

Field survey and modelling of irrigation water quality indices in a Mediterranean island catchment: A comparison between spatial interpolation methods

Nikolaos Malamos^{1*} and Demetris Koutsoyiannis²

¹*Department of Agricultural Technology, Technological Educational Institute of Western Greece, Amaliada, Greece*

²*Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Zographou, Greece*

[*nmalamos@teiwest.gr](mailto:nmalamos@teiwest.gr)

Abstract A biannual survey of physico-chemical quality indices of 104 irrigation-water wells located in a cultivated plain of a Mediterranean island catchment was conducted using a multi-parameter probe. The campaign was planned so as to differentiate between the dry and wet seasons. The acquired data constituted the test bed for evaluating the results and the features of four spatial interpolation methods, i.e. ordinary kriging, universal kriging, inverse distance weighted and nearest neighbours, against those of the recently introduced bilinear surface smoothing (BSS). In several cases, BSS outperformed the other interpolation methods, especially during the two-fold cross-validation procedure. The study emphasizes the fact that both *in situ* measurements and good mathematical techniques for studying the spatial distribution of water quality indices are pivotal to agricultural practice management. In the specific case studied, the spatio-temporal variability of water quality parameters and the need for monitoring were evident, as low irrigation water quality was encountered throughout the study area.

Keywords irrigation; water quality; specific conductance; pH; spatial interpolation; bilinear surface smoothing; kriging; inverse distance weighted; nearest neighbours; cross-validation

Introduction

Sustainable agricultural development and management are based on adequate quantity and quality of water resources. Pollution and unreasonable use of water threaten development and demand implementation of measures and policies relating to both quality and demand management, along with quality assessment depending on each particular water use.

Specifically, agriculture can cause extensive degradation of the soil–water system as well as underlying aquifers, when good agricultural practices are not implemented. The main problems associated with agriculture are: the increase in the concentration of salts, nitrates and agrochemical pollution, and the often insufficient water quantity causing unsustainable exploitation of aquifers. The quality of irrigation water is assessed on the basis of its effect on soil and on agricultural management practices (Malamos and Nalbantis 2005). Within this framework, there is a need for water quality monitoring in order to scientifically address the impacts of both existing and future degradation of water and soil quality, and to provide the basis for action at all administrative levels.

In monitoring activities (Bartram and Ballance 1996, Chapman 1996), water quality is assessed by sampling or *in situ* point measurements in selected locations within the area of interest. Generalization of the results throughout the study area is realized by implementing spatial interpolation methods, as reported in several studies about the quality of water. For example, Gong *et al.* (2014) dealt with the comparison between kriging and inverse distance weighting methods, in estimating groundwater arsenic concentrations in rural areas; Kourgialas *et al.* (2017) analysed the dissolution and transport of excess quantities of the major and trace elements of fertilizers to the groundwater, which deteriorate the quality of drinking and irrigation water, by

implementing a GIS decision support system; Mir *et al.* (2017) dealt with the spatial monitoring of chemical parameters of water in dry and wet years in order to follow the variations in the water quality and determine the most suitable sites to extract potable and irrigation water by mapping these parameters; Murphy *et al.* (2010) performed comparison of three spatial interpolation methods – inverse distance weighting, ordinary kriging, and universal kriging – for water quality evaluation; while Yidana *et al.* (2012) classified groundwater quality control parameters and determined the quality for domestic and commercial irrigation purposes, using multivariate statistical methods and geographic information systems.

The predictive performance of spatial interpolation methods is affected by interrelated factors: (a) sampling design and spatial distribution of samples, and (b) the nature and quality of data (Li and Heap 2014). Several studies, in various disciplines, have dealt with the evaluation of different spatial interpolation methods (Burrough and McDonnell 1998, Goovaerts 2000, Price *et al.* 2000, Vicente-Serrano *et al.* 2003, Stahl *et al.* 2006, Kis 2016, Malamos and Koutsoyiannis 2016b, Malamos *et al.* 2017).

Here, spatial interpolation methods are assessed based on a biannual survey of the physico-chemical quality indices of irrigation water, such as: electrical conductivity, pH, dissolved oxygen (DO), temperature, turbidity and oxidation–reduction potential, along with the variability of the water level inside irrigation wells of a cultivated plain of a Mediterranean island catchment, which was conducted using a multi-parameter probe. This survey resulted in data collection from 104 water wells, constituting a dataset capable of addressing the spatial variation of the aforementioned variables.

Apart from the obvious importance of the acquired data considering agricultural practice management during the dry and wet seasons, they constituted the basis for testing the special interpolation methods. The features of four well-established spatial

interpolation methods, i.e. ordinary kriging (OK), universal kriging (UL), inverse distance weighting (IDW) and nearest neighbours (NN), were evaluated against those of the recently introduced bilinear surface smoothing (BSS; Malamos and Koutsoyiannis 2016a). The performance of each method was assessed by means of two different kinds of cross-validation procedures, implementing several statistical criteria. The mathematical derivation of the BSS's leave-one-out cross-validation residuals from the existing mathematical framework is also depicted.

Study area

General characteristics

The study area is the plain of Komi-Kalloni (Fig. 1) located on the northeast side of the Greek island of Tinos (37.6°N, 25.15°E; WGS84–EPSG: 4326), in the Aegean Sea. Tinos is the fourth largest island in the Cyclades archipelago, after the islands of Naxos, Andros and Paros, with an area of approx. 195 km². It has about 9000 inhabitants across 62 settlements. Tinos is located in the northern Cyclades, southeast of Andros island and northwest of Mykonos island.

[Figure 1]

The plain of Komi-Kalloni (268.1 ha), located at the estuary of the largest water basin of Tinos island, has an area of 3821.6 ha and a total stream length of about 90 km (Fig. 1). The average altitude of the plain is 27.5 m a.s.l., with minimum and maximum altitudes of 12 and 96 m a.s.l., respectively. Three villages, named Komi, Kalloni and Kato Klesma, are located on the southern (upstream) edges of the plain and are populated by farmers.

From a soil perspective, the study area consists of well-drained alluvial deposits of neutral reaction, with loamy sand texture without CaCO₃, and shale bedrock (Shahabi and Anagnostopoulos 1973). There are soil salinity problems, as a direct consequence of the vicinity of the sea, caused by saline irrigation water and sea salt aerosols due to the presence of strong north winds all year round. Because of these strong winds, the farmers grow windbreaks from common reed (*Phragmites australis*) at the edges of every field, in order to protect their crops.

The agricultural production of the plain is essential to the island's sustainability, since it provides the majority of edible fruits and vegetables consumed by both the inhabitants and tourists. The main cultivation consists of horticultural crops such as potatoes, tomatoes, leafy vegetables and several kinds of fruits, along with citrus trees, mainly lemon trees. The cultivation of artichokes (*Cynara cardunculus*) increases downstream due to their tolerance to salinity, while the area occupied by citrus trees remains stable and is upstream. There are also some vineyards and olive groves located on the edges of the plain. This farming pattern, which is characterized by high quality agricultural products in limited volumes, is typical of the Cyclades islands.

The climate of the region is Mediterranean and, according to the Köppen-Geiger classification, is characterized as CSa, i.e. warm temperate steppe with a hot summer (Kottek *et al.* 2006). The average annual rainfall does not exceed 460 mm; the average temperature is 18.0°C, with average maximum and minimum temperatures of 30.8 and 7.3°C reported in July and February, respectively. The crops are irrigated almost exclusively by water wells during the dry period of the year, which lasts from early April to late September.

The monthly variations in the basic climatic parameters, acquired from the climatic atlas of Greece (Mamara *et al.* 2016, 2017) are presented in Figure 2.

[Figure 2]

Field survey

To acquire the necessary information about the spatio-temporal variation in the irrigation water quality, a field survey (Bartram and Ballance, 1996; Chapman, 1996) was planned and conducted, consisting of *in situ* measurements of several physico-chemical properties of every well located in the study area at the end of the dry and wet periods respectively, i.e. twice a year.

There are three types of monitoring activities that distinguish between long-term, short-term and continuous monitoring programmes. To perform the field survey, the location of all the water wells in the study area was identified. The methodology for collecting and recording the wells' locations into a geographic information system (GIS) consisted of the following steps:

1. Initial location of the wells, using a GIS and remote sensing data together with interviews of local farmers. Farmers were able to identify well locations on a computer screen with the help of a thematic map of the study area that included: (a) very high resolution colour satellite images; (b) a digital elevation model (DEM); (c) different points of interest, such as place names, churches, streams and roads.
2. On-the-spot visits with the participation of farmers as guides, to carry out the first round of measurements while simultaneously verifying the position of the wells using a geographical positioning system (GPS) device.

3. Registration of the actual positions of the wells into the GIS using the projected coordinate system for Greece, i.e. Greek Grid (EPSG: 2100) and updating the thematic map with all the necessary information for the next measurement cycle.
4. A second round of measurements was carried out without involving farmers as guides, by just using the GPS device with the actual positions of the wells.

In this way, the 104 water wells were identified and included in both measurement rounds (Fig. 3).

[Figure 3]

A multi-parameter water quality probe (TROLL 9000E, In-Situ Inc. 2005) connected with a 20-m-long cable to a PDA was used for acquiring water quality measurements and storing the data. The probe is equipped with several sensors to measure, e.g. pH, DO, conductivity (and specific conductance, salinity, total dissolved solids, resistivity), temperature, turbidity and oxidation–reduction potential. All the sensors, except for the turbidity sensor, were calibrated using a standardized solution from In-Situ Inc. (Quick Cal). This combination of sensors is capable of providing an assessment of water quality (Bartram and Ballance 1996).

Materials and methods

Following the preliminary data analysis, spatial interpolation and mapping of the parameters affecting irrigation water quality was performed, applying four of the most commonly used methods, i.e. ordinary kriging (OK), universal kriging (UK), inverse distance weighted (IDW), nearest neighbours (NN), against the recently introduced bilinear surface smoothing (BSS; Malamos and Koutsoyiannis 2016a, 2016b).

The results of all five methodologies were compared, in terms of several performance evaluation criteria, along with two different types of cross-validation, i.e. leave-one-out and two-fold cross-validation.

The implementation of the leave-one-out cross-validation procedure, using the BSS mathematical framework, is also presented.

Bilinear surface smoothing

The non-parametric mathematical framework of BSS (Malamos and Koutsoyiannis 2016a) incorporates smoothing terms with adjustable weights, defined by means of the angles formed by consecutive bilinear surfaces into a piecewise surface regression model with known break points. An alternative implementation of the main methodology is the bilinear surface smoothing with explanatory variable (BSSE) that incorporates, in an objective manner, an explanatory variable available from measurements in a considerably denser dataset than the initial main variable.

The mathematical framework of BSS suggests that fitting is meant in terms of minimizing the total square error among the set of original points $z_i(x_i, y_i)$ for $i = 1, \dots, n$ and the fitted bilinear surface, that in matrix form, can be written as:

$$p = \|\mathbf{z} - \hat{\mathbf{z}}\|^2 \quad (1)$$

where $\mathbf{z} = [z_1, \dots, z_n]^T$ is the vector of known applicates of the given data points with size n (the superscript T denotes the transpose of a matrix or vector) and $\hat{\mathbf{z}} = [\hat{z}_1, \dots, \hat{z}_n]^T$ is the vector of estimates with size n .

Both BSS and BSSE have the following features, as outlined in Malamos and Koutsoyiannis (2016a):

- BSS is *univariate*, while BSSE is *multivariate*.

