Water Quality Map Extraction from Field Measurements Targetting Robotic Simulations

A Anderson, J Martin, J Mougin, Noury Bouraqadi, E Duviella, L Etienne, Luc Fabresse, K Langueh, G Lozenguez, C Alary, et al.

Water Quality Map Extraction from Field Measurements Targetting Robotic Simulations

A. Anderson [*,**] **J.G. Martin** [***] **J. Mougin** [****] **N. Bouraqadi** [*]
E. Duviella [*] **L. Etienne** [*] **L. Fabresse** [*] **K. Langueh** [*] **G. Lozenguez** [*]
C. Alary [*] **G. Billon** [****] **P.J. Superville** [****] **J.M. Maestre** [***]

[*] *IMT Nord Europe, Univ. Lille, F-59000 Lille, France*
[**] *Instituto de Desarrollo Tecnológico para la Industria Química (INTEC), Consejo Nacional de Investigaciones cientficas y tecnicas (CONICET) and Universidad Nacional del Litoral (UNL), Santa Fe, Argentina*
[***] *Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, C/ Camino de los Descubrimientos, s/n., 41092 Sevilla, Spain*
[****] *Univ. Lille, CNRS, UMR 8516 - LASIRE - Laboratoire Avancé de Spectroscopie pour Les Intéractions La Réactivité et L'Environnement, 59000, Lille, France*

Abstract: The assessment of the quality of water can be shortly defined as the analysis of its physical, chemical and biological characteristics in order to determine the degradation of freshwater resources. In this context, one of the latest technological methods for real-time data acquisition comes from the use of unmanned vehicles (aerial, surface and underwater). Therefore, the development of control strategies to perform environmental missions is crucial to manage water resources in an efficient and effective way. Prior to the actual implementation, some in-silico experiments are needed to test the proposals, which is one of the purposes of this work. This proposal, based on real experiments in a lake, presents a novel method for the construction of a water quality map based on polygons. The result is compared with a classical data generation method showing positive outcomes. The generated limnological map has a twofold purpose: to test set-based predictive controllers in simulation scenarios with an aquatic robot and to determine if there is a source of contamination in the analyzed region of the lake.

Keywords: Unmanned Surface Vehicle, Environmental Monitoring, Water Quality, Kriging

1. INTRODUCTION

Water resource management is an integral aspect of the preventive management of aquatic ecosystems health and drinking-water quality. A comprehensive monitoring program is becoming a necessity in order to safeguard public health and to protect the freshwater resources as required by the Water Framework Directive (WFD) (Edition, 2011; Kannel et al., 2007).

Water quality assessment requires real-time measurement of many physical and chemical parameters. Usually, the monitoring of these parameters exploits highly reliable instrumentation which is deployed only in few locations. Nevertheless, the vast collection of data necessary to define the conservation status of a region represents a set of challenges to be solved. For instance, the size of the monitored area will require the deployment of a number of detection devices ranging from a single device to a fleet of unmanned vehicles covering several points. The second challenge comes from the remoteness of the sites to be monitored. In this case, the sensing devices must be provided with some kind of data transmission system. Finally, to be efficient, human intervention must be kept to a minimum (Madeo et al., 2020).

To fulfill all the previous objectives, recent works propose the use of Unmanned Surface Vehicles (USV) for the collection of data (Sinisterra et al., 2017; Siyang and Kerdcharoen, 2016). The implementation of autonomous vehicles offers the flexibility for a variety of intervention plans and it provides an energy efficient and robust solution for specific tasks in sophisticated environments (*see* Wang et al. (2021) and the references therein).

In this preliminary work, the SPYBOAT® vessel - equipped with specific sensors to monitor turbidity, dissolved oxygen (DO), pH, conductivity, temperature from water surface - was tested in several environmental scenarios. Some of these scenarios were intended to guarantee the quality of the data by means of a proper validation using data provided by a Laboratory deployed in the field (Mougin, 2021) in coordinated experiments. Another experimental scenario was performed on the wider area of the Heron lake in Villeneuve d'Ascq, France. The monitored area is located close to the arrival of contaminated waters from storm overflows and was expected to show interesting gradients.

The main purpose of this paper is to provide a novel data approximation method to extract the maximum amount of information possible from a minimum amount of data. This approximation method is used to generate a limnological profile of the complete monitored region of the Heron lake aforementioned. The result is compared with a classical data generation method showing positive outcomes. The generated map has a two-fold purpose, to create a structural reliability framework to test predictive controllers in simulation scenarios with an aquatic robot and to determine if there is a significant source

of pollution in the analyzed region of the lake by considering the spatial direction in which the parameters rises most quickly.

