Estuarine, Coastal and Shelf Science 91 (2011) 411–422

Contents lists available at ScienceDirect

Estuarine, Coastal and Shelf Science

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

Exploring the long-term and interannual variability of biogeochemical variables in coastal areas by means of a data assimilation approach

Stefano Ciavatta [a,b,*], Roberto Pastres [c]

[a] Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, United Kingdom
[b] EuroMediterranean Centre for Climate Change, Consorzio Venezia Ricerche, Via della Libertà 12, 30175 Venezia, Italy
[c] Department of Environmental Sciences, Informatics and Statistics, University Cà Foscari of Venice, Dorsoduro 2137, 30123 Venezia, Italy

ARTICLE INFO

Article history:
Received 21 March 2010
Accepted 28 November 2010
Available online 3 December 2010

Keywords:
time series analysis
Kalman filter
water quality monitoring
lagoon of Venice

ABSTRACT

Dynamic Harmonic Regression (DHR) models are applied here to the investigation of the interannual changes in the trend and seasonality of biogeochemical variables monitored in coastal areas. A DHR model can be regarded as a time-series component model, where the phases and amplitudes of the seasonal component, as well as the trend, are parameters that vary with time, reflecting relevant changes in the evolution of the biogeochemical variables. The model parameters and their confidence bounds are estimated by data assimilation algorithms, i.e. the Kalman filter and the Fixed Interval smoother. The DHR model structure is here identified by a preliminary spectral analysis and a subsequent minimization of the Bayesian Information Criterion, thus avoiding subjective choices of the frequencies in the seasonal component. The methodology was applied to the investigation of the long-term and interannual variability of ammonia, nitrate, orthophosphate and chlorophyll-a monitored monthly in the lagoon of Venice (Italy) during the years 1986–2008. It was found that the long-term evolutions of the biogeochemical variables were characterized by non-linear patterns and by statistically significant changes in the trend. The seasonal cycles of all the variables were characterized by a marked interannual variability. In particular, the changes in the seasonality of chlorophyll and nitrate were significantly related to the changes in the seasonality of water temperature at the study site and of nutrient concentrations in river discharges, respectively. These results indicate that the methodology could be a sound alternative to more traditional approaches for investigating the impacts of changes in environmental and anthropogenic forcings on the evolution of biogeochemical variables in coastal areas.

1. Introduction

In recent decades, the efforts in monitoring biogeochemical variables in coastal areas, e.g. chlorophyll and nutrient concentrations, have increased worldwide, fuelled by concerns about global warming (Smetacek and Cloern, 2008; Ducklow et al., 2009) or responding to environment protection legislation, such as the Water Framework Directive in Europe or the Ocean Act in USA, Australia and Canada (Borja et al., 2008).

Decadal time series of biogeochemical variables are currently available at several coastal sites such that appropriate time series analysis tools could be applied for estimating both multi-annual trends and systematic seasonal fluctuations.

Indeed, changes in the trend as well as interannual variations of the seasonal component could be related quantitatively to changes of the climatic forcings (see for example Villate et al., 2008) or changes of the anthropogenic pressures (see for example Guadayol et al., 2009 and Aravena et al., 2009), that may trigger changes and regime-shifts of the ecosystem (Folke et al., 2004; Viaroli et al., 2008; Zaldivar et al., 2008; Widdicombe et al., 2010). Therefore, univariate time series (TS) models that decompose the time series of monitoring data into long-term, seasonal and random components have proved to be valuable tools in the context of coastal-system investigations. Component TS models can support different analytical approaches such as ecosystem modelling (de Vries et al., 1998), transfer function models (Villate et al., 2008; Aravena et al., 2009), and wavelet analysis (Nezlin and Li, 2003; Kromkamp and Van Engeland, 2010).

Recent attempts at applying univariate models to time series of biogeochemical variables addressed their non-stationarity (Young et al., 1991), but focused separately either on the estimation of the non-linear trend or on the estimation of the seasonal

* Corresponding author. Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, United Kingdom.
E-mail address: s.ciavatta@pml.ac.uk (S. Ciavatta).

0272-7714/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ecss.2010.11.006

component. For example, non-linear trends were estimated by means of moving averages of zooplankton observations (e.g. David et al., 2005), by fitting exponential and gamma models to nutrient concentrations (Pastres et al., 2004), or fitting second-order polynomials to chlorophyll concentrations (Kromkamp and Van Engeland, 2010). On the other hand, shifts in the seasonal patterns of a plankton time series were determined by Dowd et al. (2002, 2004) by means of a Fourier model, with a single frequency, whose time-varying phase and amplitude were estimated using the Kalman filter. The seasonality of the same time series was investigated by Ikeda et al. (2008) by means of functional data analysis with a Fourier basis. This method led a higher flexibility in the choice of the model structure, allowing the subjective inclusion of a second frequency in the seasonal component, which interannual variability was investigated by means of derivative calculation and curve registration. Nevertheless, the above methods were based on the assumption that the long-term trend of the plankton time series was not significant (Dowd et al., 2002, 2004; Ikeda et al., 2008).

Thus, the mentioned approaches to time series decomposition do not address the root of the problem, as pointed out by Ikeda et al. (2008), namely the simultaneous estimation of both the trend and the seasonal component in presence of non-linear changes of the mean level and interannual shifts of the periodical component. Equally relevant is to provide uncertainty measures of the estimates, in order to evaluate the statistical significance of the changes of the biogeochemical patterns in the coastal area (Beck, 1987).

Given the above, Dynamic Harmonic Regression models (DHR; Young et al., 1999) could represent a sound approach, since they are characterized by a very flexible structure that allows the decomposition of non-linear and non-stationary time series (Young et al., 1999). Therefore, DHR models can present several advantages with respect to classical analytical methods such as ARIMA or Census models (see for example the discussions in Young et al., 1999, and Pedregal and Trapero, 2007). The DHR parameters that define the trend, the amplitude and the phase of the seasonal component are regarded as time varying and they are estimated simultaneously, by means of data assimilation algorithms that process the data in sequence. This approach allows the estimate of the trajectory in time of the parameters and of the model output, as well as those of their standard errors, even with respect to missing data (Young et al., 1999). The trajectories of the parameters – and of their standard error – could provide insights into the dynamic of environmental systems, as shown in Young (1998), and were applied to detect statistically significant changes of the trend of air quality (Becker et al., 2006), and phase shifts of air temperature (Young, 2000).

In the framework of environmental studies, the DHR modelling approach has already been applied to non-stationary time series in hydrology (e.g. Keery et al., 2007; Chappel et al., 2009; Vogt et al., 2010), climate science (e.g. Young, 1998; Young, 2000; Taylor et al., 2007) and air quality studies (e.g. Romanowicz et al., 2006; Becker et al., 2006, 2008). Nevertheless, to the authors' knowledge, the potentiality and usefulness of this approach in the framework of coastal areas studies have not been explored as yet.

The objective of the present work is to demonstrate the potential advantages of applying Dynamic Harmonic Regression models to estimate the non-linear trends and the interannual variability of the seasonal cycles of highly noisy biogeochemical data collected in coastal areas. The identification of the most adequate DHR model is a key issue, which, however, has not been fully addressed as yet from the theoretical point of view (Pedregal and Trapero, 2007; Jiang et al., 2010). Therefore, in previous applications this problem was addressed by exploiting "a priori" a hypothesis (e.g. Vogt et al., 2010) or by using an arbitrary threshold of model performance (e.g. Jiang et al., 2010). In this work, we propose an operational procedure for the identification of the adequate DHR model, based on a preliminary spectral analysis (Young et al., 1999) and on the subsequent application of a Goodness-of-Fit criterion to a set of candidate models.

The method was successfully tested on twenty-year long time series of monthly, highly noisy (inherently variable) observations of chlorophyll-a, nitrogen, ammonia and orthophosphate monitored in the shallow-water lagoon of Venice, Italy, during the years 1986–2008. This case study explores the potential advantages offered by a peculiar feature of the method – i.e. the estimation of the trajectories of the parameters characterizing the trend, the seasonal component and their standard errors – for: i) detecting statistically significant changes of the trends of the biogeochemical variables, and ii) investigating the relationships between the interannual variability of the seasonal cycles and that of the environmental forcings.

2. Methods

2.1. The Dynamic Harmonic Regression model

The Dynamic Regression (DHR) model, described in detail in Young et al. (1999), is a non-stationary univariate time series model that can be represented in the following component form:

$$y_t = T_t + S_T + e_t \quad e_t \sim N\left(0, \sigma^2\right) \tag{1}$$

where y_t is the observed time series, and T_t, S_t, and e_t, represent the trend, seasonal and stochastic components, respectively. In Eq. (1) e_t is a normally distributed random sequence with zero mean and variance σ^2, and S_t has the form of a harmonic regression model or, equivalently, of a Fourier polynomial:

$$S_T = \sum_{i=1}^{R} \left(a_{i,t}\cos(\omega_i t) + b_{i,t}\sin(\omega_i t)\right) \tag{2a}$$

where $\omega_i = (2\pi i)/s$, $i = 1, 2, \ldots, R$ are the fundamental and harmonic frequencies of the sinusoidal term i, and s is the period of the fundamental cycle. The number R of the sinusoidal components needs to be opportunely estimated when applying the DHR model, as described in Section 2.2.

Equation (2a) can be rewritten in an equivalent form that puts in evidence the meaning of the parameters $a_{i,t}$ and $b_{i,t}$, $i = 1,\ldots, R$:

$$\begin{aligned} S_T &= \sum_{i=1}^{R} \left(\sqrt{a_{i,t}^2 + b_{i,t}^2}\cos\left(\omega_i t + \tan^{-1}(b_{i,t}/a_{i,t})\right)\right) \\ &= \sum_{i=1}^{R} \left(A_{i,t}\cos\left(\omega_i t + \phi_{i,t}\right)\right) \end{aligned} \tag{2b}$$

As one can see in Eq. (2b), the parameters $a_{i,t}$ and $b_{i,t}$ define the amplitude $A_{i,t} = \sqrt{a_{i,t}^2 + b_{i,t}^2}$ and the phase $\phi_{i,t} = \tan^{-1}(b_{i,t}/a_{i,t})$ of the sinusoidal term i.

