Chapter 6

Statistical Modelling of Water Quality Time Series – The River Vouga Basin Case Study

Marco André da Silva Costa and
Magda Sofia Valério Monteiro

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/59814

Introduction

"Water is the principle of all things" Thales of Miletus.

Water quality models are essential tools for the assessment of the impact of ecosystem changes in response to variable inputs, as well as the interactions occurring within the system [1]. Taking this into account, developed countries have continually been monitoring and classifying their water. The over-abstraction of fresh water is a problem in certain parts of Europe, especially in coastal countries and islands in the Mediterranean, leading to the decrease of groundwater bodies, the deterioration of water quality and the destruction of certain habitats [2]. Water quality monitoring procedures may be used in decision-making processes for supporting policy options. For this reason, several European Union (EU) countries have developed a national water quality system considering the characteristic structures of their own rivers and have used these types of indicators to evaluate their current water quality levels. Water quality monitoring is usually conducted using physical and chemical parameters; nevertheless, biological factors such as algae, especially diatoms, benthic macro invertebrates and fish, have recently been used for this purpose [3]. Nowadays, the majority of the population in northern and central Europe is connected to wastewater treatment plants that make use of advanced treatments. This development allows for a decrease in the load of organic matter and nutrients transported by rivers and subsequently lead to an improvement of the state of water resources [4].

The management of water resources is regulated by EU directives and their transposition into national legislation. For instance, in Portugal, the Law nº58/2005 (Law of Water) ensures implementation into national law the Directive nº2000/60/CE (the Water Framework Directive,

WFD), which creates the institutional framework for the sustainable management of surface, interior, transitional, coastal and even groundwater. According to this directive, each member-state has to project, improve and recover all surface waters in order to achieve the good qualitative and quantitative status of all water bodies by 2015 [5,6]. The Decree-Law nº 77/2006 complements the WFD by characterizing the waters of a river basin. A regulatory instrument establishes the status of surface waters and groundwater and their ecological potential.

According to the directive's framework, the monitoring of water resources essentially has two purposes. The first is the evaluation of water status (surveillance monitoring) and the second is the implementation of programmes that include measures for identifying water resources at risk of failing environmental objectives (operational monitoring). If surveillance-monitoring reveals that the water in question has not reached the necessary "ecological status", a basin management plan should be implemented for the duration of one year. During this process, indicative parameters of all elements of biological, hydromorphological and general physico-chemical quality, as well as significant discharges in the basin containing pollutants, should be monitored in order to permit the classification of ecological status. In operational monitoring, measures are implemented and the status of water is evaluated as a result of the implementation of such measures [7].

In this context, every European country organizes itself using various entities that monitor water resources at various levels. The water area in Portugal consists of six large sets of entities: the trusteeship policy (national and regional), advisory bodies (national and regional), public water management (national and regional), utilities users and regulators (urban, agricultural, energy developments and multipurpose), mixed structures and associations and non-governmental entities.

The Water Portal website aims to respond to the provisions of Chapter VIII of the Water Law, especially its art. 87º, which empowers the National Water Authority – Institute of Water (Instituto da Água I.P.) to create a national information system for water and place it at the disposal "of... entities that have responsibilities [in] exercising public functions or [in] providing public services directly or indirectly related to water [to] the technical and scientific community and the public in general." [8]. The national water resources monitoring system is supported by a database that stores, prepares and publicize both hydro-meteorological and water quality (surface and groundwater) data, which have been collected by the monitoring network of the Ministry of the Environment (through the http://snirh.pt portal system). The National Information System for Water Resources (Sistema Nacional de Informação de Recursos Hídricos – SNIRH) was created by the Institute of Water in mid-1995. The monitoring network consists of automatic and conventional stations, some of which are equipped with remote transmission. The portal system also publishes monthly thematic syntheses aimed at characterizing the national water availability, technical reports, the mapping of water resources (e.g., flood zones) and maintaining technical documents and photographs related to water resources [9], following the significant economic support from European funds devoted to investments in sewage systems. Portugal has had slow growth in the area of wastewater treatment. The rate of the population connected to public wastewater treatment plants was about 56% in 2005 and 74% in 2009 (see Figure 1).

http://dx.doi.org/10.5772/59814

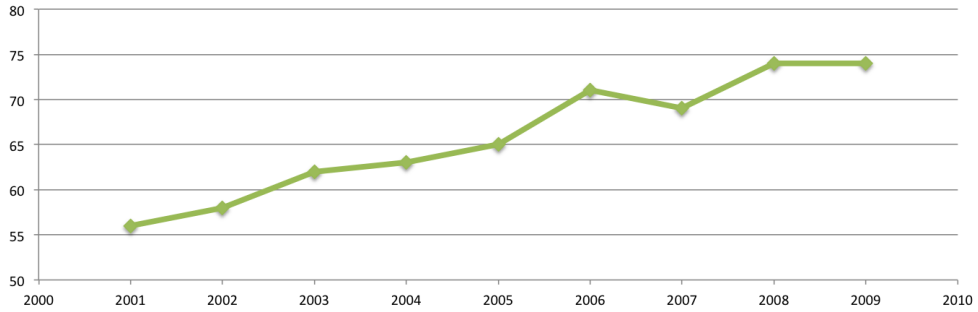

Figure 1. Evolution of the percentage of the population connected to public wastewater treatment plants in mainland Portugal between 2001 and 2009 [10].

Despite high investments at various levels of Portuguese public administration and those of economic actors, the economic crisis of recent years has led to a disinvestment in the monitoring and in the water treatment systems. Undoubtedly, the economic constraints of local authorities in Portugal have led to a reduction in investments in waste management, which may jeopardize the quality of surface water in several Portuguese hydrological basins (see Figure 2).

Figure 2. Investments in municipal waste management, in millions of euros, between 1993 and 2012 in Portugal [10].

Surface water-monitoring systems are important tools for analysing environmental processes and for identifying changes within these processes. In this context, statistical tools and methodologies are relevant for assessing water quality and for allowing the analysis of available data yielded by information systems. The temporal evolution analysis of water quality indicators is an important tool in this context, as the real-time monitoring procedures. Thus, modern statistical techniques are an integral part of any water monitoring process.

This paper presents a set of issues that analysts encounter in the analysis of data that support decision-making processes. The focus of this work is on contributions to the analysis of time series of water quality variables. The study of these time series has characteristics related to

data collection procedures, as well as to several economic and geographic aspects. Some characteristics of water quality variables pose challenges to their statistical modelling, namely, seasonality, change-point detection, the existence of outliers, etc. Some of these topics will be addressed in this chapter. The presentation of these topics is accompanied by the study of water quality data from the hydrological basin of the Vouga River and the Ria de Aveiro lagoon. In particular, statistical modelling is performed using data pertaining to the dissolved oxygen concentration variable (in *mg/l*) from the Carvoeiro water-monitoring site in the Vouga River basin. Section 2 discusses the area under study, while Section 3 presents a data description with an exploratory analysis of the primary monitoring sites of the Vouga basin. Section 4 addresses the analysis of a time series through the decomposition into its main components and their treatment, while Section 5 presents the modelling of time series through a state space model approach.

2. The study area

The Vouga River is situated at the centre of Portugal at an altitude of about 930m, near the geodesic landmark Facho da Lapa in Serra da Lapa, a mountain located in the district of Viseu. The river flows 148 km before empting into Ria de Aveiro. The watershed of the Vouga (also referred to as Vouga e Ribeiras Costeiras) is the second largest basin among the watercourses that run exclusively in Portuguese territory, comprising a total area of 3706 Km^2. More specifically, the Vouga basin is located in the transition zone between the north and south of Portugal, i.e., between the watersheds of the Douro in the north and the Mondego in the south (see Figure 3).

Several rivers flow into the Ria de Aveiro estuarine system. The hydrologic regime involves a summer low flow condition and the dynamics of the coastal lagoon are dominated by tidal oscillation. Ria de Aveiro is characterized by its rich biodiversity, as well as by increasing pressure related to anthropogenic activities near its margins, i.e., building and land occupation and agricultural and industrial activities. This has resulted in a significant change to the lagoon's morphology and in the constant input of a large volume of anthropogenic nutrients, as well as contaminant loads, which in turn has had a negative impact on water circulation and the water quality of the lagoon, [11]. The construction, management and operation of the multi-municipality system drainage of the Ria de Aveiro is the responsibility of SIMRIA – Integrated Sanitation of Municipalities of Ria, SA, a private company with a public capital majority (established by Decree-Law nº 101/97 of 26 April).

