Spatial Distribution of a Sewage Outfall Plume Observed with an AUV

P. Ramos[1,2], M. Monego[1], S. Carvalho[3]
patricia@fe.up.pt; mdmonego@fe.up.pt ; s.carvalho@aguasdooeste.com
[1]Faculty of Engineer of University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[2]Institute of Accountancy and Administration of Porto, Dept. of Mathematics
R. Jaime Lopes Amorim, 4465-004 S. M. Infesta, Portugal
[3]Águas do Oeste, S.A., Convento de S. Miguel das Gaeiras, 2510-718 Óbidos, Portugal
Phone: +351960084473, Fax: +351225081443

***Abstract*- The main purpose of this study was to examine the applicability of geostatistical modeling to obtain valuable information for assessing the environmental impact of sewage outfall discharges. The data set used was obtained in a monitoring campaign to Foz do Arelho outfall, located off the Portuguese west coast near Óbidos region, using an AUV. The Matheron's classical estimator was used the compute the experimental semivariogram which was fitted to three theoretical models: spherical, exponential and gaussian. The cross-validation procedure suggested the best semivariogram model and ordinary kriging was used to obtain the predictions of salinity and temperature at unknown locations. The generated maps show clearly the plume dispersion in the studied area. Our study suggests that an optimal design for the AUV sampling trajectory from a geostatistical prediction point of view, can help to compute more precise predictions and hence to quantify more accurately dilution. Moreover, since accurate measurements of plume's dilution are rare, these studies might be very helpful in the future for validation of dispersion models.**

I. INTRODUCTION

Outfalls are designed to promote the natural assimilative capacity of the oceans to dispose of wastewaters with minimal environmental impact. This is accomplished through the vigorous initial mixing that is followed by oceanic dispersion within spatially and temporally varying currents. Usually, those mixing processes, in conjunction to bacterial mortality, result in rapid reductions in the concentrations of contaminants and organisms present in the wastewater to near background levels. However, coastal physical, chemical and biological processes, very dynamic and complex, and intimately coupled to the concentration and content of wastewater, are in most instances, poorly understood. Consequently, how sewage disperses and how effluent modifies and is modified by coastal environment remains in many aspects unknown and unpredictable.

Effluent plumes dispersion is still a difficult problem to study *in situ* making reliable field measurements rare. Autonomous Underwater Vehicles (AUVs) already demonstrated to be very appropriate for high-resolution surveys of small features such as outfall plumes [1]. Some of the advantages of these platforms include: easier field logistics, low cost per deployment, good spatial coverage and sampling over repeated sections and capability of feature-based or adaptive sampling. While the gross properties of the plumes can be reasonably predicted by the most commonly used marine discharge models [2], there remain many aspects that cannot be, particularly the patchy nature of the wastefields. If we want to calibrate these models with real data we have to be able to quantify spatial correlations and other related characteristics.

In this paper we show the results of a monitoring campaign performed on November 28, 2007 to study the shape and dispersion of Foz do Arelho sea outfall plume, using MARES AUV.

Geostatistics was used in the spatial analysis of the environmental data gathered with the AUV in the vicinity of the outfall discharge. Geostatistical modeling has been used with success to analyze and characterize the spatial variability of soil properties [3]-[4], to obtain information for assessing water and wind resources [5]-[6], to design sampling strategies for estuarine sediments field studies [7], to study the thickness of effluent-affected sediment in the vicinity of wastewater discharges [8], to obtain information about the spatial distribution of sewage pollution in coastal sediments [9], among many others.

This paper is organized as follows. First we give a brief description of MARES AUV. Then we present the geostatistical study of the salinity and temperature measurements obtained, using ordinary kriging interpolation. In a first step the spatial structure of the observations was inspected thought a descriptive statistical analysis. Then, the degree of spatial correlation among data in the study area as function of the distance and direction was expressed in terms of the semivariogram. Finally, ordinary kriging was used to estimate salinity and temperature at unknown locations, and a map of these parameters distribution in the field was generated. Cross-validation indicators and additional model parameters helped to choose the most appropriate models.

II. MARES AUV

MARES (Modular Autonomous Robot for Environment Sampling) is a highly modular autonomous underwater vehicle built by some researchers of the Ocean Systems Group

at the Faculty of Engineer of University of Porto (http://oceansys.fe.up.pt/index.php). It has a diameter of 20 cm and is about 1.5 meters long, weighting about 32 kg in air (see Fig. 1).

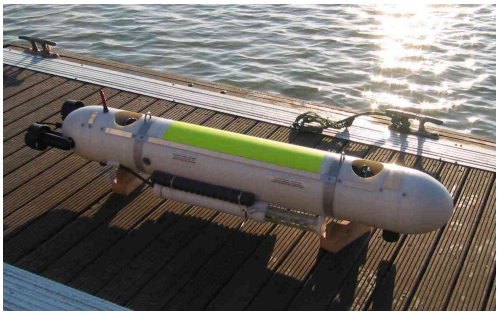

Fig. 1 MARES Autonomous Underwater Vehicle.