The structure of this paper is as follows. Section 2 provides the architecture of the USV used on the real measures and a proper data validation analysis. Section 3 describes the main experiment on the Heron lake, including the region of interest, the analyzed parameters and the sampling procedure. The main result of the paper is presented in Section 4. This section presents an estimated limnological map of the region based on data collected on the field. Two different approaches were proposed to build the map, the strategy knowing as Kriging and a novel proposal based on the specific features of the experiment. A comparison between the results and a discussion about the methods are further presented. Lastly, Section 5 presents the conclusions and perspectives of the paper.

2. THE USE OF ROBOTS FOR REAL-TIME MEASUREMENT

The exploration and exploitation of water resources has a remarkable growth and development in robotics in recent decades. Unmanned surface vehicles (USVs) have become common in marine robotics executing missions such as source seeking, environmental monitoring and water sampling (in a coordinated way with multivehicles). In what follows the architecture and specifications of the USV employed to perform the environmental missions described in this work is presented.

2.1 SPYBOAT®'s description

The CT2MC company has designed a range of vessels dedicated to answer the need of data monitoring of freshwater resources. The main feature of these vehicles consists in a flat hull and aerial propulsion system. This guarantees the realization of sampling missions and inspections without contamination of the environment. The official SPYBOAT® system used in the present study is illustrated in Fig. 1.

Fig. 1. The SPYBOAT® technology is based on the use of an aquatic drone allowing the realization of water sampling and inspections by guaranteeing the non-contamination of the environment.

The SPYBOAT® technology follows standard equipment configuration including multiple sensors (localization system, compass, sonar, camera) and is propelled by two independent actuators. Thus the heading is controlled through a differential thrust method. The architecture of the SPYBOAT® technology is described in Fig. 2.

Remark 1. *It is noteworthy that in Hervagault (2019) a model was proposed based on appropriate assumptions according to the characteristics of the SPYBOAT® vessel. Moreover, this model was identified with measurements performed on a real system. The access to an identified model is a crucial point to the development of model predictive control strategies to automatized the environment missions.*

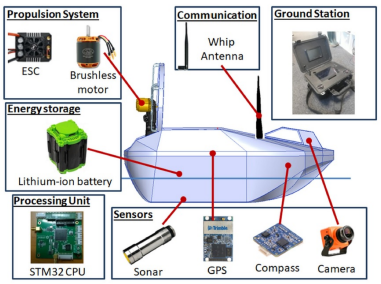


Fig. 2. Architecture of SPYBOAT® technology.

The USV is equipped with a *Hyperion* optical sensor from Valeport [1], for the measurement of the turbidity. It is also equipped with Tripod sensors from AquaLabo [2] to measure the temperature, DO, pH, and conductivity.

2.2 Data validation

The first experiment scenario - performed in the Marque River (close to Lille in France) - attempted to validate the SPYBOAT® sensors measurements. This was carried out with an intercomparison between the USV on a fixed location and a field Laboratory (static station on the side of the river) (Mougin, 2021), performing an online analysis of the physicochemical parameters of the water with a multiparameter probe (Manta+, Eureka Water Probes). The comparison was performed during a one hour monitoring period, with a one second delay between measurements for the USV and a two and a half minute delay for the laboratory.

Table 1. Comparison between sensors from USV and from field Laboratory.

	Temperature	pH	Conductivity	Turbidity	O_2
Shift	3%	6%	5%	17%	5%
Correlation	0.98	0.95	0.93	0.42	0.96

The results show a small shift in absolute value for most of the parameters between the USV and the laboratory (*see* Table 1). The small difference could be explained by some errors in the calibration and it will be necessary to give additional care to this essential step in the future. The measured variations over the hour were good as a very strong correlation between sensors (more than 95%) was found. Turbidity however was very different between the USV and the station. If calibration problems and noisy measures could be suspected, it is not impossible that both sensors are properly functionning but are not measuring in the same sample. Turbidity can indeed exhibit strong gradients with depth as particles will often come from the sediment and the lighter the particles, the higher they will ascend. The pump for the station is 1 m below the surface, 50 cm above the sediment while the USV measure a few cm under the surface. An intercomparison within the same sample or at the same depth should be done in the future.

