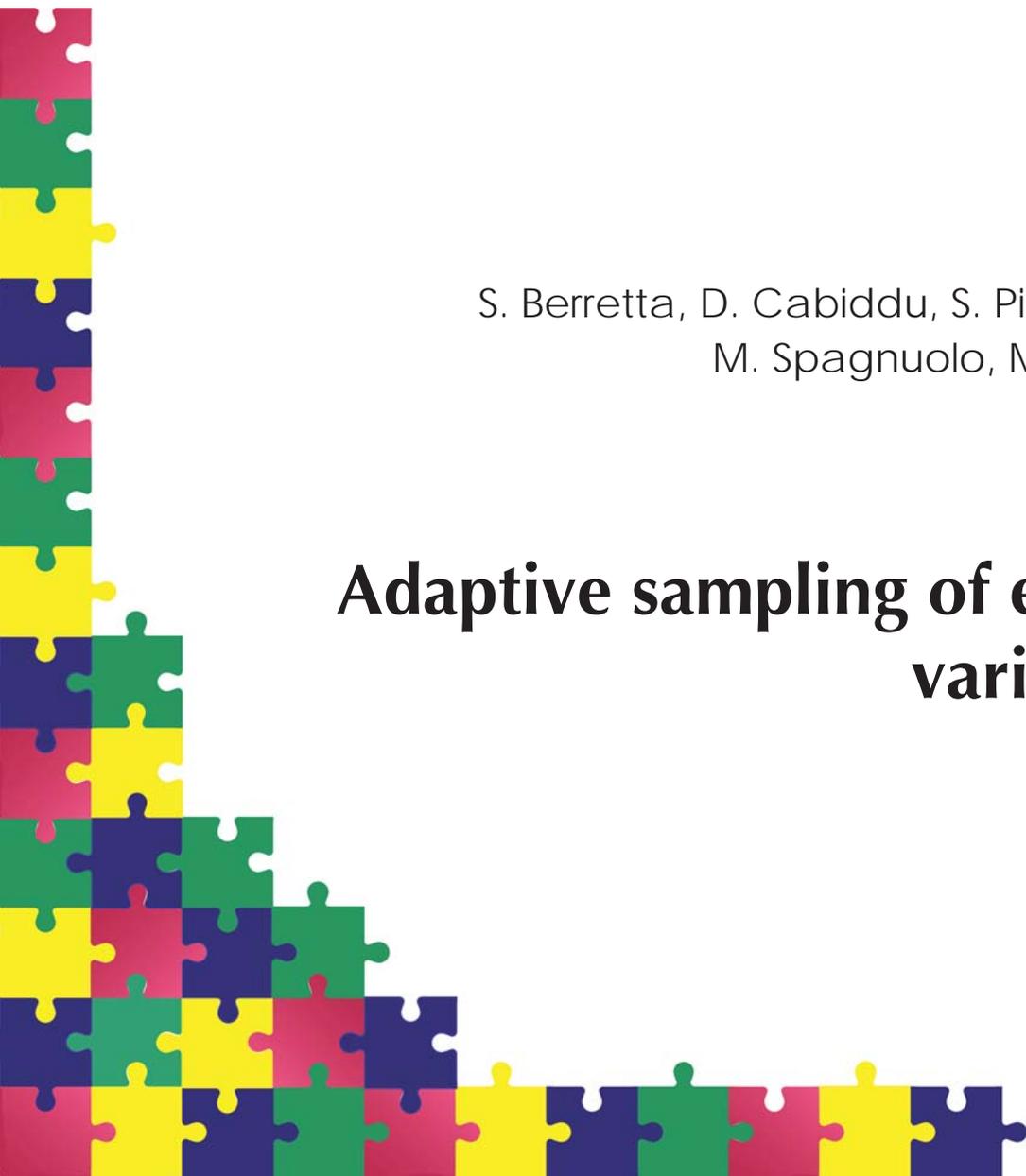


## REPORT SERIES

S. Berretta, D. Cabiddu, S. Pittaluga, M. Mortara,  
M. Spagnuolo, M. Vetuschi Zuccolini

# Adaptive sampling of enviromental variables (ASEV)



# IMATI REPORT Series

Nr. 18-06

June 2018

## **Managing Editor**

Michela Spagnuolo

## **Editorial Office**

Istituto di Matematica Applicata e Tecnologie Informatiche "*E. Magenes*"

Consiglio Nazionale delle Ricerche

Via Ferrata, 5/a

27100 PAVIA (Italy)

Email: [reports@imati.cnr.it](mailto:reports@imati.cnr.it)

<http://www.imati.cnr.it>

Follow this and additional works at: <http://www.imati.cnr.it/reports>

---

Copyright © CNR-IMATI, June 2018.

IMATI-CNR publishes this report under the Creative Commons Attributions 4.0 license.