- They are both *local* and *global*.
- They can be either *exact* or *inexact*.
- They are *stochastic*, since the proposed mathematical framework, apart from estimations, provides direct means of evaluating interpolation errors; also, it provides the leave-one-out cross-validation residuals, as demonstrated in the following sections.
- The surfaces that they produce can be either *gradual* or *abrupt* depending on the magnitude of the smoothing parameters.
- Both BSS and BSSE use a regular grid not necessarily square, since the number of bilinear surfaces along the x direction does not have to coincide with the number of bilinear surfaces along the y direction.

For explanation and analysis of the above terminology the interested reader is referred to the classification of spatial interpolation methods presented by Li and Heap (2008). A brief presentation of the method and its equations follows, while the details of the method, including the algorithms and derivations of the equations, may be found in (Malamos and Koutsoyiannis 2016a).

Let (cx_l, cy_k) , $l = 0, \dots, mx$, $k = 0, \dots, my$, be a grid of $(mx + 1) \times (my + 1)$ points on the $x y$ plane, so that the rectangle with vertices (cx_0, cy_0) , (cx_{mx}, cy_0) , (cx_0, cy_{my}) and (cx_{mx}, cy_{my}) contain all (x_i, y_i) . For simplicity, we assume that the points on both axes are equidistant, i.e. $cx_l - cx_{l-1} = \delta_x$ and $cy_k - cy_{k-1} = \delta_y$.

The general estimation function for point u on the $(x y)$ plane, according to the BSS method, is:

$$\hat{z}_u = d_u \quad (2)$$

while, according to the BSSE method it is:

$$\hat{z}_u = d_u + t_u e_u \quad (3)$$

where d_u, e_u are the values of the two bilinear surfaces at that point and t_u is the corresponding value of the explanatory variable.

Equations (2) and (3) can be more concisely written, for all given points $z_i(x_i, y_i)$ simultaneously, as:

$$\hat{\mathbf{z}} = \mathbf{\Pi} \mathbf{d} \quad (4)$$

and

$$\hat{\mathbf{z}} = \mathbf{\Pi} \mathbf{d} + \mathbf{T} \mathbf{\Pi} \mathbf{e} \quad (5)$$

where $\mathbf{d} = [d_0, \dots, d_m]^T$ is a vector of unknown applicates of the bilinear surface d , with size $m+1$ ($m = (mx + 1) \times (my + 1) - 1$); $\mathbf{e} = [e_0, \dots, e_m]^T$ is a vector of unknown applicates of the bilinear surface e , with size $m+1$; and \mathbf{T} is a $n \times n$ diagonal matrix with elements:

$$\mathbf{T} = \text{diag}(t(x_1, y_1), \dots, t(x_n, y_n)) \quad (6)$$

where $t(x_1, y_1), \dots, t(x_n, y_n)$ are the values of the explanatory variable at the given data points; and $\mathbf{\Pi}$ is a matrix with size $n \times (m + 1)$, whose ij th entry (for $i = 1, \dots, n; j = 0, \dots, m$) is:

$$\pi_{ij} = \begin{cases} \frac{(cx_l - x_i)(cy_k - y_i)}{\delta_x \delta_y}, & \text{when } cx_{l-1} < x_i \leq cx_l \text{ and } cy_{k-1} < y_i \leq cy_k \\ \frac{(cx_l - x_i)(y_i - cy_{k-1})}{\delta_x \delta_y}, & \text{when } cx_{l-1} < x_i \leq cx_l \text{ and } cy_k < y_i \leq cy_{k+1} \\ \frac{(x_i - cx_{l-1})(y_i - cy_{k-1})}{\delta_x \delta_y}, & \text{when } cx_l < x_i \leq cx_{l+1} \text{ and } cy_k < y_i \leq cy_{k+1} \\ \frac{(x_i - cx_{l-1})(cy_k - y_i)}{\delta_x \delta_y}, & \text{when } cx_l < x_i \leq cx_{l+1} \text{ and } cy_{k-1} < y_i \leq cy_k \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

The calculation of the unknown vectors \mathbf{d} and \mathbf{e} requires also the definition of matrices Ψ_x and Ψ_y with size $(m-1) \times (m+1)$ (for $i = 1, \dots, m-1$ and $j = 0, \dots, m$) and ij th entry:

$$\psi_{x_{i,j}} = \begin{cases} 2, & \text{when } i = j \text{ and } i - k(mx + 1) \notin \{1, mx + 1\} \\ -1, & \text{when } |i - j| = 1 \text{ and } i - k(mx + 1) \notin \{1, mx + 1\} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where $k = 0, \dots, my$, while:

$$\psi_{y_{i,j}} = \begin{cases} 2, & \text{when } i = j \text{ and } i - l(my + 1) \notin \{1, my + 1\} \\ -1, & \text{when } |i - j| = 1 \text{ and } i - l(my + 1) \notin \{1, my + 1\} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

with $l = 0, \dots, mx$ (note that Ψ_x and Ψ_y are identical when $mx = my$).

In the case of BSS, the solution that minimizes error has the following form:

$$\mathbf{d} = (\mathbf{\Pi}^T \mathbf{\Pi} + \lambda_x \Psi_x^T \Psi_x + \lambda_y \Psi_y^T \Psi_y)^{-1} (\mathbf{\Pi}^T \mathbf{z}) \quad (10)$$

Likewise, in the case of BSSE, the solution is:

$$\begin{bmatrix} \mathbf{d} \\ \mathbf{e} \end{bmatrix} = \mathbf{\Omega}^{-1} \begin{bmatrix} \mathbf{\Pi}^T \mathbf{z} \\ \mathbf{\Pi}^T \mathbf{T}^T \mathbf{z} \end{bmatrix} \quad (11)$$

where

$$\mathbf{\Omega} := \begin{bmatrix} \mathbf{\Pi}^T \mathbf{\Pi} + \lambda_x \Psi_x^T \Psi_x + \lambda_y \Psi_y^T \Psi_y & \mathbf{\Pi}^T \mathbf{T} \mathbf{\Pi} \\ \mathbf{\Pi}^T \mathbf{T} \mathbf{\Pi} & \mathbf{\Pi}^T \mathbf{T}^T \mathbf{T} \mathbf{\Pi} + \mu_x \Psi_x^T \Psi_x + \mu_y \Psi_y^T \Psi_y \end{bmatrix} \quad (12)$$

The minimum number of $m + 1$ points required to solve equations (10) or (11) is 6, since the minimum number points needed to define the bilinear surfaces is the number of points that define two consecutive planes oriented according to either x or y direction. Based on the above equations, we can estimate the applicate of any point that lies in the two-dimensional interval ($[cx_0, cx_{mx}] \times [cy_0, cy_{my}]$) by using either version of the proposed methodology.

Choice of parameters

The adjustable parameters required to implement each of the two versions of the methodology can be estimated by transforming the smoothing parameters λ and μ in terms of tension: τ_λ and τ_μ , whose values are restricted in the interval $[0, 1)$, for both directions (Malamos and Koutsoyiannis 2016a). This transformation provides a convenient search in terms of computational time and is based on the generalized cross-validation (GCV; Craven and Wahba 1978, Wahba and Wendelberger 1980) methodology. Thus, for a given combination of segments mx , my , the minimization of GCV results in the optimal values of $\tau_{\lambda x}$, $\tau_{\lambda y}$ and $\tau_{\mu x}$, $\tau_{\mu y}$. This can be repeated for several trial combinations of mx , my values, until the global minimum of GCV is reached.

Leave-one-out cross-validation

The leave-one-out cross-validation (LOOCV) is one of the most commonly used procedures for evaluating spatial interpolation methods, with several researchers reporting various applications in the discipline of water resources and soil science (Burrough and McDonnell 1998, Li and Heap 2008, Oliver and Webster 2014). Also, the differences between the LOOCV model predictions and observations, i.e. residuals, should be tested for normality and linear correlation with the original data points, in

order to evaluate the quality of model results (Kitanidis 1997, Malamos and Koutsoyiannis 2015).

The proposed mathematical formulation of bilinear surface smoothing easily delivers the estimation of the LOOCV residuals, in order to provide means of the method's outcome evaluation without repeating the procedure n times, one for each z_i left out.

For the BSS case, this is achieved by combining equations (4) and (10) to:

$$\hat{\mathbf{z}} = \mathbf{A} \mathbf{z} \quad (13)$$

where \mathbf{A} is a $n \times n$ symmetric matrix given by:

$$\mathbf{A} = \mathbf{\Pi} (\mathbf{\Pi}^T \mathbf{\Pi} + \lambda_x \mathbf{\Psi}_x^T \mathbf{\Psi}_x + \lambda_y \mathbf{\Psi}_y^T \mathbf{\Psi}_y)^{-1} \mathbf{\Pi}^T \quad (14)$$

while, for the BSSE case, combining equations (5) and (11), we obtain again equation (13), with \mathbf{A} being a $n \times n$ symmetric matrix, now given by:

$$\mathbf{A} = \mathbf{\Pi} \mathbf{T} \mathbf{\Pi} \mathbf{\Omega}^{-1} (\mathbf{\Pi} \mathbf{T} \mathbf{\Pi})^T \quad (15)$$

The positive-definite smoother matrices \mathbf{A} in equations (14) and (15) include all adjustable parameters: m_x , m_y , λ_x , λ_y and μ_x , μ_y and they are preserving the elements of each row sum to one. Let $\check{\mathbf{z}} = [\check{z}_1, \check{z}_2, \dots, \check{z}_n]^T$ be the vector with size n , of the estimates, \check{z} , when each data point is successively left out and predicted from the rest of the data. In order to acquire $\check{\mathbf{z}}$ using the smoother matrices presented above with their i th row and column deleted in order to be of dimension $(n-1) \times (n-1)$, we must renormalize their rows to sum to one. When the i th column is deleted, the i th row now sums to $1 - a_{ii}$, where a_{ii} are the diagonal elements of matrix \mathbf{A} . Thus, by dividing every

element of the ordinary predicted values by $1 - a_{ii}$ (for $i = 1$ to n), we get the LOOCV predicted values, i.e.:

$$\check{z}_i = \frac{1}{1-a_{i,i}} \sum_{\substack{j=1 \\ j \neq i}}^n a_{i,j} z_j \quad (16)$$

while the i th ordinary predicted value is:

$$\hat{z}_i = \sum_{j=1}^n a_{i,j} z_j \quad (17)$$

To acquire the leave-one-out residuals we multiply both sides of equation (16) with $(1 - a_{ii})$, and, after rearrangement, we obtain:

$$\check{z}_i = \hat{z}_i - a_{i,i} z_i + a_{i,i} \check{z}_i \quad (18)$$

By subtracting z_i from both sides of equation (18), we obtain:

$$z_i - \check{z}_i = z_i - (\hat{z}_i - a_{i,i} z_i + a_{i,i} \check{z}_i) \quad (19)$$

After rearrangement, the leave-one-out residuals are given by:

$$z_i - \check{z}_i = \frac{1}{1-a_{i,i}} (z_i - \hat{z}_i) \quad (20)$$

In matrix form, equation (20) can be written as:

$$\mathbf{z} - \check{\mathbf{z}} = \mathbf{S}(\mathbf{I} - \mathbf{A})\mathbf{z} \quad (21)$$

where \mathbf{S} is a $n \times n$ diagonal matrix, used for normalization, with elements:

$$\mathbf{S} = \text{diag}\left(\frac{1}{1-a_{1,1}}, \frac{1}{1-a_{2,2}}, \dots, \frac{1}{1-a_{i,i}}\right) \quad (22)$$

The formulation of equation (21) has the advantage of implementing the already computed vector $(\mathbf{I} - \mathbf{A})\mathbf{z}$ from the GCV minimization procedure to estimate the adjustable parameters of the method, as noted in the previous section and detailed in Malamos and Koutsoyiannis (2016a).