[1] https://www.valeport.co.uk/content/uploads/2021/05/0901814i-Hyperion-Optical-Sensors-Operating-Manual.pdf
[2] https://en.aqualabo.fr/

3. DESCRIPTION OF ENVIRONMENTAL EXPERIMENTATION

Pollution of surface water occurs when too much of an undesirable or harmful substance flows into a water body, exceeding its natural ability to remove it (dilute or convert it to a harmless concentration). Water pollutants are categorized as being emitted from point or nonpoint sources. Point sources are distinct and confined, such as pipes from industrial and municipal sites that discharge into streams or rivers. In general, point source pollutants from industries are controlled through on-site treatment or disposal and are regulated by authorisations.

3.1 Sampling area

The study area (*see* Fig. 3) of the experiment is a part of the Heron Lake in Villeneuve d'Ascq, France. It is an artificial lake, dug in the 70s to drain the marshy area, to receive rain water from roads and other artificialized ground and from storm overflows from mixed drainage system (rain and to a lesser extent domestic wasterwater). The water arrives at the east of the lake and when the level is too high, water is pumped out to a nearby river in the far western point. A natural remediation of the water occurs in lake so a gradient can be expected between the entrance with a high biodegradable input and the exit (Ivanovsky et al., 2018).

Fig. 3. Region Ω on the Heron lake, Villeneuve d'Ascq, where measurements are carried out.

The measures were focused in the region Ω (pink region in Fig. 3) where the gradient should be the strongest, at the interface between the small canal bringing the contaminated water and the main part of the lake where dilution and decontamination occurs. The region has been explored with the USV during close to 1 hour, getting sample each 1 second (Fig. 4).

3.2 Description of the parameters

Direct *in situ* measurement of contaminant is usually not performed as sensors are not always available or difficult to use in environmental conditions. Reliable sensors can be easily found and used for physicochemical factors such as pH, turbidity, conductivity, temperature and dissolved oxygen. These parameters can provide useful information about contamination, e.g. conductivity usually increases with input of wastewater. But also give some insight about the biological activity in the water. For instance, dissolved oxygen and pH daily cycles can be observed and are directly linked to the respiration and photosynthesis alternation. The amplitude of these cycles gives an idea about

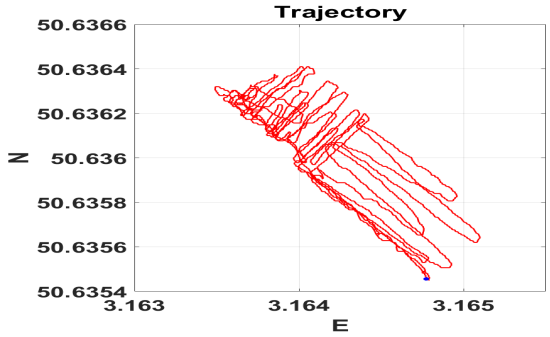


Fig. 4. Trajectory of the hand-operated vessel in region Ω with decimal GPS coordinates.

the daily growth of photosynthetic organisms and so, of the availability of nutrients in the system (Kumar and Thomas, 2019). Finally, hydrodynamic considerations can be derived from the measure of turbidity, as it is often an indication of the resuspension of sediment in the water body and so of the currents at the water - sediment interface.

4. MAIN RESULTS

The environmental mission explained in the previous section presents several challenges to improve the collection of data. The first problem is the size of the area of exploration. This means a large difference of time between the first and the last measurement. Because the natural evolution of the physical parameters to analyze, the delayed measurements are likely to result into values that are significantly different. The second problem is the regularization of the survey. Since the USV is controlled manually, the route of the vessel leaves many places unvisited (*see* Fig. 4).

In this section we discuss an approach to sort out these challenges. In order to account for the regularization of the survey, we perform an approximation of every missing data. This is accomplished thanks to a regular map of the region of interest (set Ω from Fig. 3), where every point of Ω is now associated with an estimation of its physical parameters (turbidity, DO, pH, conductivity and temperature).

In what follows, the limnological map is constructed in such a way it can be used in simulation scenarios in order to test different control strategies (for a single and a fleet of USV) to address these challenges.

4.1 Map meshing

In this section the area of interest Ω is computed by the largest convex set containing the entire data collection of Fig. 4, and a set of meshing of Ω are constructed with the aims of estimation parameters and robotic simulations.