MARES' mechanical structure follows a highly modular approach, with a central hull and several additional sections, most of them machined in acetal copolymer. Since all sections are mechanical extensions, they can be easily interchanged and it is very simple to insert new sections. Propulsion and direction are controlled by two independent horizontal thrusters. Another set of thrusters, in the vertical direction, control vertical velocity and pitch angle. The thruster arrangement permits operations in very confined areas, with virtually independent horizontal and vertical motion at velocities starting at 0 m/s.

The electronic circuits are located in the centrall hull. Energy is provided by Li-Ion batteries, with a total energy of 600 Wh. Depending on vehicle velocity, these batteries can last up to 10 hours, corresponding to about 40km. The main computer is a PC-104 stack, with a power supply, the CPU, a communications board and a solid-state disk.

The navigation system is based on a LBL acoustic network, with 2 NIBs being deployed in the operation area. The vehicle software continuously fuses ranges to the buoys, together with compass heading and thruster RPM to compute the estimated position.

MARES has limited room to accommodate payload, but it is simple to include additional sections and the main hull has several spare connectors to provide energy and communications.

MARES missions are programmed using a GUI software, where all mission segments are detailed (waypoints, velocity, depth, etc). During the mission, the buoys transmit vehicle location data to a support vessel, so that the AUV trajectory can be followed in real time. At the end, the vehicle starts transmitting his own position by radio, when at the surface.

For this experiment a Sea-Bird 49 FastCAT CTD (Conductivity, Temperature, Depth) was installed in MARES AUV. During the mission that last for about 1 hour, while navigating at a constant velocity of approximately 2 knots (~1m/s), the vehicle collected CTD data continuously within a frequency of 16 Hz.

III. GEOSTATISTICAL METHODS

The trajectory of the AUV and the study area in the vicinity of Foz do Arelho outfall, off the Portuguese west coast near Óbidos region, are shown in Fig. 2.

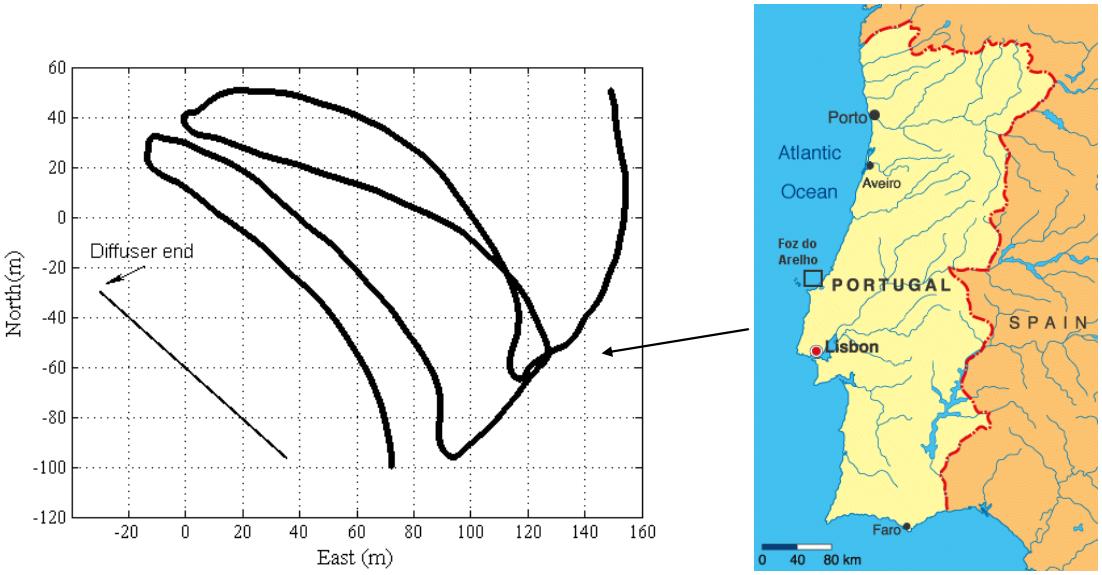

Fig. 2 (a) AUV sampling trajectory at 1.5 meters depth; (b) Study area off the Portuguese west coast near Óbidos region.

