RESEARCH ARTICLE

Comparing the performance of geostatistical models with additional information from covariates for sewage plume characterization

Maurici Del Monego · Paulo Justiniano Ribeiro Jr. · Patrícia Ramos

Abstract In this work, kriging with covariates is used to model and map the spatial distribution of salinity measurements gathered by an autonomous underwater vehicle in a sea outfall monitoring campaign aiming to distinguish the effluent plume from the receiving waters and characterize its spatial variability in the vicinity of the discharge. Four different geostatistical linear models for salinity were assumed, where the distance to diffuser, the west-east positioning, and the south-north positioning were used as covariates. Sample variograms were fitted by the Matèrn models using weighted least squares and maximum likelihood estimation methods as a way to detect eventual discrepancies. Typically, the maximum likelihood method estimated very low ranges which have limited the kriging process. So, at least for these data sets, weighted least squares showed to be the most appropriate estimation method for variogram fitting. The kriged maps show clearly the spatial variation of salinity, and it is possible to identify the effluent plume in the area studied. The results obtained show some guidelines for sewage monitoring if a geostatistical analysis of the data is in mind. It is important to treat properly the existence of anomalous values and to adopt a sampling strategy that includes transects parallel and perpendicular to the effluent dispersion.

M. Del Monego
Departamento de Química e Biologia, Universidade Tecnológica Federal do Paraná, Curitiba, Paraná CEP 81280-340, Brazil
e-mail: maurici@utfpr.edu.br

P. J. Ribeiro Jr.
Laboratório de Estatística e Geoinformacão, Universidade Federal do Paraná, Curitiba, Paraná CEP 81531-990, Brazil
e-mail: paulojus@ufpr.br

P. Ramos (✉)
INESC Technology and Science and ISCAP/Polytechnic of Porto, 4465-004 S. Mamede de Infesta, Portugal
e-mail: pramos@inescporto.pt

Introduction

The typical behavior of a sea outfall discharge is in a form of a process that can be described as follows. The wastewater is usually ejected as an array of turbulent buoyant jets from ports spaced along the outfall diffuser. These turbulent buoyant jets mix with the ambient seawater, usually resulting in rapid reductions of contaminant concentrations. The seawater around coastal outfalls is often density-stratified with density increasing with depth. The discharge, whose density is close to fresh water, is lighter than the surrounding ambient and the plumes rise due to buoyant forces until they reach a level of neutral buoyancy where the effluent spreads laterally, creating a horizontal spreading layer. Depending on the strength of seawater stratification, currents, and other variables, the horizontal spreading layer may be submerged and will not be visible on the water surface. If the receiving waters are homogeneous or weakly stratified, the plumes will reach the surface and spread horizontally away from the source (Tian et al. 2004a, b; Daviero and Roberts 2006; Tian et al. 2006; Hunt et al. 2010).

The physical and biological coastal processes that determine the values of environmental variables are complex and

still poorly understood (Niu et al. 2007; Rogowski et al. 2012). Usually, the observations are so unpredictable that the spatial distribution of these variables appears to be random (Washburn et al. 1992; Wu et al. 1994; Carvalho et al. 2002; Petrenko et al. 1998; Ramos et al. 2007; Ramos and Neves 2009). This randomness makes deterministic solutions adequate to describe only the simplest components (Goovaerts 1997; Webster and Oliver 2007). The measurements of environmental variables are usually obtained in very restricted areas and constitute a sample from a continuum that cannot be recorded everywhere. Yet, people aim to predict the spatial distribution of these variables from a more or less sparse data set. The maps made by using the classical methods of spatial interpolation usually display the spatial variation poorly. The interpolators of these methods also fail to provide any estimates of the error, which are desirable for prediction (Webster and Oliver 2007; Isaaks and Srivastava 1989). In this line of thought, geostatistical prediction is the logical solution in the sense that it provides a quantitative description of how the environmental variable varies spatially and a prediction for the places that were not sampled (Bivand et al. 2008). Additionally, geostatistics also provides estimates of the errors of these predictions, so that we can assess the uncertainty associated to the predicted plume behavior.

Methodology

REMUS AUV

Autonomous underwater vehicles (AUVs) are particularly useful to detect and map sewage plumes because of their easier field logistics, reduced cost per study, increased spatial resolution, reduced spatial aliasing effects, and adaptive sampling capabilities (Ramos and Abreu 2010). *Isurus* is a REMUS class AUV acquired in the Woods Hole Oceanographic Institution, USA (see Fig. 1). These vehicles are low-cost lightweight AUVs specially designed for coastal water monitoring. Their reduced weight and dimensions make them extremely easy to handle, requiring no special equipment for launching and recovery. Isurus has a diameter


Fig. 1 Isurus AUV

of 20 cm and is about 1.5 m long, weighing about 35 kg in air. The maximum forward speed of the vehicle is 2 m/s; however, the best energy efficiency is achieved at about 1 m/s. At this velocity, the energy provided by a set of rechargeable lithium-ion batteries may last for over 20 h (equivalent to traverse distances of over 40 nautical miles). Although small in size, this vehicle can accommodate a wide range of oceanographic sensors, according to mission objectives. For the field experiment described in this paper, two specific sensors were integrated to measure the height from the bottom: a CTD (conductivity, temperature, depth), OS200 model from Ocean Sensors, Inc., USA, and an altimeter, from Imagenex, Inc., Canada (see measuring details in Ramos and Abreu (2010)).

Study site

The S. Jacinto outfall is located off the Portuguese west coast near to the Aveiro estuary (see the map in Fig. 2). The total length of the outfall, including the diffuser, is 3378 m (the first 3135 m section has a diameter of 1600 mm and the last 243 m section has a diameter of 1200 mm). The diffuser, which consists of 72 ports alternating on each side, nominally 0.175 m in diameter, is 332.5 m long. Currently, only the last 20 of the 72 ports are working in a length of 98.2 m. These are discharging upwards at an angle of 30° to the horizontal axis; the port height is about 1.3 m. The outfall has a true bearing direction of 290° and is discharging at a depth varying between approximately 14 and 17 m. The sea floor near to the diffuser is moderately sloped, with a sandy bottom and isobaths that are parallel to the coastline. In that area, the coastline itself runs at about a 200° angle with respect to true north. Flow variation through the outfall in question is not typical of WWTPs since the effluent is mainly of industrial origin. Effluent flowrate ranges most frequently between 0.6 and 0.8 m^3/s. During the campaign, the discharge remained fairly constant with an average flowrate of 0.61 m^3/s.