Each sample is geo-referenced with decimal GPS coordinates. Based on these coordinates, new metric coordinates are generated for each sample B_i relative to the minimum longitude and minimum latitude of all samples, *i.e.* the origin denoted A. The metric coordinates (B_i^x, B_i^y) of the i_{th} sample denoted B_i in the relative reference map with origin A, are easily computed according to the distance formula between A and the projections on the horizontal and vertical axis of B_i:

$$B_i^x = cos^{-1}(sin\varphi_A.sin\varphi_{Bi} + cos\varphi_A.cos\varphi_{Bi}).r_E$$
$$B_i^y = cos^{-1}(sin\varphi_A.sin\varphi_A + cos\varphi_A.cos\varphi_A.cos\Delta\lambda).r_E$$

with latitudes φ_A, φ_{Bi}, longitudes λ_A, λ_{Bi}, $\Delta\lambda = \lambda_{Bi} - \lambda_A$ and r_E the conventional Earth radius equal to $6,378,137$ m.

The boundaries of the region Ω are determined thanks to the Matlab function `boundary`. This function returns a vector of points indices $\boldsymbol{P_B}$ that represents a compact boundary around this area. It is possible to obtain less detailed boundaries by tuning the parameter *FaceAlpha* of the function, with a value less than one. The following step aims at meshing the map.

Remark 2. *The covered area is usually very large and its meshing is required to build an adapted environment for robotic simulations. From the computational mechanics domain, the in-book chapter is of interest for the mesh generation methods and mesh adaptivity issues (George et al., 2007). This paper is dedicated to numerical solutions for computation in mechanics, but it presents some rules, advices and methodology for the surface domains.*

A regular map meshing of the region Ω could be done by different shapes of different sizes (*see* Fig. 5), which particularities are discussed in Section 4.5. Matlab software offers the library `Generate Mesh` for triangular mesh for a 2-D geometry. Another option is to perform map meshing thanks to the *Multiparametric Toolbox 3.0*[3] library (Herceg et al., 2013), that allows adjusting the size and choosing the shape of each finite element. From this library, the functions `Polyhedron`, `meshGrid` or `isInside` are available. Hence, according to the function `Polyhedron` and the previous determined vector $\boldsymbol{P_B}$, an entity *Zone* is created from the area of interest.

The meshing of *Zone*, *i.e.* the region Ω, is performed according to three steps. The first one consists in defining the shape, the size and the orientation of the finite element, *i.e.* an *elementary* entity *E* (square or hexagon in Fig. 10); the second in determining the location in the *Zone* of the center of each *E*; the third in attaching one *E* to each center.

The *elementary* entity *E* is created with the relative coordinates of the polyhedron's vertices $\boldsymbol{V_E}$. For a square entity E_{sqr} with a side length *Sp*, it is necessary to code:

$E_{sqr} = Polyhedron('lb', [-Sp/2, -Sp/2], 'ub', [Sp/2, Sp/2])$.
For any other polyhedron with a side length *Sp*, the code is:

$E_p = Polyhedron(\boldsymbol{V_E})$, with $\boldsymbol{V_E}$ the coordinates of vertices.

As an example, to define the entity E_H as a hexagon, the vertices are defined as:

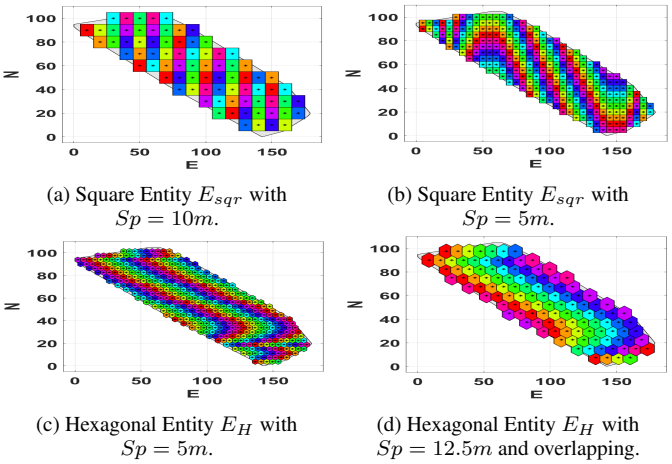
$\boldsymbol{V_{E_H}} = \{(-Sp * cosd(30)/2; Sp * cosd(60)/2), (0; Sp/2),$
$(Sp * cosd(30)/2; Sp * cosd(60)/2),$
$(Sp * cosd(30)/2; -Sp * cosd(60)/2),$
$(0; -Sp/2), (-Sp * cosd(30)/2; -Sp * cosd(60)/2)\}$.

Note: E_{sqr} is dedicated to an *elementary* entity with a square shape, E_H to a hexagonal shape, and E_p to a polyhedral shape.