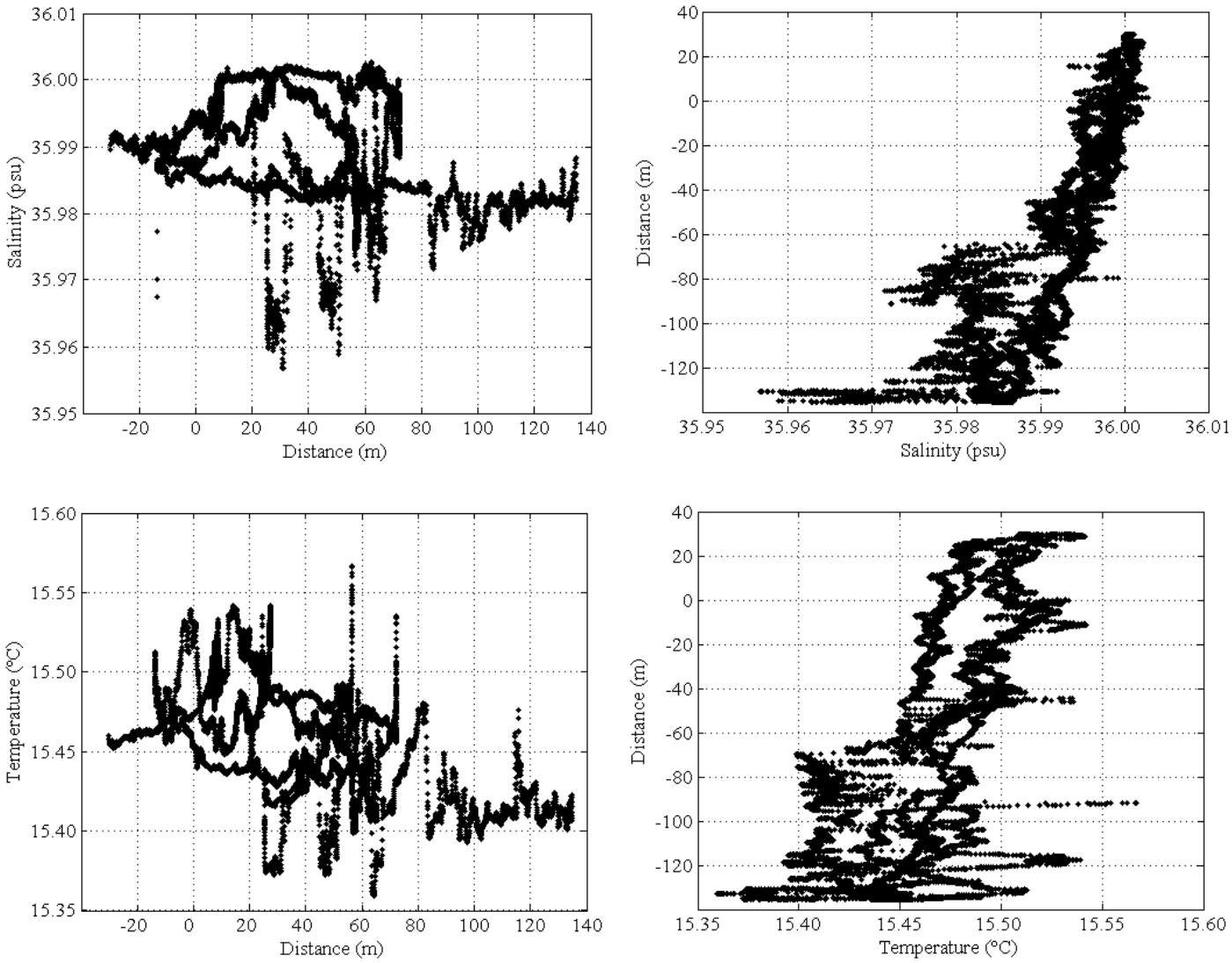

Fig. 3 Raw data of salinity and temperature measured by the AUV at 1.5 m depth.

A rectangular area of 170x150m^2 was covered. Fig. 3 shows raw data of temperature and salinity (computed from conductivity, temperature and depth) collected by the AUV at 1.5 m depth, where the plume was found established and dispersing horizontally. The direction of the current at surface was about 120º. These data were rotated - 50º due to interpolation conditions.

A. Exploratory Analysis

Table I gives the summary statistics of the salinity and temperature data set (17612 measurements). The salinity ranged from 35.9567 psu to 36.0028 psu. The mean of the salinity value of the data set was 35.9905 psu, being close to the median value that was 35.9915 psu. The temperature ranged from 15.3595 ºC to 15.5671 ºC. The mean of the temperature of the data set value was 15.4623 ºC, being close to the median value that was 15.4654 ºC.

TABLE I
SUMMARY STATISTICS OF THE SALINITY AND TEMPERATURE DATA SET

Statistics	Salinity	Temperature
Number of data	17612	17612
Minimum	35.9567 psu	15.3595 ºC
Mean	35.9905 psu	15.4623 ºC
Median	35.9915 psu	15.4654 ºC
Maximum	36.0028 psu	15.5671 ºC
Variance	0.00006	0.0011
Standard Deviation	0.0079	0.0327
Skewness	-0.6418	-0.2629
Kurtosis	3.3303	2.7747

As in conventional statistics, a normal distribution for the variable under study is desirable in linear geostatistics [10]. It can be seen from Table I that both skewness and kurtosis

values of salinity indicate a negatively skewed distribution of these data set, while both skewness and kurtosis low values of temperature indicate an approximated normal distribution of these data set.