The model in Eqs. (1) and (2) is different from a classical additive time series model (see for example the Census decomposition approach in David et al., 2005) because the parameters which define the trend, i.e. T_t itself, and the seasonal component S_t, i.e. $a_{i,t}$ and $b_{i,t}$, are modelled as Time Variable Parameters (TVPs), i.e. as stochastic variables. As a consequence, also the amplitudes $A_{i,t}$ and the phases $\phi_{i,t}$ of the harmonic components in Eq. (2b) can vary with time (Young et al., 1999).

The TVPs are indicated with X_t^j in the following, and their time evolution is modelled accordingly to a Generalized Random Walk model (GRW; Young et al., 1999):

$$\begin{aligned} X_t^j &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x_{1,t}^j \\ x_{2,t}^j \end{pmatrix} \\ \begin{pmatrix} x_{1,t}^j \\ x_{2,t}^j \end{pmatrix} &= \begin{pmatrix} \alpha^j & \beta^j \\ 0^j & \gamma^j \end{pmatrix} \begin{pmatrix} x_{1,t-1}^j \\ x_{2,t-1}^j \end{pmatrix} + \begin{pmatrix} \delta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_{1,t}^j \\ \eta_{2,t}^j \end{pmatrix} \\ &= \mathbf{F}^j \mathbf{x}_{t-1}^j + \mathbf{G}^j \boldsymbol{\eta}_{t-1}^j \end{aligned} \tag{3}$$

In Eq. (3), $X_t^j, j = 0, 1, \ldots, 2R$ represent the parameters T_t, $a_{i,t}$ and $b_{i,t}$, $i = 1, \ldots, R$ – defined in Eqs. (1) and (2) – and α, β, and γ are constants which define the type of GRW model adopted for X_t^j. For example, by setting $\alpha = \beta = \gamma = 0$ and $\delta = 1$, one obtains a simple Random Walk (RW) model, while for $\alpha = \beta = \gamma = 1$ and $\delta = 0$, a smoother Integrated Random Walk (IRW) model is obtained (Young et al., 1999).

The two stochastic state variables $x_{1,t}^j$ and $x_{2,t}^j$ can be interpreted as the changing level and the changing slope of the parameter X_t^j and their total number in a DHR model with R components is $n = 2(2R+1)$. The temporal changes of X_t^j are given by $\eta_{1,t}^j$ and $\eta_{2,t}^j$, that are zero mean, serially uncorrelated, white noise variables with equal variances $\hat{\sigma}_{\omega_i}^2$. The variances $\hat{\sigma}_{\omega_i}^2$, $i = 0, \ldots, R$ ($\hat{\sigma}_{\omega_0}^2$ refers to the trend) are called the hyper-parameter of the DHR model, and their values are collected in the block covariance diagonal matrix $\mathbf{Q}(n \times n)$. The hyper-parameters $\hat{\sigma}_{\omega_i}^2$ as well as the variance σ^2 in Eq. (1), are unknown and need to be estimated. As pointed out in Young et al. (1999) the problem of estimating $\hat{\sigma}_{\omega_i}^2$ and σ^2 can be conveniently cast by estimating the noise to variance ratios $\hat{\sigma}_{r,i}^2 = \sigma_{\omega_i}^2/\sigma^2$, i.e. $\mathbf{Q}_r = \mathbf{Q}/\sigma^2$. In order to estimate this ratios, we applied the method of estimation in the frequency domain presented in Young et al. (1999), which has been proved to be advantageous, e.g. with respect to Maximum Likelihood estimation, when applied with DHR models and environmental data (Young et al., 1999; Taylor et al., 2007; Keery et al., 2007; Vogt et al., 2010). The method requires the preliminary estimation of the empirical spectrum of the time series by means of an auto regressive model, which order is identified by using the Akaike Information Criterion (AIC). The logarithm of the empirical spectrum is then approximated by the logarithm of the pseudo-spectrum of the DHR model. To this aim, an objective function is minimized, by using a non-linear least square algorithm, which provides the estimates of the ratios $\hat{\sigma}_{r,i}^2$.

In order to estimate the TVPs X_t^j in Eq. (3) – and thus the temporal changes of the trend T_t and of the phases $\phi_{i,t}$ and amplitudes $A_{i,t}$ in Eqs. (1) and (2) – Eqs. (1)–(3) are rewritten in the following State-Space form (Young et al., 1999):

$$\text{Observation equation}: y_t = \boldsymbol{h}_t^T \boldsymbol{x}_t + e_t \tag{4a}$$

$$\text{State equations}: \boldsymbol{x}_t = \boldsymbol{F} \boldsymbol{x}_{t-1} + \boldsymbol{G} \eta_t \tag{4b}$$

In Eq. (4a) y_t is the observed time series, the vector $\boldsymbol{x}_t = (x_{1,t}^0 x_{2,t}^0 \ldots x_{1,t}^{2R} x_{2,t}^{2R})^T$ collects the $n = 2(2R+1)$ additional state variables $x_{1,t}^j$ and $x_{2,t}^j$ defined in Eq. (3); the $(n \times 1)$ vector $\boldsymbol{h}_t = (10\cos(\omega_1 t)0\sin(\omega_1 t)0\ldots\cos(\omega_R t)0\sin(\omega_R t)0)$ collects the sinusoidal terms in Eq. (2a) and accounts for the definition of X_t^j in Eq. (3); and the stochastic variable e_t, with variance σ^2, is defined as in Eq. (1).

Eq. (4) describe the temporal dynamic of the state variables $x_{1,t}^j$ and $x_{2,t}^j$ in Eq. (3), driven by the $(n \times n)$ block matrixes $\boldsymbol{F}$, $\boldsymbol{G}$ and by the $(n \times 1)$ vector $\boldsymbol{\eta}_t$, which collect, in the opportune positions, the matrixes $\boldsymbol{F}^j$, $\boldsymbol{G}^j$ and $\boldsymbol{\eta}_{1,t}^i$ defined in Eq. (3).

In the framework of the DHR modelling approach, the estimation of temporal dynamic of $x_{1,t}^j$ and $x_{2,t}^j$ in Eq. (4) is carried out by using in sequence the Kalman filter and the Fixed Interval Smoothing (FIS) algorithms (Young et al., 1999). The Kalman filter processes the time series from the first up the last observation and provides initial estimates of the TVPs evolutions, which are then refined by processing the observations singularly and in reverse order, using the FIS algorithm. In this way, the mean square error of the one-step-ahead predictions of y_t is minimized.

In the last decades, recursive algorithms like the above ones have been often referred to as "data assimilation methods", coupled with mechanistic or statistical models (see for example Cohn et al., 1994, and Romanowicz and Young, 2003), and the same terminology has been adopted in the present paper.

It is of note that the application the Kalman filter and FIS algorithm with the model in Eq. (4) allows one the simultaneous estimation of the trend and of the seasonal component, and provides estimates of the error covariance matrix $\widehat{\mathbf{P}}_t$ of the time-varying parameters (Young et al., 1999). This matrix, which cannot be estimated in such a straightforward way when the time series components are estimated separately and in sequence, allows one i) to estimate the uncertainty of the DHR model outputs, and ii) to detect statistically significant changes of model parameters that can be ecologically relevant (e.g. the slope of the trend), as it is exemplified in the case study.

2.2. Identification of the seasonal component

The identification of the seasonal component was carried out in this work according to a procedure which was found to be suitable for the analysis of noisy time series of biogeochemical data. This procedure aims to address the current lack of consensus on the criteria for selecting the number R of harmonics of the DHR model (see Eq. (2)), when dealing with noisy time series potentially characterized by very high number of relevant frequencies (Pedregal and Trapero, 2007; Jiang et al., 2010).

The approach consists in i) a preliminary spectrum analysis of the time series, and ii) a subsequent search of the minimum of a Goodness-of-Fit index – which includes a penalty factor for the model complexity – within a set of candidate models.

The preliminary spectrum analysis is part of the DHR modelling approach presented in Young et al. (1999). It consists in the estimation of the spectrum of the time series by using an Auto Regressive (AR) model, with constant parameters, and with order n identified by reference to the Akaike Information Criterion. The peaks in the AR(n) spectrum define the most relevant frequencies, which are here used for constructing a set of candidate models of the seasonal component.

The set of candidate models is obtained by starting with the model including the most relevant frequency identified in the preliminary analysis, and progressively adding the second, third, etc. most significant frequencies, according to the descending values of their spectrum power. For the sake of interpretability, the frequencies were rounded to the annual harmonic values – i.e. 1/12, 1/6, 1/4 etc. – when the numerical differences were small, i.e. less than 5%. The candidate DHR models are then applied to estimate $\hat{y}_t$, after setting the noise to variance ratio of the model parameters equal to zero. The "best" model in the set is the one that minimizes the Bayesian Information Criterion (BIC; Schwarz, 1978):

$$BIC = N \ln \left(\frac{\sum_{t=1}^{N} \left(y_t - \hat{y}_t \right)^2}{N} \right) + k \ln(N) \tag{5}$$

where N is the number of observations, y_t and $\hat{y}_t$ are, respectively, the observation and its DHR estimate at time t, and $k = 1 + (R \times 2)$ is the number of parameters when R sinusoidal components, besides the trend, are included in the model.

The model that minimizes the BIC index in Eq. (5) defines which frequencies are included in the model that is used to describe the time series under investigation. The BIC index is a consolidated statistical criterion for model selection (see for example Mestekemper et al., 2010), which decreases with the residual sum of squares $RSS = \sum_{t=1}^{N}(y_t - \hat{y}_t)^2$, and increases when the model complexity – i.e. the number of model parameters k – increases. Thus, the model which minimizes the BIC value can be regarded as a suitable compromise between accuracy and complexity.

A relevant aspect of the above procedure is that a DHR with constant parameters is used in the place of the DHR model with time variable parameters, as a consequence of setting the noise to variance ratios equal to zero. The relevance of this approximation, which is functional to the applicability of the BIC index, is addressed in the discussion.

2.3. Decomposition of the time series

Once the seasonal component has been identified, the DHR model is applied to the decomposition of the time series by modelling the trend T_t as an IRW (see eq. 3) and the seasonal parameters $a_{i,,t}$ and $b_{i,t}$ as RW, and by applying the data assimilation algorithms to their estimation. The choice of these GRW models for the TVPs aims to estimate, on the one hand, a smooth long-term component, and on the other, to empathize the interannual changes of the seasonal cycles over the time window spanned by the data (Young et al., 1999).