Figure 2 shows a plan view of the AUV's position estimate during the plume tracking survey. A rectangular area of approximately 200 × 100 m^2 was covered starting 20 m downstream from the middle point of the outfall diffuser (located at point (0,0)). Salinity measurements were obtained at depths of 2 and 4 m where the effluent plume was predicted to be horizontally dispersing. In each horizontal trajectory, the vehicle described six parallel transects that were perpendicular to the direction of the current, 200 m in length and at 20 m intervals. When performing horizontal trajectories, vertical oscillations of the AUV were less than 1 m (up and down) in the 2 m survey and less than 1 m–down and less than 1.5 m–up in the 4 m survey. In the 2 m trajectory, the average depth of the AUV was 2.0 m with a standard deviation of 0.20 m. In the 4 m trajectory, the
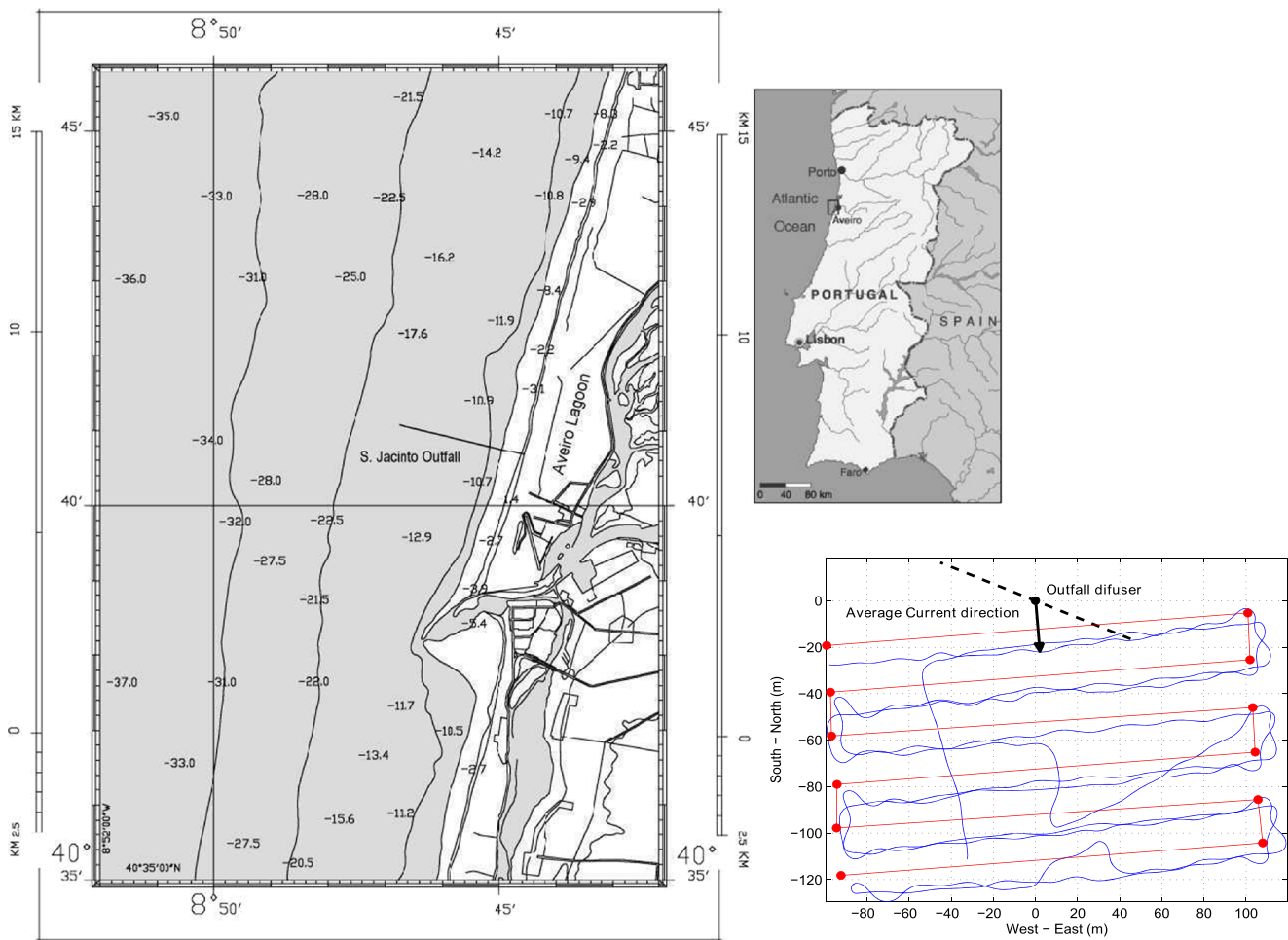
Fig. 2 Map of the study site and plan view of the AUV's position estimate during the plume tracking survey

average depth was 4.0 m with a standard deviation of 0.33 m. During the mission, the vehicle moved at a fairly constant velocity of 1 m/s (2 knots). Salinity data were recorded at a rate of 2.4 Hz. The geostatistical analysis was carried using the statistical software R programming language (R Development Core Team 2013) and the specialized package geoR (Ribeiro and Diggle 2001).

Results and discussions

Exploratory analysis

Exploratory analysis allows to identify the presence of spatial dependence, the possible occurrence of anomalous values, the existence of asymmetries in the distribution, and the presence of any trends. Table 1 shows the summary statistics of the two data sets gathered by the AUV at 2 and 4 m depths. At the 2 m depth, the salinity ranges from 35.15 to 35.61 psu. The mean value of the data set is 35.45 psu, which is very close to the median value (35.46 psu). The

skewness coefficient is relatively low (−0.53), indicating that the distribution is only slightly negatively skewed. The low value of the variation coefficient (0.183) reflects the fact

Table 1 Summary statistics of salinity measured at depths of 2 and 4 m

Summary	Salinity (2 m)	Salinity (4 m)
Samples	2400	2803
Minimum	35.15 psu	35.10 psu
Lower quartile	35.41 psu	35.54 psu
Mean	35.45 psu	35.57 psu
Upper quartile	35.50 psu	35.61 psu
Median	35.46 psu	35.57 psu
Maximum	35.61 psu	35.71 psu
Skewness coefficient	−0.53	−1.32
Kurtosis	0.22	6.43
Variation coefficient	0.183	0.147
Variance (10^{-3})	4.19	2.74
Standard deviation (10^{-2})	6.47	5.23

that the histogram does not have a tail of low values. At the 4 m depth, the salinity ranges from 35.10 to 35.71 psu. The mean value of the data set is 35.57 psu, which is equal to the median value. The skewness coefficient is a little higher than the one at the 2 m depth (−1.32), indicating that the asymmetry increased with depth. The variation coefficient is also quite low as desired (0.147).

A vertical profile of the background salinity at the vicinity of the diffuser was obtained by the AUV at the end of the campaign. When considering the maximum vertical oscillations of the AUV in performing the horizontal trajectories, the background salinity at the 2 m depth ranged from 35.52 to 35.56 psu and at the 4 m depth ranged from 35.61 to 35.66 psu. At the 2 m depth, the percentage of salinity measurements below 35.52 psu was 84.8 % and, at the 4 m depth, the percentage of salinity measurements below 35.61 psu was 75.6 %, which indicates the presence of the sewage plume at both depths and explains the negative skewness due to some lower values. When data are lognormally distributed, logarithmic transformation of the data may be advisable (Lloyd 2011). Alternatively, more general transformations such as the Box-Cox family, which include the logarithmic transformation, may be used (Diggle and Ribeiro 2007). We applied several appropriate transformations of the Box-Cox family to both data sets. However, no significant changes were found on the results, and we decided to use the raw data without any transformations.

By dividing the 2 m depth data set into quartiles and making scatterplots of the measurements position of each subset (south-north positioning versus west-east positioning), we observed that values below the lower quartile were located at the center of the survey area and values above the median

were located at the lateral edges, which suggested the existence of a spatial pattern at this depth. A similar analysis was made for the 4 m depth measurements, and we concluded that this behavior was not so evident. To explore these spatial patterns more clearly, salinity versus west-east positioning and south-north positioning was plotted (see the scatterplots in Fig. 3). The scatterplot of salinity measured at the 2 m depth versus the west-east coordinate shows clearly a spatial variability with salinity increasing from the center of the survey area to the lateral edges. In the scatterplot of salinity measured at the 2 m depth versus the south-north coordinate, salinity tends to increase as the south-north positioning decreases, i.e, as we move away from the diffuser. Although not so evident, these trends are also visible in the 4 m depth scatterplots. To investigate the relationship between salinity and distance to diffuser, three-dimensional Euclidean distances between the location of each measurement and the ends and middle point of the diffuser were computed, and the minimum of these three distances was considered. From the scatterplot, it can be seen that salinity increases as the distance to the diffuser increases and that is more evident at the 2 m depth. The values of Pearson and Spearman correlation coefficients between these two variables at the 2 m depth are 0.58 and 0.56, respectively. At the 4 m depth, these coefficients are 0.26 and 0.22, respectively. Therefore, we decided to account for such trends in the subsequent analyses.