The left-hand side of equation (21) is the vector containing the LOOCV residuals, while the right-hand side expresses these residuals in terms of the ordinary residuals. Since the leave-one-out residuals ($z_i - \check{z}_i$) are known, the leave-one-out predicted values, \check{z}_i , can be computed.

Kriging

One of the commonly-used estimators for the interpolation of spatial data is the kriging technique. Kriging selects weights so that the estimates of a regionalized variable at selected points are unbiased and the estimation variance is minimized.

The character of the spatially correlated variation is encapsulated in functions such as the variogram and the covariogram, and these provide the information for optimizing interpolation weights and search radii. Experimental variograms are computed from sample data in one, two, or three spatial dimensions. These experimental data are fitted by a theoretical variogram model, which serves to provide data for computing interpolation weights (Burrough and McDonnell 1998).

Kriging requires a large number of available data points, at least 100, according to Oliver and Webster (2014), or 50–100 according to other studies (Li and Heap 2008), in order to produce a reliable estimation of variogram. The number of the required data points depends on the kind of spatial variation encountered, though smooth surfaces require fewer points than those with irregular variation (Burrough and McDonnell 1998, Goovaerts 2000).

Currently, there are a lot of geostatistical methods incorporating different approaches of kriging, such as: simple, ordinary and universal kriging, kriging with an external drift or cokriging, which can accomplish interpolation tasks (Burrough and McDonnell 1998, Goovaerts 1997, Goovaerts 2000, Li and Heap 2008). Ordinary and

universal kriging, which are used in the present study, are gradual, local, and may or may not reproduce the measured data.

Both methods require fitting a theoretical variogram to the empirical one. The actual process of fitting a model to an empirical variogram is referred to as more of an art than a science and involves evaluation of several types of models, a procedure that is time consuming and to some extent subjective with different authorities suggesting different methods and protocols (Bohling 2005).

Ordinary kriging (OK) is similar to simple kriging and the only difference is that OK estimates the local constant mean, then performs simple kriging on the corresponding residuals, and only requires the constant mean of the local search window (Goovaerts 1997, Li and Heap 2008).

Universal kriging (UK), also known as kriging with a trend or kriging in the presence of a drift, is a multivariate extension of ordinary kriging accommodating a spatially varying trend, introduced by Matheron (1969). It can be used both to produce local estimates in the presence of trend and to estimate the underlying trend itself, if it can be modelled by simple functions. UK can be used when the stochastic field of interest does not meet the criterion of second-order stationarity necessary for kriging. Second-order stationarity suggests that the mean and variance are the same on the entire area and that the correlation between any two observations depends only on their relative position in space. If the mean is assumed not constant across the entire study area the model is said to be nonstationary. UK splits the random function into a linear combination of deterministic functions, the smoothly varying and the trend, which is also called a drift, and a random component representing the residual stochastic field. Spatial trend or a drift represents any detectable tendency for the values to change as a function of the coordinate variables. The mean can be a function of the coordinates in

linear, quadratic or higher form (Kis 2016, Oliver and Webster 2014, Tabios and Salas 1985, Vicente-Serrano *et al.* 2003).

Inverse distance weighted and nearest neighbours

The inverse distance weighting (IDW) and nearest neighbours (NN) methods were implemented as quick and exact interpolators capable of addressing the characteristics of the study area due to the large number of available data points.

The IDW method is straightforward and computationally non-intensive, and effective in many aspects (Tegos *et al.* 2015, 2017). It has been regarded as one of the standard spatial interpolation procedures in geographic information science (Burrough and McDonnell 1998) and has been implemented in almost every GIS software package. Formally, the IDW method estimates the values of an attribute at unsampled points using a linear combination of values at sampled points weighted by an inverse function of the distance from the point of interest to the sampled points. The assumption is that sampled points closer to the unsampled point are more similar to it than those farther away in their values (Li and Heap 2008).

The NN method, otherwise known as Thiessen polygons, predicts attributes at unsampled locations based on the nearest single data point. So, NN divides a region geometrically, in a way that is totally determined by the configuration of the data points, with one observation per cell. This is accomplished by triangulating all available data points into an irregular network that meets the Delaunay criterion, i.e. no point rests inside the circumcircle of any triangle. The perpendicular bisectors for each triangle edge are generated, forming the edges of the Thiessen polygons. The locations at which the bisectors intersect determine the locations of the Thiessen polygon vertices. If the data lie on a regular square grid, then the produced polygons are all equal, regular cells

with sides equal to the grid spacing; if the data are irregularly spaced, then an irregular lattice of polygons results. Obviously, since all estimates equal the values at the data points, NN is classified as an exact interpolator.

Computational implementation

The BSS method was realised in Microsoft Excel, as it provides a direct means of data visualization and graphical exploration. This was accomplished by the development of a dynamic link library in Free Pascal (Lazarus Team 2016), which was linked to Microsoft Excel. In this context, an Excel array formula acts as the main interface, with its arguments being the values and coordinates of the available points along with the coordinates of the unknown points, the number of points on the x and y axis that form the bilinear surfaces and the smoothing parameters values (Malamos and Koutsoyiannis 2016a).

The IDW, NN, OK and UK methods were performed by means of the free and open source GIS software: System for Automated Geoscientific Analyses (SAGA) version 5.0.0 (Conrad *et al.* 2015). SAGA is equipped with numerous tools, providing an almost complete collection of interpolation techniques, comprising deterministic and geostatistical kriging methods and its derivatives, with variable search radii.

One important feature of the above-mentioned SAGA modules (IDW, NN, OK and UK) is the ability to perform different kinds of cross-validation, i.e. leave-one-out, two-fold or k -fold cross-validation. The output of the two-fold and k -fold cross-validation is a set of performance metrics such as: mean square error (MSE), root mean square error (RMSE), normalized root mean square error (NRMSE) and the coefficient of determination (R^2). In the case of LOOCV, SAGA provides additional information in terms of the estimation residuals along with the leave-one-out predicted values. These

residuals can be tested for normality and linear correlation to the original data points and compared against the results of the BSS residuals analysis, presented earlier.

The computations involving the IDW and NN interpolations were relatively straightforward, while the kriging interpolation required the fitting of a theoretical variogram to the empirical one, for each parameter. The investigated variogram models were the widely-used: (a) spherical, (b) exponential, (c) Gaussian and (d) power. In the case of UK, we chose the geographical coordinates as “external drift” predictors.

Since SAGA’s variogram fitting module required the range as input, the fitting procedure was accomplished by means of the Excel XonGrid Interpolation Add-in (<http://xongrid.sourceforge.net/>). XonGrid is a free Excel library of functions to perform multidimensional interpolations from scattered data. Excel Solver was used to perform the optimization procedure by adjusting range, sill and nugget. The upper limit of range was set as the distance, h , at which the value of the model variogram reaches 95% of the sill, i.e. the 95% of the midpoint distance between the outermost data points. The sill was maintained larger than the variance of the observations, γ , while the nugget was estimated through the optimization procedure.

All four variogram models were fitted for each parameter using the procedure described above. Those finally selected for kriging performed better in terms of the statistical criteria provided by SAGA GIS during the LOOCV procedure along with tests considering residuals normality and linear correlation. So, both kriging versions were implemented using the best variogram model according to the residuals analysis of the LOOCV procedure and not the best fitted model to the experimental variogram.

Evaluation criteria

The criteria used for the evaluation of the methodologies performance are: mean bias

error (MBE), mean absolute error (MAE), root mean square error (RMSE) and modelling efficiency (EF) (Loague and Green 1991, Nash and Sutcliffe 1970, Willmott 1982). Willmott (1982) suggests that RMSE and MAE are among the best overall measures of model performance, as they summarize the mean difference in the units of observed and predicted values. RMSE provides a measure of model validity that places a lot of weight on high errors, whereas MAE is less sensitive to extreme values. The relationships that provide them are:

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (P_i - O_i) \quad (23)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |P_i - O_i| \quad (24)$$

$$\text{RMSE} = \left[\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2 \right]^{1/2} \quad (25)$$

$$\text{EF} = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (\bar{O} - O_i)^2} \quad (26)$$

where n is the number of observations, O_i are the observed values, P_i are the predicted values, and \bar{O} is the mean of the observed values. The optimum (minimum) for the MBE, MAE, RMSE statistics is 0, while the optimum (maximum) for EF is 1.

Ideal point error

The ideal point error (IPE; Domínguez *et al.* 2011) metric is calculated by identifying the ideal point, up to a five-dimensional space, against which each model should be evaluated. For the purposes of the present study, the three-dimensional vector IPE3 is implemented by normalizing RMSE, MBE and the coefficient of determination (R^2), so the individual IPE3 for each measure ranges from 0 for the best model to 1 for the worst.

The coordinates of the ideal point are: $RMSE = 0$, $R^2 = 1$, $MBE = 0$. IPE3 measures how far a model is from this ideal point by the relationship:

$$IPE3_i = \left[(1/3) \left(\left(\frac{RMSE_i}{\max RMSE} \right)^2 + \left(\frac{R_i^2 - 1}{\min R^2 - 1} \right)^2 + \left(\frac{MBE_i}{\max |MBE|} \right)^2 \right) \right]^{1/2} \quad (27)$$

In equation (27), i represents each of the models under investigation.

Results and discussion

Preliminary data analysis

As previously stated, 104 water wells were identified inside the study area, resulting in an average density of one well every 2.6 hectares. The majority of sampling points were taken up to 35 m of altitude (Fig. 3).

Two rounds of sampling were performed, the first at the end of the dry period of the hydrological year (from 28 August 2007 to 21 September 2007), and the second at the end of the following wet period (14 March 2008 to 30 March 2008 (Fig. 3). It is worth pointing out that the timing of the second round of measurements was determined by the weather, since access to the lower parts of the plain was almost impossible due to rainfall events and the subsequent floods that occurred during the first 10 days of March 2008. All measurements were executed following the procedures described in the manual of the TROLL 9000E probe. Each measurement was taken when the full length of the instrument's body was submerged, at 60 cm depth. After each measurement completion, the probe was rinsed with clear water to remove impurities on the sensors before the next measurement. The duration of the measurements depended on the time needed for the sensors to stabilise. Also, calibration was performed twice in each round

of measurements, the first at the beginning and the second in the middle of each round, using the Quick Cal solution (In-Situ Inc. 2005).

Apart from the TROLL 9000E parameters, the distance between the water surface and the ground, together with the total depth of each well, were measured. Then, the water level inside the well was determined by subtracting the first measurement from the second. Consequently, the water elevation above sea level was estimated by subtracting the distance between the water surface and the ground from the corresponding altitude provided by the DEM of the study area.

The first round of sampling took 16 days to complete, while the second was only 12 days. This is explained by the fact that the locations of the wells and the routes to reach them were not *a priori* available. As this was not the case for the second round of measurements, a greater number of samples could be taken on each day.