Several studies have been developed around this watershed and in particular, around the Ria de Aveiro lagoon. Some studies focus on the ecological systems related to Ria de Aveiro and the diversity of its flora and fauna (e.g., [12, 13]. Other studies have examined the biochemical properties of the region's environmental systems. Still others have studied the hydro-meteorological aspects of the watershed. This diversity and multidisciplinary approach to the Ria de Aveiro lagoon and its associated watershed was recently formalized by the University of Aveiro as '*Grupo uariadeaveiro*'. The group aims to monitor and contribute to the management

http://dx.doi.org/10.5772/59814

Figure 3. Hydrological basin of mainland Portugal [9].

of the Ria de Aveiro by focusing mainly on knowledge transfer activities; by doing so, it can also contribute to the enhancement of other activities within the University of Aveiro. Additionally, special attention to the activities of the primary entities and relevant stakeholders associated with the management of the Ria de Aveiro lagoon can in this way be maintained in

order to ensure the useful and effective contribution of the University of Aveiro. The University of Aveiro, through its *uariadeaveiro* group, is committed to facilitating the provision of scientific knowledge about this lagoon area and places it at the service of the region. Thus, the University of Aveiro hopes to contribute to the improvement of the protection, enhancement and management of the Ria de Aveiro.

The website http://www.ua.pt/riadeaveiro/ (see Figure 4) was established as a mechanism for the dissemination of knowledge and activities of interest pertaining to the Ria de Aveiro. This work in progress aims to establish itself as the most important collection of data about the Ria de Aveiro in the country. It is hoped that this initiative will help to create internal synergies within the scientific community focused on the Ria de Aveiro and also to strengthen bridges of communication with regional stakeholders in order to maximize the successful management of the Ria.

Figure 4. The website http://www.ua.pt/riadeaveiro/ located within the University of Aveiro site (http://www.ua.pt).

The webpage of the *uariadeaveiro* group includes five menus. The first is dedicated to the presentation of the group and its goals. The second is dedicated to the library, where access can be gained to all scientific and non-scientific publications regarding the Ria de Aveiro. The third page provides spatial information on the Ria de Aveiro and allows for access to other pages with geographic information of interest regarding the Ria systems. The fourth page is devoted to the approximation entities and users of the Ria. The fifth page discusses research projects completed and ongoing, relevant training courses and activities related to the Ria de Aveiro [14].

The average fresh water flow into the Ria de Aveiro is about 40 m^3/s. The Vouga and Antuã rivers are the main sources of fresh water, with an average annual flow of 24 m^3/s and 2.4 m^3/s; both of these rivers are part of the Vouga watershed, [15].

The main tributaries of the Vouga River are, from upstream to downstream, the River Mel, the Sul River, the Varoso, the River Teixeira, the River Arões, the River Mau and the Caima River on the right bank; on its left bank, the River Ribamá, the MarneI and the River Águeda with its major tributary, the Alfusqueiro. Officially, the catchment of the River Vouga consists in the rivers that dewater into the Ria de Aveiro with the exception of Vala da Fervença, in the Mira area, which flows to the southern end of the Ria de Aveiro [16].

A set of water monitoring sites (see Figure 5), where several variables pertaining to water quality can be collected, is available in the hydrological basin of the Vouga River. However, some problems have arisen in the context of statistical modelling, i.e., some water monitoring

http://dx.doi.org/10.5772/59814

sites and variables yield little data and/or have missing values. Furthermore, due to a lack of economic resources and other factors, data collection was discontinued at some sites.

Within the SNIRH system, 78 water-monitoring sites are registered within the hydrological basin of the Vouga River. However, data collection is not continuous and some stations had been deactivated somewhere in time. Relative to the DO concentration, 26 stations showed a significant data set, with data up to May 2013 (the last month available in the system). Hence, the statistical analysis will be focus mainly on this data set.

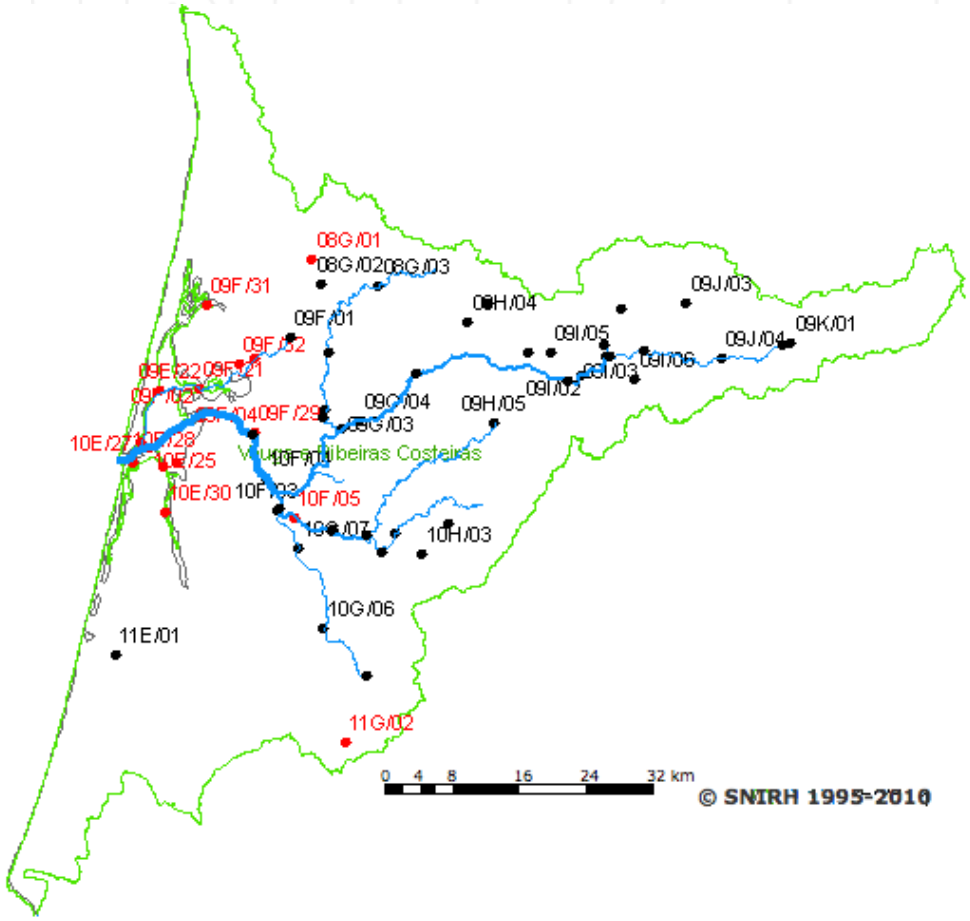

Figure 5. Water monitoring sites locations (red – inactive in the present; black – actives in the present; green – hydrological basin geographical boundaries), [9].

3. Data set description

The statistical analysis was performed using a data set related to the dissolved oxygen concentration in a large set of water monitoring sites located in the hydrological basin of the River Vouga, a basin associated with the Ria de Aveiro lagoon. The analysis focused on the DO concentration because it is one of the most important variables in the evaluation of a river's water quality [6] and because of its continuity in terms of measurements taken at all selected water quality monitoring sites under analysis.

The DO variable is widely recognized as one of the most relevant for the monitoring of water quality. DO concentration is controlled by several physical processes such as the rate of production of oxygen by algae, the nitrogen cycle and other biochemical processes [17]. At the same time, available oxygen is the primary factor affecting aquatic life in rivers. A good level of dissolved oxygen is essential for aquatic life. Oxygen concentration is therefore, the most important parameter for aquatic life and ecological states in estuaries [11]. Dissolved oxygen analysis measures the amount of gaseous oxygen dissolved in an aqueous solution. Oxygen gets into water by diffusion from the surrounding air, by aeration and as a waste product of photosynthesis. Adequate dissolved oxygen is necessary for good water quality and is a necessary element to all forms of life. Natural stream purification processes require adequate oxygen levels in order to provide for aerobic life forms. When dissolved oxygen levels in water drop below 5 mg/l, aquatic life is put under stress. With a decrease in oxygen concentration the stress on life forms increases. Oxygen levels that remain below 1-2 mg/l for a few hours can result in large numbers of fish deaths [18].

