Ecological Applications, 17(5) Supplement, 2007, pp. S185–S196
© 2007 by the Ecological Society of America

ARTIFICIAL NEURAL NETWORK ANALYSIS OF FACTORS CONTROLLING ECOSYSTEM METABOLISM IN COASTAL SYSTEMS

Emma. J. Rochelle-Newall,[1,7] Christian Winter,[2] Cristina Barrón,[3] Alberto V. Borges,[4] Carlos M. Duarte,[3] Mike Elliott,[5] Michel Frankignoulle,[4,8] Fred Gazeau,[2,4,9] Jack J. Middelburg,[6] Marie-Dominique Pizay,[2] and Jean-Pierre Gattuso[2]

[1]*Institut de Recherche pour la Développement, UR 103, Centre IRD Nouméa, BP A5, NC-98848, Nouméa, New Caledonia*
[2]*Laboratoire d'Océanographie, Université Pierre et Marie Curie-Paris 6, UMR 7093, F-06234 Villefranche-sur-Mer Cedex, France,* and *Centre Nationale de la Recherche Scientifique, UMR 7093, F-06234 Villefranche-sur-Mer Cedex, France*
[3]*Instituto Mediterráneo de Estudios Avanzados (Consejo Superior de Investigaciones Científicas–Universidad de las Islas Baleares) (IMEDEA [CSIC–UIB]), Grupo de Oceanografía Interdisciplinar, C/ Miquel Marques 21, 07190, Esporles, Islas Baleares, Spain*
[4]*Unité d'Océanographie Chimique, Interfacultary Center for Marine Research, Université de Liège, Belgium*
[5]*Institute of Estuarine and Coastal Studies, University of Hull, Hull HU6 7RX United Kingdom*
[6]*Centre for Estuarine and Marine Ecology, Nederlands Instituut voor Ecologie–Koninklijke Nederlandse Akademie van Wetenschappen (NIOO–KNAW), P.O. Box 140, 4400 NT, Yerseke, The Netherlands*

Abstract. Knowing the metabolic balance of an ecosystem is of utmost importance in determining whether the system is a net source or net sink of carbon dioxide to the atmosphere. However, obtaining these estimates often demands significant amounts of time and manpower. Here we present a simplified way to obtain an estimation of ecosystem metabolism. We used artificial neural networks (ANNs) to develop a mathematical model of the gross primary production to community respiration ratio (GPP:CR) based on input variables derived from three widely contrasting European coastal ecosystems (Scheldt Estuary, Randers Fjord, and Bay of Palma). Although very large gradients of nutrient concentration, light penetration, and organic-matter concentration exist across the sites, the factors that best predict the GPP:CR ratio are sampling depth, dissolved organic carbon (DOC) concentration, and temperature. We propose that, at least in coastal ecosystems, metabolic balance can be predicted relatively easily from these three predictive factors. An important conclusion of this work is that ANNs can provide a robust tool for the determination of ecosystem metabolism in coastal ecosystems.

Key words: artificial neural networks; coastal ecosystems; metabolic balance; primary production; respiration.

Introduction

Eutrophication, as defined by Nixon (1995), is the increase in supply of organic matter to a system. Although not all coastal systems are subject to eutrophication, many nearshore systems adjacent to large population centers are organically enriched. In a review of coastal eutrophication, Boesch (2002) reported that 67% of the combined surface area of U.S. estuaries exhibited moderate to high degrees of eutrophication, and a similar picture is emerging in regions of Europe (Vidal et al. 1999, Conley et al. 2002, Herman et al. 2005). Estuarine and coastal areas are naturally productive ecosystems; nevertheless, human activities such as urbanization, industrialization, and agricultural practices have led to dramatically increased inputs of nitrogen (N) and phosphorus (P) to coastal areas (Howarth et al. 1996, Bouwman et al. 2005). Moreover, recent estimates show that ~20% of the global human population lives within 30 km of the coast, and this percentage is likely to increase in the future (Small and Nicholls 2003).

Increases in the nutrient loading to a system can lead to dramatic changes in phytoplankton community structure and activity (e.g., Berg et al. 2003). Higher phytoplankton biomass in turn leads to increased turbidity in the water column and organic-matter supply to the sediments, often resulting in anoxic bottom waters such as in the Baltic Sea (Bonsdorff et al. 2002, Lundberg 2005) and Chesapeake Bay (Cooper and Brush 1991). Nutrient and organic-matter concentrations also influence prokaryotic community composition (Cottrell and Kirchman 2000, Bouvier and del Giorgio 2002, Makino et al. 2003). Thus, increases in nutrient loadings to a system induce changes in the structure of both the autotrophic and heterotrophic communities.

Aquatic ecosystems are categorized in terms of their net ecosystem metabolism, defined as the balance

Manuscript received 9 November 2005; revised 24 March 2006; accepted 27 April 2006. Corresponding Editor: A. R. Townsend. For reprints of this Special Issue, see footnote 1, p. S1.

[7] E-mail: emma.rochelle-newall@noumea.ird.nc
[8] Deceased, 13 March 2005.
[9] Present address: Netherlands Institute of Ecology (NIOO-KNAW), P.O. Box 140, 4400 AC, Yerseke, The Netherlands.

between the autochthonous production of organic matter (gross primary production, GPP) and the remineralization (community respiration, CR) of organic matter. Systems are considered autotrophic and, hence, a potential net sink of CO_2 when production of organic matter exceeds the demands of respiration (GPP $>$ CR). Conversely, in heterotrophic systems respiratory demand outweighs autochthonous production of organic matter (GPP $<$ CR), and thus the system is potentially a net source of CO_2. Thus, ecosystems are classified in terms of the fate of the organic matter produced within the system rather than in terms of production and respiration rates.

Environmental legislation has developed from sectoral-based or regional-based regulations to more ecosystem-based, holistic environmental directives (Apitz et al. 2006). An understanding of ecosystem metabolism is necessary for any holistic ecosystem management plan as ecosystem metabolism governs the accumulation or removal of organic matter from a system. Moreover, ecosystem metabolism has major consequences for nutrient recycling and use (Smith et al. 1989) and thus management. Having a rapid estimate of ecosystem metabolism and its response to any management options can provide important information for future decision making. For example, systems that have very low GPP:CR ratios are highly net heterotrophic and thus have high oxygen demands that may present anoxia problems. Conversely, systems with high GPP:CR ratios will have high primary production rates, which will lead to an accumulation of organic matter in the system. Thus it is clear that a rapid and robust method of determining ecosystem metabolism can provide useful data.

Many studies have focused on the causes and consequences of eutrophication in coastal waters (e.g., Kemp et al. 1997). However, none to our knowledge has adopted a comparative approach to address the factors influencing ecosystem metabolism. In this study, we present data on the metabolic balance of three markedly different European coastal ecosystems: the Bay of Palma in Mallorca Island, Spain; Randers Fjord in Denmark; and the Scheldt Estuary in Belgium and The Netherlands. The study sites were chosen to encompass gradients of light penetration, nutrient concentration and enrichment, and eutrophication. A number of additional physical (e.g., salinity), chemical (e.g., concentration of nutrients), and biological (e.g., bacterial abundance) parameters were recorded as well to identify cause-and-effect relationships. In order to have a clearer understanding of the factors potentially influencing the metabolic state of the three ecosystems studied, the data were used to develop a mathematical model based on artificial neural networks (ANNs). Conventional statistical methods used to identify relationships between parameters are restricted to data that fit a predefined model (e.g., the normal distribution). In contrast, ANNs employ a training phase to adapt the model parameters to the data and are thus not, a priori, confined to a predetermined model condition (e.g., linear or nonlinear relationships). Additionally, ANNs are very well suited to process multidimensional data in which complex nonlinear interrelationships between the parameters can be expected. Artificial neural networks have been successfully used for such diverse tasks as comparing natural microbial communities based on the analysis of 5S rRNA (Noble et al. 1997), modeling the frequency of virally infected bacterial cells and rates of viral production in the North Sea (Winter et al. 2005), and modeling phytoplankton primary production in Chesapeake Bay (Scardi and Harding 1999). Artificial neural networks were employed in this study to model the metabolic state of the three study sites as represented by the GPP:CR ratio.