## **Adaptive sampling of environmental variables (ASEV)**

Serena Berretta, Daniela Cabiddu, Simone Pittaluga, Michela Mortara,  
Michela Spagnuolo, Marino Vetuschi Zuccolini

*Corresponding author:*

Serena Berretta

Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes"

Consiglio Nazionale delle Ricerche

Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy

E-mail: [serena.berretta@ge.imati.cnr.it](mailto:serena.berretta@ge.imati.cnr.it)

Daniela Cabiddu

Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes"

Consiglio Nazionale delle Ricerche

Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy

E-mail: [daniela.cabiddu@ge.imati.cnr.it](mailto:daniela.cabiddu@ge.imati.cnr.it)

Simone Pittaluga

Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes"

Consiglio Nazionale delle Ricerche

Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy

E-mail: [simone.pittaluga@ge.imati.cnr.it](mailto:simone.pittaluga@ge.imati.cnr.it)

Michela Mortara

Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes"

Consiglio Nazionale delle Ricerche

Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy

E-mail: [michela.mortara@ge.imati.cnr.it](mailto:michela.mortara@ge.imati.cnr.it)

Michela Spagnuolo

Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes"

Consiglio Nazionale delle Ricerche

Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy

E-mail: [michela.spagnuolo@ge.imati.cnr.it](mailto:michela.spagnuolo@ge.imati.cnr.it)

Marino Vetuschi Zuccolini

DISTAV

Università degli Studi di Genova

Corso Europa 26, 16132 Genova - Italy

E-mail: [zucco@dipteris.unige.it](mailto:zucco@dipteris.unige.it)

---

## **Abstract.**

In environmental surveys a large sampling effort is required to produce accurate geostatistical maps, representing the distribution of environmental variables and the analysis of each sample is often expensive. The standard way to plan a survey is by non-adaptive sampling, whose distribution is usually completely specified prior to data-collection. In general, the sampling points are located on a regular grid, or along directions that are selected following a priori knowledge. However, the contribution of each point to the final prediction accuracy is typically unknown, and it is likely that lesser points distributed in a different manner might reach the same accuracy results eliminating sampling redundancies. Adaptive sampling uses only few initial sampling points and then follows an iterative collection of data, learning and refining the distribution of the variables in order to optimize the uncertainty of the estimates. In this technical report, we describe the design of an adaptive sampling scheme, where at each sampling step the next optimal point to be sampled is selected based on: i) an uncertainty map of the environmental parameter distribution, which is continuously updated when a new measure is acquired, and ii) the geometric and physics constraints given by the morphology of the surveyed area. The optimality criterion will take into account the trade-off between cost and prediction precision. This research is funded by the INTERREG-MATRAC-ACP project; its strategic objective is to enhance the protection of marine waters by improved real-time driving of ROVs equipped with a series of digital sensors. The project will lead to the definition of protocols for intervention in emergency situations with minimum risks for human operators, and more efficient monitoring procedures in routine conditions. Before starting experimentation on the field, a simulation study was developed to evaluate the effectiveness of this strategy.

**Keywords:** *Adaptive, Environmental, Gaussian simulation, Sampling*

---

*[page left intentionally blank]*

# Adaptive sampling of environmental variables (ASEV)

Serena Berretta<sup>1</sup>, Daniela Cabiddu<sup>1</sup>, Simone Pittaluga<sup>1</sup>, Michela Mortara<sup>1</sup>, Michela Spagnuolo<sup>1</sup>, and Marino Vetuschi Zuccolini<sup>2</sup>

<sup>1</sup>CNR - Institute of Applied Mathematics and Information Technologies

<sup>2</sup>DISTAV - University of Genova

## ABSTRACT

In environmental surveys a large sampling effort is required to produce accurate geostatistical maps, representing the distribution of environmental variables and the analysis of each sample is often expensive. The standard way to plan a survey is by non-adaptive sampling, whose distribution is usually completely specified prior to data-collection. In general, the sampling points are located on a regular grid, or along directions that are selected following a priori knowledge. However, the contribution of each point to the final prediction accuracy is typically unknown, and it is likely that lesser points distributed in a different manner might reach the same accuracy results eliminating sampling redundancies. Adaptive sampling uses only few initial sampling points and then follows an iterative collection of data, learning and refining the distribution of the variables in order to optimize the uncertainty of the estimates. In this technical report, we describe the design of an adaptive sampling scheme, where at each sampling step the next optimal point to be sampled is selected based on: i) an uncertainty map of the environmental parameter distribution, which is continuously updated when a new measure is acquired, and ii) the geometric and physics constraints given by the morphology of the surveyed area. The optimality criterion will take into account the trade-off between cost and prediction precision. This research is funded by the INTERREG-MATRAC-ACP project; its strategic objective is to enhance the protection of marine waters by improved real-time driving of ROVs equipped with a series of digital sensors. The project will lead to the definition of protocols for intervention in emergency situations with minimum risks for human operators, and more efficient monitoring procedures in routine conditions. Before starting experimentation on the field, a simulation study was developed to evaluate the effectiveness of this strategy.

## 1 Introduction

Nowadays, new lighter and cheaper sensors are available, and will be even more in the near future, yielding to an explosion of highly accurate data collection for many different purposes. This requires innovative computational solutions to make data analysis precise and fast. The INTERREG project MATRAC-ACP (Adaptive Monitoring in Real-Time with Automatization of Sampling) provides evidence of the perspective that technological evolution can bring in sampling approaches. The project started in April 2018 with the strategic objective of enhancing the protection of marine waters in ports by improving real-time monitoring procedures through highly automated monitoring protocols. The two case studies refer to the marine water in the Genova and Toulon harbours where the survey field will be defined by bathymetric data and delimited by piers. The project will lead to the definition of protocols for intervention in emergency situations with minimum risks for human operators, and more efficient monitoring procedures in routine conditions, repeatable and accurate compared to those currently in use.

The key scientific challenge of the project is to define and deploy an effective and efficient computational mechanism that implements an adaptive sampling strategy, guiding underwater robots equipped with sensors.

