DON, and DOP. The increase in salinity and the decrease in NO$_x$ could be related to the decrease of rainfall, and consequently to the reduced freshwater inputs in 2005. Apparently, the continental forcing was weaker than the marine, and this may have resulted in an increase of marine characteristics of benthic communities in 2005. Because of reduced inorganic nutrients input, chlorophyll concentrations in water were lower and zoobenthic species that usually live in oligotrophic conditions were favoured. Possibly, the drop of temperature could also have caused a decrease in ‘intermediate’ and ‘confined’ benthic species, which are adapted to higher average annual temperatures. On the contrary, habitats suitable to ‘marine’ species extended further into the lagoon, and ‘marine’ characteristics of benthic communities increased. Thus, the different extension of benthic communities between the two years may be related to the different extension of their habitats. Since we had only two years of data, we cannot speculate on if the differences between 2004 and 2005 were only due to interannual variability or if they were part of a multiannual trend, possibly resulting in an increase of ‘marine’ characteristics of the Lagoon of Venice (Sconfietti et al., 2003).

5.3. Artificial neural networks as habitat suitability models

We used the artificial neural network as a habitat suitability model for the 3 benthic communities, modelling their relationships
Fig. 8. Relationship between the membership grade of the 'confined' cluster and the four most important predictors (relative importance >10%): PO4 (upper left), DOP (upper right), RMSV (lower left) and wooden substrate (lower right).

with environmental parameters. The overall correlation coefficient between target and predicted values for the results of the BPNN was quite good, but the three clusters were predicted with different quality. The two clusters on the extremes were predicted much better than the 'intermediate' cluster, probably because the 'marine' and the 'confined' cluster have clear characteristics, as they are found at the two opposite ends of the sea-coastline gradient and the variables we used were good enough to predict them.

It should be clear, that in artificial neural networks all nodes are fully connected (Zupan and Gasteiger, 1999), i.e., all predictors concur in predicting each one of the target values. Since the profile method can give only a relative estimate of what might be the effect of one predictor on one target variable, we were not interested in the actual predicted values at different levels of the environmental variables, but only in the form of the functional response. Furthermore, BPNNs do not assume any particular functional type between predictors and targets, and indeed several of the identified relationships were linear, while others were clearly non-linear and complex (Figs. 6–8).

The 'marine' cluster had high affinity with decreasing temperature, POC concentrations in water and silt percentage in sediment. Areas near the inlets and along the main channels are those exposed to higher hydrodynamics, thus the smaller particles of the sediment are constantly removed by currents. The decreasing 'marine' cluster membership with increasing silt content in the sediment is consistent with the direct relationship between sand content of the sediment and 'marine' community observed in Bandelj et al. (2009), since silt percentage in sediment composition is roughly inversely proportional to sand percentage. Organic compounds concentrations in water are usually higher in high productive coastal and transitional systems than in marine waters, which is the case also for the Lagoon of Venice (Solidoro et al., 2004a; Bandelj et al., 2008). A decreasing preference of the 'marine' community with increasing POC concentrations in water was observed also in Bandelj et al. (2009). As for the relationship with the temperature, we pointed already to the lower temperature in the year 2005, when also the area occupied by the 'marine' cluster was larger, due to colonization of less eurivalent species. Furthermore, median temperature value of sites with maximum membership for the 'marine' cluster is lower than of sites belonging to the 'confined' and 'intermediate' clusters. A possible cause for this could be the localization of pollution sources discharging heated water near urban and industrial areas inside the lagoon (Mag. Acque-Thetis, 2003).