The centers for E_{sqr} can be easily determined thanks to the function `meshGrid`. However, the coordinates of centers for any other polyhedron has to be determined, based on the shape of the entity E_p. Hence, the set of these centers $\boldsymbol{Cp}$ that belong to *Zone* is determined thanks to the function `isInside`.

The final step consists in computing for each center $c_i \in \boldsymbol{Cp}$, a new entity $E_{p_i} = c_i + E_p$. As a result we get one entity E_{p_i} for each center c_i. Figure 5 shows different meshes with constructed by varying the entity shape (square & hexagonal),

[3] https://www.mpt3.org/



(a) Square Entity E_{sqr} with $Sp = 10m$.
(b) Square Entity E_{sqr} with $Sp = 5m$.
(c) Hexagonal Entity E_H with $Sp = 5m$.
(d) Hexagonal Entity E_H with $Sp = 12.5m$ and overlapping.

Fig. 5. Map meshing of Ω.

in *a*. $Sp = 10m$ (meters), in *b*. $Sp = 5m$, and with hexagonal Entity without overlapping in *c*. $Sp = 5m$, and with in *d*. $Sp = 10m$.

Based on the map meshing, the next step is to create the limnological map of *Zone*, *i.e.* to associate an estimated value of each physico-chemical parameter (turbidity, DO, pH, conductivity and temperature) to each of the entity E_{p_i}, and to determine the variance of these estimations if any. This represent a challenge that will be tackle in the next Section. Basically, the USV has explored the *Zone* with different velocities and several trajectories. Hence, first the USV has not taken the same number of samples for each entity E_{p_i}, second some entities E_{p_j} could not be explored because of their size of entities.

Remark 3. *In the past two decades, several approximation methods for the structural reliability analysis were proposed. The main objective of these methods is to extract the maximum amount of information from the collected data in the smallest number of experimental runs, resulting in lower material consumption and considerably less laboratory work. Among them, response surface methodology (RSM) and Kriging have drawn extensive attention specially in design, modelling, and optimization of environmental experiments (Karimifard and Moghaddam, 2018; Brus and Heuvelink, 2007). To the best of our knowledge, there exist no result about the application of these approximation methods in experimental scenarios similar to ours, where there is no simultaneous measurements. This study is addressed in the next sections.*

In the next Sections, two methods to attribute one value of each physical parameter to each entity E_{p_i} is presented. The first approach is the Kriging strategy, which results will be compared with the proposed strategy presented in Section 4.3. The two methods have been used to estimate all parameters (turbidity, DO, pH, conductivity and temperature). However, for sake of simplicity and due to the limitation of space, only DO is depicted in this paper.

4.2 Map construction: Kriging approach

Kriging is a geostatistical interpolation method that was developed for the mining field at the end of the decade of 1960 (*see* Fig. 6). This method has been widely extended in the literature to the interpolation of environmental variables, such as soil quality (Snepvangers et al., 2003), wheater tempera-

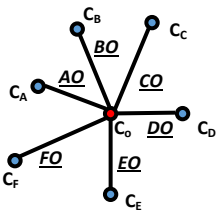


Fig. 6. Blue dots $\{A, B, ..., F\}$, represent already taken measurements, $\{C_A, C_B, ..., C_F\}$, and the red dot, O, is the point where the value is being interpolated. Thus, since distances from the measurements to the interpolated point are available, we can use kriging to interpolate C_O.

ture (Shtiliyanova et al., 2017), solar irradiance (Yang et al., 2013), and water quality (Chen et al., 2012).

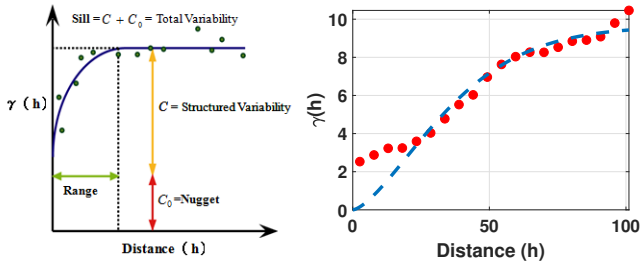
To perform these interpolations, Kriging strategy uses Variograms to relate the distance, h, between the measurements and the semivariance, $\gamma(h)$, and that can be computed from the known measurements. However, these *empirical* variograms do not provide information for all possible h (needed for kriging). Thus, they are often approximated by model function, as *circular variogram* (*see* Fig. 7a), ensuring validity (Chiles et al., 1999). Fig. 7b shows as red dots *empirical* variogram computed from the measurements taken in Ω. The theoretical model used to fit the data is represented as blue dashed line. Note that in Fig. 7b there is a large divergence between the experimental and the theoretical variagrams, specially for small distances. The reason of this divergence is that there are many measurements which, disregarding being very close in distance, are very far in time. Thus, the *real* value of the parameter can have changed a lot during the time between them.