Fig. 4 (a) shows the histogram of the salinity data set. The left tail of the histogram shows a negatively skewed distribution, which is in accordance with the negative value of the skewness parameter of Table I. This can be justified by the sampling strategy adopted, which was not completely succeeded.

Fig. 4 (b) shows the histogram of the temperature data set which has an approximated normal distribution. The small negative value of the skewness parameter of Table I is in accordance with the small left tail of this histogram.

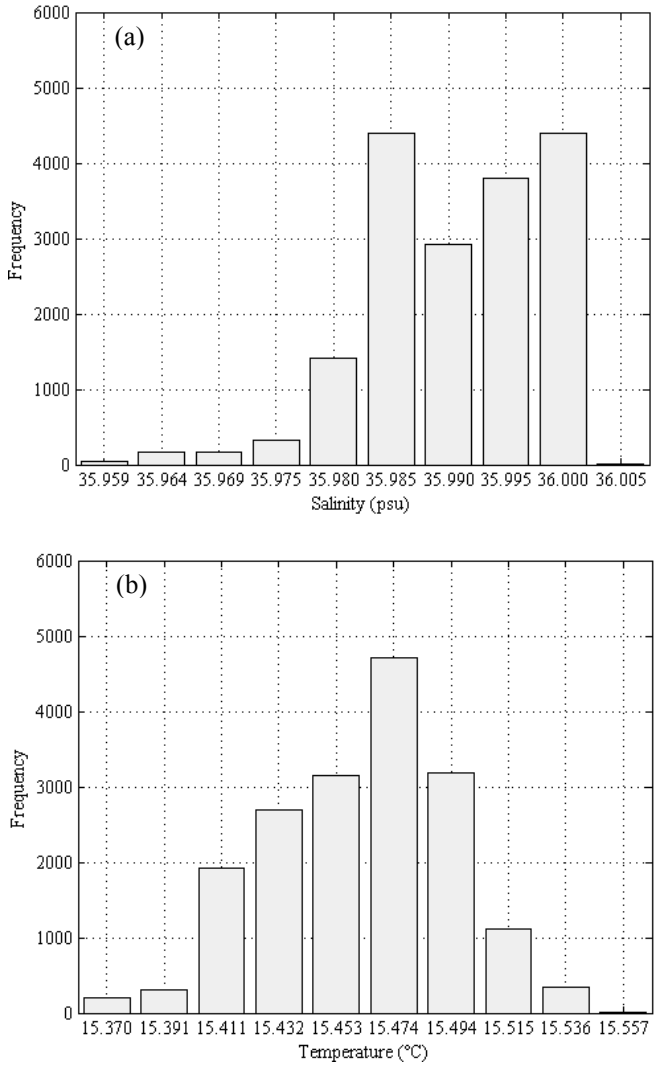

Fig. 4 (a) Histogram of the salinity data set; (b) Histogram of the temperature data set.

B. *Semivariogram*

Geostatistical methodology uses the semivariogram to quantify the spatial variation of the variable in study [11]-[12]. The semivariogram measures the mean variability between two data points as function of their distance. Matheron's classical estimator of the semivariogram was used in this study, whose computing equation is [13]:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[Z(x_i) - Z(x_i + h) \right]^2 \qquad (1)$$

where $\gamma(h)$ is the semivariogram, $Z(x_i)$ is the salinity value measured at location x_i, h is the *lag* distance and $N(h)$ is the number of pairs of measurements which are $N(h)$ distance apart. The experimental semivariogram is calculated for several lag distances. Anisotropy should be investigated by calculating the semivariogram for several directions.