Variogram modeling and estimation

When there is a large-scale component of spatial variability, that should be represented by a deterministic function in the

Fig. 3 Scatter plots of salinity at depths of 2 m (*upper panel*) and 4 m (*lower panel*) versus west-east positioning, south-north positioning, and distance to diffuser
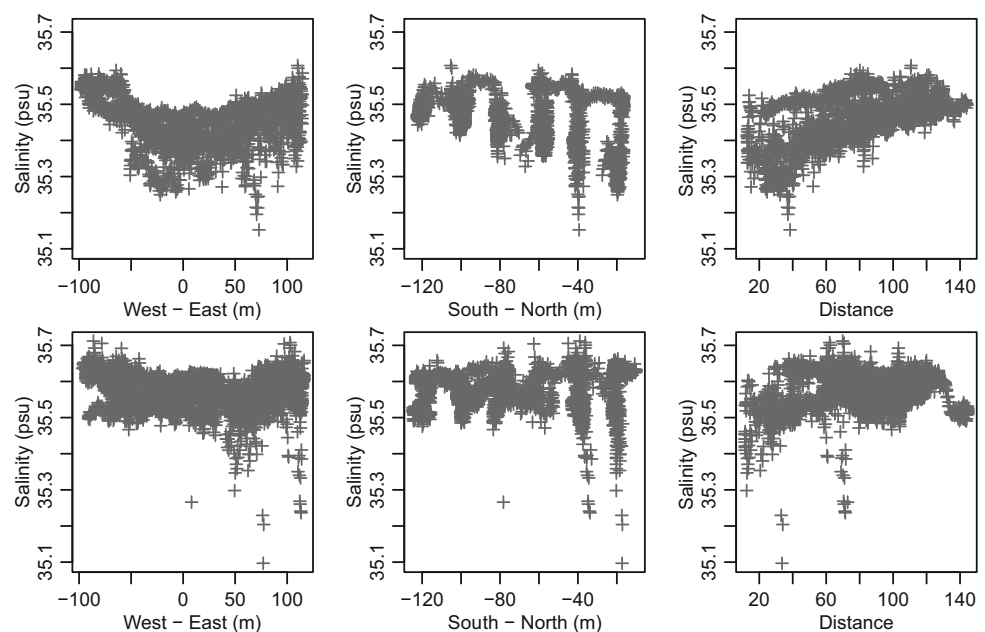
geostatistic model (Webster and Oliver 2007). The assumption of variable mean usually leads to better estimation, i.e., improved accuracy of predictions and more realistic evaluation of the mean square error (Kitanidis 1997). According to the exploratory analysis and to evaluate more clearly the effect of the distance to diffuser (DD) and the south-north and west-east positionings, the following geostatistical

$$Z(\mathbf{s}) = m(\mathbf{s}) + R(\mathbf{s}) \qquad (1)$$

where $Z(\mathbf{s})$ is salinity at location $\mathbf{s} = (x, y)$–x is the west-east coordinate and y is the south-north coordinate–$R(\mathbf{s})$ is the residual, and $m(\mathbf{s})$ is a deterministic function of the covariates DD, x and y, defined as follows:

- Model A: $m(\mathbf{s}) = m$, where m is a constant;
- Model B: $m(\mathbf{s}) = \beta_0 + \beta_1 DD$;
- Model C: $m(\mathbf{s}) = \beta_0 + \beta_1 DD + \beta_2 x + \beta_3 x^2$;
- Model D: $m(\mathbf{s}) = \beta_0 + \beta_1 DD + \beta_2 x + \beta_3 y + \beta_4 x^2 + \beta_5 y^2 + \beta_6 x y$.

Here, $\beta_0, \ldots, \beta_6$ are the drift coefficients. Model A has the simplest possible structure and was defined to serve as

a benchmarking for assessing the gains of the other models. Ordinary least squares (OLS) regression was used to estimate $m(\mathbf{s})$ (in the case of models B, C, and D), and the OLS residuals were used as input to the subsequent semivariogram (or hereafter, variogram) calculations. The influence of covariates on salinity was clearer by plotting histograms of OLS residuals of each model (B, C, and D) and scatterplots of OLS residuals of each model versus west-east and south-north coordinates. The same plots were obtained for salinity data in the case of model A. A significant difference between the plots of model A and the plots of model B for the 2 m depth measurements was observed which reinforced the inclusion of DD covariate. The histogram of the OLS residuals of model C was closer to a symmetric shape than the histogram of the OLS residuals of model B, and the scatterplot of OLS residuals versus west-east positioning of model C did not show any spatial variability, unlike the corresponding plot of model B, justifying clearly the inclusion of the quadratic function of x coordinate for the 2 m depth measurements. No significant differences were observed between the plots of model C and the

Fig. 4 Omnidirectional variograms of salinity/OLS residuals at 2 m depth fitted by the Matèrn models using WLS and ML
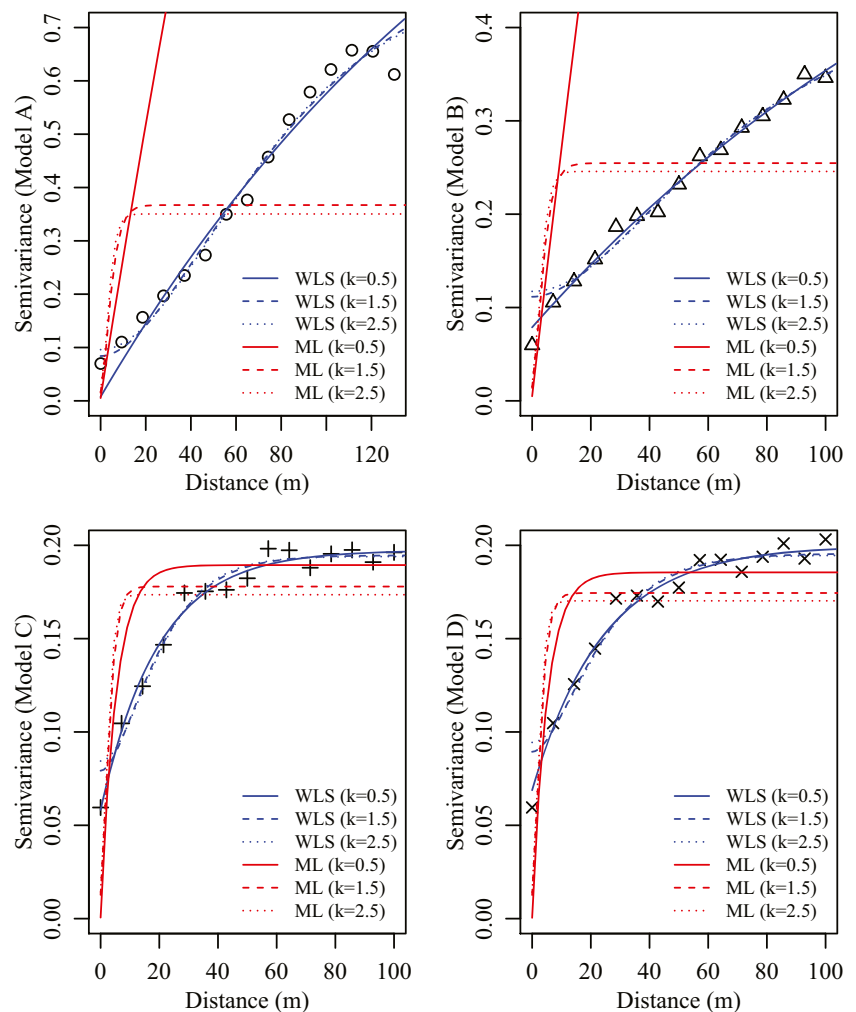
corresponding plots of model D, which suggested that the y coordinate could be an irrelevant covariate for the 2 m depth measurements. These differences were less evident for the plots of 4 m depth measurements.

For cross-validation analysis, the salinity measurements of both depths were randomly divided into a modeling set, consisting of 70 % of the samples, and a validation set, with the remaining 30 % of the samples. The modeling set was used for variogram modeling and kriging on the locations of the validation set; the measurements of the validation set were used for comparison with their predictions.

The sample variograms also enabled us to analyze the effect of covariates on spatial dependence. Sample variograms for salinity (model A) and for OLS residuals (models B, C, and D) are shown as plotted points in Fig. 4 (2 m depth) and Fig. 5 (4 m depth). Anisotropy was investigated by calculating directional variograms for four different angles: 0, 45, 90, and 135. No anisotropy effects were found. In the case of model A, this can be

explained by the few data in the south-north direction, and in the case of models B, C, and D, it can be due to the inclusion of the covariates. Therefore, all fitted models presented here are isotropic. The cutoff value for the several sample variograms was specified visually, taking into account the maximum distance between measurements and the value at which variograms leveled off. All sample variograms show evidence of spatial correlation with semivariance increasing with distance. The effect of covariates is particularly visible in the sample variograms at the 2 m depth. The inclusion of DD and x covariates has lowered the total variability and has also reduced the range of spatial dependence. No significant differences are observed between the sample variograms of models C and D.