In the field of environment analysis, innovative sampling designs are being developed in the last decades through the use of optimization techniques. New sampling schemes are defined in order to describe a phenomenon through a representation of it, which must be as precise as possible, but using only a limited amount of selected information to optimize this reliability. The choice of the optimal sampling technique is not an easy task, but it is necessary to reduce the computational cost of the analysis while getting the same information that would be obtained having the entire sample population available. In this research program, the optimization of sampling schemes is based on the measures of uncertainty. Uncertainty can be viewed as “a sort of information deficiency” (1) which may be a consequence of different factors, such as the limitation of human capacity and knowledge, the precision of the acquisition tools, the selected analytical techniques and a pre-established budget. In the past, uncertainty has often played a “passive” role and was only used retrospectively as a tool for measuring the accuracy of a model. Our adaptive sampling approach makes uncertainty an “active” part of the decision process.

In the following, we will describe the research programme that we will tackle in MATRAC-ACP, starting with a description of the preliminary synchronization mechanisms of sensor data and following then with the presentation of preliminary results obtained with the synthetic data produced to validate the technique under development.

## 2 Towards an adaptive sampling strategy

One of the main objectives of the project is to define a new real-time adaptive sampling method which exploits the precision of modern acquisition devices to improve the environmental monitoring procedures. The design of the spatial and stochastic iterative adaptive procedure is devoted to build a resilient tool for the *real-time* survey planning in a module able to define under some criteria the sequence of the points to be sampled. Our aim is to develop such an iterative adaptive sampling by considering:

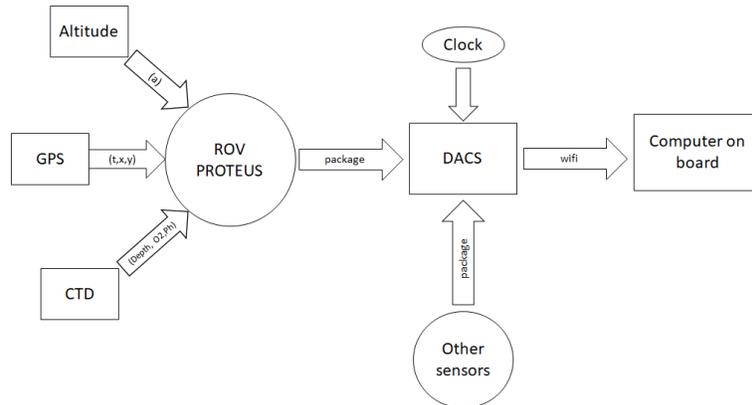
- an uncertainty map of the environmental parameter distribution, which is continuously updated;
- the geometric and physics constraints given by the morphology of the surveyed area;
- additional physical constraints (e.g., topography, accessibility, distance, chemical constraints, physical constraints on obstacles, robot battery consumption).

In the MATRAC project the instrumentation includes sensors mounted on a Remotely Operated Vehicle (ROV), each acquiring several environmental variables at a sensor-specific and variable sampling rate. The features of the sensors selected in the project, and the procedure used to synchronize the measures are described below.

### 2.1 Sensor Data Acquisition and Synchronization

In our scenario, data acquisition will be performed by a set of specialized sensors embedded into a ROV (Remotely Operated Vehicle). The interface between the user and the ROV will be a Graphical User Interface by which the user will supervise and guide the vehicle. Behind the graphical interface, a set of specialized modules will be integrated in the system and will be responsible for real-time data storage, synchronization, and analysis.

Figure 1 shows complete overview of our reference ROV and the connection among the components of the system.



**Figure 1.** Design of the communication protocol.

Our reference ROV will be equipped with a set of sensors (See Table 1). Each sensor has its own sampling rate and may be able to sample one or more environmental variables.

**Table 1.** Each sensor has different properties that describe the behaviour of acquisition. All sensors measure inside a range with a specific unit. Some sensors need to wait some time to stabilize the measurement; others sensors have a delay between the moment in which the sample has been gathered and the sensor return a value. Features of sensors: RES (resolution), FREQ (frequency).

SENSOR	TYPE OF MEASURE	SCALE	RANGE	RES	FREQ	DELAY	STABILIZATION	MEASURE IN MOTION
SALINITY	Salinity	PSU	0.0 - 40.0	-	1s	-	-	Yes
PH	Ph	adimensional	0.0 - 14.0	-	1s	-	10s	Yes
T	Temperature	celsius (C)	0.0 - 30.0	-	1s	-	-	Yes
DO	Dissolved oxygen	mg/l	0.0 - 12.0	-	1s	-	-	Yes
EH	State REDOX	mV (Volt)	-500.0 - 1300.0 (-5.0 - 1.3)	-	1s	-	-	Yes
COND.	Conductivity	micro-, milli- Siemens	0.0 - +∞	-	1s	-	-	Yes
APNA	Nutrients (NO2/NO3/NH4/PO4)	mg/l	0.0 - +∞	-	1s	10s	-	Yes
C6	Refined oil	ppb (parts per bilions)	10 - 6000	10	1s	-	50s	No
CO2	CO2	ppm (parts per millions)	0 - 600	-	1s	-	-	Yes
ISE	Elements of LIBELIUM	mol/l (mg/l)	$10^{-1} - 10^{-7}$	-	1s	-	-	Yes

## Data Acquisition

While sampling, each sensor sends the acquired data to a *DACS* (Data Acquisition Control System) which is responsible for forwarding such a data flow to the computer located on board and linked through Wi-Fi connection. Additionally, our ROV will be equipped with a machine called *PROTEUS*, on which three specialized sensors are mounted: a GPS, a CTD (Conductivity - Temperature - Depth) sensor and an altimeter. The GPS detects the position during harbour water sampling, while the CTD sensor measures the temperature of the marine water,  $O_2$ , Ph, salinity and provides information about the position (coordinates, depth from sea surface and the altitude from sea bed). Similarly to the previously described sensors, *PROTEUS* sends acquired data to the *DACS*.