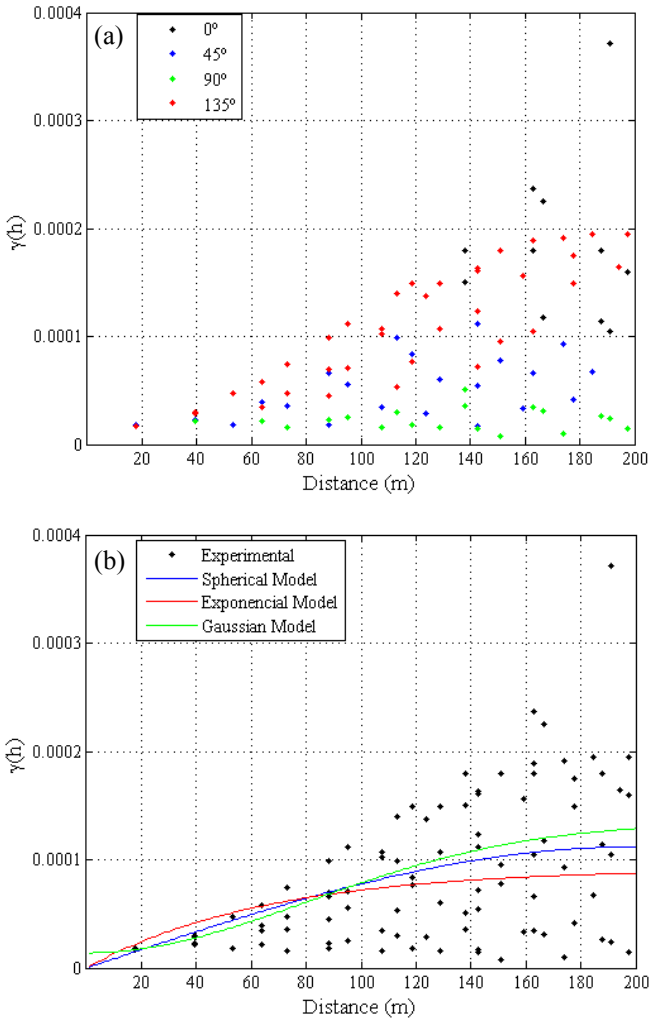

Fig. 5 (a) Directional experimental semivariograma of salinity; (b) Omnidirectional experimental semivariogram of salinity and fitted models.

Estimation of semivariances was carried out using a lag distance of 25 m and a range value of 200 m. The experimental semivariogram of both salinity and temperature data sets was calculated for 0º, 45º, 90º and 135º directions.

Fig. 5 (a) shows the directional experimental semivariogram of the salinity data set and Fig. 6 (a) shows the directional experimental semivariogram of the temperature data set. Both

phenomenons present a quite similar behavior in the four studied directions, showing both to be isotropic.

Once the experimental semivariogram is computed, the next step is to adjust it to a theoretical model. This model gives information about the structure of the spatial variation as well as the input parameters for the spatial prediction by kriging. The most commonly used theoretical models are spherical, exponential and gaussian [14].

Both semivariograms were fitted to three theoretical models: spherical, exponential and gaussian.

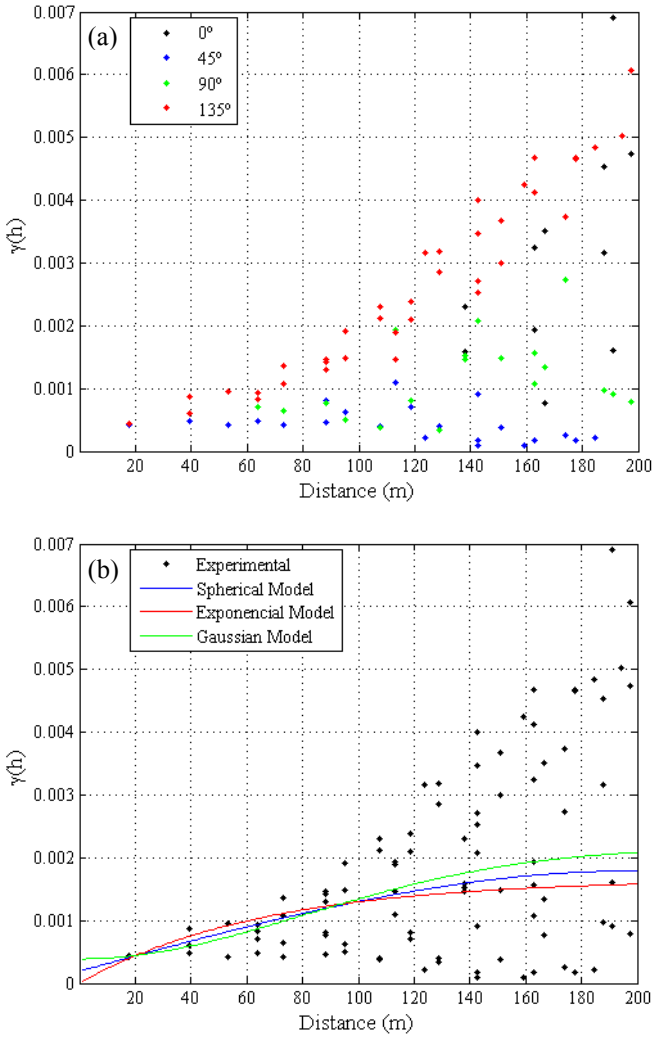

Fig. 6 (a) Directional experimental semivariograma of temperature; (b) Omnidirectional experimental semivariogram of temperature and fitted models.

Fig. 5 (b) shows the omnidirectional experimental semivariogram of salinity and the spherical, exponential and gaussian models fitted; Fig. 6 (b) shows the omnidirectional experimental semivariogram of temperature and the spherical, exponential and gaussian models fitted.