The adequacy of the DHR model in decomposing the time series is evaluated by analyzing the statistical distributions of the model residuals $\hat{e}_t = y_t - \hat{y}_t$. In particular, we tested the symmetry of their distribution by applying the stricter Lilliefors' test of normality, at the significance level of 1% (Sheskin, 2007). The model Goodness-of-Fit was evaluated by calculating the coefficient of determination $R^2 = 1 - \sum_{t=1}^{N}(y_t - \hat{y}_t)^2 / \sum_{t=1}^{N}(y_t - \bar{y}_t)^2$ where y_t are the N observations, $\bar{y}_t$ their mean value, and $\hat{y}_t$ the model estimates.

The signal-to-noise ratio (SNR), which estimates the balance between the deterministic and stochastic component of the time series, is evaluated as the mean value of the ratios between the model estimates of the data and the associated time variable standard errors (Dowd et al., 2004): $SNR = \langle \hat{y}_t/\sigma_{y,t} \rangle$ where $\sigma_{y,t}$ is provided by the error covariance matrix $\widehat{\mathbf{P}}_t$.

2.4. Case study

The methods outlined above were tested in the decomposition of time series of nitrate, ammonia, orthophosphate and chlorophyll-a monthly data collected during the years 1986–2008 in the lagoon of Venice (Italy), at the study site shown in Fig. 1. These time series were collected in the framework of several scientific and institutional monitoring activities (see Table 1), and they are among the longest and more complete available for the lagoon of Venice (Pastres et al., 2004; Penna et al., 2007; Solidoro et al., 2010).

The study site is located in a shallow area of the Central Lagoon (average depth ~1 m), which severely experienced from eutrophication in the eighties (Sfriso and Marcomini, 1996). The area is located closed to the mouth of the Naviglio Brenta river (Fig. 1), which discharges agricultural pollutants from the drainage basin, as well as the emission of a waste water treatment (WWT) plant. The influence of these pollution sources on the evolution of the biogeochemical variables at the study site have been suggested in Solidoro et al. (2010). In the present work, we exploited the available data of nitrate concentrations collected in the Naviglio Brenta river in the years 2000–2007 by the Regional Agency of Environmental Protection (Arpav, 2009), in order to discuss the outcomes of the decomposition of the biogeochemical time series collected at the study site. For the same reason, we exploited also water temperature data collected at the study site in the framework of the monitoring activity "MELa" in the years 2001–2008 (MAV, 2004, 2006, 2009). The time series of the nutrient concentrations at the river mouth and the water temperature data at the study site were decomposed by means of DHR models, following the procedure outlined in the previous sections.

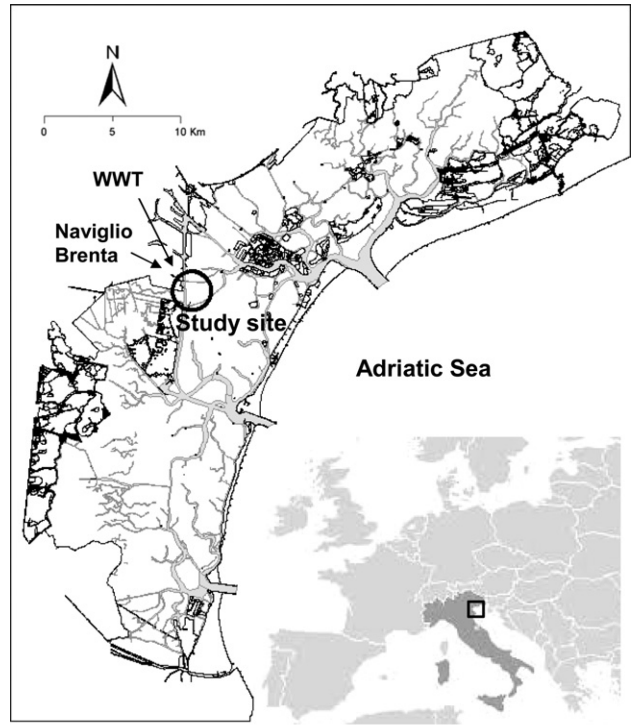

Fig. 1. The lagoon of Venice, Italy, and the study site.

The biogeochemical data were log-transformed prior of the analysis in order to stabilize their variance (Dowd et al., 2002): $y_t = \log_{10}(y_t^* + 1)$, where y_t^* indicates the original data, and y_t the time series of the log-transformed observations.

The case study was carried out in MATLAB® computing environment, by using the DHR modelling functions of the toolbox CAPTAIN developed at the Lancaster University, UK (Taylor et al., 2007).

3. Results

The results of the preliminary spectral analysis aimed to the selection of the DHR models for each biogeochemical variable are presented in Fig. 2, which shows the AR(n) spectra that minimize

Table 1
The monthly biogeochemical data collected at the study site during the years 1986–2008.

Years	Ammonia	Nitrate	Orthophosphate	Chlorophyll	Reference
Apr 86–Dec 90	X	X	X	X	Alberotanza and Zucchetta (1992)
Feb 92–Jun 94	X	X	X		Sfriso and Pavoni (1994)
Jan 95–Jul 99	X	X	X	X (from Jul 97)	MAV (1999)
Sep 00–Dec 03	X	X	X	X	MAV (2004)
Jan 04–Dec 05	X	X	X	X	MAV (2006)
Feb 07–Dec 08	X	X	X	X	MAV (2009)

the AIC index. These show that the time series are characterized by a trend, as indicated by the maxima of the power at periods comparable to the time series length. Moreover, in all cases the highest peaks corresponded to the fundamental cycle of period ~12 months, while lower spectra values were obtained, in most of the cases, at the harmonics corresponding to periods of ~6, ~4 and ~3 and ~2.4 months.

The results of the application of the BIC criterion to the set of candidate models are presented in Fig. 3. The Figure shows the BIC values, rescaled in the range 0–1, as a function of the number R of the sinusoidal terms in Eq. (2). The different number of harmonics accounted for in the BIC minimization, with respect to the different time series, reflects the number of relevant frequencies identified in the preliminary spectrum analysis, which is maximum – and equal to 10 – for orthophosphate (see also Fig. 2). As one can see in Fig. 3, the BIC minima are unambiguously found at $R_{\mathrm{amm}} = 1$, $R_{\mathrm{nit}} = 1$, $R_{\mathrm{orthoph}} = 1$, and $R_{\mathrm{chl}} = 2$, which means that the seasonal cycles of the three nutrient concentrations present only a 12 month periodicity. On the other hand, the chlorophyll model includes also the first harmonic, with period six months, but its relevance is lower with respect to the fundamental component, as indicated by the lower value of the spectrum power associated with the first harmonic (Fig. 2).

The overall performance of the DHR models in decomposing the time series is summarized in Table 2. As one can see, the estimated signal-to-noise ratios *SNR* were in all the cases greater than one, indicating the identifiability of a deterministic signal. However, the noise was not negligible, in particular for the chlorophyll time series, which had the lowest value of *SNR*. The model residuals were in all the cases symmetric and normally distributed at a 1% significance level, based on the Lilliefors' test. Moreover, the R^2 values indicate that the models can explain at least the 50% of the data variance, for ammonia concentrations, but the fraction of explained variance is as high as 83% for nitrate.

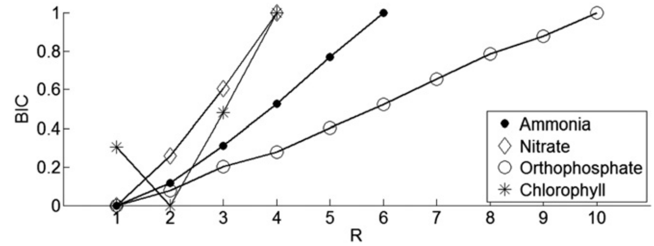

Fig. 3. The rescaled BIC values as a function of the number of sinusoidal components R.

The results of the time series decomposition are presented in Fig. 4, which shows, for each variable: i) the model output $\widehat{y}_t$ and the trend component $\widehat{T}_t$ compared to the log-transformed observations y_t (graphs on the left); and ii) the seasonal component $\widehat{S}_t$ compared to the log-transformed, detrended data: $\widehat{y}_t^{\mathrm{det}} = y_t - \widehat{T}_t$ (graphs on the right).

In Fig. 4, the dotted lines represent the one-standard-error bounds around the model estimates and the seasonal component, provided by the error covariance matrix $\widehat{\mathbf{P}}_t$.

In order to have a closer look at the interannual variations of the seasonal component S_t, in Fig. 5 we plotted the temporal evolution of the anomalies of both the amplitude and the phase around their mean values: $\Delta A_t = A_t - \overline{A}_t$ and $\Delta\phi_t = \phi_t - \overline{\phi}_t$, respectively. Such evolution may provide valuable information about the

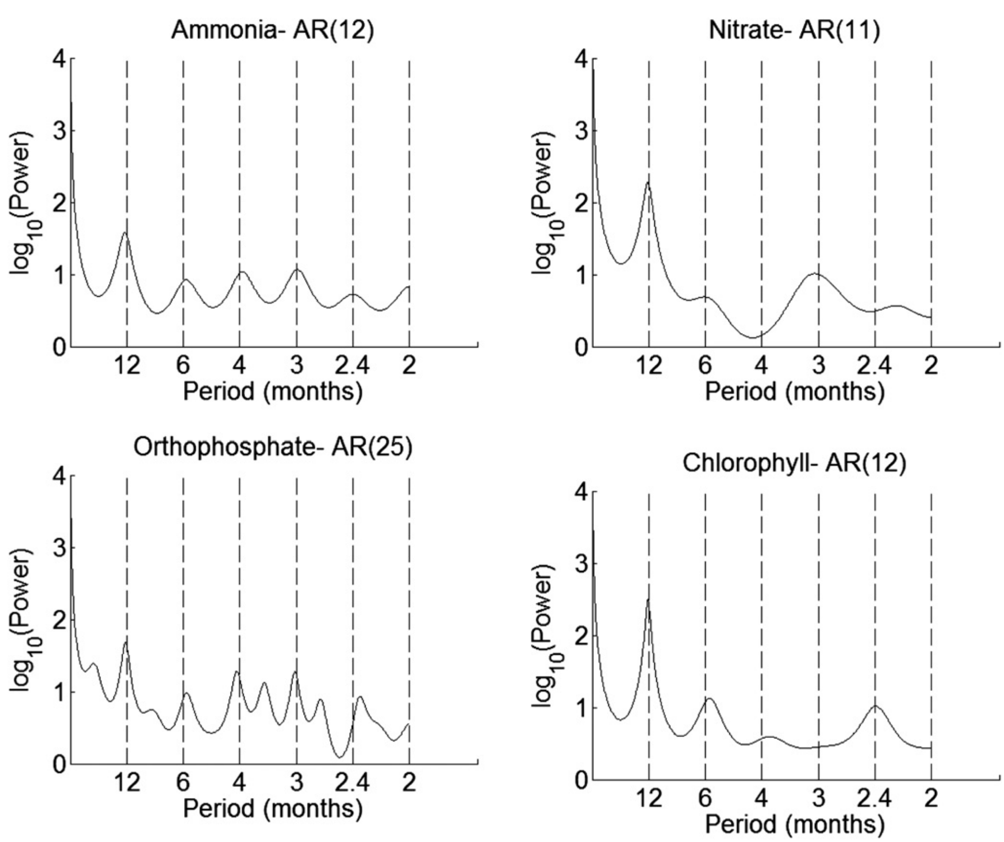

Fig. 2. The AR(n) spectra that minimize the Akaike Information Criterion.