From a very technical point of view, the *DACS* is responsible to convert data received from each sensor into a uniform, customized data format to enhance the communication stability and reduce the subaerial noise. We define our data format to be as simple and general as possible, in order to manage any kind of sensors besides those already planned.

Our data format is actually a structured string composed by 7 fields. Table 2 shows the string structure, as a composition of the following fields:

- The *timestamp* (*DACS*'s clock) is necessary to synchronize data and supervise the data acquisition;
- The *sensor type* field encodes the type of sensor from which the data is coming from (i.e. GPS, CDT, ...);
- The unique identifier of the sensor (*ID*) refers to the specific sensor;
- The acquired data.

**Table 2.** Data structure. Type of sensor can be GPS or SENSOR.

TIMESTAMP	TYPE OF SENSOR	ID NUMBER	X-GPS	Y-GPS	Z-GPS	MEASURE
-	-	-	-	-	-	-

## Data Synchronization

Independently of the complexity of deployment and environment (land, sea, or air), an environmental survey is carried out by means of:

- a set of sensors to sample the environment and to measure the environmental variables;
- a GPS system that provides the positions corresponding to the measurements.

Each sensor (including GPS) has its own sampling rate and starts measuring at a different starting time. Additionally, sensors are equipped with an internal clock in order to provide an acquisition timestamp. Nevertheless, it is not possible to have exactly synchronized timestamps between GPS and sensors. Also, it may often happens that no measurement corresponds to the GPS's timestamp. To overcome this problem we proceed to an interpolation of the sensor measurements. There are many ways to interpolate two values; in this preliminary work we select a simple linear interpolator thanks to the high density of information.

The procedure of synchronization can be summarized as follows: as soon as the data flow has been received by the computer on board as a stream of data, the synchronization procedure identifies whether the first value of the file comes from the sensor or the GPS. In the first case, values have this order: sensor, GPS, sensor, GPS, ... , and we proceed by selecting the two sensor values that precede and follow the position of the GPS for which an interpolation is required. By exploiting a simple linear interpolation based on timestamps, we compute the value of environmental variable corresponding to GPS's timestamp. In the second case, instead, the first value of the stream is eliminated since in correspondence of that position / timestamp there wouldn't have been two measurements of the sensor to interpolate (one that precedes and the other that follows the value of the GPS).

Once this position is eliminated, we can proceed as in the first case and interpolation is carried out. The same happens for the last value of the data stream: if the last value corresponds to a measure of the GPS it is not considered since the next measurement of the sensor would not be available for interpolation. Once the synchronization procedure is finished, the structure of the data stream is transformed into the one shown in the Table 3.

**Table 3.** Data structure after synchronization.

TIMESTAMP	X-GPS	Y-GPS	Z-GPS	MEASURE
-	-	-	-	-

### 3 The adaptive sampling algorithm

Our adaptive sampling approach can be described as an iterative method which involves three sequential steps: an initialization phase, a main iterative phase and a final control phase. At each iteration, our method estimates the best next point to be sampled and updates the known information with new acquired data. The expected input to our approach is a stream of data encoded according to our format shown in Table 3. Additionally, our method requires a description of the area to be sampled and an initial set of randomly sampled points. At this stage, the area to be sampled is represented as a regular grid with a user-defined resolution, while in future development will represent the actual morphology of the domain.

An important feature in our sampling is the description of the degree of spatial dependence in a field through the variogram. It depends only on the distance that separate two observations ( $h$ ). The variogram must be expressed as a mathematical function before being used for geostatistics methods. This function is expressed in terms of three parameters: *range* of spatial correlation, *sill* value and *nugget effect*. Usually, a variogram increases with distance,  $h$ , until sill value is reached, at a specific distance (range), representing the sample variance. The *nugget effect* measures discontinuity of the process.

In the initialization step, we acquire an initial set of points and verify they satisfy the requirement to be normally distributed. This is because we exploit geostatistics techniques suitable for normally distributed data. If not, we carry out the Normal Score (NSCORE) transformation to match such requirement.

The adaptive sampling procedure works as follows:

1. A variogram is fitted to the set of sampled points. In the first iteration, these are just the points sampled during the initialization phase;
2. Sequential Gaussian Simulations (SGS) are performed to compute, for each cell of the grid, a set of value estimations;
3. Based on the results of the SGS, mean and variance of each cell are calculated. This information is used to generate a spatial prediction map and the related uncertainty map, respectively;
4. by analysing the uncertainty map, the next point to be sampled is selected as the center of the grid cell having the maximum variance;
5. The next sampling position is communicated to the ROV so that it can actually move to the destination. While moving, sensors continue acquiring data and send them back to the system (see Section 4);
6. The data stream encoding new measures acquired during the displacement is received by the adaptive sampling module, which adds them to the set of samples.
7. If a stop criterion is reached, the procedure ends, and the data are transformed back with the NSCORE transformation if needed. Otherwise, a new iteration begins (step 1).