Methods

Study areas

The Bay of Palma, Mallorca, Spain, is a large, oligotrophic bay located in the northwestern Mediterranean (Fig. 1). The bay has an area of 216 km^2, with a mean depth of 31 m (Gazeau et al. 2005*a*). There are no major riverine inputs to the bay, and water residence is determined by the wind direction and intensity, with residence times varying between 2.5 and 10 d (Gazeau et al. 2005*a*). The city of Palma de Mallorca is located at the head of the bay and has a resident population of 360 000 inhabitants. However, the region is a popular tourist area, and during peak occupancy the population can reach 500 000 in the bay area. In February/March and June 2002, four stations were sampled to cover the range of productivities in the bay, and at each station four depths, covering the range of the water column, were sampled (Gazeau et al. 2005*a*).

The Randers Fjord is the longest Danish estuary with a length of 27 km, a surface area of 23 km^2, and a mean depth of 1.6 m (Nielsen et al. 2001). Freshwater inputs are dominated by the Gudenå River, which drains a large proportion of the catchment area (3200 km^2; Andersen 1999) and has a discharge of $\sim 10^9$ m^3/yr. The estuary also receives water from a minor tributary, Grund Fjord, and from the wastewater treatment plants in the catchment. The estuary empties into the Kattegat, and water residence time within the estuary varies seasonally and averages 13 d (Nielsen et al. 2001). Randers Fjord was sampled in April and August of 2001. Five stations were visited along the salinity gradient (0–29‰). This meso-eutrophic estuary exhibits a typical two-layer circulation with a well-developed pycnocline. Therefore, samples (Table 1) were collected at four discrete depths to allow sampling from both the surface mixed layer and the sub-pycnocline bottom layer (Gazeau et al. 2005*b*).

The Scheldt Estuary is an organically enriched and hyper-nutrified system located in Belgium and The Netherlands. The dominant freshwater supply to the estuary derives from the Scheldt River, with a catchment

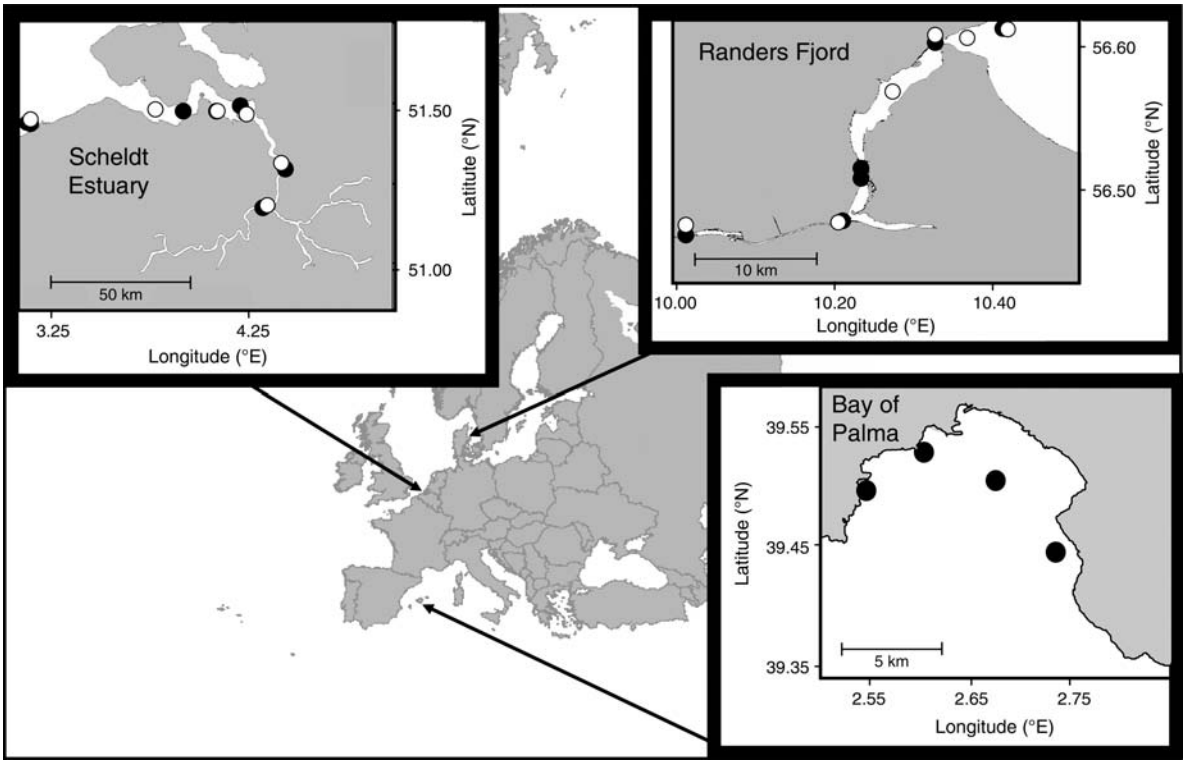

FIG. 1. Map of the three study areas: Bay of Palma in Mallorca Island, Spain; Randers Fjord in Denmark; and Scheldt Estuary in Belgium and The Netherlands. For the Bay of Palma, sampling stations are indicated by solid circles. For Randers Fjord, sampling stations in April 2001 are indicated by solid circles; sampling stations in August 2001 are indicated by open circles. For Scheldt Estuary, sampling stations in November 2002 are indicated by solid circles; sampling stations in April 2003 are indicated by open circles.

area of 19 500 km^2 (Heip 1989). The dominance of tidal exchange in the estuary relative to that of freshwater inputs results in extremely long freshwater residence times of up to 70 d in the freshwater reaches (Wollast 1988, Soetaert and Herman 1995). Strong tidal currents (up to 1.5 m/s), combined with the relatively low proportion of freshwater inputs (up to 200 times less than seawater influx) result in a well-mixed water column along the length of the estuary (Wollast 1988). The estuary and plume were sampled in November 2002 and April 2003. Five stations were sampled in November and six in April covering the salinity gradient (0–34‰), and, due to the well-mixed nature of this estuary, only the subsurface was sampled (5 m; Fig. 1) (Gazeau et al. 2005*c*). Hereafter, Scheldt Estuary refers to both the inner estuary and the dilution plume.

Sample collection and analysis

Many of the methods used are described in more detail in other papers from this study, and therefore only a brief description of each method is presented below (Rochelle-Newall et al. 2004, Veuger et al. 2004, Gazeau et al. 2005*a*, *b*, *c*). At each station, salinity and temperature were measured with a conductivity, temperature, and depth (CTD) probe (Sea-Bird Electronics, Bellevue, Washington, USA; see Plate 1) and light penetration (photosynthetically active radiation, PAR) in the water column was measured with a LI-COR spherical sensor LI-193SA (LI-COR, Lincoln, Nebraska, USA). Samples were collected as close to sunrise as possible using Niskin bottles (General Oceanics, Miami, Florida, USA).