The nugget, sill and range parameters of these three models fitted to the salinity semivariogram are shown in Table II. The nugget, sill and range parameters of these three models fitted to the temperature semivariogram are shown in Table III.

TABLE II
PARAMETERS OF THE MODELS FITTED TO THE SALINITY SEMIVARIOGRAM

Parameters	Spherical	Exponential	Gaussian
Nugget	0	0	0.0000140
Partial Sill	0.000112	0.000092	0.000120
Sill	0.000112	0.000087	0.000128
Range	197.642	197.642	197.642
Nugget/Sill	0%	0%	10.91%

TABLE III
PARAMETERS OF THE MODELS FITTED TO THE TEMPERATURE SEMIVARIOGRAM

Parameters	Spherical	Exponential	Gaussian
Nugget	0.000193	0	0.000388
Partial Sill	0.001597	0.001659	0.001779
Sill	0.001789	0.001577	0.002078
Range	197.642	197.642	197.642
Nugget/Sill	10.77 %	0 %	18.66 %

The semivariances used in the spatial prediction of salinity by kriging are then computed according to equation (2), (3), or (4), if, respectively, the spherical, exponential or the gaussian model, is selected. The semivariances used in the spatial prediction of temperature by kriging are computed according to equation (5), (6), or (7), if, respectively, the spherical, exponential or the gaussian model, is selected.

$$\begin{cases} \gamma(h) = 0 + 0.000112\left[\frac{3}{2}\left(\frac{\|h\|}{a}\right) - \frac{1}{2}\left(\frac{\|h\|}{a}\right)^3\right], 0 \le h \le a \\ \gamma(h) = 0.000112, \; h > a \end{cases} \quad (2)$$

$$\gamma(h) = 0 + 0.000092\left[1 - \exp\left(-\frac{3\|h\|}{a}\right)\right], \; h \ge 0 \quad (3)$$

$$\gamma(h) = 0.0000140 + 0.000120\left[1 - \exp\left(-3\left(\frac{\|h\|}{a}\right)^2\right)\right], \; h \ge 0 \quad (4)$$

According to [4] the spatial dependence of a variable in study can be verified through of the nugget/sill ratio. Nugget/sill ratios less than 25% suggest that the variable has a strong spatial dependence; nugget/sill ratios between 25% and 75% suggest that the variable has a moderate spatial dependence; and nugget/sill ratios above 75% suggest that the variable has low spatial dependence. As can be observed in Table II and Table III, the nugget/sill ratios of all the semivariogram models of both salinity and temperature data

sets are low and less then 25%, suggesting that both variables have a strong spatial dependence and that their local variations can be captured as expected.

$$\begin{cases} \gamma(h) = 0.000193 + 0.001597\left[\frac{3}{2}\left(\frac{\|h\|}{a}\right) - \frac{1}{2}\left(\frac{\|h\|}{a}\right)^3\right], 0 \le h \le a \\ \gamma(h) = 0.001789,\ h > a \end{cases} \quad (5)$$

$$\gamma(h) = 0 + 0.001659\left[1 - \exp\left(-\frac{3\|h\|}{a}\right)\right],\ h \ge 0 \quad (6)$$

$$\gamma(h) = 0.000388 + 0.001779\left[1 - \exp\left(-3\left(\frac{\|h\|}{a}\right)^2\right)\right],\ h \ge 0 \quad (7)$$

C. Cross-Validation

Cross-validation was used to compare the prediction performances of the three fitted semivariogram models. In the cross-validation procedure each sample is eliminated in turn and the remaining samples are used by kriging to predict the eliminated observation. The differences between the observations and the predictions are evaluated using the mean error (ME), the root mean squared error (RMSE), the root mean kriging variance (RMKV), the mean standardized error (MSE), and the root mean squared standardized error (RMSSE), computed respectively according to the following equations:

$$\mathrm{ME} = \frac{1}{N}\sum_{i=1}^{N}\left[\hat{Z}(x_i) - Z(x_i)\right] \quad (8)$$

$$\mathrm{RSME} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\hat{Z}(x_i) - Z(x_i)\right]^2} \quad (9)$$

$$\mathrm{RVKM} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i)} \quad (10)$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{\hat{Z}(x_i) - Z(x_i)}{\sigma^2(x_i)}\right] \quad (11)$$

$$\mathrm{RMSSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\hat{Z}(x_i) - Z(x_i)}{\sigma^2(x_i)}\right]^2} \quad (12)$$

where $\hat{Z}(x_i)$ is the predicted value at cross-validation point x_i, $Z(x_i)$ is the actual (measured) value at point x_i, N is the number of measurements of the data set, and $\sigma^2(x_i)$ is the kriging variance at cross-validation point x_i.

Table IV shows these indicators relative to the spherical, exponential and gaussian models used in salinity estimation.

Table V shows these indicators relative to the spherical, exponential and gaussian models used in temperature estimation.