The current version of our method considers different stop criteria, such as a total number of sampled points, or a desired maximum variance value. In the next period, we will define an optimal stop criterion specific for the actual application scenario.

### 4 Components of the Simulator

In order to simulate the whole adaptive sampling procedure in our lab before actually testing it in the real setting, we built a software ROV simulator.

The environment is represented as a synthetic three dimensional grid of values that mimics the real area to be sampled. The synthetic grid is generated by GSLIB (Deutsch and Journel , 1998). We can control the generated distribution by setting sill and range values of the variogram, and by selecting a few samples as nuggets. Additionally, in order to replicate as far as possible the real scenario, the ROV-simulator is highly configurable: it may simulate a configuration with one or more sensors embedded, and each sensor properties can be customized (i.e. frequency, delay, ... ).

#### **Remote Operated Vehicle simulator**

The ROV displacement from a source to a target point is simulated according to kinematics laws, which have been implemented to mimic the platform motion in a three dimensional space. We assume that there are not external perturbations like waves, wind or currents.

The cruise speed and the acceleration of the ROV is defined by the user, as well as the source and the target point of the course of the vehicle, which starts from a static state, accelerates to reach the cruise speed and then decelerates to stop at the target point. The ROV is embedded with an onboard clock synchronized with the (simulated) GPS time.

### Sensors simulator

The simulated ROV can be equipped with different sensors defined by user. For each simulated sensors specific sampling rate, sampling delay, startup delay can be defined. To simulate the real sampling, each sensor reads data from the synthetic three dimensional grid representing the distribution of an environmental variable to be sampled.

In our experiments, the synthetic grid is the result of both unconditional and conditional sequential gaussian simulations (SGSIM+POSTSIM). The sampling rate describes the number of samples for time unit. The sampling delay represents the time needed to get the measurement since the sample is acquired. The startup delay is the time needed to get the first stable measurement. Each sensor can be defined with different settings. In our tests, the GPS is assumed to provide the position of ROV every second, and gaussian error is added to simulate the uncertainty of real device.

Our simulator is able to simulate any complex underwater positioning system, that in a harbour is a tricky job.

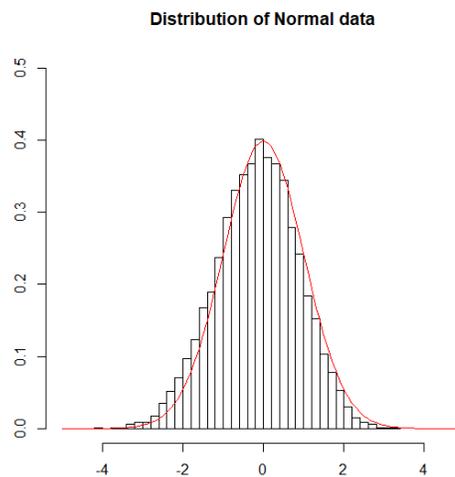
### Communication simulator

All data collected from sensors and GPS are stored and sent to the adaptive sampling module as a stream of data structured as in table 2. Actually, not all the fields will be filled at the same time, depending on the sensor. The fields `TIMESTAMP`, `TYPEOFSENSOR` and `ID NUMBER` are always determined. The fields `X-GPS`, `Y-GPS`, `Z-GPS` are set when the GPS is communicating its position, while the `MEASURE` field will be void. The opposite holds for the other sensors. We use a null placeholder encoded as a “Not-Available” (NA) value to indicate that there is no measurement associated with it.

## 5 Experimental Results

A prototypal version of our adaptive sampling approach has been implemented using the *R* software. *R* is a free software environment for statistical computing and graphics. Behind the adaptive sampling method, the ROV-simulator has been implemented in C++. Both tools are multi-platform and have been tested on different operating systems (Windows, Linus, MacOS).

In our experiments, we simulate a scenario where data satisfy the normality condition (Figure 2). The initial set of sampled

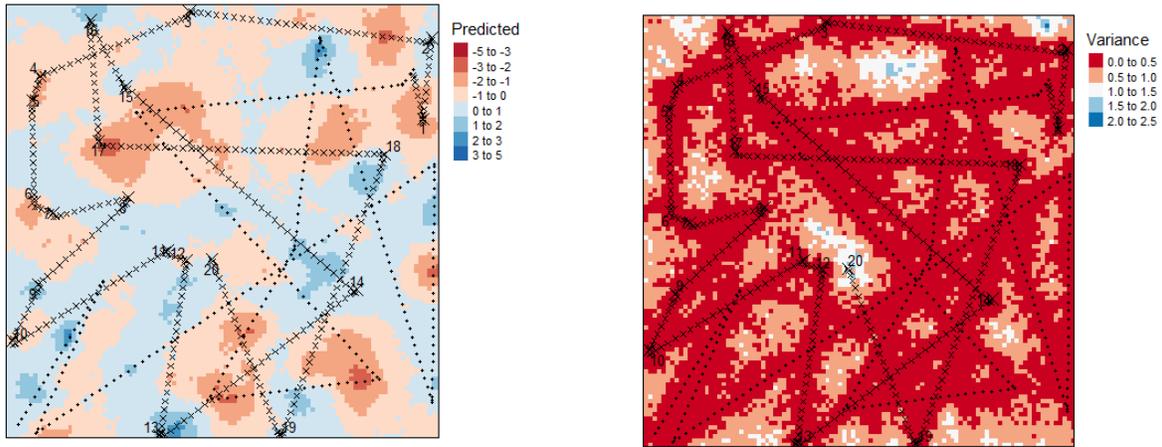


**Figure 2.** Histogram of the distribution of Normal data.

points is selected as follows: we start from a initial point and we select the next one with the criterion of minimizing uncertainty of the estimate with SGS; we continue the sampling until we have a set of data that is considered sufficient to begin the principal sampling.