The SNIRH system holds data regarding the DO concentration of 78 water-monitoring sites between April 1989 and May 2013. However, among them several sites hold little or exhibit several missing values, which are associated with large gaps between measurements. Thus, the statistical descriptive analysis will be performed primarily on the largest time series. A significant number of locations had less than 23 observations for all time periods. These time series will not be considered in the descriptive analysis. On the one hand, only 26 water-monitoring sites had at least 100 observations and these will be considered in the exploratory analysis. On the other hand, only between January 2002 and May 2013 are consistent data sets available for all locations. Thus, the main data set that will be considered consist of these 26 time series during this period.

Data collection was not equally distributed during all time periods. Sometimes there was more than one measurement in a month. Therefore, in order to not lose information, we took the average of all measurements if there was more than one observation in a month. Table 1 presents descriptive statistics for the 26 time series between January 2002 and May 2013 (137 months) and Figure 6 shows the 26 time series representations. The graphical representation clearly shows that time series exhibited seasonal behaviour, as was expected due to the nature of the data.

http://dx.doi.org/10.5772/59814

Site	Code	Abbrev.	N. obs.	Min	Max	Average	St dev	Skew	Kurtosis
Agadão	10H/03	AGA	111	5.8	11.0	8.74	1.26	-.35	-.25
Carvoeiro	09G/03	CAR	112	6.2	11.0	8.79	1.18	-.15	-.40
Alombada	09G/04	ALO	113	6.1	11.0	8.90	1.08	-.51	.41
Captação Burgães	08G/03	BUR	122	6.5	12.6	9.40	1.16	-.15	-.18
Captação Rio Ínsua	08G/02	INS	122	6.4	12.4	9.31	1.05	-.33	.65
Ponte Redonda	10G/05	RED	112	4.6	11.5	8.88	1.22	-.47	.70
Frossos	09F/04	FRO	110	4.5	11.0	8.17	1.22	-.16	-.24
Pampilhosa	11G/02	PAM	100	4.3	12.0	7.95	1.68	.03	-.35
Ponte São João de Loure	10F/04	LOU	112	5.4	11.0	8.24	1.25	-.02	.01
Ponte Vale Maior	09G/01	MAI	112	6.2	12.0	8.62	1.12	-.05	-.05
Ponte Águeda	10G/02	AGU	111	5.1	11.0	8.39	1.20	-.34	-.18
São Tomé	11E/01	TOM	118	5.0	11.0	7.88	1.16	-.07	-.37
Aç. Maeira	09K/01	MAE	115	5.6	11.0	8.50	1.20	-.16	-.41
Aç. Rio Alfusqueiro	09H/05	ALF	113	2.9	12.0	7.80	1.75	-.31	-.01
Pindelo Milagres	09J/03	MIL	110	4.6	12.0	8.16	1.42	-.14	-.09
Ponte Antim	09I/05	ANT	113	0.8	12.0	7.38	2.05	-.68	.14
Ponte Pouves	09I/03	POU	115	2.6	11.0	8.27	1.46	-.67	1.30
Ponte Vouzela	09I/02	VOZ	109	1.8	13.0	8.10	1.91	-.84	1.32
São João Serra	09H/04	SER	115	6.0	12.0	8.70	1.18	.01	-.08
São Miguel Mato	09I/06	MAT	111	4.3	12.0	8.44	1.53	-.41	.03
Vouguinha	09J/04	VOG	114	5.4	11.0	8.42	1.35	-.26	-.52
Estarreja	09F/05	EST	114	3.4	11.0	7.62	1.32	-.71	1.02
Perrães	10G/07	PER	111	4.6	9.8	7.19	1.17	.14	-.53
Ponte Canha (Vouga)	10G/06	CAN	114	2.6	10.1	6.89	1.92	-.26	-.96
Ponte Minhoteira	09F/01	MIN	112	0.7	10.0	7.73	1.50	-1.10	3.60
Ponte Requeixo	10F/03	REQ	111	3.9	11.0	7.13	1.52	.33	-.09

Table 1. Descriptive statistics for dissolved oxygen concentration between January 2002 and May 2013.

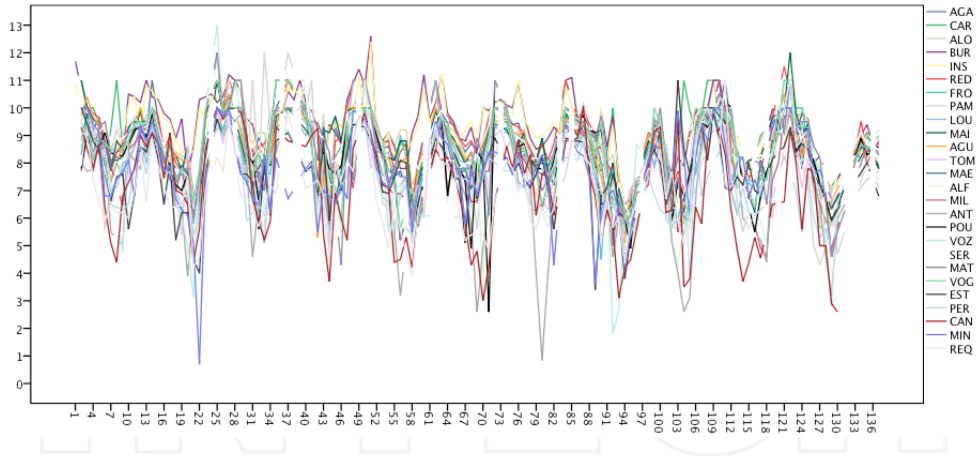

Figure 6. Time series for dissolved oxygen concentration in the 26 water-monitoring sites from January 2002 to May 2013.

An exploratory analysis showed that, in general, observations were not normally distributed. Indeed, in some locations, the distribution of observations is skewed or leptokurtic. This fact must be considered in the modelling procedures, since Gaussian distribution is a usual assumption in several statistical analyses. Figure 7 represents box-plots for DO concentrations. Boxplots are able to identify several moderate outliers in many locations. However, due to existence of temporal correlation structure in the measurements, the analyses of outliers must be carefully considered.

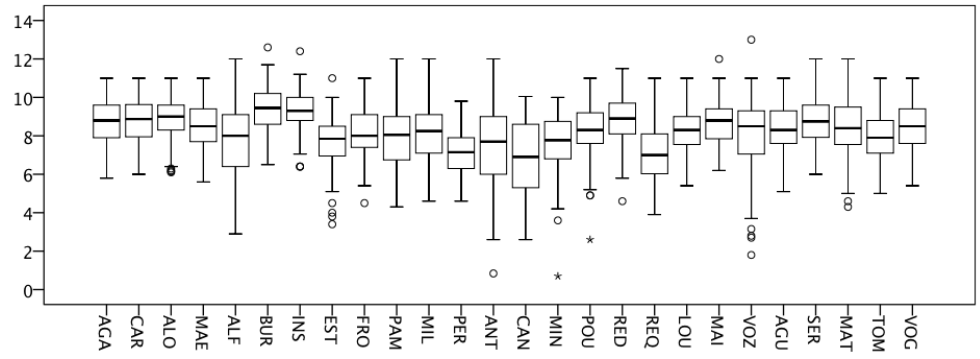

Figure 7. Boxplots for dissolved oxygen concentration from January 2002 to May 2013.

As DO concentration exhibits seasonal behaviour, it is necessary to analyse the monthly averages. Figure 8 shows the monthly averages of DO concentration during the analysed period for the 26 water-monitoring sites. These results indicated that DO concentration was

http://dx.doi.org/10.5772/59814

greater in the winter months and presented lower values in the summer months. This fact is directly related to hydro-meteorological conditions, since DO concentration is influenced primarily by precipitation amounts and temperature. The monthly standard deviations presented in Figure 9 exhibit a diffuse pattern. Although it is difficult to identify a common pattern for all sites, a graphic representation indicates that in several locations there was larger variability during the months of January, September, October and November. The lower variability of the monthly DO concentration measures in December may be justified by the reduced observations typically collected during this month.