The results are summarized in Table 1 (dry period) and Table 2 (wet period), by means of the maximum, minimum and average values, along with the percentage of measurements exceeding the average. The parameter values were within the limits found in the literature (In-Situ Inc. 2005). The values of actual conductivity were normalized to 25°C, to allow comparison between measurements made at different temperatures. This conversion required a temperature coefficient for the solution being measured. By convention, the temperature coefficient for potassium chloride (KCl) calibration standards was used. Therefore the specific conductance was calculated from (In-Situ Inc., 2005):

$$SC = \frac{AC}{[1+0.0191(T-25)]} \quad (28)$$

where AC is the actual conductivity (in $\mu\text{S}/\text{cm}$), 0.0191 is the nominal temperature coefficient for KCl solutions and T is the solution temperature (in $^{\circ}\text{C}$). According to In-

Situ Inc. (2005), specific conductance is used to estimate total dissolved solids (TDS), by multiplying the specific conductance by a factor of 0.65.

[Table 1, Table 2]

Figure 4 shows the variation in specific conductance along with the geological faults present in the study area between the two measurement periods, in terms of analogy between those acquired during the wet period against those of the dry period, while Figures 5 and 6 present the corresponding variation in the water level and the DO, respectively.

[Figure 4, Figure 5, Figure 6]

The water level fluctuation is evident along the main stream, especially upstream. In general, as the distance from the main stream increases, the water level fluctuation is reduced too. However, the water level rises by 40% during the wet period, with an average value of 4.5 m, and 46.2% of measurements exceeding this average. The average distance between the ground surface and the water inside the wells, was 2.7 m for the wet period, while it reached 5.1 m for the dry period, presenting a notable increase.

In the case of specific conductance, a decrease is obvious (Fig. 4), especially along the main stream. It should be noted that the maximum value of specific conductance measured during the wet period (7205 $\mu\text{S}/\text{cm}$, Table 2) can be explained by sea water intrusion, since measurements were done along the geological faults on the northeast side of the study area, near the sea.

The vicinity of the sea constitutes a major reason to investigate the existence of trend between the measured parameters and their location. In this context, the degree of linear dependence between the measured parameters and the projected coordinates (Greek Grid-EPSSG: 2100) of each water well was examined in terms of the coefficient

of determination (R^2) of the fitted line, when the parameters were plotted against each direction (Table 3). The y-coordinates are well correlated to the distance from the sea due to the shape of the catchment. Water elevation, specific conductance and oxidation–reduction potential are linearly correlated to the variation of y-coordinates, in both dry and wet periods, with R^2 values in the range 0.202–0.457. For the wet period, pH and DO have R^2 values of 0.263 and 0.330, respectively, with the y-coordinates, while for the dry period the corresponding values are much smaller. Finally, the water temperature has a R^2 value of 0.375 during the dry period only. Based on the above discourse, the presence of a spatial trend imposes the use of universal kriging for spatial interpolation of the selected variables.

[Table 3]

The variogram analysis described earlier showed that, for the wet season, the best variogram model for all the parameters, except water elevation, was the spherical for both kriging versions, i.e. OK and UK, while the exponential model performed better for the dry season. In the case of groundwater elevation above sea level, the power model outperformed the other models, for both seasons.

Spatial interpolation

Spatial interpolation of parameters requires representability of the point measurements, not only at the specific locations where the data were collected, but also in adjacent areas. Measurements of physical parameters, such as DO and turbidity, inside non-flowing water bodies, e.g. wells, cannot be considered as representative of the corresponding soil water properties across the study area, due to the complex external and biochemical influences, e.g. time of day of the measurement, sun exposure, algal growth and decomposition of organic matter, that occur in each well (Bartram and

Ballance 1996, Chapman 1996, Murphy *et al.* 2010). However, simple geochemical variables, such as conductivity normalized by temperature, i.e. specific conductance (or total dissolved solids) and pH, may be considered representative of the soil water in adjacent areas, thus capable of producing surfaces that describe their spatial structure. According to Chapman (1996), all three variables are used in irrigation water quality assessment.

In this context, maps of specific conductance and pH were produced using all methods, for both wet and dry periods, along with the difference of groundwater elevation between the wet and the dry periods. The grid resolution for the output maps was set to 10 m, according to the size of the study area and the number of samples (Hengl 2006).

In view of BSS implementation, the global minimum of GCV for every parameter was reached by implementing the methodology for different numbers of segments m_x and m_y ($1 \leq m_x \leq 15$, $1 \leq m_y \leq 15$, $m_x + m_y \geq 6$) and minimizing GCV for each combination, by altering each one of the adjustable parameters, as detailed in Malamos and Koutsoyiannis (2016a). Additionally, we assessed larger values of m_x and m_y up to 30 segments in either direction (i.e. $16 \leq m_x \leq 30$ and $16 \leq m_y \leq 30$) by setting each smoothing parameter to its minimum value (0.001) or its maximum value (0.99) alternately (i.e. four different combinations) in order to reduce the computational effort required to implement the GCV minimization procedure. The results of the above procedure are presented in Table 4.

[Table 4]

Figure 7 presents the bilinear surface d acquired from the solution of equation (10), along x and y axes, by applying the obtained parameters for the specific conductance of the wet period. The available measurements are indicated with stars.

Figure 8 presents the specific conductance maps of the study area, acquired from all five methods. The spatial pattern of the specific conductance variation demonstrated by all methods throughout the study area, showed increased values in the northeast, i.e. near the sea, with smaller values in the south-southeast, i.e. upstream. Also, proximity along the main stream bed does not contribute to a variation in specific conductance as much as the distance to the sea, especially in the lower parts of the study area. Considering the temporal variation of the specific conductance, the decrease of its values during the wet period is evident throughout the study area, except for a small part located in the northeast where increased values were encountered due to possible sea water intrusion, as already mentioned. The similarities between the outputs of UK, OK and IDW are noticeable.

[Figure 7, Figure 8]

Considering the spatio-temporal variation of pH (Fig. 9), a significant spatial variability throughout the study area is clear during the dry period, due to the influence of the diverse agricultural activities. However, the influence of freshwater from upstream during the wet period led to clearly distinguished areas of low and high pH values. Also, during the wet period, alkali values are shown by all methods in the eastern parts of the study area.

[Figure 9]

Figure 10 presents the difference in water elevation between the wet and dry periods, i.e. the temporal fluctuation of the water surface inside the wells, throughout the study area. All methods showed a significant increase in water level upstream, up to 8.4 m, demonstrating the influence of the surface runoff originating from rainfall in the wet period. Considering the output of each method, BSS presented very plausible,

homogeneously distributed areas, while the kriging implementations produced almost identical maps. Both IDW and NN presented either “bull’s eye” or polygon shaped artefacts, inconsistent to physical interpretation.

[Figure 10]

The results of all five methodologies (Figs 8, 9 and 10) show the significant improvement in irrigation water quality at the end of the wet season in upstream locations, with the greatest increase in water level, suggesting that the use of wells located in these areas is preferred to those located downstream. This could be achieved by the use of a centralized irrigation system, collecting the surface runoff, along with groundwater pumping from the preferred areas, thus distributing irrigation water of adequate quality across the plain.

Even though the overall assessment demonstrated low irrigation water quality throughout the study area, frequent flooding of the well-drained loamy sand soil combined with the salts leaching caused by the excess irrigation applied by the farmers, suggest limiting factors to the salinity problems and soil structure degradation.

Evaluation of spatial interpolation methods

The evaluation of spatial interpolation methods using different statistical metrics may not be representative with respect to the validity of the interpolation results in other locations, except for those incorporated in the interpolation procedure.

In order to tackle this, leave-one-out (LOOCV) and two-fold cross-validation procedures were implemented for the evaluation of the five methods efficiency, based on the already presented criteria. LOOCV was accompanied by tests regarding normality and linear correlation of the residuals to the original data points. The

statistical criteria results, concerning each parameter, are summarized in Tables 5, 6 and 7.

[Table 5, Table 6, Table 7]

Based on the IPE3 metric (Tables 5, 6 and 7), the kriging implementations outperformed the rest in the cases of the specific conductance (wet and dry periods) and pH (dry period), while BSS performed better in the remaining cases. The IDW and NN methods produced the worst results in the specific conductance and water elevation cases.

Figure 11 demonstrates the normal probability plots of the empirical distribution function of the residuals for the pH estimates from UK (dry period) and BSS (wet period). For comparison, the corresponding theoretical normal distribution functions were also plotted. As can be seen, the residuals follow the normal distribution in both cases. Similar results, not shown here for brevity, were obtained for the residuals of all parameters in both periods, except for the residuals obtained by NN that presented a sizable deviation from the corresponding normal distribution. In contrast, the NN residuals were those with the smallest coefficient of determination (R^2) values amongst the other four methods, presenting negligible linear correlation between measured values and residuals. This can be explained by the fact that the predictions of the NN method at unsampled locations are those of the nearest single data point, so in the leave-one-out procedure only the nearest point to the one omitted is used, probably having a similar value, thus resulting in more uncorrelated residuals than the other methods delivered, using all the remaining points as predictors.

Furthermore, a two-fold cross-validation was performed as an extreme case of evaluation performance, where half of the dataset is used to estimate the other half and vice versa. For the BSS case, the dataset consisting of 104 data points was randomly

divided to two equal parts and, using each part successively as input, we obtained the estimations of the parameters at the locations of the other part. Each of the two validation rounds produced 52 estimations that were combined to produce the two-fold cross-validation output of the BSS. For the remaining methods, the two-fold cross-validation was performed by means of SAGA GIS while keeping the same analysis extent.

[Table 8]

Table 8 shows the RMSE values obtained by the two-fold cross-validation procedure. BSS clearly outperformed all other methods confirming the findings of (Malamos and Koutsoyiannis 2016b) that in the case of scarce data, the bilinear surface smoothing mathematical framework provides consistent results. The kriging implementations gave a similar performance, followed closely by IDW. The NN method presented the poorest results apart from the case of water level.

Conclusions

A survey of physico-chemical quality indices of irrigation water in a Mediterranean island catchment irrigated by water wells resulted in a dense dataset collected *in situ* in two rounds, one at the end of the dry season and one at the end of the wet season, using an In-Situ Inc. Multi-Parameter TROLL 9000E probe. During the first round, the water wells were located using a geographical information system and remote sensing data, together with interviews with local farmers. During the second round, knowledge of the exact locations and routes led to an optimal allocation of time and resources and thus a reduction in the time required to cover the study area, with an increased number of samples taken per day.

The acquired data constituted a test set for assessing the performance and the features of four well-established spatial interpolation methods, i.e. ordinary kriging, universal kriging, inverse distance weighting and nearest neighbours, against those of the recently introduced bilinear surface smoothing.

Furthermore, the derivation of the leave-one-out cross-validation residuals of BSS from the existing mathematical framework is also described, along with information concerning the implementation of the other four methods, in the open-source geographical information system SAGA.

The performance evaluation of each methodology took place on the basis of two types of cross-validation, i.e. leave-one-out and two-fold. The results of the leave-one-out procedure were evaluated using the IPE3 criterion, which is a combined evaluation vector comprising three traditional metrics. The IPE3 values revealed that the kriging implementations, especially universal kriging, performed better for specific conductance (wet and dry periods) and pH (dry period), while bested by BSS in the remaining cases. The IDW and NN methods produced the worst results in the specific conductance and water level cases.

For the two-fold cross-validation, BSS produced very good results, which, based on the acquired RMSE values, outperformed those of the other interpolation methods. Thus, BSS constitutes a good alternative to kriging in cases of data scarcity.

In every case, the spatio-temporal variability of water quality parameters and the need to monitor them at least twice a year were evident. Also, low irrigation water quality was reported throughout the study area as a direct consequence of vicinity to the sea. The results of all five methodologies show a significant improvement of all indices at the end of the wet season, especially in upstream locations with the greatest increase in water level. This finding suggests a centralized irrigation system should be used that

collects surface water runoff along with groundwater from the preferred areas and distributes it across the plain.