Making use of this theoretical Variogram we can interpolate the mesh of points inside Ω, carried out with hexagonal shape of $Sp = 5m$. Fig. 8 shows the boundary of explored area in red (set Ω), the explored entity *Zone* with grey face color, the centers and each hexagonal entities E_H and the real trajectory of the USV in blue. Note that in the top left corner, it can be seen that the USV explored several times the same entities. Meanwhile, in the bottom right corner there are some non-explored entities. With the center of each entity E_{H_i}, a limnological map is obtained for every physical parameter analyzed in this work. Fig. 9a shows the map for the DO.

Furthermore, by assessing Fig. 9a, it can be seen that the maximal rate of DO in this region comes from $1.1mg/l$ (dark blue) to $9.09mg/l$ (yellow). The DO concentration presents smaller values in the top left corner of the set Ω, which corresponds to the river mouth where there is a suspicion of a source of pollution. By means of the generated maps, the gradient of each parameter can be assessed in order to determine if there is a significance pollution source. Lastly, the variance of the measures are depicted in Fig. 9b, with values from 0 to 0.51. The variance takes the maximum values at the non-explored area of Ω (*see* Fig. 8), *i.e.*, the farthest points to the trajectory.

4.3 Map construction: Proposed approach

The creation of the measurement map is achieved in two steps. The first step aims at determining, thanks to the function `isInside`, the set of samples made in each constructed entity E_{p_i}. The average of these samples leads to the estimation of one value for each entity E_{p_i}, and the uncertainty associated



(a) Theoretical variogram (circular model).

(b) Experimental variogram (red dots) and fitting to a theoretical model.

Fig. 7. Kriging Variograms. Circular variograms like the one in Fig. 7a are one of the most use mathematical models. However, there are other models as Gaussian, exponential, spherical, or lineal variagrams among others. Note that many of the terms used came from the mining field as *Nugget* (that models the uncertainty of the measuring instruments) or *Sill* (which model the maximum uncertainty). In Fig. 7b we have computed the empirical variogram according to the gathered data and we have fitted it to a theorical model. In this case the mathematical model which better fits the gathered data is the Gaussian variogram.

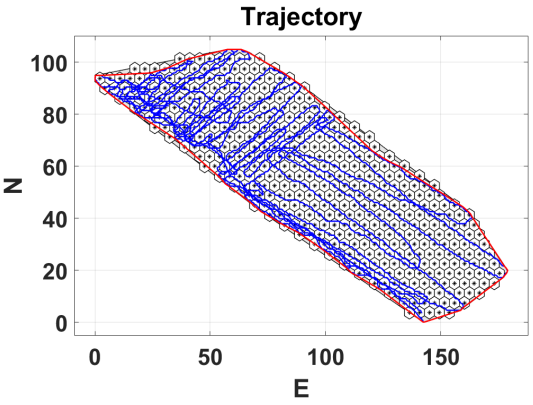
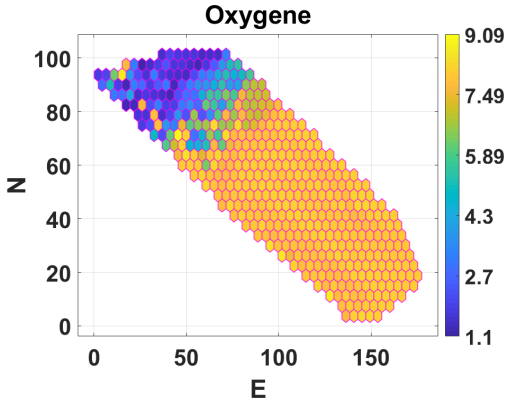


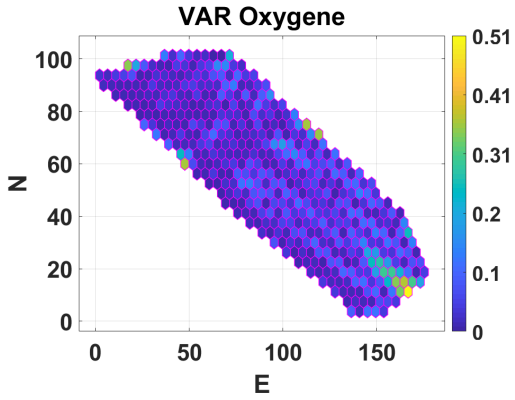
Fig. 8. Map meshing with hexagonal Entity E_H with $Sp = 5m$.

to each entity E_{p_i} estimation is determined by computing the variance between the involved measurements.