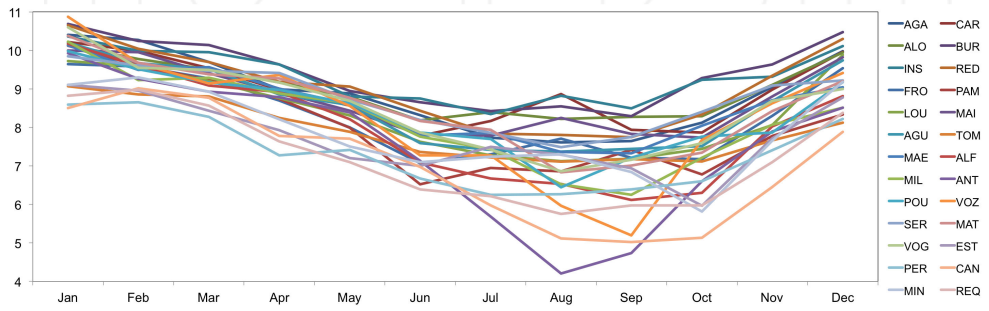

Figure 8. Monthly averages of DO concentration for the 26 water-monitoring sites between January 2002 and May 2013.

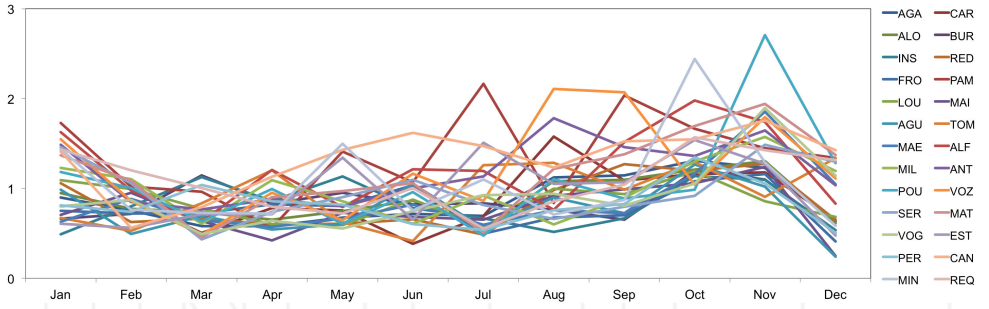

Figure 9. Monthly standard deviations in DO concentration for the 26 water-monitoring sites between January 2002 and May 2013.

4. Time series analysis

This section addresses several issues in the statistical modelling of time series water quality variables. Water quality variables present characteristics that pose some challenges to model-

ling procedures. By their nature, these variables have properties that must be incorporated in order for statistical techniques to yield good performance.

There are different ways of dealing with these features based on different approaches. One way of modelling a time series with components as either a trend or a seasonal/stochastic behaviour is by employing the time series decomposition method. This method consists of sequentially and individually modelling each component so that trends and/or seasonal components can be estimated and subtracted from the data to generate a noise sequence.

In the present work, we present, from a practical point of view, the most common methodologies applied to the time series of DO concentration in the Carvoeiro water-monitoring site. The choice of this monitoring site is related to its location in the main river, in the middle section, near to a point of abstraction of water for human consumption. Furthermore, the time series data in this location is long – the data is available between April 1989 and September 2012 – and it will used to apply the statistical procedures.

4.1. Heteroscedasticity and trend modelling

Generally, environmental time series can be non-stationary in a variety of ways. The most common case occurs when data present a trend and/or heteroscedasticity. When data has variance in terms of time, i.e., $var(Y_t)=\sigma_t^2$, the most commonly applied procedure is the application of data transformation.

In several areas there are models that accommodate heteroscedasticity within their structure. For example, in econometrics, ARCH models are commonly employed in the modelling of financial time series that exhibit time-varying volatility. When an autoregressive moving average model (ARMA) is assumed for the error variance, the model is a generalized autoregressive conditional heteroscedasticity [19] model.

Heteroscedasticity in an environmental time series is commonly treated using the Box-Cox transformation [20]

$$y_\lambda^* = \frac{y^\lambda - 1}{\lambda}$$

which is indexed by λ, where y_λ^* is the transformation of the original observation y. Note that for $\lambda=0$, the transformation is the natural logarithm of y. On the one hand, the Box-Cox transformation is not suitable for all heteroscedasticity types. On the other hand, this transformation, among others, changes the data magnitude hindering the interpretation of the model.

As will be shown later, this work considers a class of models (state space models) that is able to incorporate some types of heteroscedasticity and does not require data transformation. A time series frequently has a trend, i.e., it is not an observed process with a constant mean. In general, the trend is modelled using a deterministic function $f(t)$, usually a polynomial

http://dx.doi.org/10.5772/59814

function, power function or logarithmic function. The adjustment of the function $f(t)$ to data is generally performed using the least square method:

$$\min \sum_t \left[y_t - f(t) \right]^2 .$$

Nevertheless, the most common function is of the type $f(t)=\alpha+\beta t$. However, the simplicity of this function is not always appropriate to model real data, see for instance Figure 10. The DO concentration data for the Carvoeiro water-monitoring site represented in Figure 10 corresponds to an extended period of time. The graphical representation clearly indicates that the time series is not stationary in mean and the variability does not follow a monotonic function in time. Furthermore, there is no unique function among those referred above that globally fit the time series in order to model the trend. Thus, we can identify a trend function for the time series defined by sub-periods of time.

Figure 10 suggests that there is a time where the trend function changes its expression, possibly with a non-linear function prior to that time and a linear form following it. For simplicity and according to the exploratory analysis, we considered the most suitable function to the first part of the series to be a quadratic function and for the second part, a constant.

After fitting a regression model with a quadratic trend up to a certain time t_0, to all possible times t_0 and a constant after that instant, we chose the value of t_0 with the smallest residual sum of squares [21]. This procedure led to $t_0 = 139$, i.e., October 2000 and the estimated trend was:

$$\hat{T}_t = \begin{cases} -0.0014t^2 + 0.01986t + 6.8991 & \text{if} \quad t \le 139 \\ 9.128 & \text{if} \quad t > 139. \end{cases}$$

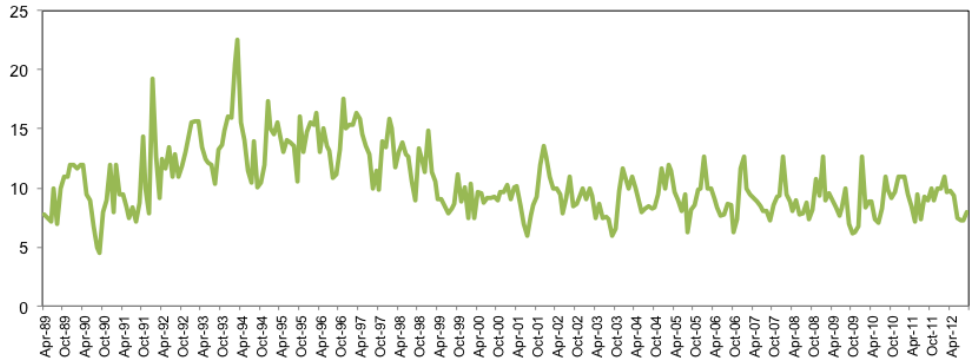

Figure 10. Monthly DO concentration in the Carvoeiro water-monitoring site from April 1989 to September 2012.

Figure 11 presents the series resulting from the subtraction of the adjusted trend to the original data, $Y_t - \hat{T}_t$. As is shown, this new series does not present a deterministic trend; however, as is the norm for monthly environmental data, seasonal behaviour is present.

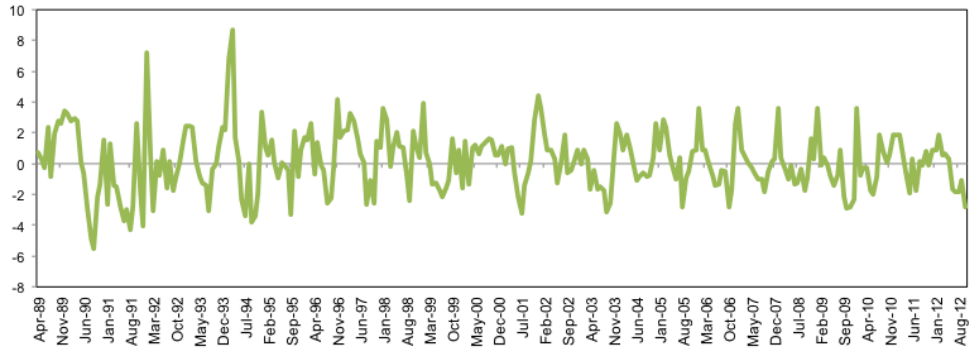

Figure 11. Residual series from the trend adjustment to the Carvoeiro DO concentration time series.