Dissolved inorganic nutrients ($NO_3^- + NO_2^-$, NH_4^+, dissolved inorganic phosphorus [DIP]) were measured using automated colorimetric techniques after filtration through GF/F filters (Veuger et al. 2004). Dissolved inorganic nitrogen (DIN) is defined as the sum of NH_4^+ and $NO_3^- + NO_2^-$; DIP is the concentration of soluble reactive phosphate. Dissolved silicate (DSi) was measured after filtration through 0.45-μm Millipore filters. Total dissolved nitrogen (TDN) and total dissolved phosphorus (TDP) were measured using standard methods. Chlorophyll *a* (chl *a*) and pigment concentrations were measured on GF/F filters that were stored frozen until extraction and analysis by high-performance liquid chromatography (Barranguet 1997). Suspended particulate material (SPM), as well as the particulate organic carbon and nitrogen content, were measured using a carbon–hydrogen–nitrogen (CHN) analyzer (Nieuwenhuize et al. 1994).

Planktonic GPP and CR were measured using the light-dark O_2 incubation method at each of the three

sites (Gazeau et al. 2005*a*, *b*, *c*). In Randers Fjord, planktonic metabolism during the two cruises was estimated using five replicates incubated in situ at four depths from the sunrise to sunset in transparent and dark 60-mL biological oxygen demand (BOD) bottles. In the Bay of Palma, planktonic metabolism was estimated using in situ incubations in March and June 2002 at four stations. In the Scheldt Estuary planktonic metabolism was measured at five and six stations in November 2002 and April 2003, respectively. Samples from the Scheldt were incubated in flowing seawater in an onboard incubator (at 100%, 50%, 25%, and 12% incident irradiance; Gazeau et al. 2005*c*). Daily planktonic respiration rate (CR) was calculated by multiplying the hourly dark rates by 24, therefore assuming a constant rate over 24 hours. Hourly GPP were multiplied by day length to estimate daily planktonic GPP.

Samples for dissolved organic carbon (DOC) collected in Randers Fjord and the Scheldt Estuary were filtered through glass fiber filters in precombusted 10-mL glass ampoules, preserved with 12 μL 85% phosphoric acid (H_3PO_4) and flame-sealed (Rochelle-Newall et al. 2004). Due to the low particulate organic carbon concentration in the Bay of Palma, total organic carbon (TOC) samples were collected to minimize the risk of contamination from filtration. Dissolved organic carbon concentration was measured on a Shimadzu 5000 TOC analyzer (Shimadzu, Kyoto, Japan), using potassium phthalate calibration standards over the measurement range 0–800 μmol/L C. Certified reference materials (Hansell Laboratory, University of Miami, Miami, Florida, USA) were also used to assess the performance of the instrument on and between measurement days. The machine blank, calculated as the difference between the measured concentration of the external standards and the "real" concentration of the external values, was between 8 and 12 μmol/L C for the measurement days.

Bacterial abundance (BA) was measured using the 4, 6-diamidine-2-pheylindole, dihydrochloride (DAPI)-staining technique of Porter and Feig (1980). Bacterial cells were enumerated in at least 20–30 randomly selected fields of view under 1250× magnification (Rochelle-Newall et al. 2004).

Bacterial production (BP) was estimated by the ^{3}H-leucine incorporation method (Smith and Azam 1992). The amount of radioactivity contained in the particulate fraction was measured in a scintillation counter using Ultima Gold scintillation cocktail (Packard Instruments, Downers Grove, Illinois, USA). Uptake of ^{3}H-leucine was transformed into micromoles of C per liter per day using the conversion factors of Simon and Azam (1989).

Modeling GPP:CR ratio using ANNs

Basheer and Hajmeer (2000) provide a good introduction to ANN theory. Feed-forward ANNs with one layer of hidden neurons and one output neuron were implemented in Mathematica 5.2 using the Neural Networks application package (Wolfram Research, Champaign, Illinois, USA). A bias term with a fixed value of one was included in the input and the hidden layer. Before training, the parameters of the networks were initialized using the option "LinearParameters" in order to randomize the initial values of the nonlinear parameters within the range of the input data and to completely randomize the linear parameters. We used the Levenberg-Marquardt algorithm (Marqhardt 1963, Haykin 1999) to train the ANNs for 100 iterations, employing the sigmoid function as the activation function of the hidden neurons. Prior to training, the data were scaled to a mean of zero and unity variance. Progress of the training procedure was monitored using the root mean square error (RMSE) of the networks.

The first phase of modeling was designed to determine suitable sets of input parameters for modeling the GPP:CR ratio. We considered the following 23 input parameters: bacterial abundance and production, bottom depth, the concentrations of chlorophyll *a*, *b*, and *c*, sampling depth, DOC concentration, NH_4, NO_2, NO_3, DIP, DSi, oxygen saturation, PAR, and the percentage of PAR at the sampling depth, salinity, SPM, percentage of carbon and nitrogen in SPM, TDN, TDP, and water temperature.

PLATE 1. A CTD (conductivity, temperature, and depth) probe, shown here at the Scheldt Estuary, was used at each site to measure the physical parameters of the water column. Photo credit: F. Gazeau.

TABLE 1. Summary of physicochemical data (minimum and maximum in ranges) from the three study sites (the Bay of Palma in Mallorca Island, Spain; Randers Fjord in Denmark; and the Scheldt Estuary in Belgium and The Netherlands) for both sampling periods.

Parameter	Randers Fjord		Bay of Palma		Scheldt Estuary	
	Apr 2001	Aug 2001	Mar 2002	Jun 2002	Nov 2002	Apr 2003
Mean depth (m)	1.6[a]	1.6[a]	31[b]	31[b]	9.7[c]	9.7[c]
Maximum depth (m)	10	8.5	37	33	17	17
Temperature (°C)	6.6–9.9	17.5–19.8	14.2–14.4	18–23.5	11–12.7	7.7–11.4
Salinity	0.2–29.5	0.2–28.1	37.4–37.6	37.7–37.8	0.55–30.4	1.2–34.0
Surface irradiance (μmol photons·m^{-2}·s^{-1})	205–738	183–588	242–761	960–1110	114–331	257–810
Light attenuation coefficient (m^{-1})	0.6–1.3	0.2–1.3	0.11–0.15	0.1–0.2	0.9–8.2	1.7–4.3
Chlorophyll *a* (μg/L)	0.6–19.7	0.3–16.3	0.2–4.2	0.04–0.6	1.7–16.9	2.9–13.9
SPM (mg/L)	1.5–43.1	0.6–12.1	0.4–60.3	0.3–40	22–118	29–179
SPM (%C)	2.7–16.9	3.1–24	0.5–20	1.4–20	5–6	0.4–0.7
SPM (%N)	0.3–1.8	1.5–3.4	0.5–4.6	0.2–2.4	0.4–0.7	0.6–1.8
POC (μmol/L)	16–232	5–89	2–538	2–17	87–592	344–1384
DOC (μmol/L)	209–470	244–392	57–76	63–97	98–510	195–428
DSi (μmol/L)	1.3–113	6.4–283	···	···	14–227	17–131
DIP (μmol/L)	0.05–0.4	0.1–2.9	···	···	0.64–6.0	0.6–5.2
NO_3 (μmol/L)	0.06–134[d]	0.03–9.2[d]	001–0.5	03–0.6	6.0–297	29.1–354
NH_4 (μmol/L)	0.9–7.2[d]	0.9–66.6[d]	01–0.4	0.1–0.3	2.8–88	0.5–129
TDN (μmol/L)	21.5–144	0.6–104	2.6–4.6	3.1–4.0	21.5–448	219–491
TDP (μmol/L)	0.08–3.9	0.3–3.9	0.17–0.34	0.24–0.39	0.4–4.3	0.5–3.4
BA (millions of cells/mL)	2.1–6.0	1.4–6.8	0.5–0.8	0.3–1.0	1.1–7.1	2.9–7.4
BP (μmol C·L^{-1}·d^{-1})	1.4–7.2	0.24–9.6	0.1–1.4	0.17–1.44	0.24–4.8	1.7–7.2
Surface GPP (μmol O_2·L^{-1}·d^{-1})	5.5–44.3[a]	13.7–80[a]	2.4–5.4[b]	3.3–6.3[b]	1.7–5.9[c]	3.7–55.6[c]
CR (μmol O_2·L^{-1}·d^{-1})	−11–−0.5[a]	−11.2–−2.4[a]	−3–0.1[b]	−7.1–−1[b]	−21.8–−0.3[c]	−14.8–−2.4[c]