Although the BSS concept is simple, the overall performance against the other methods is quite satisfactory, indicating its applicability to provide factual information on the spatial distribution of water quality indices, even with scarce datasets. Also, further research considering the form of the variograms between wet and dry seasons will contribute to further assessment of the geostatistical properties of the aforementioned variables. This information, combined with appropriate agricultural practices, such as effective drainage and salts leaching from excess irrigation, suggest limiting of salinity problems and soil structure degradation and contribute to optimal irrigation water management.

Acknowledgements

We wish to kindly acknowledge the Editor Attilio Castellarin and the two anonymous reviewers for their thoughtful and thorough reviews which have considerably helped us to improve our manuscript during revision. Also, we acknowledge the contribution of George Paganelis and the farmers of Komi, Kalloni and Kato Klesma villages of Tinos. The dataset is available from the authors upon request. The maps included in the present study were produced using the free and open-source geographical information system QGIS (<http://www.qgis.org>).

References

- Bartram, J. and Ballance, R., 1996. *Water Quality Monitoring: A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes*. Taylor & Francis.
- Bohling, G., 2005. Introduction to geostatistics and variogram analysis. *Kansas geological survey*, 1–20.

- Burrough, P.A. and McDonnell, R.A., 1998. *Principles of Geographical Information Systems*. New York: Oxford University Press.
- Chapman, D. V, 1996. *Water Quality Assessments: A guide to the use of biota, sediments and water in environmental monitoring, Second Edition*. Taylor & Francis.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8 (7), 1991–2007.
- Craven, P. and Wahba, G., 1978. Smoothing noisy data with spline functions. *Numerische Mathematik*, 31 (4), 377–403.
- Domínguez, E., Dawson, C.W., Ramírez, A., and Abrahart, R.J., 2011. The search for orthogonal hydrological modelling metrics: a case study of 20 monitoring stations in Colombia. *Journal of Hydroinformatics*, 13, 429.
- Gong, G., Mattevada, S., and O’Bryant, S.E., 2014. Comparison of the accuracy of kriging and IDW interpolations in estimating groundwater arsenic concentrations in Texas. *Environmental Research*, 130, 59–69.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press.
- Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228 (1–2), 113–129.
- Hengl, T., 2006. Finding the right pixel size. *Computers and Geosciences*, 32 (9), 1283–1298.
- In-Situ Inc., 2005. *Multi-Parameter TROLL 9000, Operator’s Manual*. Ft. Collins, USA.
- Kis, I.M., 2016. Comparison of Ordinary and Universal Kriging interpolation techniques on a depth variable (a case of linear spatial trend) , case study of the Šandrovac Field. *The Mining-Geology-Petroleum Engineering Bulletin*, 41–58.
- Kitanidis, P.K., 1997. *Introduction to Geostatistics: Applications in Hydrogeology*. New York: Cambridge University Press.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F., 2006. World Map of the

- Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15 (3), 259–263.
- Kourgialas, N.N., Karatzas, G.P., and Koubouris, G.C., 2017. A GIS policy approach for assessing the effect of fertilizers on the quality of drinking and irrigation water and wellhead protection zones (Crete, Greece). *Journal of Environmental Management*, 189, 150–159.
- Lazarus Team, 2016. Lazarus: The professional Free Pascal RAD IDE. Version 1.6.
- Li, J. and Heap, A.D., 2008. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Geoscience Australia. GPO Box 378, Canberra, ACT 2601, Australia: Geoscience Australia.
- Li, J. and Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling and Software*, 53, 173–189.
- Loague, K. and Green, R.E., 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology*, 7 (1–2), 51–73.
- Malamos, N. and Koutsoyiannis, D., 2015. Broken line smoothing for data series interpolation by incorporating an explanatory variable with denser observations: application to soil-water and rainfall data. *Hydrological Sciences Journal*, 60 (3), 468–481.
- Malamos, N. and Koutsoyiannis, D., 2016a. Bilinear surface smoothing for spatial interpolation with optional incorporation of an explanatory variable. Part 1: Theory. *Hydrological Sciences Journal*, 61 (3), 519–526.
- Malamos, N. and Koutsoyiannis, D., 2016b. Bilinear surface smoothing for spatial interpolation with optional incorporation of an explanatory variable. Part 2: Application to synthesized and rainfall data. *Hydrological Sciences Journal*, 61 (3), 527–540.
- Malamos, N. and Nalbantis, I., 2005. *Analysis of water demand management practices*. The ODYSSEUS Project - integrated management of water systems in conjunction with an advanced information system, Vol. 15, operational program ‘Competitiveness’. Athens.
- Malamos, N., Tsirogiannis, I.L., Tegos, A., Efstratiadis, A., and Koutsoyiannis, D.,

2017. Spatial interpolation of potential evapotranspiration for precision irrigation purposes. *European Water*, (59), 303–309.
- Mamara, A., Anadranistakis, M., Argiriou, A.A., Szentimrey, T., Kovacs, T., Bezes, A., and Bihari, Z., 2017. High resolution air temperature climatology for Greece for the period 1971-2000. *Meteorological Applications*, 24 (2), 191–205.
- Mamara, A., Argiriou, A.A., and Anadranistakis, M., 2016. Recent trend analysis of mean air temperature in Greece based on homogenized data. *Theoretical and Applied Climatology*, 126 (3–4), 543–573.
- Matheron, G., 1969. *Le krigeage universel. Cahiers du Centre de Morphologie Mathématique*. École nationale supérieure des mines de Paris, Fontainebleau.
- Mir, A., Piri, J., and Kisi, O., 2017. Spatial monitoring and zoning water quality of Sistan River in the wet and dry years using GIS and geostatistics. *Computers and Electronics in Agriculture*, 135, 38–50.
- Murphy, R.R., Curriero, F.C., and Ball, W.P., 2010. Comparison of Spatial Interpolation Methods for Water Quality Evaluation in the Chesapeake Bay. *Journal of Environmental Engineering*, 136 (2), 160–171.
- Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, 10 (3), 282–290.
- Oliver, M.A. and Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *CATENA*, 113, 56–69.
- Price, D.T., McKenney, D.W., Nalder, I. a., Hutchinson, M.F., and Kesteven, J.L., 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agricultural and Forest Meteorology*, 101 (2–3), 81–94.
- Shahabi, S. and Anagnostopoulos, K., 1973. Soil map of Tinos Island.
- Stahl, K., Moore, R.D., Floyer, J. a., Asplin, M.G., and McKendry, I.G., 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and Forest Meteorology*, 139 (3–4), 224–236.
- Tabios, G.Q. and Salas, J.D., 1985. A Comparative Analysis Of Techniques For Spatial

- Interpolation Of Precipitation. *Journal of the American Water Resources Association*, 21 (3), 365–380.
- Tegos, A., Malamos, N., Efstratiadis, A., Tsoukalas, I., Karanasios, A., and Koutsoyiannis, D., 2017. Parametric Modelling of Potential Evapotranspiration: A Global Survey. *Water*, 9 (10).
- Tegos, A., Malamos, N., and Koutsoyiannis, D., 2015. A parsimonious regional parametric evapotranspiration model based on a simplification of the Penman–Monteith formula. *Journal of Hydrology*, 524, 708–717.
- Vicente-Serrano, S., Saz-Sánchez, M., and Cuadrat, J., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate Research*, 24, 161–180.
- Wahba, G. and Wendelberger, J., 1980. Some New Mathematical Methods for Variational Objective Analysis Using Splines and Cross Validation. *Monthly Weather Review*, 108 (8), 1122–1143.
- Willmott, C.J., 1982. Some Comments on the Evaluation of Model Performance. *Bulletin of the American Meteorological Society*, 63 (11), 1309–1313.
- Yidana, S.M., Banoeng-Yakubo, B., Aliou, A.-S., and Akabzaa, T.M., 2012. Groundwater quality in some Voltaian and Birimian aquifers in northern Ghana—application of multivariate statistical methods and geographic information systems. *Hydrological Sciences Journal*, 57 (6), 1168–1183.

Table 1. Synopsis of dry period results (August 2007). DO: dissolved oxygen.

| | Maximum | Minimum | Average | % of measurements exceeding the average |
|--|---------|---------|---------|---|
| Water elevation (m) | 101.8 | 10.6 | 19.2 | 24.0 |
| Specific conductance ($\mu\text{S}/\text{cm}$) | 4769.5 | 1159.3 | 1757 | 36.5 |
| Temperature ($^{\circ}\text{C}$) | 22.8 | 15.2 | 18.2 | 47.1 |
| Turbidity (NTU) | 16.6 | 0.4 | 1.0 | 21.2 |
| Oxidation–reduction potential (mV) | 193 | –274 | 70.1 | 62.5 |
| pH | 8.22 | 6.78 | 7.56 | 63.5 |
| DO ($\mu\text{g}/\text{L}$) | 9676 | 608 | 3756 | 48.1 |

Table 2. Synopsis of wet period results (March 2008).

| | Maximum | Minimum | Average | % of measurements exceeding the average |
|--|---------|---------|---------|---|
| Water elevation (m) | 101.9 | 11.5 | 21.6 | 32.7 |
| Specific conductance ($\mu\text{S}/\text{cm}$) | 7205 | 851.6 | 1684.6 | 32.7 |
| Temperature ($^{\circ}\text{C}$) | 19.19 | 12.21 | 15.6 | 53.8 |
| Turbidity (NTU) | 2.9 | 0 | 0.4 | 35.6 |
| Oxidation–reduction potential (mV) | 228 | –159 | 93.7 | 57.7 |
| pH | 8.65 | 5.81 | 7.45 | 43.3 |
| DO ($\mu\text{g}/\text{L}$) | 10398 | 569 | 4296.6 | 46.2 |

Table 3. Coefficient of determination (R^2) of the linear regression between measurements and projected geographical coordinates.

| | Dry period | | Wet period | |
|--|--------------|--------------|--------------|--------------|
| | x coordinate | y coordinate | x coordinate | y coordinate |
| Water elevation (m) | 0.002 | 0.331 | 0 | 0.457 |
| Specific conductance ($\mu\text{S}/\text{cm}$) | 0 | 0.202 | 0 | 0.228 |
| Temperature ($^{\circ}\text{C}$) | 0.046 | 0.375 | 0.167 | 0.01 |
| Turbidity (NTU) | 0.006 | 0.013 | 0.044 | 0 |
| Oxidation–reduction potential (mV) | 0.01 | 0.266 | 0.236 | 0.117 |
| pH | 0.037 | 0.062 | 0.123 | 0.263 |
| DO ($\mu\text{g}/\text{L}$) | 0.008 | 0.038 | 0.003 | 0.330 |

Table 4. BSS optimal parameter values and performance indices.

| Parameter / Period | Number of segments, mx | Number of segments, my | $\tau_{\lambda x}$ | $\tau_{\lambda y}$ | Global minimum GCV |
|----------------------|------------------------|------------------------|--------------------|--------------------|----------------------|
| Specific conductance | | | | | |
| Wet | 7 | 27 | 0.99 | 0.001 | 3.4×10^5 |
| Dry | 4 | 12 | 0.99 | 0.001 | 2.3×10^5 |
| pH | | | | | |
| Wet | 6 | 3 | 0.001 | 0.001 | 1.4×10^{-1} |
| Dry | 1 | 13 | 0.001 | 0.001 | 8.0×10^{-2} |
| Water elevation | | | | | |
| Wet | 4 | 2 | 0.001 | 0.001 | 1.5 |
| Dry | 4 | 2 | 0.001 | 0.001 | 2.1 |

Table 5. Leave-one-out cross-validation statistics for both periods – specific conductance.

| Interpolation method | MBE ($\mu\text{S}/\text{cm}$) | MAE ($\mu\text{S}/\text{cm}$) | RMSE | EF | IPE3 |
|----------------------|---------------------------------|---------------------------------|-------|-------|------|
| <i>Dry period</i> | | | | | |
| BSS | 3.0 | 293.6 | 457.0 | 0.402 | 0.68 |
| OK | 10.8 | 267.3 | 437.2 | 0.453 | 0.65 |
| UK* | 9.9 | 255.6 | 429.7 | 0.472 | 0.63 |
| IDW | 44.6 | 255.3 | 432.8 | 0.464 | 0.85 |
| NN | 21.9 | 281.8 | 569.5 | 0.072 | 0.86 |
| <i>Wet period</i> | | | | | |
| BSS | 8.3 | 354.6 | 602.0 | 0.514 | 0.74 |
| OK* | 16.8 | 288.8 | 534.9 | 0.616 | 0.62 |
| UK | 18.0 | 292.6 | 538.0 | 0.612 | 0.63 |
| IDW | 64.6 | 293.8 | 561.7 | 0.577 | 0.75 |
| NN | 98.3 | 368.9 | 745.7 | 0.255 | 0.99 |

* denotes best performance according to IPE3.