Table 2
Results of the statistical analysis of the DHR model residuals of the biogeochemical time series: number of residuals (N), mean value and standard deviation of the residuals; significance of the Lilliefors' test for normality (p). The last two columns present the coefficient of determination of the DHR model (R^2), and the values of the signal-to-noise ratio (SNR) of the time series.

Variable	N	Mean value	Std. deviation	p	R^2	SNR
Ammonia	223	2.7 E to 08	2.9 E − 01	<0.01	0.47	5.2
Nitrate	209	−1.2 E to 09	1.5 E − 01	<0.01	0.83	10.1
Orthophosphate	221	−3.1 E to 08	2.3 E − 01	<0.01	0.69	4.4
Chlorophyll	172	3.9 E to 08	1.7 E − 01	<0.01	0.77	1.7

relationships between changes in the ecosystem dynamics and forcings, as it will be shown in the discussion.

Figs. 4 and 5 show that, in general, the evolutions of the biogeochemical variables were characterized by non-linear trends and by a marked interannual variability of the seasonal cycles during the investigated twenty years.

In particular, the trends of ammonia and nitrate present some analogies. Their mean levels were fairly constant during the first ten years and then dropped to a lower plateau at the end of the nineties. At the same time, the annual cycles of the two nitrogen forms also changed. Indeed, the estimated annual oscillations of both variables were rather flat in the first decade, in particular in the years 1996–1998, and afterwards they became more pronounced in the last decade. These changes of the seasonal cycles are highlighted in Fig. 5 by the shift of the phases occurred in 1996–1998 and by the positive anomalies of the amplitude during the last decade.

The trend of the orthophosphate had an evident decrease during the first years of the investigated period, and stabilized to a fairly constant value after the year 1990. A rather well defined seasonal signal was evident during most of the years. Nevertheless, the anomalies of the phase and of the amplitude frequently change signs in two subsequent years, e.g. in the biennia 97–98 and 01–02, indicating that the magnitude of the orthophosphate peaks as well as the time of their occurrence can change rather markedly year-by-year.

With respect to chlorophyll, the trend indicates that the mean values were slightly lower in the last quinquennium rather than in the eighties. Nevertheless, relatively high values of chlorophyll where observed through the years 2001–2002. This change was driven by the relative high summer peaks, observed in this period, and it is highlighted by the positive anomalies of the amplitude of the fundamental component (Fig. 5). The amplitude of the first harmonic was lower with respect to the fundamental component – as expected from the lower value of its spectrum power (Fig. 2) – and Fig. 5 shows that the shifts of the harmonic were in general

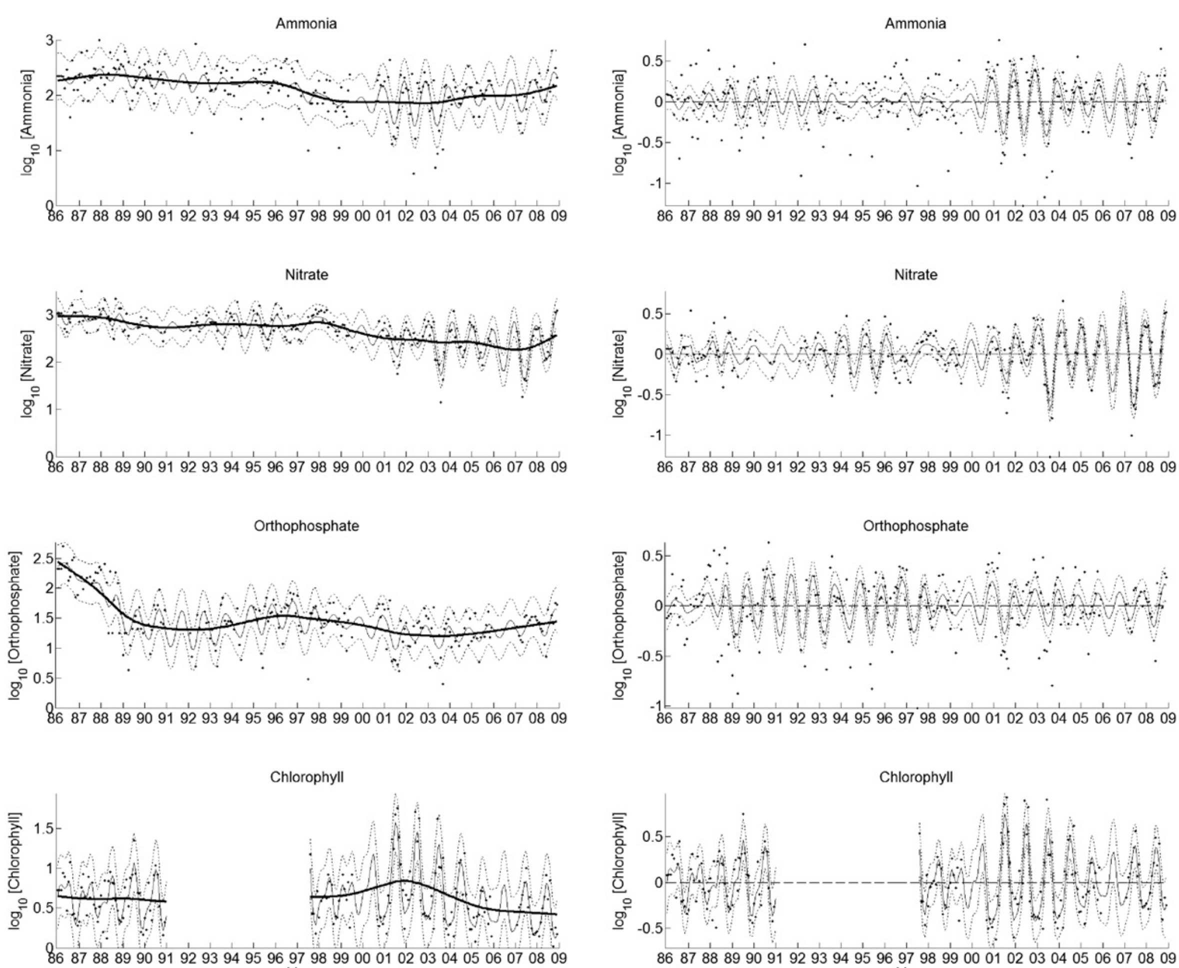

Fig. 4. Results of the decompositions of the time series of the biogeochemical time series. The graphs on the left show the log-transformed data (points), the DHR estimates $\hat{y}_t$ (continuous thin line), the one-standard error band (dotted lines), and the trend $\hat{T}_t$ (bold line); the graphs on the right show the log-transformed, detrended data $\hat{y}_t^{\mathrm{det}}$ (points), the evolution of the seasonal component $\hat{S}_t$ (continuous line) and the one-standard error band (dotted lines).

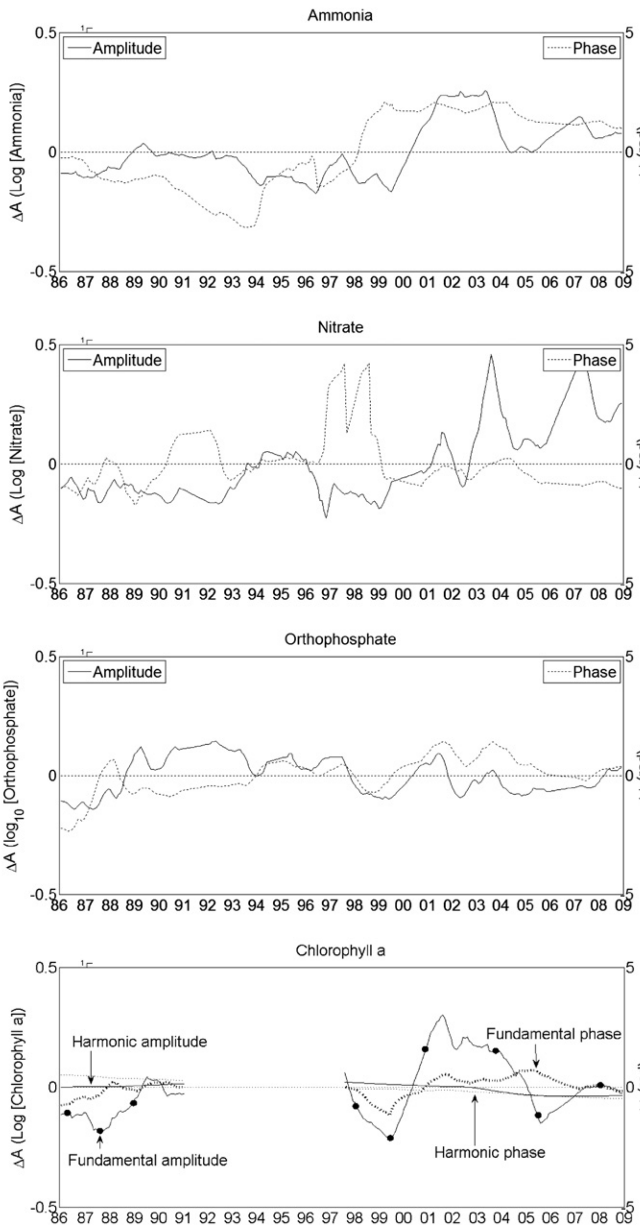

Fig. 5. Anomalies of the phase and of the amplitude of the seasonal components of ammonia, nitrate, orthophosphate and chlorophyll.

negligible with respect to the ones of the fundamental component. The most relevant shifts of the phase were observed for the fundamental term, in particular in the last decade of the time series.