Notes: Abbreviations are: SPM, suspended particulate matter; POC, particulate organic carbon; DOC, dissolved organic carbon; DSi, dissolved silicate; DIP, dissolved inorganic phosphorus; TDN, total dissolved nitrogen; TDP, total dissolved phosphorus; BA, bacterial abundance; BP, bacterial production; GPP, gross primary production; CR, community respiration. References are indicated by superscript letters: a, Gazeau et al. (2005*b*); b, Gazeau et al. (2005*a*); c, Gazeau et al. (2005*c*); d, Veuger et al. (2004). Ellipses indicate values below the detection limit.

Only parameters collected in situ at 5-m depth were considered in the Scheldt Estuary due to its well-mixed condition. The input parameters were used alone and in combination with up to two other parameters to train ANNs with one to five hidden neurons. The initial values of the weights of the network can have a profound influence on the outcome of the training procedure (Haykin 1999). Thus, for each set of input parameters and number of hidden neurons, 100 ANNs were initialized and trained as described above. Artificial neural networks are capable of memorizing the data used for training when trained over too many iterations (over-training), particularly when the networks have a large number of hidden neurons. Over-training makes predictions on the basis of new input data inaccurate. In order to prevent over-training, we performed cross-validation (Haykin 1999). The data set was split into a training data set (75% of the data) and a validation data set (25% of the data) by assigning every fifth data vector to the validation data set. While the training data set is used for adjusting the networks parameters in order to decrease the networks error, the validation data set is only used to validate the network at each iteration during training without interfering in parameter adjustment. Training was assumed to have converged when the sum of the RMSE of the training and validation data sets reached a minimum.

The networks were reconstituted at the iteration of the minimum of the combined errors by retrieving the corresponding network parameters from the training record. The best performing ANNs for each set of input parameters and number of hidden neurons were determined by the smallest sum of the RMSE of the training and validation data set at convergence of training. In this first phase we initialized, trained, and screened a total of 254 000 ANNs.

The second phase of modeling used only the sets of input parameters yielding good results in the first phase and was targeted at finding the most suitable set of input parameters and the best performing network structure. The number of hidden neurons was increased stepwise to a maximum of 12. For each set of input parameters, as identified in the first phase of model development and number of hidden neurons (6–12 hidden neurons), 1000 ANNs were initialized, trained, and screened as described above. The best performing set of input parameters and network structure was found by searching for the minimum of the combined RMSE of the training and validation data set between the sets of input parameters and number of hidden neurons. The ANN was used to perform a simulation of the GPP:CR ratio by varying the input parameters within the ranges found in the three study sites.

Statistical analyses

The Kruskal-Wallis and the Mann-Whitney tests were used to test whether the variables differed between the study sites. To avoid any bias introduced by data from

Scheldt Estuary, only the in situ values from the 5-m depth (the sampling depth) were applied. Spearman rank correlation coefficients were employed to determine the degree of correlation between parameters. The slope of linear regression analyses was tested against a hypothetical value of one by calculating a t value according to the formula $t = (|b_{yx} - B_{yx}|)/Sb_{yx}$ where t is the t value, b_{yx} is the slope of the linear regression, B_{yx} the hypothetical growth of 1, and Sb_{yx} the standard deviation of the increase of the linear regression. The P values are given for the two-tailed t distribution. Stepwise multiple regression analysis was performed as a comparison to the ANN-based model of the GPP:CR ratio developed in this study. The results of the statistical tests were assumed to be significant at $P \leq 0.05$.

Results

Table 1 provides a summary of the physicochemical data collected at the three study sites. Temperature varied over a relatively small range within each field campaign. The widest range was observed in the Bay of Palma in summer due to the presence of cooler bottom water at the deepest station (18°C at 30 m compared to 22°C at the surface). Salinity varied very little in the Bay of Palma, reflecting the lack of riverine influence in this environment. In contrast, large salinity gradients were observed in the two estuaries. Turbidity, incident irradiance, and water column light penetration, as represented by the light attenuation coefficient, also varied between the three sites, with the Bay of Palma having significantly higher light penetration (lower light attenuation coefficients) than the other two sites (Kruskal-Wallis, $P < 0.0001$).

Inorganic and organic nutrients were significantly lower in the oligotrophic Bay of Palma than in the two estuaries (Kruskal-Wallis, $P < 0.0001$). This was also the case for chl *a* and BA and BP (Kruskal-Wallis, $P < 0.0001$). Interestingly, although Randers Fjord had significantly lower inorganic and organic nutrients and DOC than the Scheldt (Mann-Whitney, $P < 0.05$), bacterial abundance and production rates were not significantly different between the two estuaries (Mann-Whitney, $P > 0.05$). Suspended particulate matter concentrations were significantly higher in the Scheldt ($P < 0.05$) than in the Bay of Palma and the Randers Fjord. Moreover, the lack of significant difference between bulk SPM concentration in these two latter sites was also manifested in the percentage of carbon or nitrogen content of the SPM. In contrast, the percentage of carbon and nitrogen content of SPM in the Scheldt was significantly lower than in Randers ($P < 0.05$) and in both the Randers Fjord and the Bay of Palma for percentage of carbon and percentage of nitrogen, respectively.

In summary, the chemical and biological parameters in the Bay of Palma were all significantly lower than those in the estuaries. The two estuaries differed from one another in terms of nutrient concentrations, with the Scheldt generally having higher concentrations than the Randers Fjord. Thus, there were significant differences among the sites with regard to physicochemical and biological gradients.

The principal objective of this research was to determine whether there are robust factors that could be used to predict the metabolic status of an estuarine or coastal marine site. We therefore tested whether any of the above biological and physicochemical factors were correlated with the GPP:CR ratio. Although several parameters were very weakly correlated (e.g., PAR and percentage of PAR), the correlation coefficients did not exceed 0.5 and thus were considered too weak to be of relevance (Appendix A).