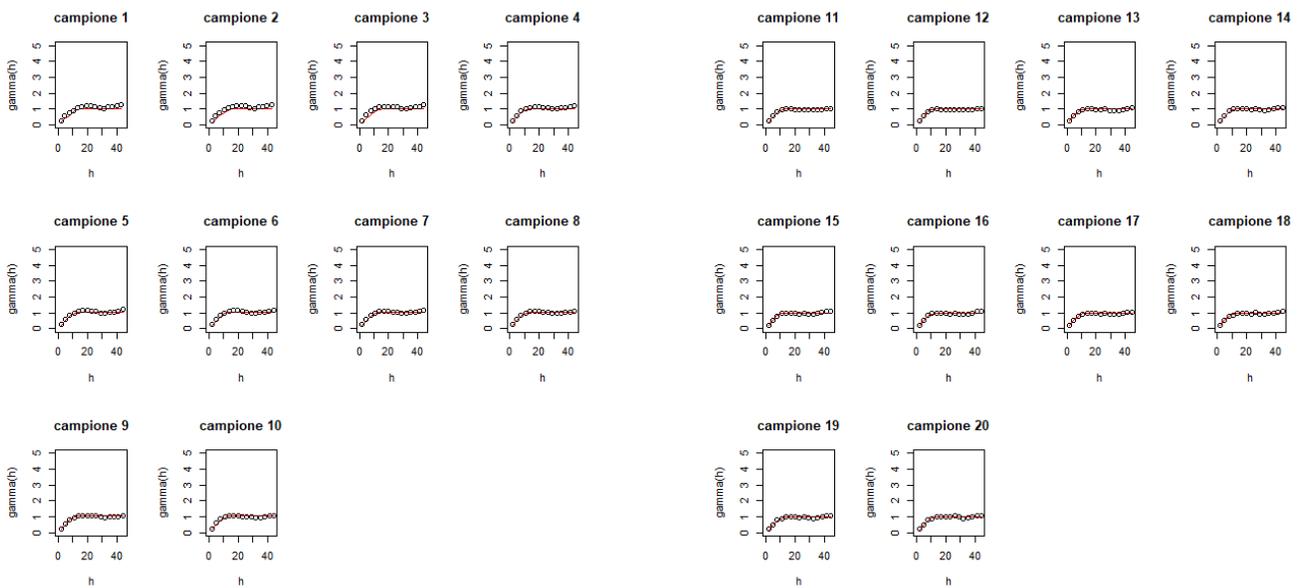
In this experiment, we set the stop criterion to be a desired number of adaptively selected points. Note that, since the sensors continuously acquire data while moving, the total number of sampled points is much greater than the number of the adaptive selected points. Figure 3 shows the spatial prediction map (on the left) and the uncertainty map (on the right) of the simulated field. At each iteration, the next point to be sampled is selected as the one having the maximum uncertainty, but trying to avoid too long displacements. This is done by selecting the new point as the one with the maximum uncertainty among a set of points which are closest to the current position. As an example, see the uncertainty map in Figure 3 where the cell with the greater

value of variance is at the top right of the map, while the twentieth adaptive sampled point is actually at the center of map. This is because the selected point is a local maximum in terms of variance.



**Figure 3.** Spatial prediction map (left) and Uncertainty map (right).

Figure 4 shows the variograms and their fitting for each iteration of adaptive sampling. The automatic fitting calculated by function *autofitVariogram* is satisfied.