The results of the decomposition of the time series of nutrient concentrations collected at the Naviglio Brenta river in the years 2000–2007, and of the water temperature data at the study site in the years 2001–2008 are synthesized in Table 3. The table shows that the DHR models were effective in decomposing the above time series, since the model residuals were normally distributed ($p < 0.01$) and the explained variance was at least 60% for the nutrient concentrations, up to 93% for the water temperature.

4. Discussion

The results presented in Section 3 indicate that the Methods here applied can be effective in decomposing the time series of biogeochemical data collected in highly dynamic costal areas, such as the Lagoon of Venice. Indeed, the DHR models proved to be adequate to estimate the non-linear trends and the changes of the seasonal cycles of ammonia, nitrate, orthophosphate and chlorophyll during the years 1986–2008 despite the relatively low signal-to-noise ratios. The adequacy is indicated by the symmetric, normal distribution of the residuals, and by the high values of the variance explained by the models (see Table 2 and Fig. 4).

Table 3
Results of the statistical analysis of the DHR model residuals of the forcing data time series (nutrient concentrations at the Naviglio Brenta river and water temperature at the study site): number of residuals (N), mean value and standard deviation of the residuals; significance of the Lilliefors' test for normality (p). The last column presents the coefficient of determination of the DHR model (R^2).

Variable	N	Mean value	Std. deviation	p	R^2
Ammonia	89	−2.5 E to 08	0.28	<0.01	0.65
Nitrate	95	−3.7 E to 08	0.06	<0.01	0.84
Orthophosphate	86	−9.7 E to 08	0.12	<0.01	0.62
Water temperature	83	−2 E to 07	1.89	<0.01	0.93

In the following, we discuss the critical aspects of the procedure that we proposed for identifying the seasonal component of the DHR models. The results of the lagoon case study are then exploited to illustrate that the outcomes of the DHR modelling approach – and in particular the estimates of the time variable model parameters – can be straightforwardly used: i) to detect statistically significative changes of the trend of the biogeochemical variables and ii) to investigate the relationship between the interannual variability of the biogeochemical variables and of the environmental forcings.

4.1. The model identification

The identification of the seasonal component, i.e. the selection of the R harmonics in Eq. (2), is a relevant and critical step in the time series decomposition, which has mostly been dealt using a-priori knowledge and subjective criteria in the framework of biogeochemical studies (see for example Dowd et al., 2002, 2004; David et al., 2005; Ikeda et al., 2008), apart from few exceptions (see for example Dejak et al., 1993). Nevertheless, the lack of "a priori" knowledge or wrong assumptions can lead to miss cyclical signals that can be relevant for characterizing the particular ecosystem under investigation (Ikeda et al., 2008).

In the lagoon case study, we selected the harmonics of the seasonal component of the DHR models by means of an original procedure (see Section 2.2), which has the advantage of providing reproducible results, despite it being an approximation of a rigorous method of model identification, as argued in the following.

The procedure responds effectively to the need for sub-sampling the most relevant frequencies, amongst the whole set identified by the AR(n) spectrum, when dealing with highly variable biogeochemical data. Indeed, it has been pointed out that the AR spectrum can peak out frequencies with relatively low power, which can over-complicate the DHR model with respect the brought benefits, when applied with noisy data (Pedregal and Trapero, 2007; Jiang et al., 2010). That is evident, for example, in the nitrate spectrum in Fig. 2, where the power corresponding to the three-month period is two orders of magnitude lower with respect to the main peak at twelve months. On the other hand, Fig. 2 shows that several frequencies of ammonia data had comparable, relatively high powers. In this case, the BIC criteria allowed us to exclude those frequencies that did not lead to sufficient increases of the goodness-of-fit if compared to the increase of the model complexity. Overall, the sub-sampling procedure based on the BIC allowed us to simplify the model structures, avoiding subjective choices adopted in some previous applications of DHR with environmental data (see for example Vogt et al., 2010), or the use of arbitrary thresholds of model performance adopted in Jiang et al. (2010).

Nevertheless, the procedure we adopted is an approximation of a rigorous identification method, because we used a DHR model with constant parameters for identifying the seasonal component of a DHR model with time variable parameters. This approximation is functional to the applicability of the BIC index (Schwarz, 1978). In our opinion, the following considerations can support the use of the approximation in practical applications with time series of biogeochemical data collected in coastal areas.

Firstly, our procedure partly resembles the identification procedure commonly adopted for the identification of a different type of regressive model with time variable parameters, i.e. the Dynamic Auto Regressive model (DAR, Young et al., 1991). Indeed, in that context, the Akaike Information Criterion is used with a DAR model with constant parameters to identify the order n of the DAR model with time variable parameters. Such a method was successfully applied with environmental data in Young et al. (1991), in the forecast of atmospheric CO_2 evolution.

Secondly, the case study presented in this work indicate the effectiveness of the procedure with highly noisy biogeochemical data collected in a complex coastal area, as indicated by the statistical analysis of the model residuals in Table 2. Moreover, the outcomes of the identification procedure were qualitative coherent with literature findings on the seasonal cycles of the biogeochemical variables in temperate coastal zones, and in the Lagoon of Venice in particular, as it is discussed in Section 4.3.

4.2. Detecting significant changes of the trends

As an alternative approach to the DHR model, one could develop the time series analysis as suggested by Ikeda et al. (2008), i.e. by estimating separately non-linear long trend components (as for example in Pastres et al. (2004); Aravena et al. (2009)), and the seasonal component of the detrended data, by means of Fourier analysis (as in Dowd et al., 2004), functional data analysis (as in Ikeda et al., 2008), or wavelet transforms (Kromkamp and Van Engeland, 2010). However, those approaches do not provide a straightforward way for estimating the confidence band for the predictions and for the time series components: this complicates the ability to establish, on a statistical basis, whether changes in trends and seasonal cycles are significant. The appeal of the DHR approach is that non-linear trend and the seasonal component parameters are estimated simultaneously in a unified data assimilation framework. In this framework, the uncertainties on the model parameters are tracked in time by propagating the error covariance matrix $\widehat{\mathbf{P}}_t$ of the state vector, which includes the model parameters.

As exemplified in the following, $\widehat{\mathbf{P}}_t$ can be exploited for investigating the statistical significance of changes of the long-term component of the biogeochemical time series. Indeed, the graphs in Fig. 4 showed that the DHR models led to the identification of non-linear trends in the period 1986–2008. In particular, the analysis detected changes of the mean levels of the nutrient concentrations, which occurred at the end of the 1980s for orthophosphate, and at the end of the 1990s for ammonia and nitrate. In all cases, the trend decreased and stabilized at lower levels. On the other hand, chlorophyll level increased during the 2-year period 2001–2002, and decreased in the subsequent years.

The statistical significance of these trend changes can be easily assessed with the DHR models from the trajectory of the parameter $x^j_{2,t}$ in Eq. (3), which defines the changing slope of the trend component (Becker et al., 2006). The evolution of this parameter is shown in the graphs of Fig. 6, together with the standard deviation band to the 95% confidence limit, drawn by the error covariance matrix $\widehat{\mathbf{P}}_t$. Fig. 6 shows that the trend shifts of the nutrients mentioned above were in all the cases statistically significant, because the corresponding values of the slope parameter were different from 0, at a confidence level of 95%. The negative values of the slopes indicate significant decreases of ammonia around year 1997, of nitrate in 1999, of orthophosphate in the years 1986–1990. It is of note that the mean values of nitrate increased significantly in the last period of the time series, as highlighted by the significant positive values of the slope in the year 2008.

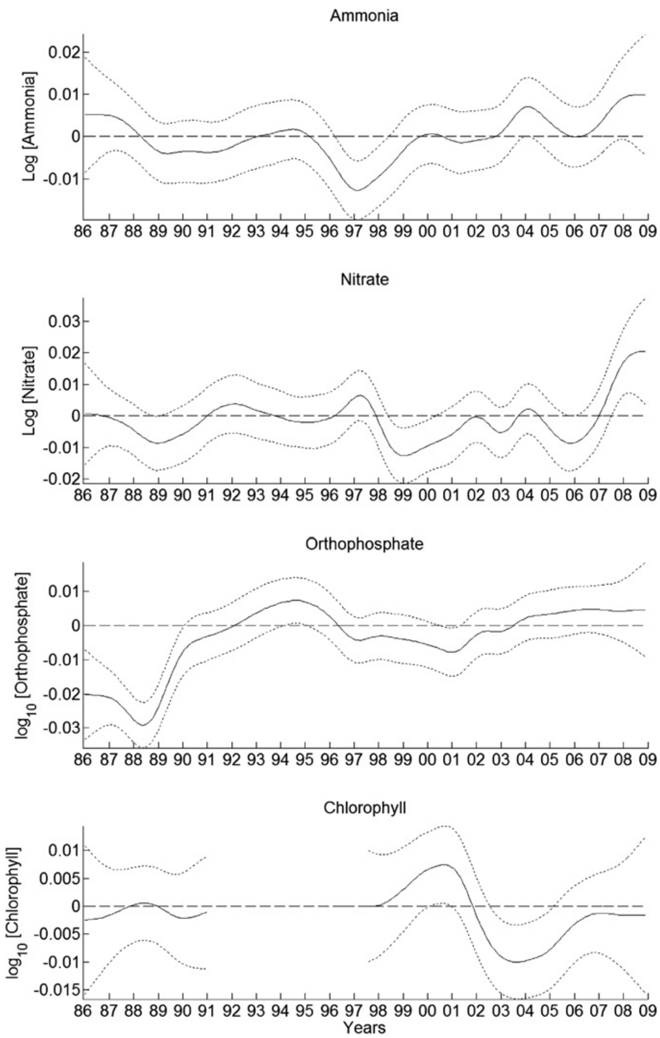

Fig. 6. Time evolution of the slope of the trends during the years 1986–2008 (continuous line). The dotted lines represent the standard deviation band to the 95% confidence limit; the dashed line indicates the zero slope, i.e. no rate of change of the trend.

The results of the trend analysis presented in Fig. 4 and supported by Fig. 6 are partially consistent with the findings in Pastres et al. (2004), which estimated exponentially decreasing trends of ammonia and orthophosphate at the same study site, using data of the period 1976–1999. Nevertheless, Pastres et al. (2004) estimated a constant-value model to represent the trend of nitrate up to 1999. The analysis here presented, based on the recent data of the monitoring activity “MELa” (MAV, 2004, 2006, 2009), showed that also the trend of nitrate is non-linear.