In the first phase of neural network development, the following five combinations of input parameters had a combined RMSE < 0.4: sampling depth, DOC, temperature; NO_3, oxygen saturation, TDP; DSi, SPM, temperature; SPM, percentage of carbon in SPM, TDP; percentage of carbon and nitrogen in SPM, TDP. Initial results indicated that the number of hidden neurons could still be increased to further improve network performance without over-fitting the data, i.e., the combined RMSE of the training and validation data set had not yet reached a minimum when using five hidden neurons (data not shown). Increasing the number of hidden neurons in the ANNs improved the performance of all five combinations of input parameters (data not shown). However, the GPP:CR ratio was best modeled using the sampling depth, DOC, and temperature as input parameters to an ANN with 11 hidden neurons (RMSE = 0.25). The values of the weights between the input and hidden layers as well as between the hidden and output layers are presented in Appendix B and the mean and standard deviation of actual data used for model development are presented in Appendix C. The GPP:CR ratio, as predicted by the ANN corresponded very well with the observed GPP:CR ratio, and 99% of the variation could be explained by the ANN-based model (Fig. 2). The slope of the regression analysis was not significantly different from one ($t = 0.09$, $P > 0.05$).

When simulating the GPP:CR ratio by varying the input parameters over the entire range found in the three study sites, the output of the ANN exceeded the maximum GPP:CR of 50 observed in this study for specific ranges of the input parameters (Fig. 3). Lower temperatures (5–15°C) had a stronger effect on the GPP:CR ratio, as predicted by the ANN-based model, than did higher temperatures (20–25°C). For example, at a temperature of 5°C, the GPP:CR ratio increased dramatically at DOC concentrations below 150–250 μmol/L, whereas at 10°C GPP:CR increased strongly below a DOC concentration of 150 μmol/L and depth values above 15 m (Fig. 3). Particularly for higher temperatures (20–25°C), the model predicted increasing GPP:CR values when increasing the concentration of

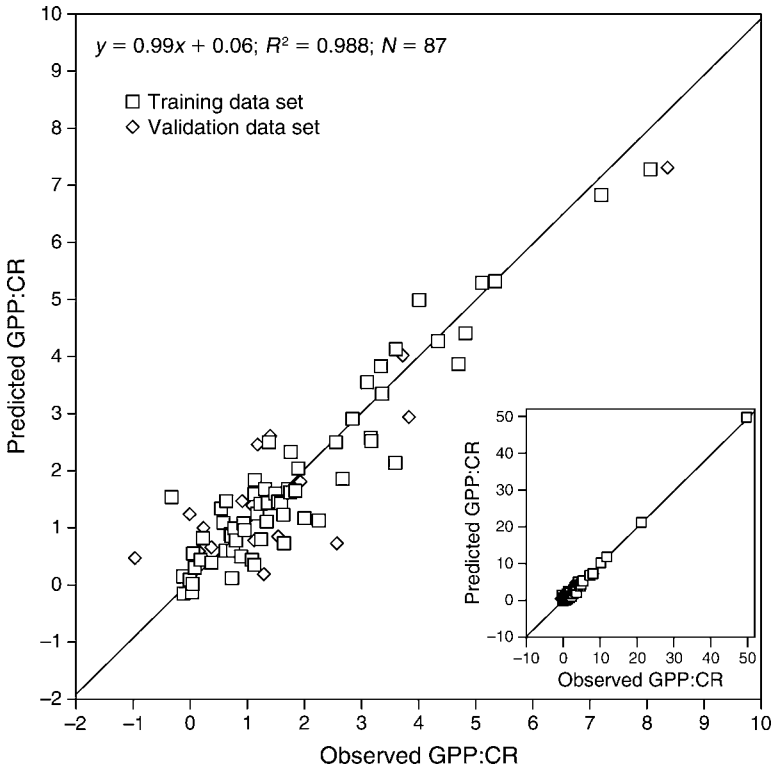

FIG. 2. Comparison between predicted and observed values of the ratio of gross primary production to community respiration (GPP:CR). Values were calculated by dividing the daily planktonic gross primary production by the daily community respiration. The predictions were computed using the artificial neural network (ANN) developed in this study. Additionally, the equation and the coefficient of determination (R^2) of the linear least-squares fit to the data are shown. The inset depicts the entire range of the data.

DOC and depth, resulting in a hump-shaped feature (Fig. 3). This relationship is also visible for temperatures ranging from 5° to 15°C, but appears to be modified substantially at these temperatures.

The same parameters used for the development of the ANN-based model were also used in stepwise multiple regression analysis to obtain a multiple linear regression model of the GPP:CR ratio (Table 2). Photosynthetically active radiation explained 27% of the variability of GPP:CR. The addition of the concentration of DOC increased the performance of the model slightly. However, the effect of the concentration of the DOC on the model was not significant ($P = 0.7576$; Table 2). Overall, the ANN-based model of the GPP:CR ratio was by far superior to the multiple linear regression model.

DISCUSSION

Prediction of GPP:CR by an ANN-based model

In this study a comparative approach was adopted in order to address the factors influencing ecosystem metabolism. Analyzing data across ecosystems is difficult to accomplish using conventional statistical methods due to the often large variability of parameters, and as a consequence, important links can often be overlooked. These problems can be especially important when data sets from different sampling strategies or methodologies are compared. However, even for data collected within the same program, simple correlations often do not suffice to explain the determining factors due to the wide variability in the parameters measured. The three ecosystems studied differed widely in organic matter and nutrient concentration as well as other physicochemical properties. This is probably one of the reasons why none of the parameters identified as being important in determining the GPP:CR ratio by the ANN-based modeling approach were indicated by correlation or stepwise multiple regression analysis (Table 2; Appendix A). In reducing ecological systems to just two simple cause-and-effect parameters, we fail to take into account the often considerable variability in the data set. In our data set, the parameter that was best correlated with GPP:CR was PAR, and yet PAR explained <50% of the variability in GPP:CR. This is in comparison to the R^2 value of over 0.98 for the network when predicted and observed values were compared.

We therefore used the available data to develop an ANN-based model of the GPP:CR ratio as a measure of ecosystem metabolism. The 23 input parameters initially used for the model development were not used to