All sample variograms were fitted by the Matèrn models with different options for the smoothing parameter and using weighted least squares (WLS) and maximum likelihood (ML) estimation methods as a way to detect eventual

Fig. 5 Omnidirectional variograms of salinity/OLS residuals at 4 m depth fitted by the Matèrn models using WLS and ML
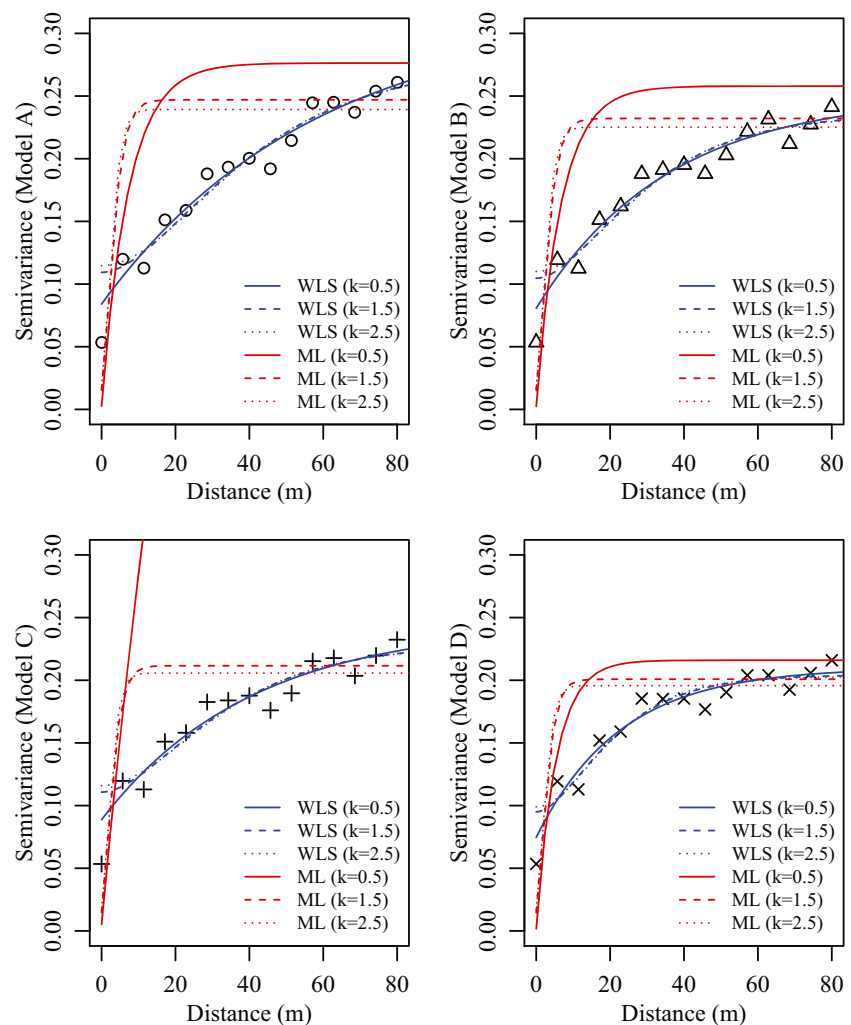
Table 2 Parameters of variogram models for salinity at 2 m depth

Model	k	Weighted least squares				Maximum likelihood					
		τ^2	σ^2	ϕ	A	τ^2	σ^2	ϕ	A	AIC	BIC
A	0.5	0.0076	1.449	200.66	601.13	0.0053	5.401	200.50	600.64	−19.20	1.162
	1.5	0.0840	0.766	44.82	212.64	0.0151	0.352	2.574	12.21	79.90	100.3
	2.5	0.0960	0.694	30.32	179.45	0.0169	0.333	1.54	9.13	137.7	158.1
B	0.5	0.0788	0.6641	186.96	560.07	0.0050	5.090	186.6	559.04	−18.13	7.318
	1.5	0.1115	0.3212	37.78	179.24	0.0140	0.241	2.067	9.81	13.05	38.50
	2.5	0.1173	0.2896	25.60	151.54	0.0155	0.230	1.29	7.62	50.90	76.35
C	0.5	0.0582	0.1390	19.66	58.90	0.0006	0.189	5.23	15.67	−93.87	−58.24
	1.5	0.0792	0.1154	10.72	50.85	0.0128	0.165	1.64	7.76	−62.80	−27.17
	2.5	0.0844	0.1096	8.20	48.55	0.0143	0.159	1.07	6.32	−42.83	−7.197
D	0.5	0.0688	0.1310	24.26	72.67	0.0005	0.185	5.08	15.22	−89.55	−38.65
	1.5	0.0894	0.1062	12.73	60.40	0.0127	0.162	1.61	7.64	−61.01	−10.11
	2.5	0.0944	0.1002	9.67	57.28	0.0142	0.156	1.06	6.25	−41.93	8.97

discrepancies. The family of the Matèrn models is defined by Webster and Oliver (2007) and Diggle and Ribeiro (2007):

$$\gamma(h) = \left(\tau^2 + \sigma^2\right)\left\{1 - \frac{1}{2^{k-1}\Gamma(k)}\left(\frac{h}{\phi}\right)^k K_k\left(\frac{h}{\phi}\right)\right\} \quad (2)$$

where τ^2 is the nugget variance, σ^2 is the structured component of the variogram (excluding the nugget variance), ϕ is the range-related parameter, k is the smoothness parameter, K_k is the modified Bessel function of the third kind of order k, Γ is the gamma function, and $\gamma(h)$ is semivariance for the distance h. The Matèrn family is very attractive due to its flexibility provided by the smoothness parameter (Pardo-Igúzquiza et al. 2009; Diggle and Ribeiro 2007). With respect to the estimation methods, WLS is preferred to unweighted ordinary least squares since, in WLS, the weights are proportional to the number of pairs at each lag. Thus, lags with many pairs have greater influence in the fitting of a model (Lloyd 2011). Estimating model parameters by maximizing the likelihood function under an assumed model provides, under very general conditions, estimators which are unbiased and efficient when applied to large samples (Diggle and Ribeiro 2007). Therefore, the use of ML estimation has become widespread not only amongst geostatisticians. Variograms estimated by likelihood methods can differ quite substantially from the ones estimated

Table 3 Parameters of variogram models for salinity at 4 m depth

Model	k	Weighted least squares				Maximum likelihood					
		τ^2	σ^2	ϕ	A	τ^2	σ^2	ϕ	A	AIC	BIC
A	0.5	0.0841	0.2337	57.89	173.41	0.0027	0.2736	7.317	21.92	−79.80	−58.29
	1.5	0.1094	0.1672	21.81	103.48	0.0153	0.2317	1.848	8.77	−8.66	12.86
	2.5	0.1147	0.1556	15.73	93.13	0.0174	0.2218	1.191	7.05	28.77	50.29
B	0.5	0.0807	0.1699	35.51	106.38	0.0025	0.2554	6.72	20.12	−82.88	−55.98
	1.5	0.1046	0.1312	16.31	77.36	0.0151	0.217	1.77	8.41	−22.19	4.713
	2.5	0.1100	0.1230	12.13	71.78	0.0172	0.208	1.15	6.81	11.71	38.61
C	0.5	0.0887	0.1563	40.50	121.32	0.0053	1.1960	37.56	112.53	−31.50	6.16
	1.5	0.1108	0.1181	18.06	85.68	0.0148	0.1968	1.668	7.91	−42.58	−4.918
	2.5	0.1160	0.1102	13.45	79.63	0.0168	0.1889	1.093	6.47	−13.59	24.07
D	0.5	0.0746	0.1354	22.74	68.12	0.0018	0.2142	5.392	16.15	−91.25	−37.45
	1.5	0.0950	0.1092	11.53	54.69	0.0146	0.1862	1.614	7.65	−51.88	1.917
	2.5	0.0990	0.1038	8.61	50.98	0.0166	0.1790	1.064	6.30	−25.36	28.44