TABLE IV
CROSS-VALIDATION PARAMETERS OF THE SALINITY SEMIVARIOGRAM MODELS

Parameters	Spherical	Exponential	Gaussian
ME	0.0000001686	0.000000169	-0.0001952
RMSE	0.0001863	0.0001863	0.002621
RMKV	0.0002058	0.0002642	0.004125
MSE	0.0008716	0.0006802	-0.04765
RMSSE	1.109	0.8641	0.6486

TABLE V
CROSS-VALIDATION PARAMETERS OF THE TEMPERATURE SEMIVARIOGRAM MODELS

Parameters	Spherical	Exponential	Gaussian
ME	-0.0001517	-0.0000093	-0.0000429
RMSE	0.008131	-0.0000093	0.01375
RMKV	0.01609	0.001122	0.02157
MSE	-0.1003	-0.0009009	-0.00436
RMSSE	0.5189	0.8232	0.6465

For a model that provides accurate predictions the ME should be close to zero, indicating that the predictions are unbiased. The RMSE should be as small as possible, indicating that the predictions are close to the measured values. The RMKV should be naturally as small as possible. If the kriging variances are accurate then the RMSSE should be close to 1 [10]. If it is higher, the kriging prediction is too optimistic about the variability of the estimatives.

The results of Table IV and Table V suggest that both semivariogram exponential models (equations (3) and (6)) give the best predictions and should then be use to estimate salinity and temperature of the studied area.

D. Ordinary kriging

After a variogram model is selected, kriging is applied to estimate the value of the variable at unsampled locations, using data from surrounded sampled points.

The estimation is based on the semivariogram model and therefore takes it account the spatial variability of the variable in study. The kriging method belongs to the best linear unbiased estimators (BLUE) family. The kriging estimator is said to be linear because the estimated value is a linear combination of the measurements, and is written in the form of:

$$\hat{Z}(x_0) = \sum_{i=1}^{M} \alpha_i Z(x_i), \qquad (13)$$

where $\hat{Z}(x_0)$ is the estimated value for location x_0, M is the number of observations in the neighborhood of x_0 used in the estimative, and α_i are the correspondent weights.

Ordinary kriging is used when the mean value of the variable in study is unknown. For this estimator to be unbiased, for any value of the mean, it is required that $\sum_{i=1}^{M} \alpha_i = 1$. The estimated value is obtained by minimizing the kriging variance with the help of the Lagrange multipliers in order to impose the unbiased condition [11]-[14].

IV. RESULTS

The prediction contours map of salinity distribution in the vicinity of Foz do Arelho sea outfall discharge, obtained using the exponential semivariogram model (eq. (3)) is shown in Fig. 7. The prediction contours map of temperature distribution in the vicinity of Foz do Arelho sea outfall discharge, obtained using the exponential semivariogram model (eq. (6)) is shown in Fig. 8. Both contours maps show clearly the spatial variation of salinity and temperature in the studied area. In these maps it is possible to identify unambiguously the effluent plume east-southeast of the diffuser, which is in accordance with the current direction at the surface of about 120º measured. The plume appears as a region of lower salinity and lower temperature compared with the surrounding ocean waters at the same depth. The plume position which is almost at the surface is probably due to the very low stratification usually found at this time of the year. We may say that the high spatial resolution of the measurements provided a clear view of the plume position and concentration. But we are certain that better results may be found if a more efficient sampling strategy is adopted.

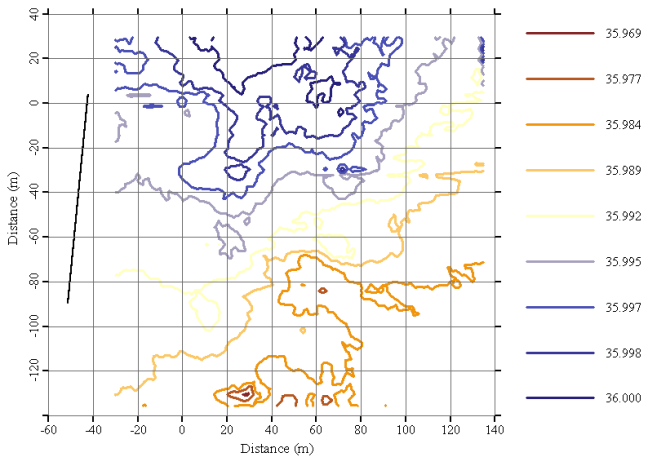

Fig. 7 Prediction contours map of salinity distribution in the vicinity of Foz do Arelho sea outfall discharge.