**Figure 4.** Variograms and their fits (red line) for each iteration of entire sampling.

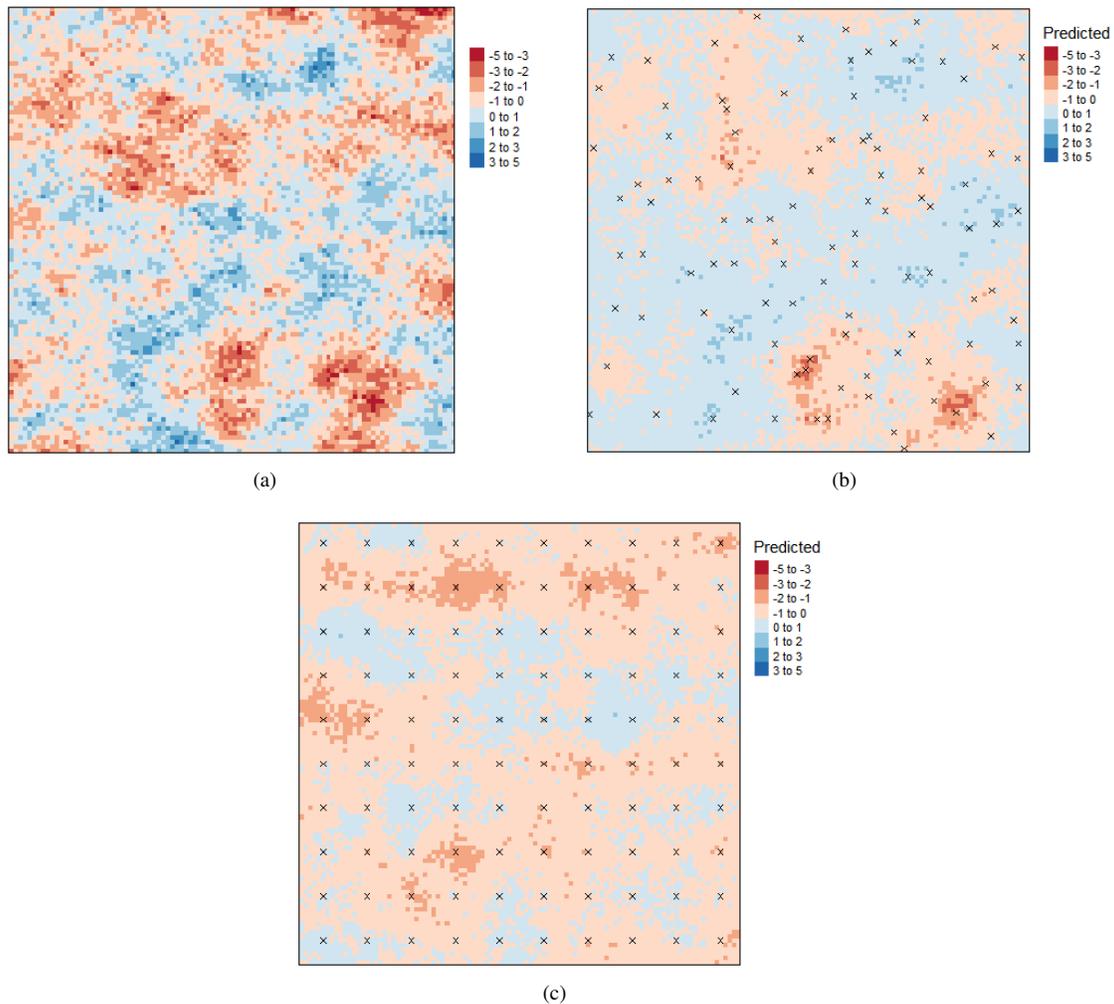
### 5.1 Validation

Conversely to adaptive sampling in the real world, our simulated environment enables the possibility to compare our results with the “real data”. This is possible since our “real world” is the synthetic grid providing the actual data for each cell. We use this information, which is usually unknown, to evaluate the reliability of the new approach. Reliability is measured as Mean Square Error (MSE) between reality and estimated values:

$$MSE = \sum_{i=1}^n (\hat{Y}_i - Y_i)^2$$

where  $n$  is the number of cells in the grid,  $\hat{Y}$  are the estimated values from SGS and  $Y$  are the real data.

Our simulated field is shown in Figure 5(a). At each iteration of our adaptive sampling, we calculate the MSE to evaluate the reliability improvement of the estimates. We recall that, when a new point is reached, not only such a point is added to data, but also all those that are sampled during the route of the ROV.



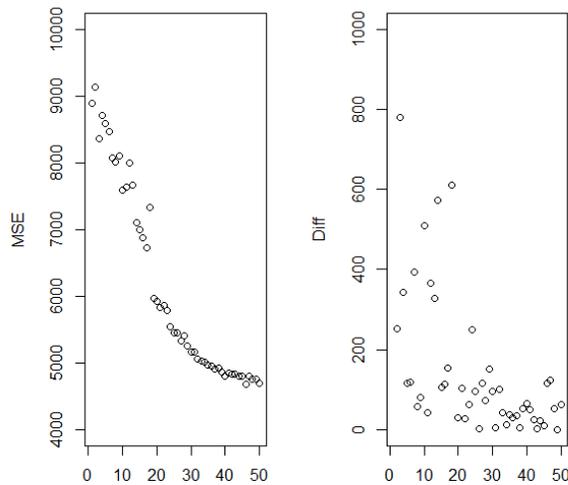
**Figure 5.** (a) Synthetic variable distribution representing reality in our experiment. (b) Spatial prediction map with adaptive sampling ( $MSE = 8838.204$ ) and (c) sampling on regular grid ( $MSE = 13757.02$ ).

In Figure 6 two graphs are shown. The graph on the left underlines a decreasing trend of the MSE while increasing the number of samples. However, when a sufficiently good number of sampled point is reached (i.e. thirty in our case), the MSE seems to be stable and its decrease is not relevant anymore. Similarly, the graph on the right shows the difference between MSE in two subsequent iterations. In other words, it measures the improvement provided by a new adaptively sampled point. Again, after about thirty adaptive samples this difference tends to zero.

These two graphs suggest the choice of a feasible stop criterion. Indeed, adding new samples after a certain number does not add enough information compared to the costs of sampling. Nevertheless, in the real case, the real data is unknown and, thus, MSE with respect to the reality cannot be evaluated this way.

## 5.2 Comparison with traditional sampling

The standard way to plan a survey to evaluate environmental parameters is by non-adaptive sampling completely specified prior to data-collection. In general, the sampling points are located on a regular grid, or along directions that are selected following a priori knowledge. To demonstrate that adaptive sampling actually allows to get an informative result as precise as the previous methods while reducing sampling efforts, we compare the prediction maps and measure their reliability with respect to reality (which in the case of the simulations is known). Again, we use MSE to measure reliability.



**Figure 6.** Behavior of the MSE for each iteration (left) and of the differences between two successive iterations (right).