The second step concerns the non-explored entities E_{p_j} for which a neighborhood propagation of the values is proposed. To this end, a neighborhood matrix of the entities E_p is built. It aims at determining which entities E_{p_i} have common edge. Two examples are provided in Fig. 10, with square and hexagonal entities. For the square Entity E_{sqr} in Fig. 10a, and the hexagonal Entity E_H in Fig. 10b, the neighborhood matrices N_{sqr} and N_H are given in Eq. (1), respectively. In N_{sqr}, all the 1 values from the fifth line show that the entity E_{sqr_5} has as neighbours the entities E_{sqr_2}, E_{sqr_4}, E_{sqr_6} and E_{sqr_9}. In N_H, the 1 from the second line shows that neighbours of the entity E_{H_2} are: E_{H_1}, E_{H_3}, E_{H_7} and E_{H_8}.

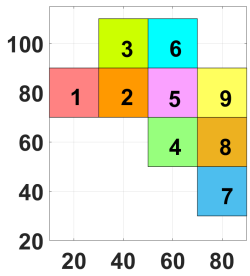


(a) Final Kriging map of DO.

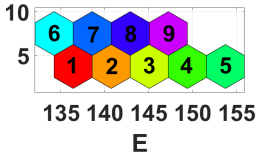


(b) Variance of measures of DO based on Kriging map.

Fig. 9. Kriging Map Generation.



(a) Square Entity E_{sqr}.

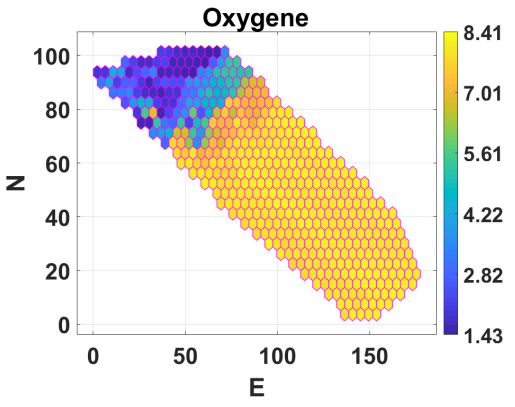


(b) Hexagonal Entity E_H.

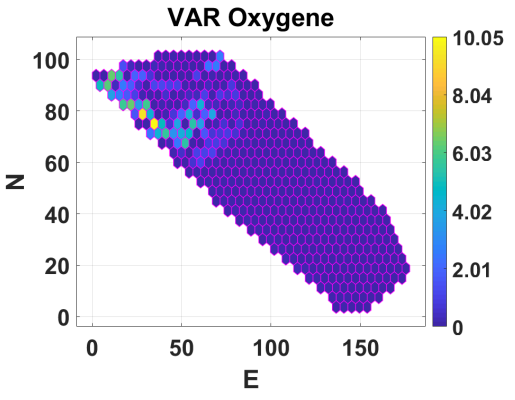
Fig. 10. Map meshing examples.

$$N_{sqr} = \begin{bmatrix} 1&1&0&0&0&0&0&0&0 \\ 1&1&1&0&1&0&0&0&0 \\ 0&1&1&0&0&1&0&0&0 \\ 0&0&0&1&1&0&0&1&0 \\ 0&1&0&1&1&1&0&0&1 \\ 0&0&1&0&1&1&0&0&0 \\ 0&0&0&0&0&0&1&1&0 \\ 0&0&0&1&0&0&1&1&1 \\ 0&0&0&0&1&0&0&1&1 \end{bmatrix} \quad N_H = \begin{bmatrix} 1&1&0&0&0&1&1&0&0 \\ 1&1&1&0&0&0&1&1&0 \\ 0&1&1&1&0&0&0&1&1 \\ 0&0&1&1&1&0&0&0&1 \\ 0&0&0&1&1&0&0&0&0 \\ 1&0&0&0&0&1&1&0&0 \\ 1&1&0&0&0&1&1&1&0 \\ 0&1&1&0&0&0&1&1&1 \\ 0&0&1&1&0&0&0&1&1 \end{bmatrix} \tag{1}$$

The value that is associated to a non-explored entity E_{p_j} is estimated as the average value of each of its neighbours. It is computed as the multiplication of the vector of measurement with the j_{th} line of the associated neighborhood matrix.