The easiest way to handle seasonality is to subtract from January's data the overall January average, from February's data the overall February average, etc. [22]. An alternative method involves estimating both trend and seasonality coefficients together. In this case, it is necessary to define a set of dummy variables as,

$$x_{ti} = \begin{cases} 1 & \text{if } t = (j-1)P + i, \text{for some } j = 1,\ldots,n \\ 0, & \text{otherwise} \end{cases}$$

where P is the period (in our case $P=12$). If the trend component is polynomial, as previously considered, the model shows unique representation through a linear regression model. As it is generally appropriate that the sum of the seasonal coefficients must be zero, the linear model should be re-parameterized in order to overcome the linear dependence of variables x_{ti}. Thus, the model can be presented as:

$$Y_t = \beta_1 t + \beta_2 t^2 + \ldots + \beta_q t^q + \sum_{i=1}^{p} x_{ti} s_i + \varepsilon_t$$

where ε_t is white noise. The vector of coefficients $(\beta_1, \beta_2, \ldots, \beta_q, s_1, s_2, \ldots, s_P)$ is estimated by the least squares method. In order to facilitate the interpretation of results, we can consider the model with a constant parameter β_0, which leads to the following transformations:

http://dx.doi.org/10.5772/59814

$$\beta_0 = \overline{s} \; and \; s_i^* = s_i - \overline{s}, i = 1, 2, ..., P$$

In general, these two approaches produce similar estimates. The first method is easier to implement and less laborious than the second. Furthermore, the regression model provides optimal estimates in certain conditions that assume that the errors ε_t are not time correlated, an assumption that is not valid for the data in the current analysis. Hence, for simplicity and considering the data characteristics, we adopted the first method, that is, the decomposition procedure that estimates trends and seasonal separately. Table 2 presents the estimates for the seasonal coefficients of the Carvoeiro data after the subtraction of the adjusted trend component.

Month	Seasonal coefficient
Jan	2.23
Feb	1.50
Mar	1.03
Apr	0.31
May	-0.24
Jun	-0.92
Jul	-1.01
Aug	-1.05
Sep	-1.40
Oct	-0.41
Nov	0.31
Dec	1.24

Table 2. Estimates for seasonal coefficients.

As expected, from the DO perspective, the water quality was better in the winter months and worse during the summer months. The DO concentration is associated with weather conditions, particularly with precipitation amounts [23].

The Box-Cox transformation is associated, for example, with multiplicative models where the time series components are multiplicative as $Y_t = T_t \cdot S_t \cdot \varepsilon_t$. However, there are also other types of heterocedasticity. Figure 12 shows for each month both the overall monthly averages and standard deviations of the DO concentration in Carvoeiro. This analysis indicated that variability had not been constant throughout the year. Indeed, the linear correlation coefficient between averages and standard deviations was 0.44, which showed that when the monthly average was large, the standard deviation tended to be large as well. This type of heterosce-

dasticity will be taken into account in the modelling procedure via the state space model approach.

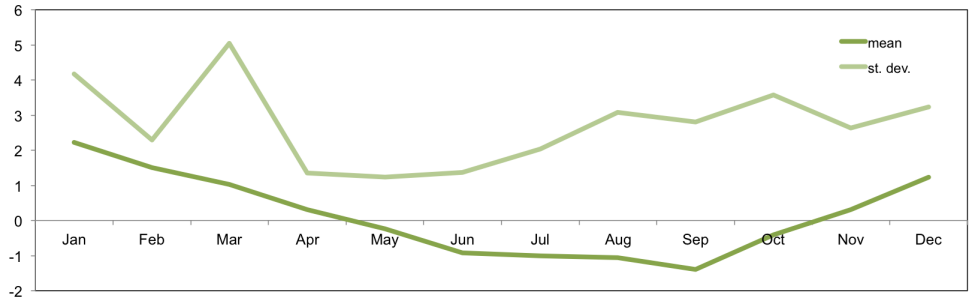

Figure 12. Monthly averages and standard deviation of DO concentration, subtracted according to trend, in the Carvoeiro water-monitoring site from April 1989 to September 2012.

4.2. Temporal dependence

It is well known that there exists a certain level of persistence in the behaviour of nature [22]. Indeed, environmental data is influenced by meteorological conditions that may persist from one month to another [21]. The series of residuals pertaining to the DO concentration in Carvoeiro does not have the characteristics of white noise. Figure 12 shows that there is a significant temporal autocorrelation according to an autoregressive process of order 1, an AR(1). Thus, stochastic modelling of residuals is needed in order to accommodate the serial correlation.

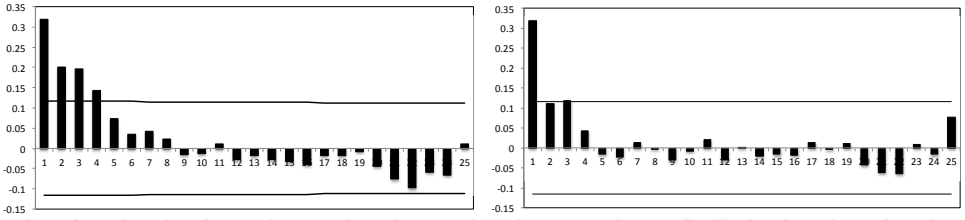

Figure 13. Sample autocorrelation function (ACF – left) and sample partial autocorrelation function (PACF – right) of the residuals.

When modelling a time series using the decomposition method approach, the procedure that is generally adopted for finding the best fit of an ARIMA model to past values of a time series is the Box-Jenkins methodology (for more details, see [24]. The modelling of the dependence of environmental data significantly influences decisions in the application of other modelling statistical procedures such as, for example, the existence of non-stationar-

ities in a series or in change point detection procedures. Therefore, in section 5, we consider a class of models that allows for incorporating several components as the temporal dependence in a dynamic manner and which will be used to model the DO concentration in the water-monitoring of Carvoeiro.

5. Time series modelling

State space models are a class of extremely versatile time series models. Their popularity is mainly due to the Kalman filter. This algorithm, [25], has been used in many different areas to describe dynamic systems evolution. The primary goal of the Kalman filter algorithm is to find estimates of unobservable variables based on observable variables related through a set of equations designated by a state space model (SSM).

In general, such models are defined by the following equations:

$$Y_t = H_t \beta_t + e_t$$
$$\beta_t = \Phi \beta_{t-1} + \varepsilon_t$$

The first equation is the measurement equation and relates the $n \times 1$ vector of observable variables, Y_t, with the $m \times 1$ vector of unobservable variables, β_t, referred to as *states*. The $n \times m$ matrix H_t is a matrix of known coefficients and e_t is a white noise $n \times 1$ vector, called the measurement error, with covariance matrix $E(e_t e_t') = \Sigma_e$.

Furthermore, the vector of states β_t varies in time according to the second equation, the state equation, where Φ is a $m \times m$ matrix of autoregressive coefficients and ε_t is a white noise $m \times 1$ vector with covariance matrix $E(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon$. The disturbances e_t and ε_t are assumed to be uncorrelated, that is, $E(e_t \varepsilon_s') = 0$ for all t and s.

A subclass of these models with a particular interest arises when the state vector is a stationary process with mean μ, $E(\beta_t) = \mu$. In this case, to ensure that the state process and the state equation follows a stationary VAR(1) with a mean μ, it is assumed that the eigenvalues of the autoregressive matrix Φ are inside the unit circle. For more details, see [26].

5.1. A brief description of the Kalman filter algorithm

As the state β_t is unobservable, it is necessary to obtain its predictions. The Kalman filter algorithm uses an orthogonal projection of the state based on the available information up to the moment. Assuming that the parameters of a state space models are known, the Kalman filter recursions provide the best linear predictors for filtering and forecasting the vector of states.