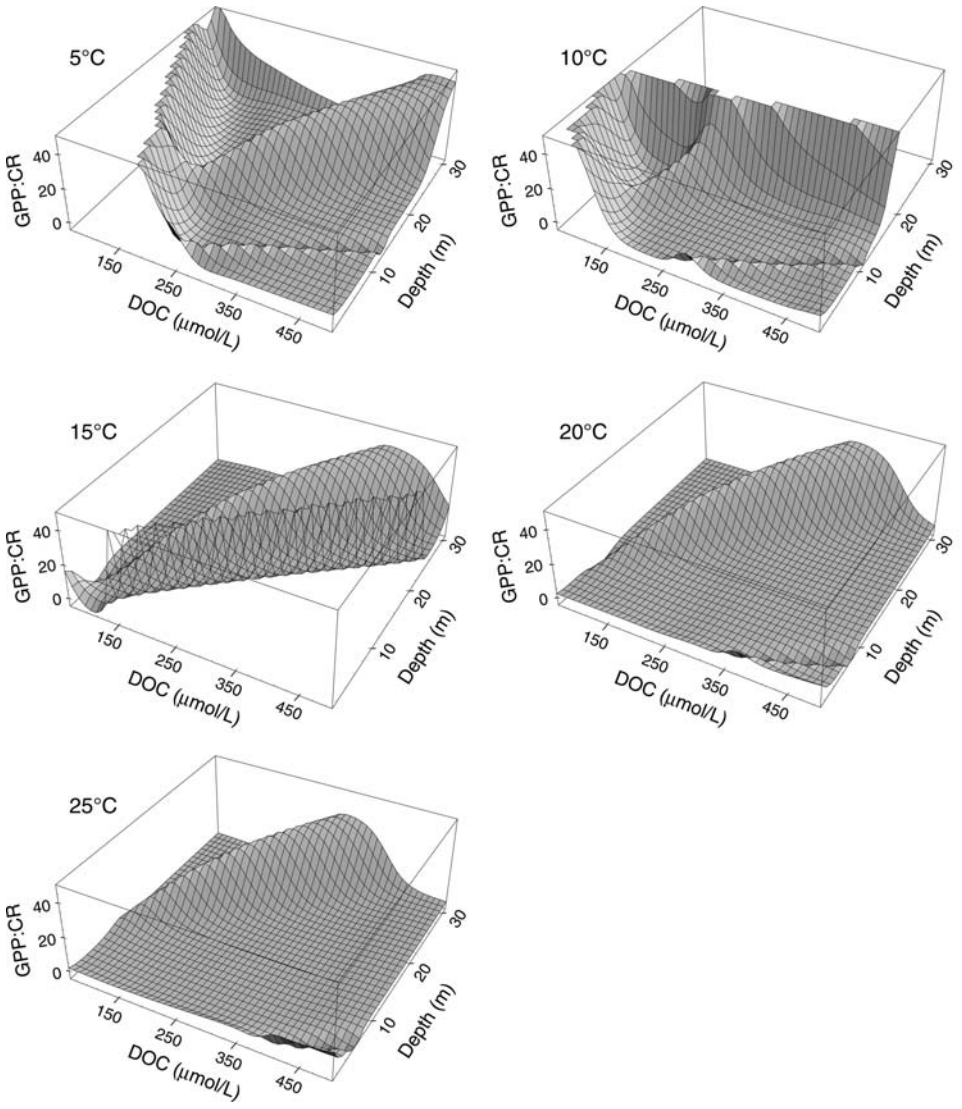

FIG. 3. Simulation of the ratio of gross primary production to community respiration (GPP:CR) at five different temperatures, 5–25°C. The model developed in this study, based on an artificial neural network, was used to simulate GPP:CR at dissolved organic carbon (DOC) concentrations ranging from 50 to 500 μmol/L and depth from 0 to 30 m.

calculate GPP:CR, and as correlation analysis revealed no relationships between any of the parameters and GPP:CR, the choice for these input parameters was a pragmatic one based on the number of data points available. Thus, parameters that missed >20% of the data points were not considered for model development. Although this approach required considerably more computing time than if a more restricted data set had been used, it ensured that all possible relationships within the data set using up to three parameters as input to the ANNs were explored. The predicted GPP:CR ratio corresponded very well with the observed values (Fig. 2), indicating that suitable input parameters for modeling GPP:CR were identified. As a comparison to the ANN-based model we performed stepwise multiple regression analysis of GPP:CR (Table 2). The performance of the ANN-based model was by far superior to the multiple linear regression model. This indicates that ANNs are a powerful modeling tool suitable for studies revealing cross-ecosystem relationships between parameters that otherwise remain hidden from conventional statistical analysis.

The strategy used in model development involved splitting the data into a training and validation data set. This cross-validation approach (Haykin 1999) is useful in preventing over-training and over-parameterization of the ANN-based model. Thus, the model developed in this study should have captured the influence of the input parameters on GPP:CR. The good fit between the observed and predicted values (Fig. 2) indicates that the ANN-based model is a good predictor for GPP:CR in these three coastal ecosystems. Moreover, the model

TABLE 2. Stepwise multiple linear regression analysis of the ratio of gross primary production to community respiration (GPP:CR).

		Model			Regression			Intercept	
Parameter	n	R^2	F	P	Coefficient	SE	P	Value	SE
PAR	**70**	**0.27**	**25.2**	**<0.0001**	**0.02**	**0.004**	**0.0386**	**−0.59**	**1.01**
PAR	70	0.31	15.4	<0.0001	0.02	0.004	0.0636	2.26	1.69
DOC					−0.01	0.005	0.7576		

Note: The best-performing model of GPP:CR is indicated in boldface type.

could be used to get a fast estimate of GPP:CR on the basis of new input data for parameters easily measured using routine methods in monitoring programs.

Although Fig. 3 shows the model predictions of GPP:CR over the entire range of the input parameters found in the three study sites (sensitivity analysis), it is important to realize that not all combinations of values of the input parameters were actually found. For example, DOC values below 150 μmol/L at a temperature of 5°C or below were not encountered in any of the three study sites. These unencountered combinations of values of the input parameters could result in model predictions of GPP:CR surpassing the maximum measured in this study. Thus, it is necessary to interpret these graphs (Fig. 3) in the context of the values of the input parameters observed in situ and to disregard those areas of the graphs that were not observed.

The model accurately predicts metabolic status with only a relatively small database. It can be argued that the metabolic estimates used in the model are based on short-term incubations obtained during two intensive two-week-long sampling periods at each of the three sites and that metabolic balance concepts are based on longer-term metabolic estimates. However, as demonstrated by Gazeau et al. (2005*a*), the net metabolic balance estimated from bottle incubations in Randers Fjord was in good agreement with estimates based on the LOICZ budgeting approach, which integrates data over a much longer spatial scale (weeks to months). Moreover, the output from the model can be used to provide an estimate of integrated water column metabolism; i.e., by selecting the output from each measured combination of depth, DOC, and temperature, an integrated measure of water column metabolism can be calculated, similar to that estimated from raw data. However, as is evident from the scatter in Fig. 2, it is clear that if one wants to precisely determine the metabolic balance of a particular site, then it is better to directly measure the balance between GPP and CR. Moreover, although our results show that ANNs can be used to accurately predict metabolic balance in widely contrasting coastal ecosystems, it is difficult to envision how the ANN-based model would perform in fundamentally different aquatic environments. Since the data for model development were exclusively derived from coastal and estuarine ecosystems, it is unlikely that the model would perform well in open ocean environments.

Ecological interpretation of the ANN-based model of GPP:CR

In terms of an ecological context and application, one of the major benefits of modeling techniques is that they provide answers to questions that would have otherwise required considerably more manpower, time, and money, as well as providing a means to investigate cause-and-effect relationships. Of course, this can only be the case if the model results can be logically placed in an ecological context. We found that the GPP:CR ratio can be best described by three parameters: sampling depth, temperature, and DOC concentrations. Two of these parameters are easily measured by a CTD probe, and the third is a relatively straightforward chemical measurement. It is therefore interesting to examine how each of these parameters might affect ecosystem metabolism.

In our sampling strategy, sampling depth also corresponds to incubation depth for the respiration and primary production measurements in two of the sites (Palma and Randers; Gazeau et al. 2005*a*, *b*). Interestingly, although the sampling depth is clearly related to light penetration in the water column, light penetration was a less significant factor in the analysis. This is probably due to the fact that the GPP:CR ratio is dependent on not only physicochemical characteristics, such as nutrients or light field, of the sample but also on the community composition present as species diversity is likely to affect both GPP and CR (Rivkin 1989).

The direct effect of temperature on phytoplankton production has been known for a long time, especially in algal cultures; however, the effect of temperature on GPP:CR is more complex and, as a result, less well known (Lomas et al. 2002). Given that a large proportion of community respiration is due to the smaller size fractions (Smith and Kemp 2001, Lemée et al. 2002) and that temperature and bacterial respiration are correlated (Rivkin and Legendre 2001), it is obvious that temperature can exert considerable control on community respiration. Temperature also affects primary production and community structure in phytoplankton (Agawin et al. 2000), both of which can entrain changes in community metabolism. However, while community respiration has Q_{10} values of more than two, the Q_{10} values of primary production are more variable and are generally less than two (Lomas et al.