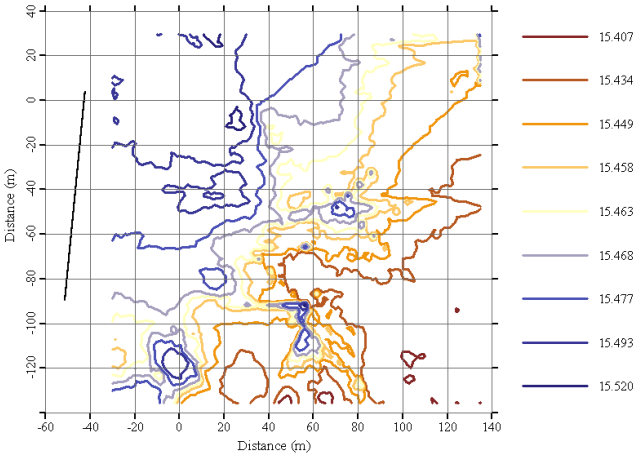

Fig. 8 Prediction contours map of temperature distribution in the vicinity of Foz do Arelho sea outfall discharge.

V. CONCLUSIONS

Geostatistical analysis of salinity and temperature obtained with an AUV in a monitoring campaign to an ocean outfall was able to produce a kriged map of the sewage dispersion in the field.

The spatial variability of the sampled data was analyzed previously calculating the classic statistical indicators. The results indicated an approximated normal distribution of the temperature data samples, which is desirable, and a negatively skewed distribution of the salinity data set.

Then, the Matheron's classical estimator was used to compute the experimental semivariogram for several directions. No effect of anisotropy could be shown.

The semivariogram was fitted to three theoretical models: spherical, exponential and gaussian. The cross-validation indicators suggested the exponential models as being the best semivariogram models among the candidates.

Finally, the predictions of salinity and temperature at unknown locations were obtained by ordinary kriging.

From the prediction contours maps of salinity and temperature distribution in the vicinity of Foz do Arelho sea outfall discharge, it is possible to identify unambiguously the effluent plume dispersion in the field. The plume was found mostly east-southeast of the diffuser at 1.5 m depth. The high spatial resolution of the measurements provided a clear view of the plume position and concentration.

Our study demonstrates that geostatistical analysis can provide good estimates of effluents dispersion very valuable for environmental impact assessment and management of sea outfalls. Moreover, since accurate measurements of plume's dilution are rare, these studies might be very helpful in the future for validation of dispersion models.

REFERENCES

[1] P. Ramos, "Advanced Mathematical Modeling for Outfall Plume Tracking and Management using Autonomous Underwater Vehicles based Systems", PhD Thesis, 2005, Faculty of Engineer. University of Porto.

[2] P.J.W. Roberts, “Environmental Hydraulics”, V. P. Singh and W. H. Hager Edn., Kluwer Academic Press Publishers, Printed in Netherlands, Chapter Sea Outfalls, 1996, pp. 63-110.
[3] N. Saby, D. Arrouays, L. Boulonne, C. Jolivet, A. Pochot, “Geostatistical assessment of Pb in soil around Paris, France”, Science of the Total Environment, 367, 2006, pp. 212-221.
[4] H. Wei, L. Dai, L. Wang, “Spatial distribution and risk assessment of radionuclides in soils around a coal-fired power plant: A case study from the city of Baoji, China”, Environmental Research, 2007, 104, pp. 201-208.
[5] T. Shoji, “Statistical and geostatistical analysis of wind. A case study of direction statistics”, Computers & Geocienses, 32, 2006, pp 1025-1039.
[6] T. Shoji, H. Kitaura, “Statistical and geostatistical analysis of rainfall in central Japan”, Computers & Geosciences, 32, 2006, pp. 1007-1024.
[7] S. Caeiro, M. Painho, P. Goovaerts, H. Costa, S. Sousa, “Spatial sampling design for sediment quality assessment in estuaries”, Environmental Modelling & Software, 18, 2003, pp. 853–859.
[8] C.J. Murray, H.J. Leeb, M.A. Hampton, “Geostatistical mapping of effluent-affected sediment distribution on the Palos Verdes shelf”, Continental Shelf Research, 22, 2002, pp. 881–897.
[9] K.F. Poon, R.W. H. Wong, M.H.W. Lam, H.Y. Yeung, T.K.T. Chiu, “Geostatistical modelling of the spatial distribution of sewage pollution in coastal sediments”, Water Research, 34, 2000, pp. 99-108.
[10] H. Wackernagel, “Multivariate geostatistics: An introduction with applications”, Berlin, Springer, 2003, 291p.
[11] N. Cressie, “Statistics for spatial data”, A Wiley Interscience Publication, New York, 1993, 900p.
[12] E.H. Isaaks, R.M. Srivastava, “Applied Geostatistics”, Oxford University Press, New York, 1989, 561p.
[13] G. Matheron, “Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature”, Paris. France: Masson; 1965, 305p.
[14] P. Kitanidis, “Introduction to geostatistics: Applications in hydrogeology”, New York (USA), Cambridge University Press, 1997, 249p.