Table 6. Leave-one-out cross-validation statistics for both periods – pH.

| Interpolation method | MBE | MAE | RMSE | EF | IPE3 |
|----------------------|-------|------|------|-------|------|
| <i>Dry period</i> | | | | | |
| BSS | 0.00 | 0.18 | 0.27 | 0.325 | 0.82 |
| OK | 0.02 | 0.15 | 0.24 | 0.488 | 0.88 |
| UK* | 0.00 | 0.15 | 0.23 | 0.511 | 0.66 |
| IDW | 0.01 | 0.15 | 0.24 | 0.478 | 0.77 |
| NN | 0.00 | 0.14 | 0.26 | 0.358 | 0.76 |
| <i>Wet period</i> | | | | | |
| BSS* | 0.01 | 0.27 | 0.37 | 0.749 | 0.61 |
| OK | 0.00 | 0.24 | 0.43 | 0.690 | 0.65 |
| UK | 0.01 | 0.25 | 0.43 | 0.681 | 0.68 |
| IDW | 0.00 | 0.24 | 0.41 | 0.718 | 0.62 |
| NN | -0.03 | 0.25 | 0.52 | 0.550 | 1.00 |

* denotes best performance according to IPE3.

Table 7. Leave-one-out cross-validation statistics for both periods – groundwater elevation.

| Interpolation method | MBE (m) | MAE (m) | RMSE | EF | IPE3 |
|----------------------|---------|---------|------|-------|------|
| <i>Dry period</i> | | | | | |
| BSS* | -0.33 | 1.76 | 6.13 | 0.659 | 0.57 |
| OK | -0.52 | 1.65 | 7.05 | 0.549 | 0.72 |
| UK | -0.52 | 1.65 | 7.04 | 0.550 | 0.72 |
| IDW | -1.16 | 1.90 | 8.14 | 0.399 | 0.99 |
| NN | -1.18 | 1.85 | 6.95 | 0.562 | 0.85 |
| <i>Wet period</i> | | | | | |
| BSS* | -0.39 | 1.61 | 6.12 | 0.685 | 0.61 |
| OK | -0.44 | 1.55 | 6.45 | 0.650 | 0.66 |
| UK | -0.44 | 1.55 | 6.43 | 0.652 | 0.66 |
| IDW | -1.12 | 1.75 | 7.85 | 0.482 | 1.00 |
| NN | -1.12 | 1.75 | 6.75 | 0.616 | 0.86 |

* denotes best performance according to IPE3.

Table 8. Two-fold cross-validation RMSE values for all parameters and both periods.

| Interpolation method | Specific conductance | | pH | | Water elevation | |
|----------------------|----------------------|-------|------|------|-----------------|------|
| | Dry | Wet | Dry | Wet | Dry | Wet |
| BSS* | 289.2 | 401.3 | 0.28 | 0.26 | 2.24 | 0.99 |
| OK | 489.6 | 577.7 | 0.29 | 0.40 | 6.44 | 6.07 |
| UK | 489.6 | 589.5 | 0.30 | 0.40 | 6.43 | 6.06 |
| IDW | 501.2 | 584.2 | 0.29 | 0.40 | 8.22 | 7.94 |
| NN | 693.1 | 824.7 | 0.35 | 0.53 | 7.10 | 6.90 |

* denotes best performance according to RMSE.

Figure 1. Study area.

Figure 2. Monthly variation of maximum (T_{\max}), average (T_{aver}), and minimum (T_{\min}) temperature, and rainfall of the study area.

Figure 3. Sampling points and altitude of the study area.

Figure 4. Ratio of specific conductance between the two measurement periods (wet *versus* dry).

Figure 5. Ratio of water level between the two measurement periods (wet *versus* dry).

Figure 6. Ratio of dissolved oxygen between the two measurement periods (wet *versus* dry).

Figure 7. Bilinear surface d fitted to the 104 specific conductance data points (stars) (minimum GCV: $m_x = 7$, $m_y = 27$) for the wet period.

Figure 8. Maps of specific conductance for both periods and all methods: (a, b) BSS; (c, d) universal kriging; (e, f) ordinary kriging; (g, h) inverse distance weighting; and (i, j) nearest neighbours.

Figure 9. Maps of pH for both periods and all methods: (a, b) BSS; (c, d) universal kriging; (e, f) ordinary kriging; (g, h) inverse distance weighting; and (i, j) nearest neighbours.

Figure 10. Fluctuation of groundwater elevation above sea level between periods for all methods: (a) BSS, (b) UK, (c) OK, (d) IDW, and (e) NN.

Figure 11. Normal probability plots of the pH empirical distribution functions of: (a) UK and (b) BSS residuals, using Weibull plotting positions against the corresponding normal distribution functions.

Figures

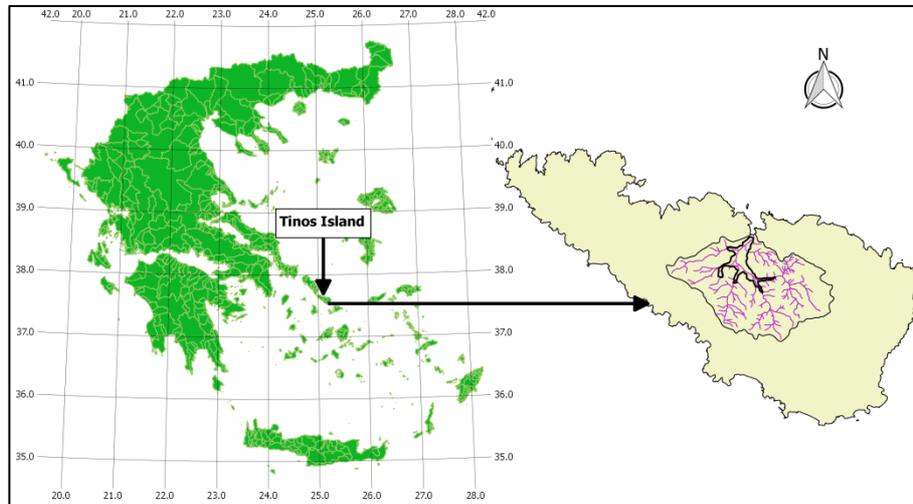


Figure 1. Study area

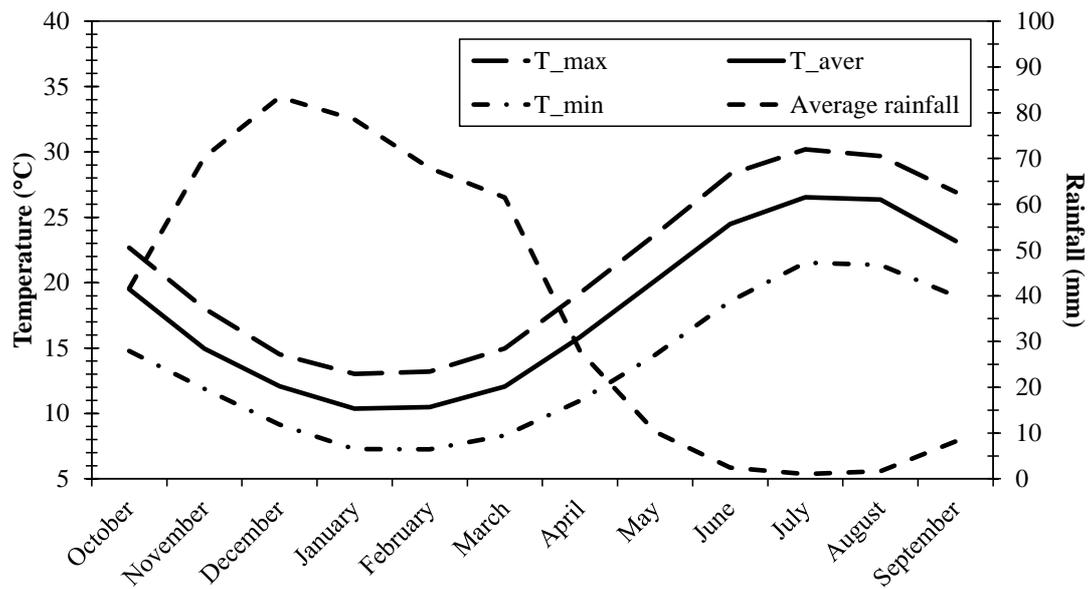


Figure 2. Monthly variation of maximum (T_{max}), average (T_{aver}), minimum (T_{min}) temperature and rainfall of the study area