2002). This may be one of the reasons why we observe a humpback-shaped response in the model predictions.

The third parameter, DOC, "is the soil of the sea—a large, biochemically resistant reservoir of organic matter providing substrate for life" (Ducklow 2002:xv). One of the major autochthonous sources of DOC is from direct production of DOC during active photosynthesis. Moreover, DOC production rates average 19% of primary production rates in coastal regions (Morán et al. 2002, Marañón et al. 2004) and can represent a significant source of potentially bioavailable DOC, and, interestingly, although GPP tends to decrease with depth, in many cases the relative percentage of DOC production actually increases with depth (Marañón et al. 2004). Dissolved organic carbon concentration and lability are key factors controlling bacterial processes (Carlson et al. 1994, Findlay 2003). Dissolved organic carbon therefore provides the energy source for the production of biomass as well as the respiratory requirements of bacterial heterotrophs. The lability and concentration of DOC can exert a strong influence on bacterial respiration (del Giorgio and Davis 2003, Rochelle-Newall et al. 2004) and, given the importance of bacterial respiration in community respiration, it is clear that changes in DOC concentration and lability can lead to changes in CR. Thus, each of the three factors selected as being important for determining GPP:CR during model development (temperature, depth, and DOC concentration) has the potential to significantly influence ecosystem metabolism.

One of the central paradigms of aquatic ecology is that increased concentrations of inorganic nutrients lead to increased organic matter in the system. Thus, given that nutrient concentration influences organic matter production, it is generally accepted that nutrients will exert a non-negligible influence on ecosystem metabolism. However, nutrient concentrations were not important for the determination of GPP:CR in the developed model, although it can be argued that DOC cannot be regarded as being strictly independent of organic nutrient concentrations because it is unrealistic to think that DOC concentrations represent a pool of organic matter that is solely comprised of organic carbon and thus is lacking other elements such as N or P (Hedges 2002). Furthermore, it is known that increases in nutrients tend to lead to increases in DOC in a system, and so it may well be that DOC is a proxy for increased nutrient concentrations. However, particulate organic carbon did not appear to be one of the major controlling parameters, either. Thus it appears that dissolved organic matter rather than particulate organic matter is more important in determining ecosystem metabolism.

Finally, as noted by Herman et al. (2005), although increased inorganic and organic matter inputs to the system can lead to fundamental changes in ecosystem functioning in terms of metabolic balance, community structure, and biodiversity, it is the hydrological and environmental conditions of the system that will determine the magnitude of this response. If it is indeed the case that physical parameters such as temperature and depth can exert a stronger effect on metabolic status than can nutrients, this poses some interesting questions as to the response of the metabolic status of the coastal ocean to future climate change and in turn, on how to manage those changes.

Acknowledgments

This paper is dedicated to the memory of our friend and coauthor, Michel Frankignoulle, who is sorely missed. We thank all Eurotrophers for their camaraderie during the campaigns and the crews of the various research vessels involved in sampling. This research was funded by the European Union (EU) in the framework of the EUROTROPH project (contract number EVK3-CT-2000-00040), CNRS post-doctoral funding to E. J. Rochelle-Newall, and FNRS where A. V. Borges is a research associate. This is MARE publication number 86, publication number 3841 of the Netherlands Institute of Ecology (NIOO-KNAW), and a contribution to the Integrated Project CarboOceans supported by the EU (511176).

Literature Cited

Agawin, N. S. R., C. M. Duarte, and S. Agusti. 2000. Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. Limnology and Oceanography 45:1891.

Andersen, J. M. 1999. 25 years of pollution abatement and environmental improvement in the Gudenaa—a Danish lake–river–estuary system. Publication number 87-7906-050-1. Gudenaa Committee, Højbjerg, Denmark.

Apitz, S. E., M. Elliott, M. Fountain, and T. S. Galloway. 2006. European environmental management: moving to an ecosystem approach. Integrated Environmental Assessment and Management 2:80–85.

Barranguet, C. 1997. The role of microphytobenthic primary production in a Mediterranean mussel culture area. Estuarine Coastal and Shelf Science 44:753–765.

Basheer, I. A., and M. Hajmeer. 2000. Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods 43:3–31.

Berg, G. M., M. Balode, I. Purina, S. Bekere, C. Bechemin, and S. Y. Maestrini. 2003. Plankton community composition in relation to availability and uptake of oxidized and reduced nitrogen. Aquatic Microbial Ecology 30:263–274.

Boesch, D. F. 2002. Challenges and opportunities for science in reducing nutrient over-enrichment of coastal ecosystems. Estuaries 25:886–900.

Bonsdorff, E., C. Ronnberg, and K. Aarnio. 2002. Some ecological properties in relation to eutrophication in the Baltic Sea. Hydrobiologia 475:371–377.

Bouvier, T. C., and P. A. del Giorgio. 2002. Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. Limnology and Oceanography 47:453–470.

Bouwman, A. F., G. Van Drecht, J. M. Knoop, A. Beusen, and C. R. Meinardi. 2005. Exploring changes in river nitrogen export to the world's oceans. Global Biogeochemical Cycles 19:GB1002. [doi: 1010.1029/2004GB002314.]

Carlson, C. A., H. W. Ducklow, and A. F. Michaels. 1994. Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature 371:405–408.

Conley, D. J., S. Markager, J. Andersen, T. Ellermann, and L. M. Svendsen. 2002. Coastal eutrophication and the Danish National Aquatic Monitoring and Assessment Program. Estuaries 25:848–861.

Cooper, S., and G. S. Brush. 1991. Long-term history of Chesapeake Bay anoxia. Science 254:992–996.

Cottrell, M. T., and D. L. Kirchman. 2000. Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. Applied and Environmental Microbiology 66:1692–1697.

del Giorgio, P., and J. Davis. 2003. Patterns in dissolved organic matter lability and consumption across aquatic ecosystems. Pages 399–424 *in* S. Findlay and R. L. Sinsabaugh, editors. Aquatic ecosystems: interactivity of dissolved organic matter. Academic Press, Amsterdam, The Netherlands.

Ducklow, H. W. 2002. Foreword. Pages xv–xix *in* D. Hansell and C. A. Carlson, editors. Biogeochemistry of marine dissolved organic matter. Academic Press, Amsterdam, The Netherlands.

Findlay, S. 2003. Bacterial response to variation in organic matter. Pages 363–379 *in* S. Findlay and R. L. Sinsabaugh, editors. Aquatic ecosytems: interactivity of dissolved organic matter. Academic Press, Amsterdam, The Netherlands.

Gazeau, F., A. V. Borges, C. Barrón, C. M. Duarte, N. Iversen, J. J. Middelburg, M.-D. Pizay, M. Frankignoulle, and J.-P. Gattuso. 2005*a*. Net ecosystem metabolism in a micro-tidal estuary (Randers Fjord, Denmark): evaluation of methods and interannual variability. Marine Ecology Progress Series 301:23–41.

Gazeau, F., C. M. Duarte, J. P. Gattuso, C. Barrón, N. Navarro, S. Ruíz, Y. T. Prairie, M. Calleja, B. Delille, M. Frankignoulle, and A. V. Borges. 2005*b*. Whole-system metabolism and CO_2 fluxes in a Mediterranean Bay dominated by seagrass beds (Palma Bay, NW Mediterranean). Biogeosciences 2: 43–60.