Table 4 Cross-validation results for salinity measurements at 2 m depth

Model	k	Weighted least squares			Maximum likelihood		
		ME	RMSE	MSSE	ME	RMSE	MSSE
A	0.5	−0.000578	0.015929	0.154744	−0.000551	0.016057	0.087115
	1.5	−0.000763	0.025802	0.075615	−0.000472	0.016793	0.093718
	2.5	−0.000569	0.027319	0.075496	−0.000398	0.017167	0.094371
B	0.5	−0.000704	0.021424	0.050009	−0.000549	0.016070	0.087855
	1.5	−0.000635	0.026497	0.060614	−0.000434	0.016846	0.095537
	2.5	−0.000458	0.027735	0.063836	−0.000352	0.017189	0.097027
C	0.5	−0.000755	0.018957	0.047713	−0.000511	0.304011	0.092423
	1.5	−0.000866	0.023801	0.065481	−0.000433	0.016941	0.097545
	2.5	−0.000889	0.024918	0.068699	−0.000365	0.017205	0.099211
D	0.5	−0.000965	0.020107	0.047518	−0.001385	0.016414	0.091359
	1.5	−0.000207	0.024610	0.062637	−0.001489	0.016910	0.096950
	2.5	0.000070	0.025506	0.064761	−0.001441	0.017154	0.098626

by ordinary least squares, since they are not based on the empirical variograms but on observed data (Lloyd 2011). Figures 4 and 5 show the omnidirectional variograms fitted by the Matérn models with $k = 0.5, 1.5$, and 2.5 using WLS and ML. The models fitted by the two estimation methods are substantially different, with WLS fittings closer to the variogram points as expected. Maximum likelihood estimation does not attempt to fit a model to an empirical variogram but, instead, fits a multivariate Gaussian model to the original data. In the case of ML, the goodness of the fit of models to the variogram can be compared through the use of the Akaike information criterion (AIC) and/or the Bayesian information criterion (BIC). The Akaike information criterion is based on the likelihood

function penalized by the number of parameters included in the model. Bayesian information criterion includes the data number in the penalty term. For both criteria, lower values indicate best models, i.e., the one with the fewest parameters that still provides an appropriate adjustment to the data (Diggle and Ribeiro 2007).

The parameters of fitted models for the 2 and 4 m depth modeling sets and the respective values of AIC and BIC are presented respectively in Tables 2 and 3 (note that A is the practical range). The values of model parameters estimated by WLS and ML are quite different on both data sets as expected, but they behave similarly with the inclusion of covariates. From Table 2, it can be observed that the inclusion of DD covariate (model B) reduces

Table 5 Cross-validation results for salinity measurements at 4 m depth

| Model	k	Weighted least squares			Maximum likelihood		
		ME	RMSE	MSSE	ME	RMSE	MSSE
A	0.5	0.000395	0.298441	0.089067	0.000168	0.022651	0.163033
	1.5	0.000510	0.033518	0.098134	0.000077	0.022984	0.166071
	2.5	0.000436	0.034462	0.100071	0.000066	0.023213	0.165642
B	0.5	0.000392	0.029065	0.088512	0.000168	0.022644	0.162915
	1.5	0.000584	0.032945	0.098239	0.000073	0.022961	0.166064
	2.5	0.000590	0.033854	0.099888	0.000058	0.023183	0.165729
C	0.5	0.000688	0.029776	0.086034	0.000467	0.377699	0.142657
	1.5	0.000804	0.033440	0.095708	0.000423	0.022991	0.166906
	2.5	0.000736	0.034332	0.097490	0.000406	0.023193	0.166728
D	0.5	0.001324	0.030124	0.094750	0.001059	0.024469	0.179719
	1.5	0.001562	0.033417	0.104845	0.000988	0.024600	0.187174
	2.5	0.001612	0.034129	0.106761	0.000977	0.024733	0.186888

Fig. 6 Spatial prediction of salinity (psu) at 2 m depth obtained using geostatistical model C with a variogram fitted by the Matèrn models with $k = 0.5$, 1.5, and 2.5 (*from up to down*) using WLS (*left panels*) and ML (*right panels*)
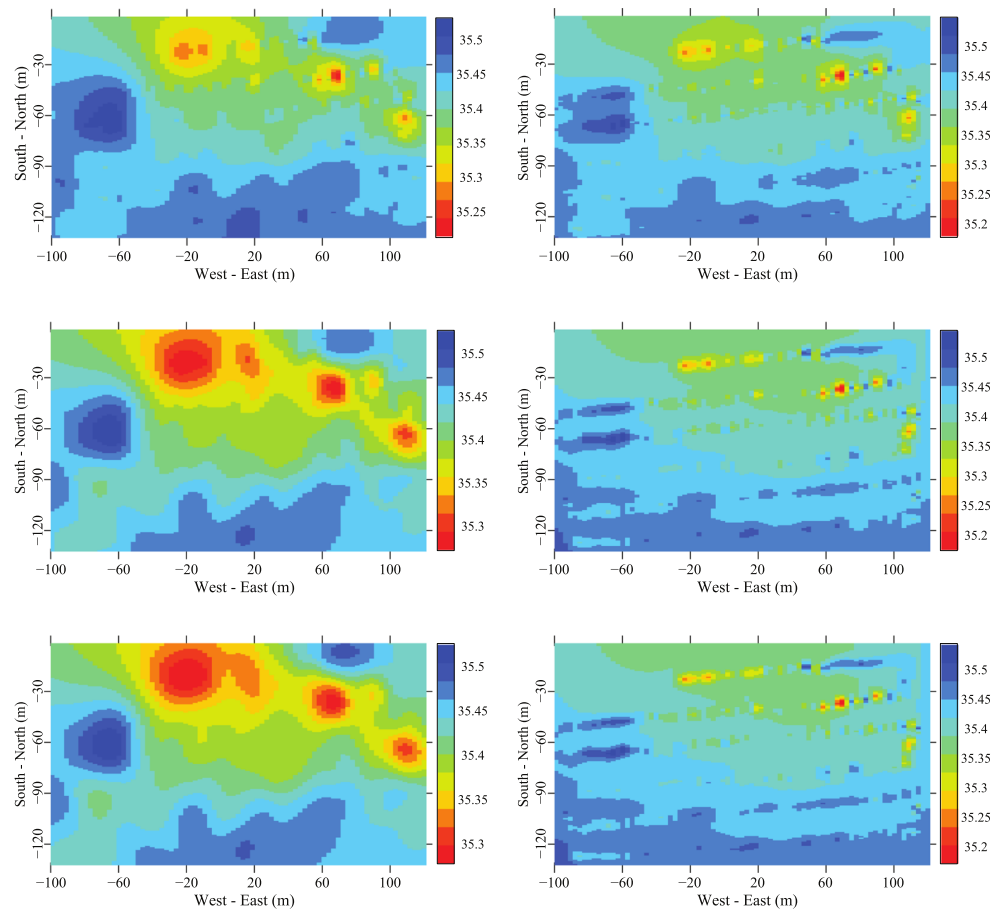
substantially the variability, showing that this covariate is definitely relevant for the study, but the range of spatial dependence still remains quite high and, in some cases, undefined relatively to the survey area. With the inclusion of the quadratic function of x coordinate as covariate (model C), an additional portion of the variability of the data is explained as well as some large-scale trends, and the range estimates are reduced to values consistent with the survey area. No significant differences are observed between estimates of model C and model D, so the more parsimonious model C is considered as more appropriate for the 2 m depth data set for both WLS and ML, which is in accordance with

the results of exploratory analysis. The values of AIC and BIC also point to model C as the best of the four models and $k = 0.5$ as its best smoothing parameter. The results of Table 3 show that in the case of the 4 m depth modeling set, the inclusion of the covariates also reduces the variability of the variable but not so significantly. The variation of range in this case may be due to sample fluctuations. This reflects a usual pattern in the data analysis, in which long-range dependencies are often due to variations that follow the positioning of the data within the area. Unlike the 2 m depth modeling set, the values of AIC and BIC do not give a strong evidence of the best model.