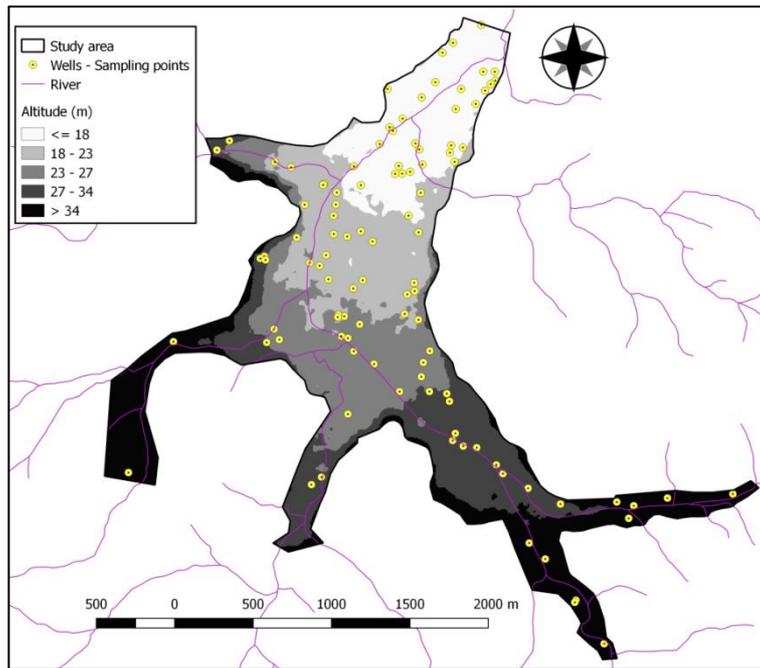


Figure 3. Sampling points and altitude of the study area

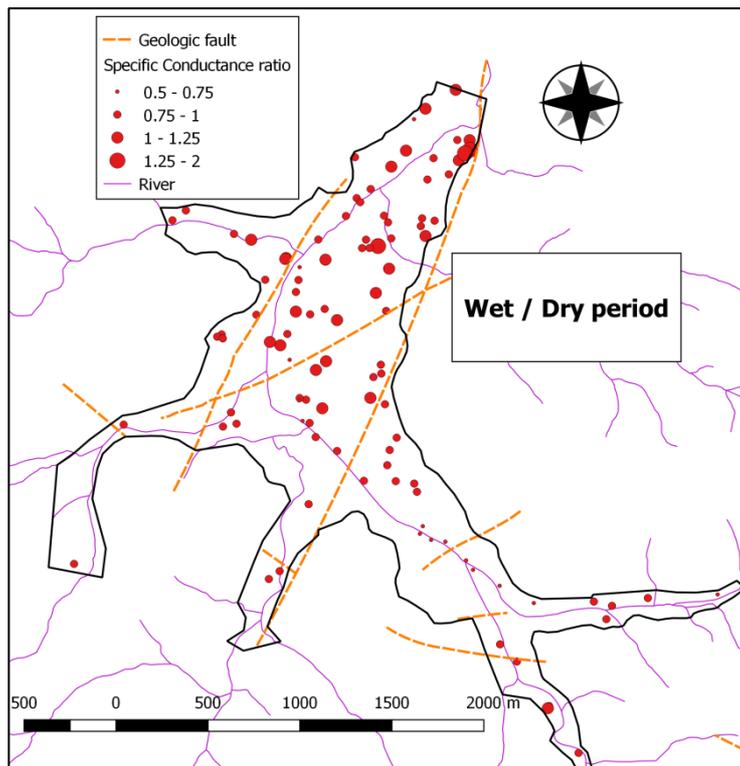


Figure 4. Ratio of specific conductance between the two measurement periods (wet versus dry)

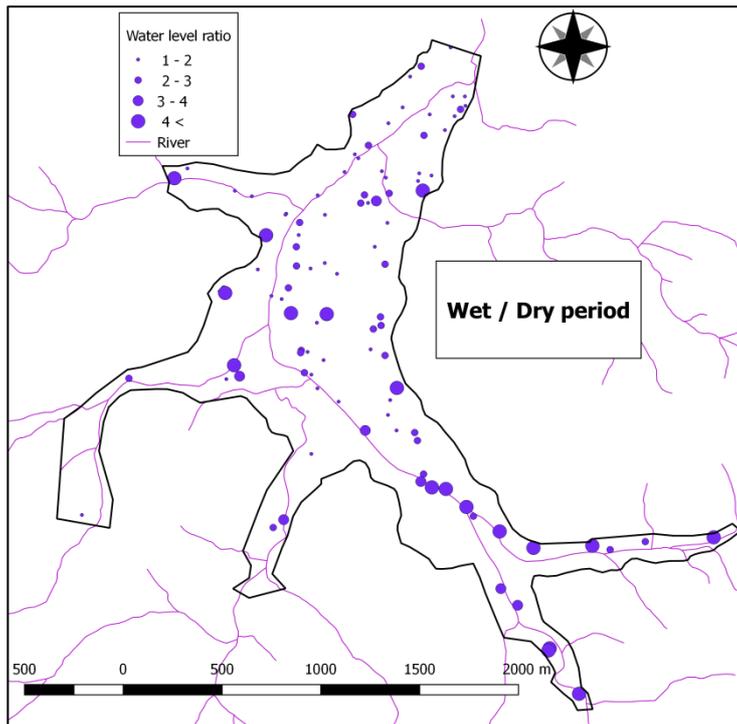


Figure 5. Ratio of water level between the two measurement periods (wet versus dry)

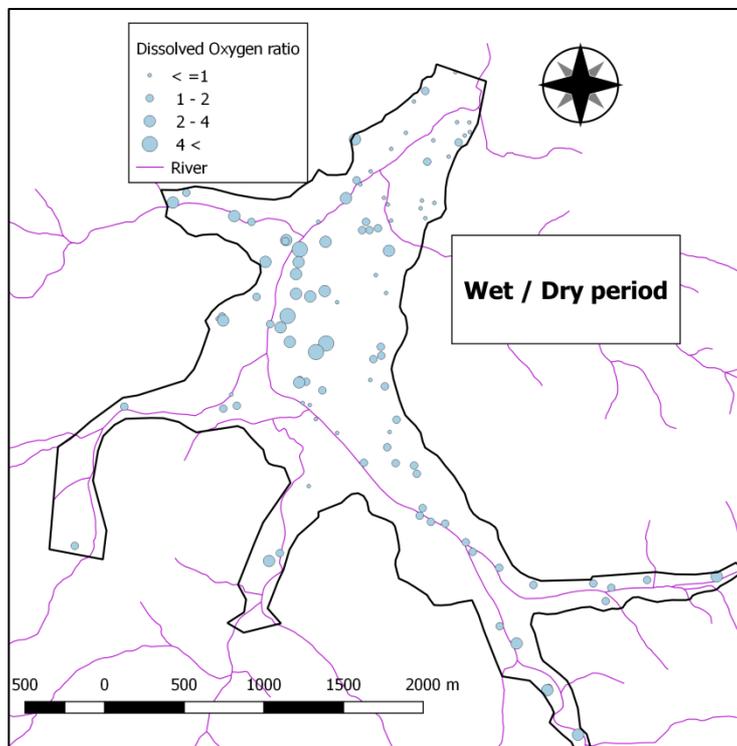


Figure 6. Ratio of dissolved oxygen between the two measurement periods (wet versus dry)

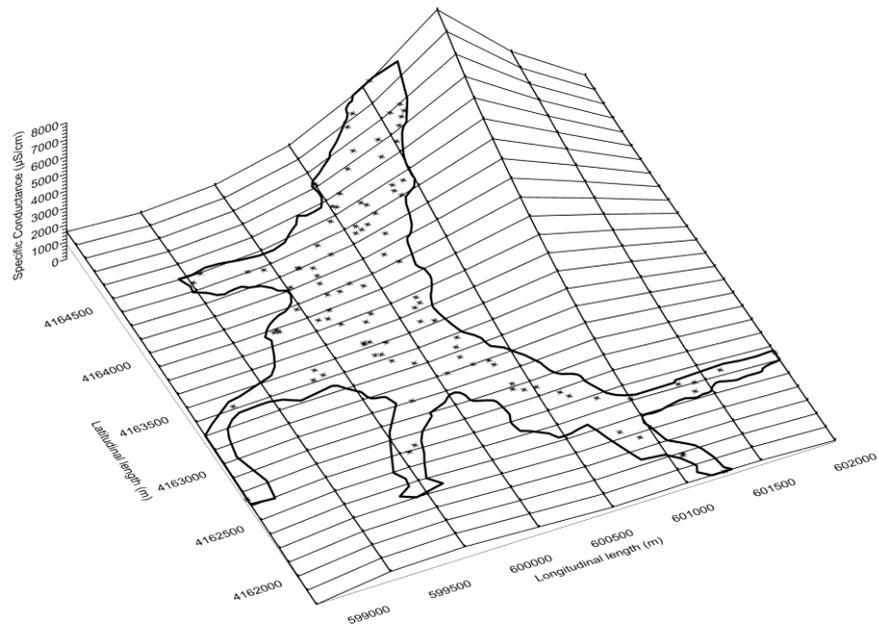


Figure 7. Bilinear surface d fitted to the 104 specific conductance data points (stars) (minimum GCV: $mx = 7$, $my = 27$) for the wet period

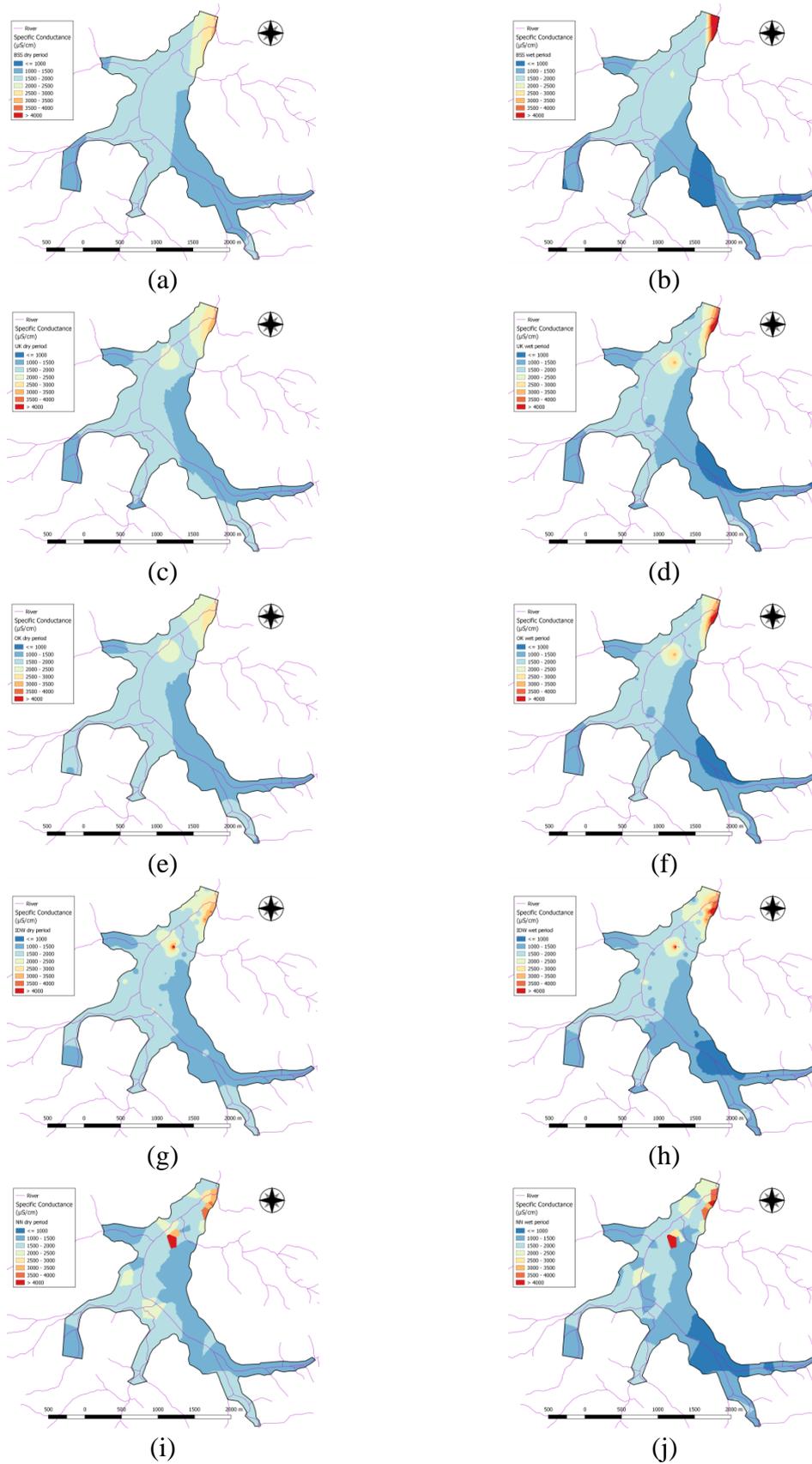


Figure 8. Specific conductance maps for both periods and all methods (a, b: BSS; c, d: UK; e, f: OK; g, h: IDW, i, j: NN)