Gazeau, F., J.-P. Gattuso, J. J. Middelburg, N. Brion, L.-S. Schiettecatte, M. Frankignoulle, and A. V. Borges. 2005*c*. Planktonic and whole system metabolism in a nutrient-rich estuary (The Scheldt Estuary). Estuaries 28:868–883.

Haykin, S. 1999. Neural networks. Prentice Hall, Upper Saddle River, New Jersey, USA.

Hedges, J. I. 2002. Why dissolved organic matter? Pages 1–34 *in* D. Hansell and C. A. Carlson, editors. Biogeochemistry of marine dissolved organic matter. Academic Press, Amsterdam, The Netherlands.

Heip, C. 1989. The ecology of the estuaries of Rhine, Meuse and Scheldt in the Netherlands. Topics in marine biology. Scientia marina 53:457–463.

Herman, P. M. J., T. Ysebaert, and C. Heip. 2005. Land–ocean fluxes and coastal ecosystem. Pages 21–49 *in* J. Vermaat, B. Bouwer, K. Turner, and W. Salomons, editors. Managing European coasts: past, present and future. Springer, Berlin, Germany.

Howarth, R. W., et al. 1996. Regional nitrogen budgets and riverine N&P fluxes for the drainages to the North Atlantic Ocean: natural and human influences. Biogeochemistry 35: 75–139.

Kemp, W. M., E. M. Smith, M. Marvin-DiPascale, and W. R. Boyton. 1997. Organic carbon balance and net ecosystem metabolism in Cheseapeake Bay. Marine Ecology Progress Series 150:229–248.

Lemée, R., E. Rochelle-Newall, F. Van Wambeke, M. D. Pizay, P. Rinaldi, and J. P. Gattuso. 2002. Seasonal variation of bacterial production, respiration and growth efficiency in the open NW Mediterranean Sea. Aquatic Microbial Ecology 29: 227–237.

Lomas, M. W., P. M. Glibert, F. K. Shiah, and E. M. Smith. 2002. Microbial processes and temperature in Chesapeake Bay: current relationships and potential impacts of regional warming. Global Change Biology 8:51–70.

Lundberg, C. 2005. Conceptualizing the Baltic Sea ecosystem: an interdisciplinary tool for environmental decision making. Ambio 34:433–439.

Makino, W., J. B. Cotner, R. W. Sterner, and J. J. Elser. 2003. Are bacteria more like plants or animals? Growth rate and resource dependence of bacterial C:N:P stoichiometry. Functional Ecology 17:121–130.

Marañón, E., P. Cermeño, E. Fernández, J. Rodríguez, and L. Zabala. 2004. Significance and mechanisms of photosynthetic production of dissolved organic carbon in a coastal eutrophic ecosystem. Limnology and Oceanography 49: 1652–1666.

Marqhardt, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal of Applied Mathematics 11:431–441.

Morán, X. A. G., M. Estrada, J. M. Gasol, and C. Pedros-Alio. 2002. Dissolved primary production and the strength of phytoplankton bacterioplankton coupling in contrasting marine regions. Microbial Ecology 44:217–223.

Nielsen, K., N. Risgaard-Petersen, B. Sømod, S. Rysgaard, and T. Bergø. 2001. Nitrogen and phosphorus retention estimated independently by flux measurements and dynamic modelling in the estuary, Randers Fjord, Denmark. Marine Ecology Progress Series 219:25–40.

Nieuwenhuize, J., Y. E. M. Maas, and J. J. Middelburg. 1994. Rapid analysis of organic carbon and nitrogen in particulate materials. Marine Chemistry 45:217–224.

Nixon, S. 1995. Coastal marine eutrophication: a definition, social causes, and future concerns. Ophelia 41:199–219.

Noble, R. A., K. A. Bidle, and M. Fletcher. 1997. Natural microbial community compositions compared by a back-propagating neural network and cluster analysis of 5S rRNA. Applied and Environmental Microbiology 63:1762–1770.

Porter, K. G., and Y. S. Feig. 1980. The use of DAPI for identifying and counting aquatic microflora. Limnology and Oceanography 25:943–948.

Rivkin, R. B. 1989. Influence of irradiance and spectral quality on the carbon metabolism of phytoplankton. I. Photosynthesis, chemical composition and growth. Marine Ecology Progress Series 55:291–304.

Rivkin, R. B., and L. Legendre. 2001. Biogenic carbon cycling in the upper ocean: effects of microbial respiration. Science 291:2398–2400.

Rochelle-Newall, E. J., M.-D. Pizay, J. J. Middelburg, H. T. S. Boschker, and J.-P. Gattuso. 2004. Degradation of riverine dissolved organic matter by seawater bacteria. Aquatic Microbial Ecology 37:9–22.

Scardi, M., and L.W. Harding. 1999. Developing an empirical model of phytoplankton primary production: a neural network case study. Ecological Modelling 120:213–223.

Simon, M., and F. Azam. 1989. Protein content and protein synthesis rates of planktonic marine bacteria. Marine Ecology Progress Series 51:201–213.

Small, C., and R. J. Nicholls. 2003. A global analysis of human settlement in coastal zones. Journal of Coastal Research 19: 584–599.

Smith, D. C., and F. Azam. 1992. A simple, economical method for measuring bacterial protein synthesis rates in seawater using 3H-leucine. Marine Microbial Food Webs 6:107–114.

Smith, E. M., and W. M. Kemp. 2001. Size structure and the production/respiration balance in a coastal plankton community. Limnology and Oceanography 46:473–485.

Smith, S. V., J. T. Hollibaugh, S. J. Dollar, and S. Vink. 1989. Tomales Bay, California: a case for carbon-controlled nitrogen cycling. Limnology and Oceanography 34:37–52.

Soetaert, K., and P. M. J. Herman. 1995. Estimating estuarine residence times in the Westerschelde (The Netherlands) using a box model with fixed dispersion coefficients. Hydrobiologia 311:215–224.

Veuger, B., J. J. Middelburg, H. T. S. Boschker, J. Nieuwenhuize, P. van Rijswijk, E. J. Rochelle-Newall, and N. Navarro. 2004. Microbial uptake of dissolved organic and inorganic nitrogen in Randers Fjord. Estuarine Coastal and Shelf Science 61:507–515.

Vidal, M., C. M. Duarte, and M. C. Sanchez. 1999. Coastal eutrophication research in Europe: progress and imbalances. Marine Pollution Bulletin 38:851–854.

Winter, C., A. Smit, T. Szorke-Dénes, G. J. Herndl, and M. G. Weinbauer. 2005. Modelling viral impact on bacterioplankton in the North Sea using artificial neural networks. Environmental Microbiology 7:881–893.

Wollast, R. 1988. The Scheldt estuary. Pages 183–193 *in* W. Salomons, B. L. Bayne, E. K. Duursma, and U. Förstner, editors. Pollution in the North Sea. Springer-Verlag, Berlin, Germany.

APPENDIX A

A table of Spearman correlation coefficients for the parameters used in model development (*Ecological Archives* A017-064-A1).

APPENDIX B

A table of values of the weights from the input layer to the hidden layer and from the hidden layer to the output layer of the model of the ratio of gross primary production to community respiration (GPP:CR) developed in this study based on an artificial neural network (ANN) (*Ecological Archives* A017-064-A2).

APPENDIX C

A table of means and standard deviations of actual data used for model development (*Ecological Archives* A017-064-A3).