Fig. 7 Spatial prediction of salinity (psu) at 2 m depth obtained using geostatistical model C with a variogram fitted by the Matèrn models with $k = 0.5$ using WLS (*left*) and ML (*right*) plotted on the same scale
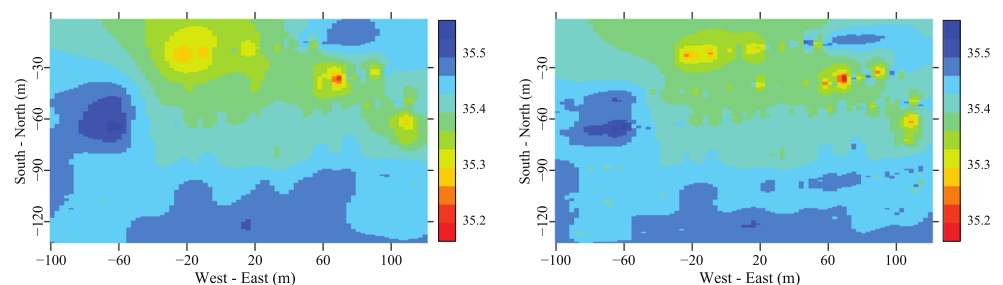
Cross-validation

The comparative performances of the fitted variogram models can also be assessed through the examination of the cross-validation errors. Cross-validation compares observed values and corresponding predictions. The cross-validation errors were calculated using the Eqs. 3–5 and are shown in Tables 4 and 5 (Wackernagel 2003):

$$\text{ME} = \frac{1}{m} \sum_{i=1}^{m} \left[Z(\mathbf{x}_i) - \hat{Z}(\mathbf{x}_i) \right] \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left[Z(\mathbf{x}_i) - \hat{Z}(\mathbf{x}_i) \right]^2} \tag{4}$$

$$\text{MSSE} = \frac{1}{m} \sum_{i=1}^{m} \frac{\left[Z(\mathbf{x}_i) - \hat{Z}(\mathbf{x}_i) \right]^2}{\sigma_{K(\mathbf{x}_i)}^2} \tag{5}$$

where $Z(\mathbf{x}_i)$ and $\hat{Z}(\mathbf{x}_i)$ are, respectively, the measurement and prediction at location $\mathbf{x}_i$, m is the number of measurements in the validation set, and $\sigma_{K(\mathbf{x}_i)}^2$ is the kriging variance at location $\mathbf{x}_i$. The availability of a large number of observations allows for dividing the data set into two subgroups:

one to characterize the behavior of the variable estimating the parameters of the geostatistical model and another for the validation process, comparing observations with predictions obtained by the fitted model. To model salinity, data were organized in order to incorporate any relevant covariates. These were separated of the set of data, and the spatial pattern of the residuals was modeled. From Tables 4 and 5, it can be seen that the cross-validation of indicators does not provide a consistent information to the best choice. Frequently, different accuracy measures lead to different results as to which model is best. Combining the variogram fitting results, the likelihood analysis given by AIC and BIC, the exploratory analysis, and the cross-validation results, we conclude that model C is the most appropriate for both data sets.

Mapping

Kriging with covariates by using model C was used to predict salinity at the 2 and 4 m depths. Figure 6 shows the spatial prediction of salinity at the 2 m depth, with a variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 using WLS and ML, and Fig. 8 shows the spatial
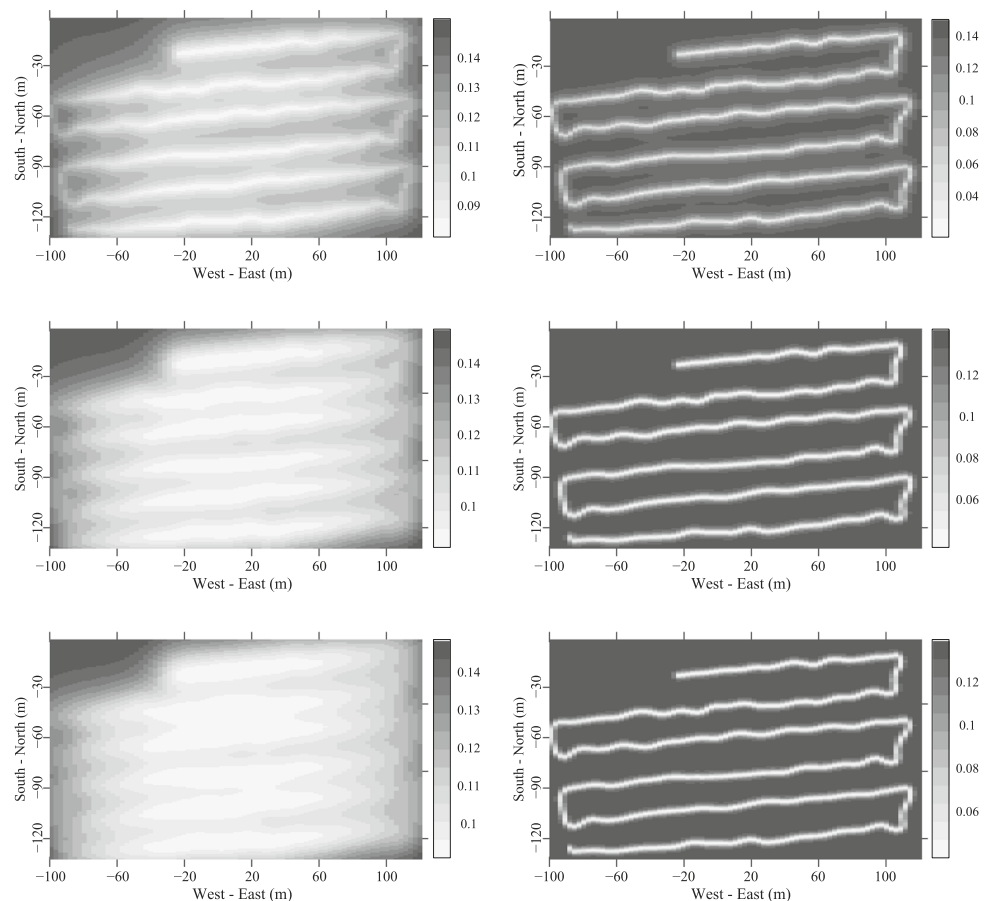