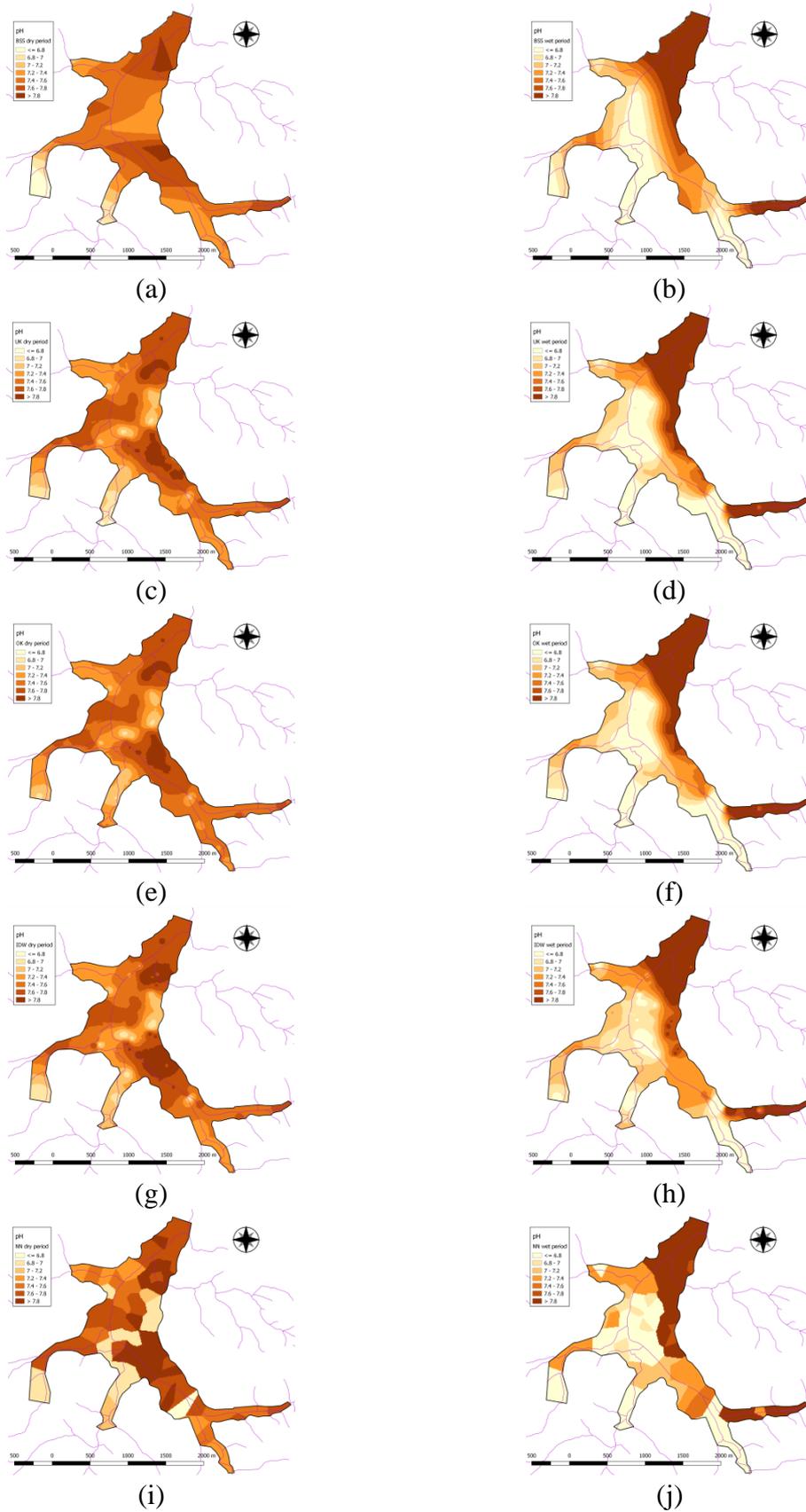


Figure 9. pH maps for both periods and all methods (a, b: BSS; c, d: UK; e, f: OK; g, h: IDW, i, j: NN)

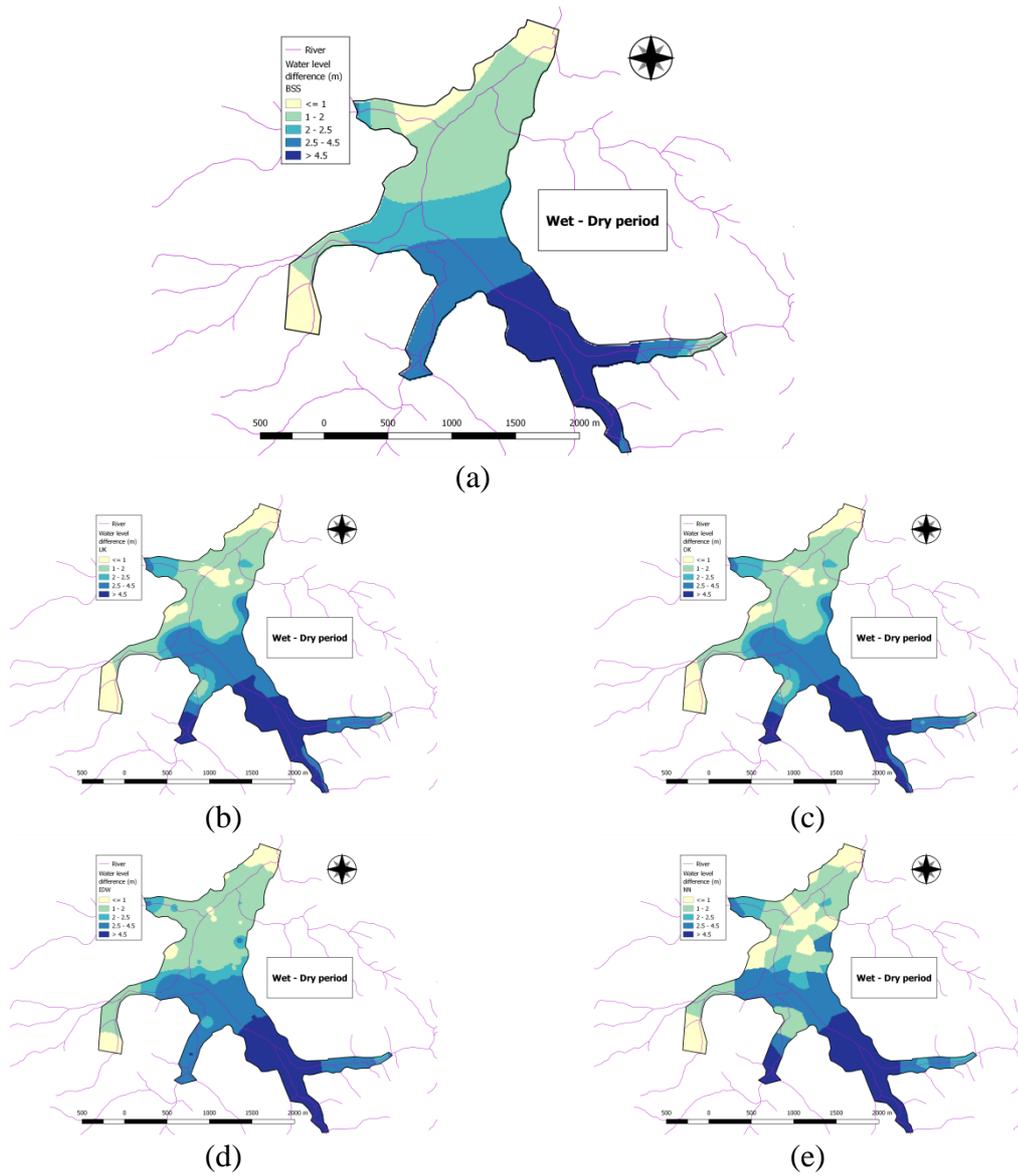
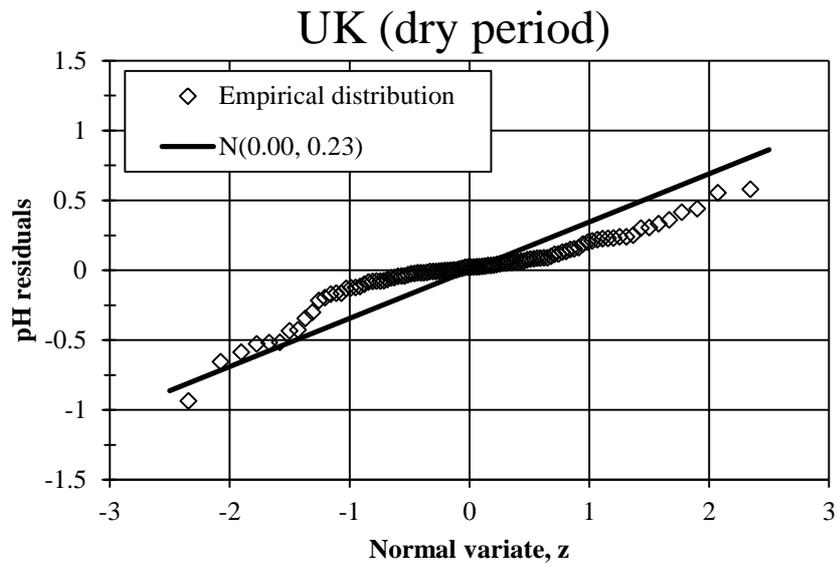
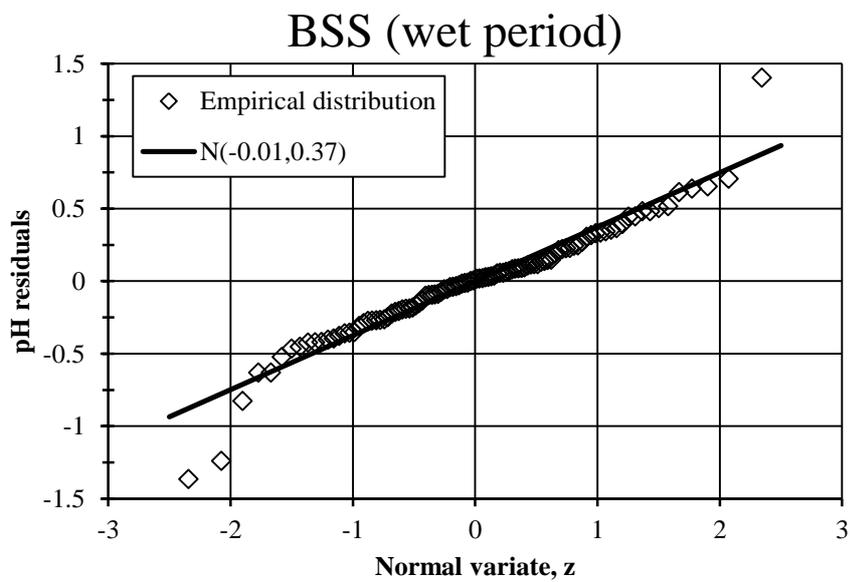


Figure 10. Fluctuation of groundwater elevation above sea level between periods for all methods (a: BSS; b: UK; c: OK; d: IDW, e: NN)



(a)



(b)

Figure 11. Normal probability plots of the pH empirical distribution functions of: (a) UK and (b) BSS residuals using Weibull plotting positions against the corresponding normal distribution functions