Let $\hat{\beta}_{t|t-1}$ represent the predictor of β_t based on the information up to time $t-1$ and $P_{t|t-1}$ be its mean square error (MSE). As the orthogonal projection is a linear estimator, the predictor for the observable variable, Y_t, is given by

$$\hat{Y}_t = H_t \hat{\beta}_{t|t-1}.$$

When at time t, Y_t is available, the prediction error or *innovation*, $\eta_t = Y_t - \hat{Y}_t$, is used to update the prediction of β_t, the filter prediction, through the equation

$$\hat{\beta}_{t|t} = \hat{\beta}_{t|t-1} + K_t \eta_t$$

where $K_t = P_{t|t-1} H_t' (H_t P_{t|t-1} H_t' + \Sigma_e)^{-1}$ is called the Kalman gain matrix. Furthermore, the MSE of the update predictor $\hat{\beta}_{t|t}$ verifies the relation $P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1}$. In turn, at time t, the forecast for the state vector β_{t+1} is given by the equation

$$\hat{\beta}_{t+1|t} = \Phi \hat{\beta}_{t|t}$$

with MSE matrix $P_{t+1|t} = \Phi P_{t|t} \Phi' + \Sigma_\varepsilon$.

When the disturbances are Gaussian, the Kalman filter provides the minimum mean square estimate for β_t.

5.2. Parameter estimation

The application of the Kalman filter to predict the unknown variables requires the estimation of the model's parameters.

The vector of parameters $\Theta = \{\Phi, \Sigma_e, \Sigma_\varepsilon\}$ and the mean vector μ if the state is stationary must be estimated from the data. When the Gaussian distribution is suitable to the disturbances e_t and ε_t the conditional log-likelihood of a sample $Y_1, Y_2 \ldots, Y_n$ is given by:

$$\begin{aligned} \ln L(\Theta; Y_1, Y_2, \ldots, Y_n) &= \sum_{t=1}^{n} \ln f_{\hat{Y}_{t|t-1}} \left(Y_t \mid Y_1, Y_2, \ldots, Y_{t-1} \right) \\ &= -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{n} \ln\left(|\Omega_t| \right) - \frac{1}{2} \sum_{t=1}^{n} \eta_t' \Omega_t^{-1} \eta_t \end{aligned}$$

where $\Omega_t = H_t P_{t|t-1} H_t' + \Sigma_e$. The maximum likelihood estimates are obtained by maximizing the log-likelihood, i.e.,

$$\hat{\Theta}_{ML} = \arg\max \ln L(\Theta; Y_1, Y_2, \ldots, Y_n).$$

http://dx.doi.org/10.5772/59814

The state space model's parameters are estimated by the maximum likelihood via numerical procedures.,The Newton-Raphson method can be adopted or, more often, it is employed the EM algorithm [24].

If the Gaussian distribution is not appropriate or the optimal properties are not required in the modelling procedure, other estimators can be considered. For example, distribution-free estimators can be adopted, which do not assume any distribution for the disturbances (see [27, 23]).

5.3. A mixed-effect state space model

In order to incorporate the temporal correlation and to accommodate the heterogeneity of the variances of the sub-series of each month in terms of DO concentration in Carvoeiro, we adopted a mixed-effect state-space model. The model is defined as:

$$\begin{aligned} Y_t &= \left(T_t + S_t\right)\beta_t + e_t \\ \beta_t &= \mu + \phi\left(\beta_{t-1} - \mu\right) + \varepsilon_t \end{aligned}$$

where T_t is the trend component, S_t is a periodic function with 12 averages as previously mentioned, the process $\{\beta_t\}$ is a stationary AR(1) with mean μ and Gaussian errors ε_t and e_t is the measurement error, assumed to be a Gaussian white noise process.

The model has two primary components: a regression structure, which incorporates both the trend and the seasonality, and an unobservable process $\{\beta_t\}$, the state. It is assumed that the state process will calibrate the deterministic structure $T_t + S_t$ previously estimated by $\hat{T}_t$ and $\hat{S}_t$ in presented in Table 2.

Another important feature of the model is that by its own formulation, it allows the existence of heterogeneity of variances previously identified. That is, the stochastic calibrator factor allows for the dynamic modelling of the heteroscedasticity during the year and incorporates the time dependence identified through the ACF and PACF functions in Figure 12.

Table 3 presents the maximum likelihood estimates of $\Theta=\{\mu, \phi, \sigma_\varepsilon^2, \sigma_e^2\}$ and their respective standard errors. As expected, the mean of the calibration factor was close to 1. This meant that the deterministic component $T_t + S_t$ represented the global behaviour of data. On the other hand, the calibration factor followed a stationary autoregressive process AR(1) process since $|\hat{\phi}| < 1$.

$\hat{\mu}$	$\hat{\phi}$	$\hat{\sigma}_\varepsilon^2$	$\hat{\sigma}_e^2$
1.00077	0.43001	0.01790	0.24292
(0.00085)	(0.00503)	(0.00024)	(0.01854)

Table 3. Maximum likelihood estimates of the mixed-effect state space model (standard errors in brackets).

Once the model had been well identified, the Kalman filter was run in order to obtain the forecasts and filtered predictions of the states β_t. The Kalman filter provided predictions to the states and their respective MSE estimates. Figure 13 represents the filtered predictions $\hat{\beta}_{t|t}$ of the calibration factors for each month and their respective empirical confidence intervals at a 95% level obtained by:

$$\beta_t = \hat{\beta}_{t|t} \pm 1.96\sqrt{\hat{P}_{t|t}}.$$

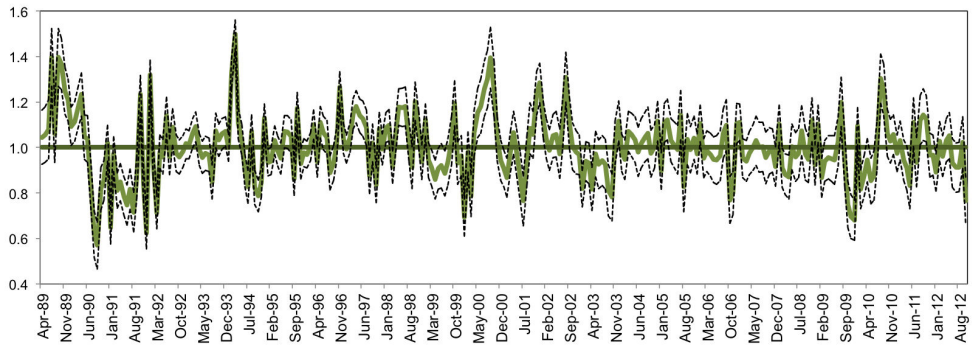

Figure 14. Representation of filtered predictions $\hat{\beta}_{t|t}$ and their respective empirical confidence intervals.

In order to assess the adjustment of the model, one step-ahead forecasts were computed with respective confidence intervals at 95% (see Figure 14):

$$Y_t = \hat{Y}_{t|t-1} \pm 1.96\sqrt{\widehat{MSE}_{t|t-1}}.$$

where $\widehat{MSE}_{t|t-1} = (\hat{T}_t + \hat{S}_t)^2 \hat{P}_{t|t-1} + \hat{\sigma}_e^2$.

The percentage of observations outside of the respective empirical confidence interval was 6.36%. The residuals analysis showed that there were some outliers; the biggest residual occurred in January 1992.

Normality was rejected at a 5% significant level, considering the Kolmogorov-Smirnov test, since the p-value was 2.8%. However, the Gaussian distribution was not rejected (K-S p-value=20.0%) when the largest outliers were not considered. As state space models are associated to the Kalman filter, in general, the impact of the outliers turns out not to be significant, since they are attenuated through the filtering procedure.

As can be seen in Figure 14, a large proportion of the observations outside the respective empirical confidence interval belonged to the first years of the series. This may indicate the

http://dx.doi.org/10.5772/59814

existence of different structures in the observable variables, which can be found through the analysis of change points in the state variables. This issue is discussed in the following section.

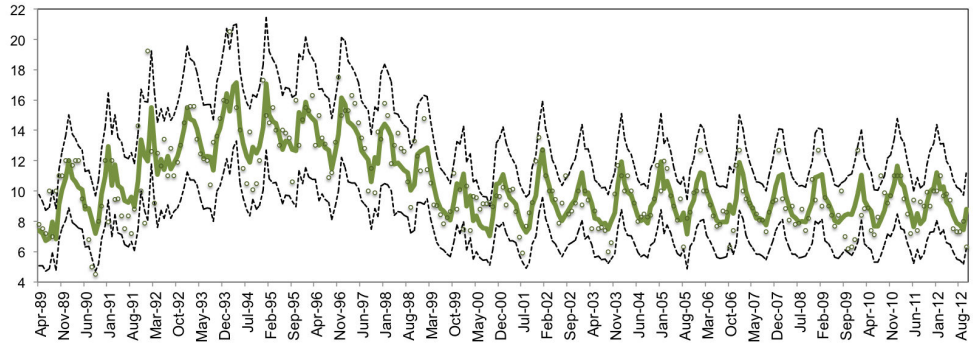

Figure 15. Dissolved oxygen concentration and the step-ahead forecast with respective confidence intervals at 95%.