Moreover, the estimated trends at the study agree qualitatively with the evolution of the anthropogenic pressures and ecosystem changes described by local studies (see Solidoro et al. (2010) for a review). In particular, the decreasing trend of orthophosphate during the years 1986–1989 follows the legislative intervention of progressively banning the phosphorus from detergents during the eighties. The significant decrease in nitrate and ammonia concentrations at the end of the 1990s can be ascribed to a series of legislative and management interventions that led to marked reduction in the nitrogen discharges from the waste water treatment plant nearby the study site (Regione Veneto, 2000). Analogous negative trends of nutrient concentrations led by the policy of load reduction have been recently demonstrated for the Scheldt estuary (Soetaert et al., 2006) and for the Danish coast (Carstensen et al., 2006), by analyzing time series of annual nutrient concentrations collected by long-term monitoring activities.

The trend of chlorophyll, that is a proxy of phytoplankton density (Trees et al., 2000), resulted unrelated to the nutrient trends at the study site; this confirms that direct relationships among nutrient loads, concentrations and phytoplankton density are difficult to establish in coastal areas, due to the complexity of the interacting responses of the ecosystem to nutrient enrichment (Cloern, 2001; Soetaert et al., 2006; Carstensen et al., 2006; Kromkamp and Van Engeland, 2010). Nonetheless, a significant positive trend of chlorophyll was observed in the year 2000 (Fig. 6), which led to the relatively high concentrations in the years 2001–2002. The latest high values have been put in relation with the fast assimilation of the nutrients supplied by rivers during the exceptionally rainy winter 2001 and spring/summer 2002 (Solidoro et al., 2004; MAV, 2004). On the other hand, the relatively dry summers in the years 2003–2004, and the consequent relatively low inputs of nutrients, could partly explain the subsequent significant decrease of the chlorophyll concentrations (MAV, 2006). The relevance of episodic nutrient discharges related to the sequence of droughts and meteorological events have been recently pointed out also by Guadayol et al. (2009), related to the Blanes Bay in the Mediterranean sea. By exploring a larger set of biogeochemical time series – i.e. including turbidity, organic dissolved nutrients, and phytoplankton measurements – these authors highlighted that meteorological and biogeochemical events, such as those described here for the Venice lagoon, can temporarily change a coastal ecosystem from heterotrophy to autotrophy.

4.3. *The interannual variability of the seasonal cycles in relation to the forcings*

The results of the case study indicate that the DHR model allowed the estimation of the interannual variability of the seasonal cycles of the biogeochemical variables at the lagoon site, as highlighted by the evolution of S_t in Fig. 4 and the by the model parameter evolutions in Fig. 5. As illustrated in the following, the interannual variability in the years 1986–2008 can be qualitatively linked to the changes of the seasonal forcings, on the basis of literature studies. On the other hand, the decomposition of the time series of nutrient discharges and water temperature data available for the years 2001–2008 (see Table 3), allows us to show that the trajectory of the DHR model parameters can be used to investigate quantitatively the relationships with the forcings.

The changes of the seasonal cycle in each year around the mean value in the period 1986–2008 are redrawn from Fig. 3 in the graphs of Fig. 7. The figure indicates that the mean seasonal cycles of each variable (bold lines) agreed with the typical pattern observed in temperate coastal areas (see for example Cloern, 2001; Villate et al., 2008; Aravena et al., 2009; Kromkamp and Van Engeland, 2010) and in the Venice lagoon (Dejak et al., 1993), and that they can be in turn related to the seasonal cycles of the environmental forcings (Cloern, 2001; Solidoro et al., 2004). Indeed, Fig. 7 shows that the highest values of nutrient concentrations were observed on the average in winter and autumn, when

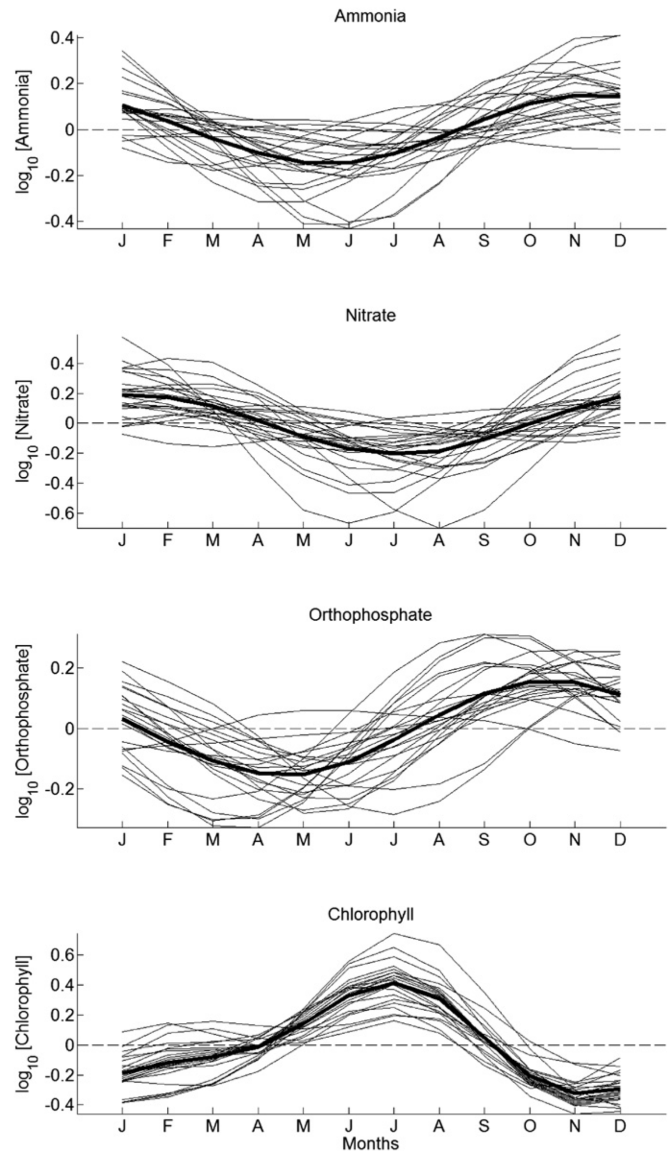

Fig. 7. The thin lines represent the seasonal cycles of the state variable during each year in the period 1986–2008; the bold line represents the mean seasonal cycle over the whole period.

precipitations drive relevant nutrient discharges from the drainage basin to the lagoon (Solidoro et al., 2004), and lower concentration were measured during the drier, spring-summer seasons. In these latter seasons, the values of chlorophyll increase, due to the phytoplankton growth that is stimulated by the higher irradiation and that contributes to reducing the nutrient concentrations by assimilation (Cloern 2001; Solidoro et al., 2004). Nevertheless, chlorophyll peaks can occur in February–March at some years (Fig. 7), as a consequence of winter diatom blooms influenced by water temperature anomalies and nutrient discharges in the study area (Socal et al., 1999). The sporadic occurrence of such high winter values explains the BIC selection of a six month harmonic in the seasonal model of chlorophyll (see Fig. 3).

Nevertheless, Fig. 7 shows also that, at some years, the seasonal cycles can change relatively to the multi-annual mean. These changes are highlighted by the changes in the amplitude and phase values shown in Fig. 5. In particular, a relative change in these two parameters was detected for ammonia at the end of the nineties. During the years 1992–1998, the amplitude of the seasonal

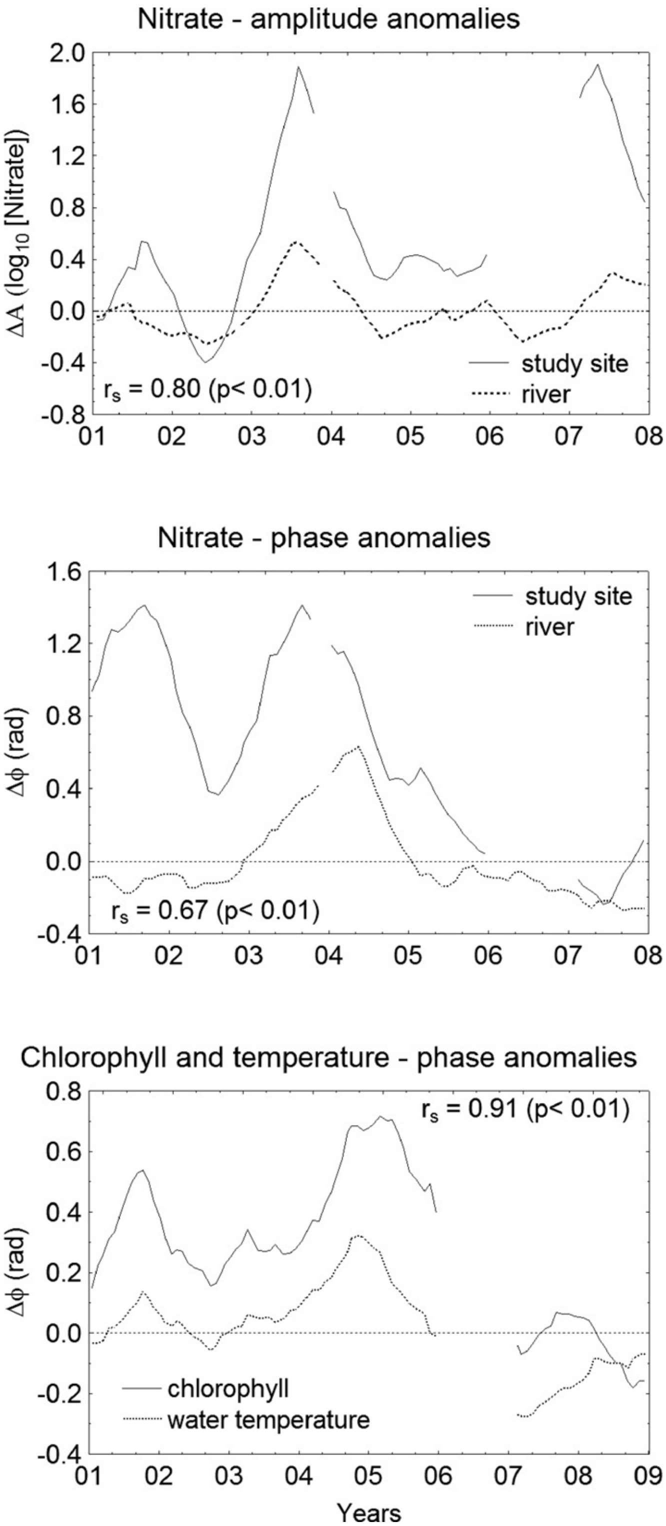

Fig. 8. Evolutions and cross-correlations (r_s Spearman, $p < 0.01$) between the anomalies of the phase and of the amplitude characterizing the evolution of the biogeochemical variables and the evolution of the environmental forcings, i.e. the nutrient concentration in the river and the water temperature at the study site.

oscillations of ammonia was relatively low and the seasonal peaks were estimated to occur in September/October (see Fig. 4). The subsequent change of the amplitude (Fig. 5) is reflected in Fig. 4 by the increase of the range of the oscillations in the period 2001–2008, while the phase change lead the seasonal peaks of ammonia to move towards November/December. These changes may be reasonably related to the decrease in the urban and industrial nitrogen loads from the Naviglio Brenta river, due to the upgrading of the waste water treatment plant in the 90's (see Section 4.2). This management intervention may have led the agricultural loads – driven by the seasonal precipitation – to become a relevant driver of the seasonal pattern of ammonia in the last decade.