Fig. 8 Spatial prediction of salinity (psu) at 4 m depth obtained using geostatistical model C with a variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 (*from up to down*) using WLS (*left panels*) and ML (*right panels*)
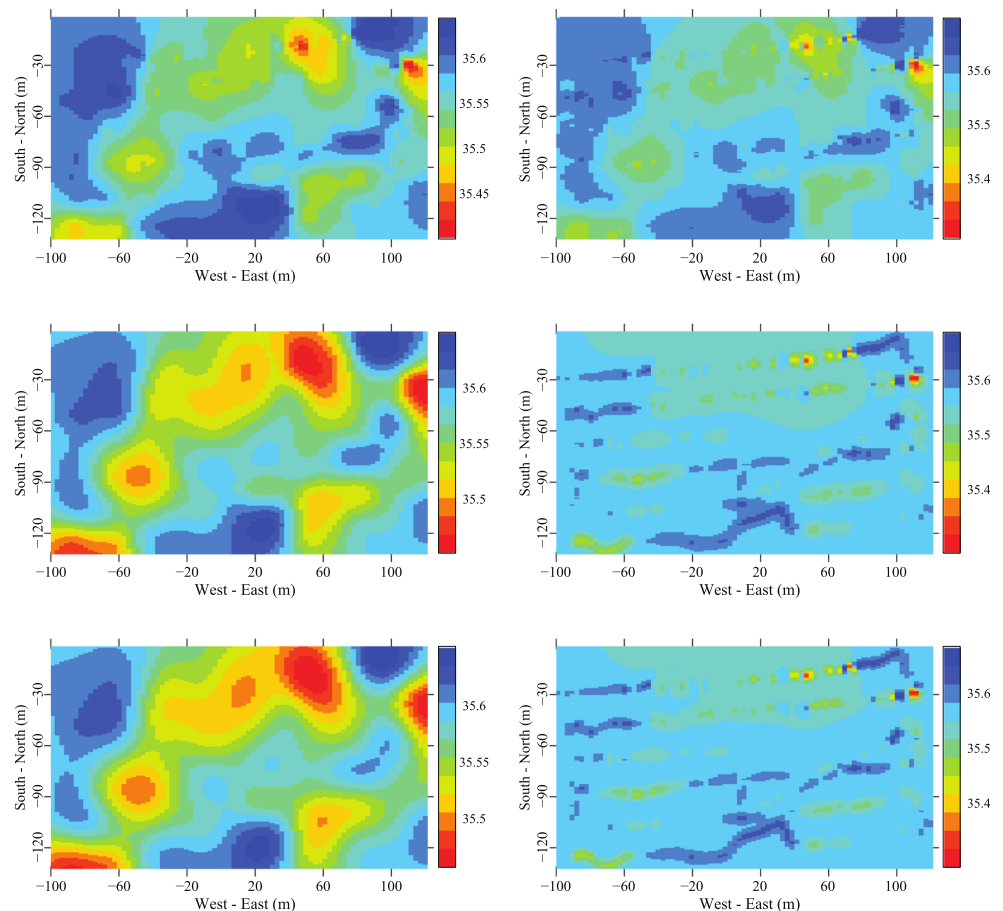
prediction of salinity at the 4 m depth, with a variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 using WLS and ML. It can be seen that the two estimation methods show differences in the predictions of salinity which was expected since their fittings to the variogram are substantially different. The amplitude of salinity values predicted using ML is larger than the amplitude of salinity values predicted using WLS, but the maps obtained using WLS show more details of salinity distribution in the field. In the case of 2 m depth kriged maps, WLS predictions range from 35.23 to 35.52, 35.28 to 35.52, and 35.29 to 35.52 psu, respectively for $k = 0.5, 1.5$, and 2.5; ML predictions range from 35.19 to 35.53 psu for $k = 0.5, 1.5$, and 2.5. In the case of 4 m depth kriged maps, WLS predictions range from 35.41 to 35.52, 35.46 to 35.64, and 35.47 to 35.65 psu, respectively for $k = 0.5, 1.5$, and 2.5; ML predictions range from 35.31 to 35.68 and 35.30 to 35.67 psu, respectively, for $k = 0.5$ and for $k = 1.5$ and 2.5. In general, the ML method estimated very low ranges, typically below 20 m, the space between parallel transects. Consequently, in this case, when the prediction location is close to one transect, the observations used in its prediction are only those from that transect and therefore its value is similar to the closest measurements. This fact is visible on

the maps in which the ML method was used, where some predictions form the shape of the vehicle trajectory. It can also be seen on these maps that as the distance to one transect increases, the predictions of salinity approach rapidly to the average value. Exceptionally, in the 4 m depth kriged map for $k = 0.5$, these effects are not visible since the estimated range by the ML method is quite high (112.53 m) and much higher than the estimated ranges by the ML method for all other cases (7.91, 6.47, and 15.67 m, 7.76 m, 6.32 m). The ML method may have estimated much lower ranges than the WLS method due to the presence of anomalous values in the data or local anomalies. Based on this analysis, we conclude that the very low ranges estimated by the ML method have limited the kriging process in the prediction computation. So, at least for these data sets, the ML method is not the most appropriate method for the variogram fitting. This problem could possible be minimized by making some changes on the sampling strategy, namely by considering also transects parallel to the current. In this case, much more data would be available to characterize the variability of the plume dispersion in that direction. Just for reference, Fig. 7 shows the spatial prediction of salinity at the 2 m depth with the variogram fitted by the Matèrn models with $k = 0.5$ using WLS and ML plotted on the same


Fig. 9 Standard error of estimates at 2 m depth using geostatistical model C with a variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 (*from up to down*) using WLS (*left panels*) and ML (*right panels*)

color scale (Fig. 8 shows the spatial prediction of salinity at the 4 m depth). When comparing these maps with their counterparts of Fig. 6, we observe that certain details of salinity distribution have been lost due to the use of the same color scale. Therefore, to observe more clearly these details, a different color scale for salinity was adopted for each estimation method. Figure 9 shows the standard error of kriging estimates at the 2 m depth using geostatistical model C with variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 using WLS and ML, and Fig. 10 shows the standard error of kriging estimates at the 2 m depth using geostatistical model C with a variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 using WLS and ML. The amplitude of standard errors obtained using ML is larger than the amplitude of standard errors obtained using WLS. In the case of 2 m depth maps, standard errors associated with the WLS method range from 0.08 to 0.15 psu and 0.09 to 0.15, respectively for $k = 0.5$ and $k = 1.5, 2.5$; standard errors associated with the ML method range from 0.02 to 0.14 psu and 0.04 to 0.14, respectively, for $k = 0.5$ and $k = 1.5, 2.5$. In the case of 4 m depth maps, standard errors associated with the WLS method range from 0.09 to 0.17 psu and 0.10 to 0.16, respectively for $k = 0.5$

and $k = 1.5, 2.5$; standard errors associated with the ML method range from 0.01 to 0.40 psu and 0.04 to 0.15, respectively, for $k = 0.5$ and $k = 1.5, 2.5$. It can be observed that the standard errors are smaller close to the sampling points, i.e., along the trajectory of the vehicle. As the distance to one transect increases, the standard errors approach rapidly to the average value, in the case of maps in which the ML method was used, an effect that was also observed in the prediction maps of salinity distribution, a consequence of the very low range estimates. In the vicinity of the diffuser, the water column was weakly stratified due to both low temperature and salinity variations. The total difference in density over the water column was about 0.40σ units. This relatively weak stratification explains the plume spreading near the surface, as predicted by a prediction model used in the field to specify the AUV survey. The 2 and 4 m depth maps with a variogram fitted using WLS show the spatial variation of salinity at these depths in the area studied. From these maps, it is possible to identify the effluent plume and its dispersion downstream in the direction of the current. It appears as a region of lower salinity when compared to the surrounding ocean waters at the same depth. When taking the standard errors into account, at the 2 m depth, the