5.4. Change point detection

The detection of structural changes in environmental data is one of the most interest issues in the applied statistical analysis. In the past decade, there has been significant progress within change point detection. However, these methodologies still require more complex statistical procedures or computational requirements. For example, in [28], a simple BIC-like multiple change point penalty is proposed that is based on the total number of change points. For further recent developments regarding change point detection in environmental applications, [29,30].

The state space modelling approach adopted in this work allows for applying standard change point detection procedures, since this model incorporates several data characteristics. Thus, as these properties of environmental data are taking into account in the modelling procedure, simple change-point detection techniques can be applied if adequately chosen.

This subsection deals with change point detection in the state process instead of the observed process Y_t. This detection is made through the filtered predictions of the calibration factors, $\hat{\beta}_{t|t}$, obtained in the filter procedure. It is reasonable to investigate changes in the structure of β_t that may be relevant to the water-monitoring process. For this purpose, the filtered predictions of the state process can be analysed in relation to the existence of one change (or more) in its mean as the best prediction of an AR(1) process.

To do so, we used maximum type test statistics, referred in [31], alongside their correct adaptation to an autoregressive process. The basic test evaluates the existence of a change point in the mean of independent and identically distributed Gaussian variables in an unknown time.

The hypotheses of the basic test are:

$H_0: \beta_1, \beta_2, \ldots, \beta_n$ are independent and distributed according to the same Gaussian distribution $N(\mu, \sigma^2)$

vs.

$H_0: \exists k \in \{1, \ldots, n-1\}: \beta_1, \ldots, \beta_k$ are distributed according to $N(\mu_1, \sigma^2)$ and $\beta_{k+1}, \ldots, \beta_n$ are distributed according to $N(\mu_2, \sigma^2)$, with $\mu_1 \neq \mu_2$.

The test for the existence of the change point k can be performed by a statistic test known as the maximum type. Since the variance is generally unknown, this approach consists of a sequence of t tests and the analysis of the significance of their maximums. Thus, the test statistic is given by:

$$T_n = \max_{1 \leq k < n} |T_k| = \max_{1 \leq k < n} \sqrt{\frac{(n-k)k}{n}} \frac{\left|\overline{\beta}_{k|k} - \overline{\beta}^*_{k|k}\right|}{s_k}$$

where

$$\overline{\beta}_{k|k} = \frac{1}{k}\sum_{i=1}^{k}\hat{\beta}_{i|i}, \overline{\beta}^*_{k|k} = \frac{1}{n-k}\sum_{i=k+1}^{n}\hat{\beta}_{i|i} \text{and } s_k = \sqrt{\frac{1}{n-2}\left[\sum_{i=1}^{k}\left(\hat{\beta}_{i|i} - \overline{\beta}_{k|k}\right)^2 + \sum_{i=k+1}^{n}\left(\hat{\beta}_{i|i} - \overline{\beta}^*_{k|k}\right)^2\right]}.$$

The null hypothesis is rejected if the statistics T_n are larger than a corresponding critical value. However, the calculation of the exact critical value is not easy, due to the test statistic being the maximum of the sequence $T_1, T_2, \ldots, T_{n-1}$. Nevertheless, approximated critical values can be obtained by different methods, namely:

i. the Bonferroni inequality and its improvement;

ii. asymptotic distribution;

iii. through simulation.

Table 4 replicates the approximated critical values obtained through simulation and presented in [22]. However, these critical values assume that the observations are uncorrelated and according to [32], it is necessary to correct them. Thus, in order to obtain the corrected critical value that takes into account the time correlation of an AR(1), the initial critical value must be multiplied by $[(1+\hat{\phi})/(1-\hat{\phi})]^{1/2}$.

The simplest method for detecting multiple change points is through binary segmentation.

The first work to propose binary segmentation within a stochastic process setting was [33]. Binary segmentation is a generic technique for multiple change-point detection in which, initially, the entire dataset is searched for one change-point, typically using a CUSUM-like procedure. When a change-point is detected, the data are split into two (hence, 'binary') sub-

http://dx.doi.org/10.5772/59814

segments, defined by the detected change-point. A similar search is then performed on each sub-segment, possibly resulting in further splits. The recursion on a given segment continues until a certain criterion is satisfied [34].

n	5% critical value	1% critical value
50	3.15	3.76
100	3.16	3.71
200	3.19	3.72
300	3.21	3.73
500	3.24	3.73

Table 4. Approximated critical values obtained by simulation.

In order to apply the change point detection procedure to the state filtered prediction of Carvoeiro, it was necessary to interpolate the critical value at 5% to $n=283$, which is 3.207. Taking into account the autoregressive parameter estimate presented in Table 3, the corrected critical value at 5% was noted as 5.079. Figure 15 represents the observed values of the statistics T_k, $1 \leq k < 283$, as well as both corrected and uncorrected critical values at 5% in the first step of the binary segmentation procedure. As can be seen, the null hypothesis is rejected, i.e., a statistically significant change point that can be identified as April 1990.

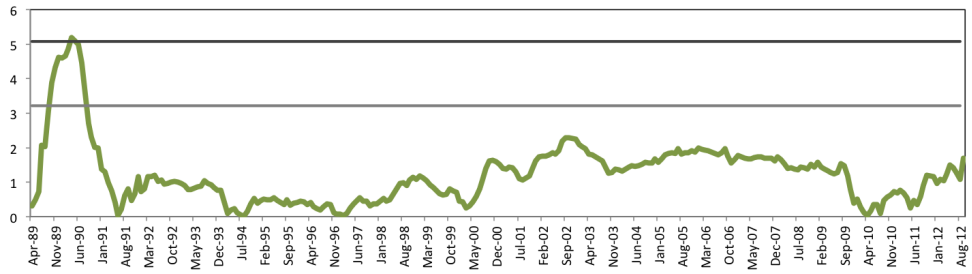

Figure 16. Observed values of the statistic T_k, $1 \leq k < 283$; horizontal lines represent both critical values at 5% corrected and uncorrected to an AR(1).

Table 5 presents the results from the complete binary segmentation procedure, which ended when the observed value of the test statistics was less than the respective corrected critical value (the critical value was obtained by interpolation according to the sub-sample size). This procedure identified two change points: April 1990 and December 1991. Note that in the final step of the binary segmentation procedure, the largest value of the statistic tests was 3.202, which would lead to the detection of another change point if the correction to the AR(1) process had not been considered, since the uncorrected critical value was 3.200. However, considering

the AR(1) correction, the critical value was 5.069, i.e., the observed statistic test was not statistically significant.

n_i	$T(n_i)_{obs}$	Month/Year	Critical value	Critical value
			uncorrected	corrected
283	5.1893	April 1990	3.2066	5.0790
270	5.7600	December 1991	3.2040	5.0749

Table 5. Results from the binary segmentation procedure.

The complete binary segmentation procedure identified two significant change points that corresponded to three different levels for the mean of the process $\{\beta_t\}$. In chronological order, the mean levels of the process $\{\beta_t\}$ were 1.190, 0.835 and 1.005. This means that during the first period of time, the DO variable was on average 19% higher than those predicted by the regression component. In the second period of time, the observations were on average 16.5% lower, while in the last period of time, observations were on average close to the regression model's predictions.

The identification of statistically significant change points provides information that can be useful to environmental investigators or technicians. Nevertheless, from a statistical point of view, this information allows for obtaining more accurate models with better adjustments. Thus, we must consider the state space model defined in subsection 5.3, but only where the state process $\{\beta_t\}$ has three average levels, one for each sub-series identified in the change point procedure.