The relationships between the changes of the seasonal cycles of the biogeochemical variables and of the environmental forcings could be extended for the years 2001–2008. Indeed, for these later years, we could compare the estimates of the phase and amplitude evolutions in Fig. 5, with the parameter trajectories obtained in the decomposition the time series of nutrient discharges and water temperature data available for the years 2001–2008 (see Table 3).

This comparison is exemplified in Fig. 8, which shows that the amplitude and the phase of the nitrate concentrations at the study site and at the Naviglio Brenta river were significantly correlated in the years 2001–2008 (Spearman correlation r_s, $p < 0.01$). This outcome indicates that the interannual variability of the seasonal cycle of nitrate was significantly modulated by the year-by-year variability of the concentrations in the river, and provides valuable insights on the bio-physical processes that influence the dynamic of nitrate at the study site. Indeed, the rapid temporal response of the concentrations at the study site to the pulses from the river – i.e. the phase correlation – and the correspondence of the absolute values of the pulses – i.e. the correlation of the amplitudes – indicate that straight transport and dilution of the river loads were the main drivers of nitrate at the study site during the investigated period. The physical processes in this case dominated over the biogeochemical processes of nitrate utilization and regeneration at the land/coast interface (see Soetaert et al., 2006), and that can be explained by the presence of a deep canal that conveys the freshwater to the study area (see Fig. 1), in accordance with the findings of Solidoro et al. (2004). On the other hand, relevant correlations were not found between the parameter trajectories of orthophosphate and ammonia concentrations in the river and at the study site. This outcome indicate that these nutrients were involved in more complex biogeochemical dynamics, in agreement with Solidoro et al. (2004), as typically observed in coastal systems (e.g. Cloern, 2001; Soetaert et al., 2006).

Fig. 8 shows also that a significant and high cross-correlation was found between the phases of chlorophyll and water temperature, while no significant cross-correlations were found between their amplitudes (not shown). These outcomes indicate that: i) shifts in the timing of the plankton activity were significantly linked to shifts of the seasonal cycle of the water temperature during the years 2001–2008; but ii) the absolute values of the plankton density were not significantly influenced by the peak values of the physical forcing. This result agrees with the findings of Sfriso and Marcomini (1996) and Socal et al. (1999), which highlighted the relevance of the winter/spring anomalies of temperature in determining shifts of the primary community at the study site. Interestingly, Kromkamp and Van Engeland (2010) related qualitatively the phase shifts of the phytoplankton biomass in the Scheldt estuary – explored with wavelet transforms – to changes of the water temperature driven by global warming. That suggests the potential application of the phase estimates provided by the DHR model for investigating quantitatively the phenological responses of phytoplankton to climate changes in coastal areas.

5. Conclusions

The advantages of applying DHR models (Young et al., 1999) to investigate non-linear trends and changes in the seasonal cycle of biogeochemical variables monitored in coastal areas was evaluated. The effectiveness of the methodology was demonstrated by applying it to the analysis of time series of ammonia, nitrate, orthophosphate and chlorophyll data collected at one site of the lagoon of Venice, during the years 1986–2008. Indeed, the DHR models – selected in this work by using an operational and reproducible procedure – were in all cases adequate to hindcast the monitoring data, as confirmed by the statistical analysis of the model residuals.

The application of this approach in the framework of coastal area studies is a useful alternative, or support, with respect to other methods of time series analysis. It allows the simultaneous estimation of the time-varying trend and seasonal component by means of well established data assimilation algorithms, which also provide estimates of the model uncertainty (Young et al., 1999).

The data assimilation algorithms are currently available in academic and commercial software (e.g. the CAPTAIN toolbox by Taylor et al. (2007)) that allows one to overcome the difficulties of their application by non-statistical practitioners, pointed out by Ikeda et al. (2008). Thus, the methods here applied can be a useful tool to support stakeholders that carry out monitoring activities with the purposes of the surveillance and safeguard of coastal areas. Interestingly, DHR models have been recently applied with remotely sensed land data (Jiang et al., 2010), indicating their potential use with time series of ocean colour observations, which are retrievable in coastal ecosystems (e.g. Ruddick et al., 2010).

The results of the lagoon case study suggest future applications, behind the objectives of the present paper, based on a peculiar outcome of the methodology, i.e. the estimates of the time variable model parameters. Indeed, we found some significant cross-correlations between the changes of the model parameters that characterize, on the one hand, the seasonal evolution of the biogeochemical variables at the study site and, on the other hand, the seasonal evolution of the environmental forcings. This suggests the possibility of investigating interrelationships between the biogeochemical variables and the anthropogenic and environmental forcings, by modelling the time variability of the parameter as a function of the state variables, i.e. by representing the parameters as State-Dependent according to the Data Based Model approach (see for example Young, 1998; Young and Parkinson, 2002; Taylor et al., 2007; Lin and Beck, 2007). These further applications are potentially useful to modelling the effects of climate changes on coastal areas.

Acknowledgements

This work was partially funded by EuroMediterranean Centre for Climate Changes (CMCC) and by the Natural Environment Research Council (UK) through the National Centre for Earth Observation (NCEO). Part of the data was provided by Ministero delle Infrastrutture-Magistrato alle Acque di Venezia, through the concessionary Consorzio Venezia Nuova, and by the Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (Arpav). The authors thank all the participants of the project "MELa" for their support of this work, Dr. Giorgio Dall'Olmo for his useful comments and Prof. Peter Young and his colleagues of the Lancaster University for assisting us with the application of the toolbox CAPTAIN.

References

Alberotanza, L., Zucchetta, G., 1992. Caratteristiche delle acque della laguna di Venezia (Bacino centrale). Dati relativi al monitoraggio effettuato nel corso degli anni 1986–1990. Consorzio Comunale Impianti di Depurazione – Consiglio Nazionale delle Ricerche, Istituto per lo studio della Dinamica delle Grandi Masse, Venezia. Technical Report (in Italian).

Aravena, G., Villate, F., Uriarte, I., Iriarte, A., Ibanez, B., 2009. Response of *Acartia* populations to environmental variability and effects of invasive congenerics in the estuary of Bilbao, Bay of Biscay. Estuarine, Coastal and Shelf Science, 621–628.

Arpav, 2009. Bacino scolante nella Laguna di Venezia – Rapporto sullo stato ambientale dei corpi idrici – Anni 2003–2004. Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto, Venezia, Technical Report (in Italian), online, http://www.arpa.veneto.it.

Beck, M.B., 1987. Water quality modelling: a review of the analysis of uncertainty. Water Resources Research 23 (8), 1393–1442.

Becker, S., Halsall, C.J., Tych, W., Hung, H., Attewell, S., Blanchard, P., Li, H., Fellin, P., Stern, G., Billeck, B., Friesen, S., 2006. Resolving the long-term trends of polycyclic aromatic hydrocarbons in the Canadian Arctic atmosphere. Environmental Science and Technology 40 (10), 3217–3222.

Becker, S., Halsall, C.J., Tych, W., Kallenborn, R., Su, Y., Hung, H., 2008. Long term trends in atmospheric concentrations of α- and γ-HCH in the Arctic provide insight into the effect of legislation and climatic fluctuations on contaminant levels. Atmospheric Environment 42, 8225–8233.

Borja, A., Bricker, S.B., Dauer, D.M., Demetriades, N.T., Ferreira, J.G., Forbes, A.T., Hutchings, P., Jia, X., Kenchington, R., Marques, J.C., Zhu, C., 2008. Overview of integrative tools and methods in assessing ecological integrity in estuarine and coastal systems worldwide. Marine Pollution Bulletin 56 (9), 1519–1537.

Carstensen, J., Conley, D.J., Andersen, J.H., Aertebjerg, G., 2006. Coastal eutrophication and trend reversal. Limnology and Oceanography 51, 398–408.

Chappel, N.A., Discenza, A.R., Tych, W., Whittaker, J., Bidin, K., 2009. Simulating hourly rainfall occurrence within an equatorial rainforest, Borneo Island. Hydrological Sciences 54 (3), 571–581.

Cloern, E.J., 2001. Our evolving conceptual model of the coastal eutrophication problem. Marine Ecology Progress Series 210, 223–253.

Cohn, S.E., Sivakumaran, N.S., Todling, R., 1994. A fixed-lag Kalman smoother for retrospective data assimilation. Monthly Weather Review 122, 2838–2867.

David, V., Sautour, B., Chardy, P., Leconte, M., 2005. Long-term changes of the zooplankton variability in a turbid environment: the Gironde estuary (France). Estuarine, Coastal and Shelf Science 64, 171–184.

de Vries, I., Duin, R.N.M., Peeters, J.C.H., Los, F.J., Bokhorst, M., Laane, R.W.P.M., 1998. Patterns and trends in nutrients and phytoplankton in Dutch coastal waters: comparison of time series analysis, ecological model simulation, and mesocosm experiments. Journal of Marine Science 55, 620–634.

Dejak, C., Franco, D., Pastres, R., Pecenik, G., Solidoro, C., 1993. An informational approach to model time series of environmental data through negentropy estimation. Ecological Modelling 67, 199–220.

Dowd, M., Martin, J.L., Legresley, M.M., Hanke, A., Page, F.H., 2002. Interannual variability in a plankton time series. Environmetrics 14, 73–86.