We set the number of adaptive sampled points to a fixed amount (i.e. 100 in our experiment). Also, we slightly modify the adaptive approach to make the two techniques comparable. Specifically, we discard values acquired during the ROV displacement and keep only measures at sample points. Figure 5(b) shows the prediction map referred to our adaptive sampling scheme, while 5(c) shown the prediction map of traditional sampling on regular grids. By looking at the MSE values, it is evident that, for the same number of samples, adaptive sampling is more reliable than traditional methods. The regular grid sampling indeed studies the environmental phenomenon at a predefined resolution: the cell dimension corresponds to the minimum sampling interval. If this value is not small enough, we risk to loose important information about the behavior of the variable between the sample points, and the estimate will be rough in these zones.

## 6 Conclusions

We propose a new approach to adaptively sample environmental variables. Using the uncertainty map as a criterion for optimization, the adaptive sampling provides the same level of information of traditional methods with a minor number of samples. Thanks to the uncertainty map, our method detects those points where more information is needed (i.e. the more uncertain ones) and enables the possibility to iteratively enhance precision and reliability by adding to the model information related to all the sampled points.

We implemented a simulator to carry out preliminary testing of the system. In our experiment, adaptive sampling showed an excellent performance for environmental survey with respect to the traditional techniques. This result supports the choice of an adaptive system to monitor water quality in the MATRAC-ACP case study and motivates further research towards the definition of new efficient and cost-effective environmental surveys.

In the near future, we will simulate a asymmetric scenario to evaluate the robustness of the new adaptive method to sample several type of spatial data distributions. Moreover, additional stop criteria (i.e. a certain variance threshold is reached) are going to be defined and tested. Finally, we will extend the approach and the implementation of our tools to support 3D fields. So far, we simulated the sampling over a bi-dimensional domain; in the real case, in the harbour, it will be very interesting to examine the behavior of the environmental variable on the third component too (depth). We are going to put an additional effort also in defining and developing advanced visualization tools for representing the uncertainty in 3D data.

## Acknowledgements

This research is financed by the project "Monitoraggio Adattivo in Tempo reale con Automatizzazione del Campionamento – Aree Costiere Portuali – MATRAC-ACP", funded by the EU Interreg V-A Italia Francia Marittimo 2014-2020 - Asse prioritario del Programma 2 - Protezione e valorizzazione delle risorse naturali e culturali e gestione dei rischi - Obiettivo specifico della Priorità d'Investimento 6C2-Accrescere la protezione delle acque marine nei porti.

## References

1. Fedrizzi, M., Kacprzyk, J., Yager, R. R. *et al.* *Advances in the Dempster-Shafer theory of evidence* (Wiley, 1994).
2. Cressie, N. *Statistics for spatial data* (John Wiley & Sons, 2015).
3. Marchant, B. & Lark, R. Adaptive sampling and reconnaissance surveys for geostatistical mapping of the soil. *Eur. journal soil science* **57**, 831–845 (2006).
4. Marchant, B. & Lark, R. Adaptive schemes for geostatistical sampling and survey. In *MODSIM 2003: International Congress on Modelling and Simulation*, 1709–1714 (Citeseer, 2003).
5. McBratney, A., Webster, R. & Burgess, T. The design of optimal sampling schemes for local estimation and mapping of regionalized variables—i: Theory and method. *Comput. & Geosci.* **7**, 331–334 (1981).
6. Pebesma, E. J. Gstat user's manual. *Dept. Phys. Geogr. Utrecht Univ. Utrecht, The Neth.* (2001).
7. Rocchini, D. *et al.* Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Prog. Phys. Geogr.* **35**, 211–226 (2011).
8. Swan, A. Deutsch, C. V. & Journel, A. G. 1998. gslib. geostatistical software library and user's guide, 2nd ed. x 369 pp. cd-rom. oxford, new york: Oxford university press. price £45.00 (hard covers). isbn 0 19 510015 8. *Geol. Mag.* **136**, 83–108 (1999).
9. Van Groenigen, J., Stein, A. & Zuurbier, R. Optimization of environmental sampling using interactive gis. *Soil Technol.* **10**, 83–97 (1997).
10. Vckovaki, A. *Interoperable and distributed processing in GIS* (CRC Press, 2014).
11. Webster, R. & Oliver, M. A. *Geostatistics for environmental scientists* (John Wiley & Sons, 2007).

---

**Recent titles from the IMATI-REPORT Series:****2018**

**18-01:** *Arbitrary-order time-accurate semi-Lagrangian spectral approximations of the Vlasov-Poisson system*, L. Fatone, D. Funaro, G. Manzini.

**18-02:** *A study on 3D shape interaction in virtual reality*, E. Cordeiro, F. Giannini, M. Monti, A. Ferreira.

**18-03:** *Standard per la gestione e procedure di validazione di dati meteo da sensori eterogenei e distribuiti*, A. Clematis, B. Bonino, A. Galizia.

**18-04:** *Strumenti e procedure per la gestione e la visualizzazione di dati meteo prodotti da sensori eterogenei e distribuiti tramite interfacce web basate su mappe geografiche*, L. Roverelli, G. Zereik, B. Bonino, A. Gallizia, A. Clematis.

**18-05:** *TopChart: from functions to quadrangulations*, T. Sorgente, S. Biasotti, M. Livesu, M. Spagnuolo.

**18-06:** *Adaptive sampling of environmental variables (ASEV)*, S. Berretta, D. Cabiddu, S. Pittaluga, M. Mortara, M. Spagnuolo, M. Vetuschi Zuccolini.