Thus, taking three means μ_1=1.190, μ_2=0.835 and μ_3=1.005 according to the identified change points of April 1990 and December 1991, the percentage of observations outside of the respective empirical confidence interval, at 95% level, for the forecasted one step-ahead of the DO concentration in the new model is 5.30% instead of 6.36% in the first model. On the other hand, the final model exhibited a determination coefficient equal to R^2=0.711 instead of R^2=0.689 in the first model.

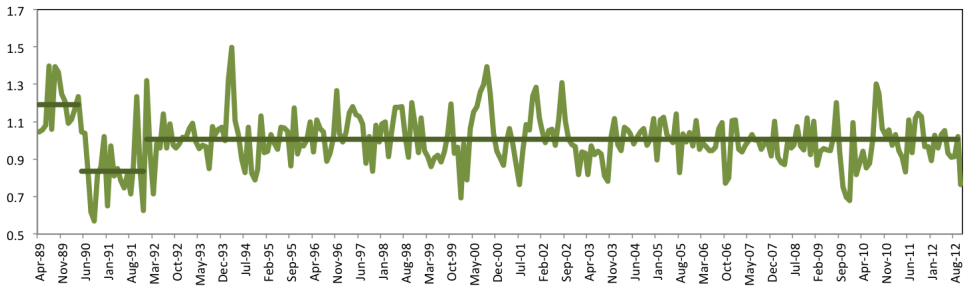

Figure 17. Representation of change points and the three mean levels of the state process.

http://dx.doi.org/10.5772/59814

6. Conclusions

The statistical analysis of water quality data is a relevant tool for the monitoring of ecosystems and water systems for human consumption. The information contained in water variable time series allows studying the past but also enables monitoring the present and planning the future. Biochemical analyses of water must be accompanied by statistical studies in order to identify patterns and analyse the evolution of water quality in its different forms.

This work presented a characterization of the River Vouga Basin in Portugal and focused primarily on the dissolved oxygen concentration variable. The complexity of databases renders more difficult to perform consistent statistical analyses of water quality. Statistical issues were presented and discussed in this paper to make their implementation easier for other researchers. Several problems were discussed and some solutions were provided, always keeping simplicity and practical relevance in mind.

Acknowledgements

Authors were partially supported by Portuguese funds through CIDMA – Center for Research and Development in Mathematics and Applications and the Portuguese Foundation for Science and Technology (FCT – Fundação para a Ciência e a Tecnologia") within project PEst-OE/MAT/UI04106/2013.

Author details

Marco André da Silva Costa* and Magda Sofia Valério Monteiro

*Address all correspondence to: marco@ua.pt

School of Technology and Management of Águeda, Center for Research & Development in Mathematics and Applications, University of Aveiro, Portugal

References

[1] Lopes JF. Silva CI. Cardoso AC. Validation of a water quality model for the Ria de Aveiro lagoon. Portugal. Environmental Modelling & Software 2008;23 479-494.

[2] AEA - Agência Europeia do Ambiente. Os recursos hídricos da Europa: Uma avaliação baseada em indicadores (Síntese). Copenhaga. Serviço das Publicações Oficiais da União Europeia; 2003.

[3] Solak CN. A`cs E. Water Quality Monitoring in European and Turkish Rivers Using Diatoms. Turkish Journal of Fisheries and Aquatic Sciences 2011;11 329-337.

[4] Andersen MS. Effectiveness of urban wastewater treatment policies in selected countries: an EEA pilot study. European Environment Agency EEA Report n. 2/ 2005. Office for Official Publications of the European Communities. Luxembourg. 2005.

[5] Machado A. Silva M. Valentim H. Contribute for the evaluation of water bodies status in Northern Region. Revista Recursos Hi ́dricos 2010;31(1) 57-63.

[6] Costa M. Gonçalves AM. Combining Statistical Methodologies in Water Quality Monitoring in a Hydrological Basin – Space and Time Approaches. In: Voudouris K. and Voutsa D. (ed.) Water Quality Monitoring and Assessment. Croatia: InTech; 2012. p121-142.

[7] Henriques V. Monitorização da qualidade da a ́gua na bacia Henriques hidrogra ́fica do Vouga. Master thesis. Universidade de Aveiro; 2010.

[8] Portal da Água. Instituto da Água. I.P. (INAG). http://portaldaagua.inag.pt (accessed 10 July 2014).

[9] Sistema Nacional de Informação de Recursos Hídricos (SNIRH). Instituto da Água. I.P. (INAG). http://snirh.pt (accessed 10 July 2014).

[10] INE – Instituto Nacional de Estatística. http://www.ine.pt

[11] Lopes JF. Silva CI. Temporal and spatial distribution of dissolved oxygen in the Ria de Aveiro lagoon. Ecological Modelling 2006;197 67-88.

[12] Ahmad I. Mohmood I. Coelho J.P. Pacheco M. Santos M.A. Duarte A.C. Pereira E. Role of non-enzymatic antioxidants on the bivalves' adaptation to environmental mercury: Organ-specificities and age effect in Scrobicularia plana inhabiting a contaminated lagoon [ria]. Environmental Pollution 2012;163 218-225.

[13] Serodio J. Cartaxana P. Coelho H. Vieira S. Effects of chlorophyll fluorescence on the estimation of microphytobenthos biomass using spectral reflectance indices. Remote Sensing of Environment 2009; 113 1760-1768.

[14] Grupo uariadeaveiro. Universidade de Aveiro. http://www.ua.pt/riadeaveiro/

[15] http://maretec.mohid.com/Estuarios/MenusEstuarios/Descri%C3%A7%C3%A3o/descricao_RiaAveiro.htm (accessed 12 August 2014)

[16] Lopes A. - In: "Aveiro e o seu distrito ", n.º 23/25, 1977/78, pp. 9-13.

[17] Booty WG. David CL. Freshwater ecosystem water quality modelling. In: Davies. A.M. (ed.) Modelling Marine Systems. vol. II. Boca Raton: CRC Press; 1989. p387-431.

[18] Shifflett DS. Water and Sustainability. http://www.unc.edu/~shashi/TablePages/dissolvedoxygen.html (accessed 16 July 2014)

[19] Bollerslev T. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 1986;31 307–327.

[20] Box GEP. Cox DR. An analysis of transformations. Journal of the Royal Statistical Society Series B 1964;26 211-252.

[21] Alpuim T. El-Shaarawi A. Modeling monthly temperature data in Lisbon and Prague. Environmetrics 2009;20 835-852.

[22] Jarušková D. Some problems with applications of change-point detection methods to environmental data. Environmetrics 1997;8 469-483.

[23] Gonçalves A. Costa M. Predicting seasonal and hydro-meteorological impact in environmental variables modelling via Kalman filtering. Stochastic Environmental Research and Risk Assessment 2013; 27 1021-1038.

[24] Shumway RH, Stoffer D. Time series analysis and its applications: with R examples. New York, Springer; 2006

[25] Kalman RE. A new approach to linear filtering and prediction problems. Transactions of the ASME, Journal of Basic Engineering 1960; 82 (Series D) 35-45.

[26] Harvey A.C. Forecasting structural time series models and the Kalman filter. Cambridge, Cambridge University Press; 1996

[27] Costa M. Alpuim T. Parameter estimation of state space models for univariate observations 2010; 140(7) 1889–1902

[28] Caussinus H. Mestre O. Detection and correction of artificial shifts in climate series. Journal of the Royal Statistical Society 2004; 53 (Series C) 405-425

[29] Lu Q. Lund R.B. Lee T.C.M. An MDL Approach to the Climate Segmentation Problem. Annals of Applied Statistics 2010; 4 299-319.

[30] Li S. Lund R.B. Multiple Changepoint Detection via Genetic Algorithms. Journal of Climate 2012; 25 674-686.

[31] Costa M. Goncalves A. Application of Change-Point Detection to a Structural Component of Water Quality Variables. In Theodore et al. (ed.) Numerical Analysis and Applied Mathematics ICNAAM 2011, AIP Conference Proceedings 1389, American Institute of Physics; 2011 New York, p. 1565-1568.

[32] Jarusková D. Change-point detection meteorological measurement. Monthly Weather Review 1996; 124 1535-1543

[33] Vostrikova L. Detecting 'disorder' in multidimensional random processes. Soviet Mathematics Doklady 1981;24 55-59.

[34] Fryzlewicz P. Wild binary segmentation for multiple change-point detection. Preprint, London School of Economics, http://stats.lse.ac.uk/fryzlewicz/wbs/ (accessed 12 August 2014).

INTECH