Fig. 10 Standard error of estimates at 4 m depth using geostatistical model C with a variogram fitted by the Matèrn models with $k = 0.5, 1.5$, and 2.5 (*from up to down*) using WLS (*left panels*) and ML (*right panels*)
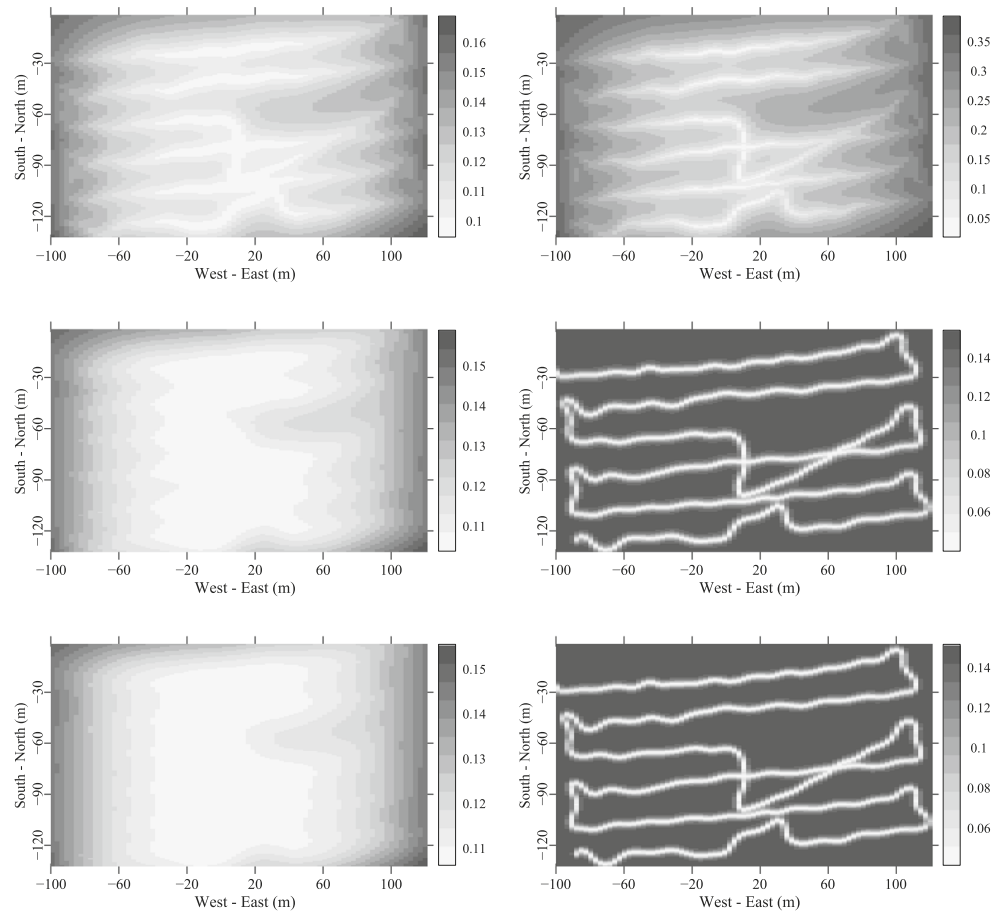
plume may be identified by the regions where salinity is less than 35.4 psu and, at the 4 m depth, by the regions where salinity is less than 35.5 psu. The plume exhibits a considerably more complex structure than the compact shape of the classical picture of the buoyant plume and as patchy as in previous studies (Petrenko et al. 1998; Ramos and Neves 2009). The small plume-related anomalies in the local salinity of Figs. 6 and 8 are evidence of the rapid mixing process. Due to the large differences in density between the rising effluent plume and ambient ocean waters, entrainment and mixing processes are vigorous and the properties within the plume change rapidly (Hunt et al. 2010). These results therefore confirm that large gradients in background salinity and the small differences in salinity between the effluent plume and the ambient waters can easily obscure the signature of the plume. The results obtained in this study suggest some improvements to the sampling strategy of the AUV if a geostatistical analysis of the data is in mind. It is important to prevent the existence of anomalous values, namely low salinity measurements usually found very close to the outfall discharge. If these values are present, they should be correctly identified and treated properly in the analysis. Otherwise, they can induce a false spatial correlation between the data. A more competitive comparison between the WLS and ML methods would probably be possible if the AUV sampling strategy would have combined transects parallel and perpendicular to the effluent dispersion. In that case, the y coordinate would probably be considered a relevant covariate, since more data would be available in the direction of the dispersion, and better results would be obtained.

Conclusions

Through a geostatistical analysis of salinity measured by an AUV at the 2 and 4 m depths in an ocean outfall monitoring campaign, it was possible to obtain kriged maps of the sewage dispersion in the field. Based on an exploratory analysis, four different geostatistical linear models for salinity were assumed, where the distance to the diffuser, the west-east positioning, and the south-north positioning were used as covariates. Sample variograms of the raw data and of the OLS residuals for both depths were fitted by the Matèrn models, considering different values of the smoothing parameter and using weighted least squares and maximum likelihood estimation methods, as a way to detect eventual discrepancies. The performance of each competing model was compared via AIC and BIC and using the split-sample approach. The results point to the geostatistical model that includes the distance to diffuser and the west-east coordinate as covariates, with the smoothing parameter $k = 0.5$ as the most appropriate. Typically, the ML method

estimated very low ranges which have limited the kriging process in the prediction computation. So, at least for these data sets, WLS showed to be the most appropriate estimation method for the variogram fitting. The 2 and 4 m depth kriged maps with the variogram fitted using WLS show the spatial variation of salinity in the area studied, and it is possible to identify the effluent plume that appears as a region of lower salinity when compared to the surrounding ocean waters dispersing downstream in the direction of the current. The results obtained in this study suggest some improvements to the sampling strategy of the AUV if a geostatistical analysis of the data is in mind. It is important to prevent the existence of anomalous values, and if present, they should be correctly identified and treated properly; otherwise, they can induce a false spatial correlation between the data. A more competitive comparison between the WLS and ML methods would probably be possible if the AUV sampling strategy would have combined transects parallel and perpendicular to the effluent dispersion. In that case, the north-south coordinate would possibly be considered a relevant covariate and better results would be obtained.

References

Bivand RS, Pebesma EJ, Gómez-Rubio V (2008) Applied spatial data analysis with R. ISBN: 97-0-387-78170-9

Carvalho JLB, Roberts PJW, Roldão J (2002) Field observations of the Ipanema Beach outfall. J Hydraul Eng ASCE 128(2):151–160

Daviero GJ, Roberts PJW (2006) Marine wastewater discharges from multiport diffusers. III: stratified stationary water. J Hydraul Eng 132(4):404–410

Diggle PJ, Ribeiro Jr. PJ (2007) Model based geostatistics

Goovaerts P (1997) Geostatistics for natural resources evaluation. Applied geostatistics series. ISBN13: 9780195115383, ISBN10: 0195115384

Hunt CD, Mansfield AD, Mickelson MJ, Albro CS, Geyer WR, Roberts PJW (2010) Plume tracking and dilution of effluent from the Boston sewage outfall. Mar Environ Res 70:150– 161

Isaaks EH, Srivastava RM (1989) Applied geostatistics. New York Oxford, ISBN 0-19-505012-6-ISBN 0-19-505013-4 (pbk.)

Kitanidis P (1997) Introduction to geostatistics: applications in hydrogeology. New York (USA)

Lloyd C (2011) Local models for spatial analysis

Niu H, Husain T, Veitch B, Bose N, Adams S, He M, Lee K (2007) Ocean outfall mapping using an autonomous underwater vehicle. In: MTS/IEEE Oceans

Pardo-Igúzquiza E, Mardia KV, Chica-Olmo M (2009) Mlmatern: a computer program for maximum likelihood inference with the spatial Matérn covariance model. Comput Geosci 35(6):1139–1150. doi:10.1016/j.cageo.2008.09.009. http://www.sciencedirect.com/science/article/pii/S0098300408002756

Petrenko AA, Jones BH, Dickey TD (1998) Shape and initial dilution of sand island, Hawaii sewage plume. J. Hydraul. Eng. ASCE 124(6):565–571

R Development Core Team (2013) R: a language and environment for statistical computing. R version 3.00. http://www.R-project.org/

Ramos P, Abreu N (2010) Environmental monitoring of wastewater discharges using an autonomous underwater vehicle. In: Workshop on Robotics for Environmental Monitoring, IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS2010, Taipei

Ramos P, Neves MV (2009) Environmental impact assessment and management of sewage outfall discharges using AUVs. In: Underwater vehicles

Ramos P, Neves MV, Pereira FL (2007) Mapping and initial dilution estimation of a sewage outfall plume using an autonomous underwater vehicle. Cont Shelf Res 27:583–593

Ribeiro PJ, Diggle PJ (2001) geoR: a package for geostatistical analysis

Rogowski P, Terrill EJ, Otero MP, Hazard L, Middleton WF (2012) Mapping ocean outfall plumes and their mixing using autonomous underwater vehicles. J Geophys Res Oceans 117

Tian X, Roberts PJW, Daviero GJ (2004a) Marine wastewater discharges from multiport diffusers. I: unstratified stationary water. J Hydraul Eng 130(12):1137–1146

Tian X, Roberts PJW, Daviero GJ (2004b) Marine wastewater discharges from multiport diffusers. II: unstratified flowing water. J Hydraul Eng 130(12):1147–1155

Tian X, Roberts PJW, Daviero GJ (2006) Marine wastewater discharges from multiport diffusers. IV: stratified flowing water. J Hydraul Eng 132(4):411–419

Wackernagel H. (2003) Multivariate geostatistics: an introduction with applications. Berlin

Washburn L, Jones BH, Bratkovich A, Dickey TD, Chen MS (1992) Mixing, dispersion, and resuspension in vicinity of ocean wastewater plume. J Hydraul Eng ASCE 118(1):38–58

Webster R, Oliver M. (2007) Geostatistics for environmental scientists, 2nd ed. ISBN-13: 978-0-470-02858-2(HB)

Wu Y, Washburn L, Jones BH (1994) Buoyant plume dispersion in a coastal environment: evolving plume structure and dynamics. Cont Shelf Res 14(9):1001–1023