Dowd, M., Martin, J.L., Legresley, M.M., Hanke, A., Page, F.H., 2004. A statistical method for the robust detection of interannual changes in plankton abundance: analysis of monitoring data from the Bay of Fundy, Canada. Journal of Plankton Research 26 (5), 509–523.

Ducklow, H.W., Scott, C.D., Steinberg, D.K., 2009. Contributions of long-term research and time series observations to marine ecology and biogeochemistry. Annual Review of Marine Science 1, 279–302.

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., Holling, C.S., 2004. Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review of Ecology, Evolution, and Systematics 35, 557–581.

Guadayol, Ò, Peters, F., Marrasé, C., Gasol, J.M., Roldán, C., Berdalet, E., Massana, R., Sabata, A., 2009. Episodic meteorological and nutrient load events as drivers of coastal planktonic ecosystem dynamics: a time series analysis. Marine Ecology Progress Series 381, 139–155.

Ikeda, T., Dowd, M., Martin, J.L., 2008. Application of functional data analysis to investigate seasonal progression with interannual variability in plankton abundance in the Bay of Fundy, Canada. Estuarine, Coastal and Shelf Science 78, 445–455.

Jiang, B., Liang, S., Wang, J., Xiao, Z., 2010. Modelling MODIS LAI time series using three statistical methods. Remote Sensing of Environment 114, 1432–1444.

Keery, J., Binley, A., Crook, N., Smith, J.W.N., 2007. Temporal and spatial variability of groundwater-surface water fluxes: development and application of an analytical method using temperature time series. Journal of Hydrology 336 (1–2), 1–16.

Kromkamp, J.C., Van Engeland, T., 2010. Changes in phytoplankton biomass in the Western Scheldt estuary during the period 1978–2006. Estuaries and Coasts 33, 270–285.

Lin, Z., Beck, M.B., 2007. Understanding complex environmental systems: a dual approach. Environmentrics 18, 11–26.

MAV, 1999. Interventi per l'arresto del degrado del connesso alla proliferazione delle macroalghe in Laguna di Venezia 1997–1998 – Trattamento ed analisi dei dati mediante un modello di qualità dell'acqua - Rapporto Finale. Magistrato Alle Acque di Venezia – Consorzio Venezia Nuova, Venezia. Technical Report (in Italian).

MAV, 2004. Progetto Mela1. Attività A.2: Trattamento ed analisi dei dati, elaborazioni in linea ed interpretazioni spazio temporali. Rapporto Tecnico sull'Analisi statistica dei dati di qualità dell'acqua raccolti nel triennio 2001–2002–2003. Magistrato alle Acque di Venezia – Consorzio Venezia Nuova, Venezia. Technical Report (in Italian).

MAV, 2006. Progetto MELa 3. Attività A: Prosecuzione delle attività di monitoraggio della qualità delle acque lagunari e loro integrazione con le attività di monitoraggio e controllo condotte dalla Sezione Antinquinamento del Magistrato alle Acque. Rapporto tecnico finale sulle attività di monitoraggio della qualità delle acque. Magistrato alle Acque di Venezia – Consorzio Venezia Nuova, Venezia. Technical Report (in Italian).

MAV, 2009. Progetto MELa 4. Monitoraggio di mantenimento delle conoscenze sullo stato delle acque a del macrobenthos. Rapporto tecnico del secondo anno di attività qualità delle acque. Magistrato alle Acque di Venezia - Consorzio Venezia Nuova, Venezia. Technical Report (in Italian).

Mestekemper, T., Windmann, M., Kauermann, G., 2010. Functional hourly forecasting of water temperature. International Journal of Forecasting 26 (4), 684–699.

Nezlin, N.P., Li, B.-L., 2003. Time-series analysis of remote-sensed chlorophyll and environmental factors in the Santa Monica-San Pedro Basin off Southern California. Journal of Marine Systems 39, 185–202.

Pastres, R., Solidoro, C., Ciavatta, S., Petrizzo, A., Cossarini, G., 2004. Long-term changes of inorganic nutrients in the Lagoon of Venice (Italy). Journal of Marine Systems 51, 179–189.

Pedregal, D.J., Trapero, J.R., 2007. Electricity prices forecasting by automatic dynamic harmonic regression models. Energy Conversion and Management 48, 1710–1719.

Penna, G., Barbanti, A., Bernstein, A.G., Boato, S., Casarin, R., Ferrari, G., Montobbio, L., Parati, P., Piva, M.G., Vazzoler, M., 2007. Institution monitoring of the Venice Lagoon, its watershed and coastal waters: a continuous updating of their ecological quality. In: Fletcher, C.A., Spencer, D.C. (Eds.), Flooding and Environmental Challenges for Venice and Its Lagoon. State of Knowledge. Cambridge University, Cambridge, United Kingdom, pp. 517–528.

Regione Veneto, 2000. Piano Direttore 2000-Piano per la prevenzione dell'inquinamento e il risanamento delle acque del bacino idrografico immediatamente sversante nella laguna di Venezia. Regione Veneto, Venezia. Technical Report (in Italian), online, http://www.regione.veneto.it.

Romanowicz, R., Young, P., 2003. Data assimilation and uncertainty analysis of environmental assessment problems – an application to Stochastic transfer function and genealised likelihood uncertainty estimation techniques. Reliability Engineering and System Safety 79, 161–174.

Romanowicz, R., Young, P., Brown, P., Diggle, P., 2006. A recursive estimation approach to the spatio-temporal analysis and modelling of air quality data. Environmental Modelling & Software 21 (6), 759–769.

Ruddick, K., Brockmann, C., Doerffer, R., Lee, Z., Brotas, V., Fomferra, N., Groom, S., Krasemann, H., Martinez-Vicente, V., Sa, C., Santer, R., Sathyendranath, S., Stelzer, K., Pinnock, S., 2010. The Coastcolour project regional algorithm round robin exercise. Proceedings SPIE 7858, 785807, doi:10.1117/12.869506.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6 (2), 461–464.

Sfriso, A., Marcomini, A., 1996. Decline of Ulva growth in the lagoon of Venice. Bioresource Technology 58, 299–307.

Sfriso, A., Pavoni, B., 1994. Macroalgae and phytoplankton competition in the central Venice lagoon. Environmental Technology 15, 1–14.

Sheskin, D.J., 2007. Handbook of Parametric and Nonparametric Statistical Procedures, fourth ed. Chapman & Hall.

Smetacek, V., Cloern, J.E., 2008. On phytoplankton trends. Science 319, 1346–1348.

Socal, G., Bianchi, F., Alberighi, L., 1999. Effects of thermal pollution and nutrient discharges on a spring phytoplankton bloom in the industrial area of the lagoon of Venice. Vie et Milieu 49 (1), 19–31.

Soetaert, K., Middelburg, J.J., Heip, C., Meire, P., Van Damme, S., Maris, T., 2006. Long-term changes in dissolved inorganic nutrients in the heterotrophic Scheldt estuary (Belgium, The Nederlands). Limnology and Oceanography 51, 409–423.

Solidoro, C., Pastres, R., Cossarini, G., Ciavatta, S., 2004. Seasonal and spatial variability of water quality parameters in the lagoon of Venice. Journal Marine System 51, 7–18.

Solidoro, C., Bernardi Aubry, F., Camatti, E., Ciavatta, S., Cossarini, G., Facca, C., Franzoi, P., Libralato, S., Melaku Canu, D., Pastres, R., Pranovi, F., Raicevich, S., Socal, G., Sfriso, A., Sigovini, M., Tagliapietra, D., Torricelli, P., 2010. Response of lagoon of Venice ecosystem to natural and anthropogenic pressures over the last 50 years. In: Kennish, M.J., Paerl, H.W. (Eds.), Coastal Lagoons: Critical Habitats of Environmental Change. CRC Press, pp. 487–516.

Taylor, C.J., Pedregal, D.J., Young, P.C., Tych, W., 2007. Environmental time series analysis and forecasting with the Captain toolbox. Environmental Modelling & Software 22, 797–814.

Trees, C.C., Clark, D.K., Bidigare, R.R., Ondrusek, M.E., Mueller, J.L., 2000. Accessory pigments versus chlorophyll a concentrations within the euphotic zone: a ubiquitous relationship. Limnology and Oceanography 45 (5), 1130–1143.

Viaroli, P., Bartoli, M., Giordani, G., Naldi, M., Orfanidis, S., Zaldivar, J.M., 2008. Community shifts, alternative stable states, biogeochemical controls and feedbacks in eutrophic coastal lagoons: a brief overview. Aquatic Conservation: Marine and Freshwater Ecosystems 18, S105–S117.

Villate, F., Aravena, G., Iriarte, A., Uriarte, I., 2008. Axial variability in the relationship of chlorophyll a with climatic factors and the North Atlantic Oscillation in a Basque coast estuary, Bay of Biscay (1997–2006). Journal of Plankton Research 30 (9), 1041–1049.

Vogt, T., Schneider, P., Hahn-Woernie, L.H., Cirpka, O.A., 2010. Estimation of seepage rates in a losing stream by means of fiber-optic high-resolution vertical temperature profiling. Journal of Hydrology 380 (1–2), 154–164.

Widdicombe, C.E., Eloire, D., Harbour, D., Harris, R.P., Somerfield, P.J., 2010. Long-term phytoplankton community dynamics in the Western English Channel. Journal of Plankton Research 32 (5), 643–655.

Young, P.C., 1998. Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. Environmental Modelling & Software 13, 105–122.

Young, P.C., 2000. Analysis of nonstationary climatic time series: central England and Antartic Air Temperature Series, Centre for Research on Environmental System and Statistics, CRES Report Number TR/182, Lancaster University.

Young, P.C., Parkinson, S., 2002. Simplicity out of complexity. In: Beck, M.B. (Ed.), Environmental Foresight Models: A Manifesto. Elsevier Science Ltd, United Kingdom, pp. 251–301.

Young, P.C., Ng, C.N., Lane, K.E., Parker, D., 1991. Recursive forecasting, smoothing and seasonal adjustment of non-stationary environmental data. Journal of Forecasting 10, 57–89.

Young, P.C., Pedregal, D.J., Tych, W., 1999. Dynamic Harmonic Regression. Journal of Forecasting 18, 369–394.

Zaldívar, J.M., Strozzi, F., Dueri, S., Marinov, D., Zbilut, J.P., 2008. Characterization of regime shifts in environmental time series with recurrence quantification analysis. Ecological Modelling 210 (1–2), 58–70.