The 'confined' cluster memberships increased with decreasing RMSV, and in fact it was mainly present in areas with limited water exchange. A similar result was obtained in Bandelj et al. (2009), were RMSV was inversely correlated with indicator taxa of the 'con-
...fined’ community. The type of substrate affects organisms that can attach and survive on it (Relini and Faimali, 2003), thus it proved to be very important especially for the ‘confined’ cluster membership, because in the inner part of the lagoon many of the sampled sites had a wooden or iron substrate. The relationships between the ‘confined’ cluster membership and the two form of phosphate, inorganic as PO₄ and dissolved organic phosphorus (DOP), were inverse, since the ‘confined’ cluster memberships decreased with increasing inorganic phosphorus and increased with increasing DOP. The main sources of inorganic phosphorus in the Lagoon of Venice are urban and industrial wastewaters near the city of Venice and near the industrial area of Porto Marghera (Solidoro et al., 2004a). The ‘confined’ community was not present in these areas neither of the two years, while it was mainly observed in areas far from the sources of inorganic phosphorus and where trophic food webs likely rely on remineralized phosphorus. Relationships between high dissolved organic compounds in water (DON and DOC) and the prevalence of the ‘confined’ community (or its indicator taxa) were observed also in Bandelj et al. (2009).

Two variables showed high importance as predictors of the ‘intermediate’ cluster membership: temperature and salinity. Temperature values between 15 and 16.5 °C seem the most favourable for ‘intermediate’ community species. Sites in the ‘intermediate’ cluster showed the highest median temperature and this could favour organisms living in these areas, such as the Ulva macroalgae (Sfriso et al., 2002; Curiel et al., 2001). Increasing salinity was related to increasing ‘intermediate’ cluster membership, but did not have an appreciable effect on memberships of the ‘marine’ nor ‘confined’ community.

The median salinity was highest in the ‘marine’ cluster and lowest in the ‘confined’, as could be expected. Nevertheless, the variability of salinity was high for all three clusters. Our results seem to indicate, that median salinity is a good indicator of the areas in which the three communities develop, but also that the majority of species is tolerant to wide ranges of salinity, and thus that salinity does not strongly affect species’ distribution in lagoons. This is in agreement with the findings of Guéguerot and Perthusoit (1992) and do not support the conclusion that salinity is the fundamental ecological parameter of the distribution of benthic communities in lagoons (Segestrale, 1959; d’Ancona, 1959; Vatova, 1963).

Several relationships found in this work confirmed those obtained with different methods using only 2004 data (Bandelj et al., 2009), yet in some cases the functional forms as modelled by BPN appeared to be more complex than simple linear relationships. This is a clear advantage of neural networks over statistical methods which a priori impose a functional form. Other relationships found in this work were not identified with the redundancy analysis models in Bandelj et al. (2009). The reason may be that some variables (e.g., temperature) influence the prediction of the membership grades of single clusters, but they are not significant for the simultaneous ordination of the membership grades of all three communities.

5.4. Assessing the use of the neural network model for detection of anomalies

The neural network model related benthic communities to environmental conditions as they are described by our set of environmental parameters. The cases where the observed benthic community differed from the model predictions represent deviations from the modelled behaviour which should be investigated.

Many anomalies were from sites with wooden substrates (e.g., sites 14, 38, 64) or on low-depth bottoms (e.g., 28, 65). A reduced water column may favour finer sediments resuspension because of meteorological, animals or humans’ disturbances, and particles may reduce photosynthesis by phytothebenthos or make difficult
possibly needed to improve the predictions. The most important environmental parameters turned out to be temperature, POC concentrations in water and silt percentage in sediment composition for the ‘marine’ community; temperature and salinity for the ‘intermediate’ community; PO4 and DOP concentrations in water, mean water energy and type of substrate for the ‘confined’ communities. The neural network identified both linear (e.g., PO4 and DOP effects on the ‘confined’ cluster) and non-linear (e.g., temperature effect on the ‘intermediate’ cluster) relationships. This is a clear advantage of the use of artificial neural networks over other more traditional methods, such as direct gradient analyses, which a priori assume a specific functional type for the target’s responses to predictors.

The analysis of model results permitted us to highlight cases which significantly deviated from the expected response of benthic communities, and to discuss possible causes for these anomalous behaviours.

Acknowledgements

Thanks to Dr. D. Melaku Canu, for the modelling of hydrodynamic variables, and to Dr. G. Cossarini, for the interpolation of physical and chemical data and of data on sediment composition.

The study was funded by Magistrato alle Acque - ConsorzioVenezia Nuova. Authors wish to thank Dr. L. Montobbio and the team of Servizio Ambiente for support.

Appendix A. Supplementary data


References