(a) Final Average map of DO.



(b) Variance of measures of DO based on Average map.

Fig. 11. Average Map Generation.

Based on the map in Fig. 8, the limnological map is generated with one mean value of the measured DO concentration for each explored entity E_{H_i}. Values for the non-explored entities E_{H_j} have been estimated. The final map of DO is depicted in Fig. 11a. Each entity is associated to a value with rate of DO from $1.43mg/l$ (dark blue) to $8.41mg/l$ (yellow). The DO is less concentrated in the top left corner because this area corresponds to the outlet of the river that runs through Lille; this river can be loaded with pollutants. Finally, Fig. 11b shows the variance of the measures given an indication of the uncertainties on the sample. The variance for the DO varies from 0 to 10. It is maximum in the top left corner. As well as the Kriging strategy, the average method also show a higher concentration of DO where a source of pollution is suspected. In the next Section, the results obtained by both methods are discussed.

4.4 Comparison of both strategies

The two approaches based on average map generation (AMG) and Kriging map generation (KMG) are compared considering the computational load, the values of DO, the variance of measurement. To generate the final map, the computational loads are $74s$ and $14s$ for AMG and KMG respectively [4] . The difference comes from AMG which consists in considering each entity E_H as surface, instead of KMG which takes into account only dots. This computational load is not really an issue in this work because the map is generated after collecting all the sample and not any dynamical mapping has to be performed.

[4] These durations were measured on a PC laptop with Intel(R) Core(TM) i7-10850H with a 2.70GHz clock, 32 GB of RAM. The Operating System is a 64 bits Windows 10 Professionnel version 21H1. Both implementations were developed using Matlab R2021b

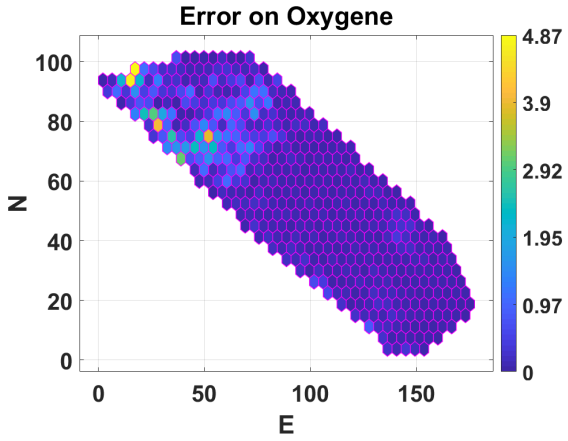


Fig. 12. Absolute errors between estimated values from the average and the Kriging approaches for each hexagonal entity.

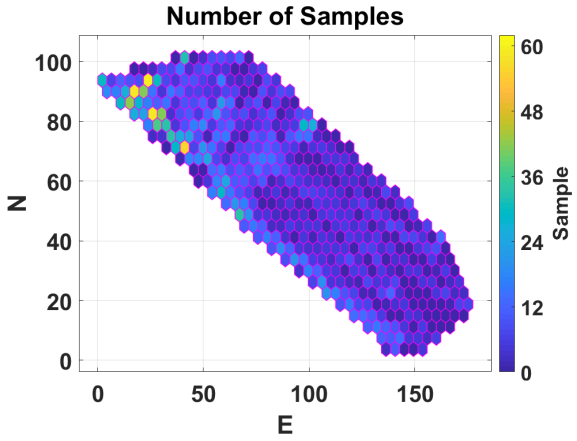


Fig. 13. Number of samples collected during the exploration per hexagonal entity.

The interval of DO from KMG is a little bigger than AMG, but the difference between both maps is not very big, from $0mg/l$ to $4.87mg/l$, with the exception of some E_H (*see* Fig. 12). The maximum rate of error is located in the top left corner. It is possible that it is due to the number of samples considered in each E_H (*see* Fig. 13), and also the duration between measurements of the first and the last exploration of each E_H (*see* Fig. 14). Indeed, the number of samples and the duration of exploration of these E_H are bigger.

If the duration between two measurement times is *high*, the water characteristics may change. This is very likely in this zone of experimentation, with a gradient of concentration due to punctual wastewater output. Depending of the weather or of the river flow, differences in sample values can appear quickly.

Also, the number of samples and the duration of exploration of each E_H have a direct impact on the computed measurements variance. The variance of the measurement from the AMG is maximum in the top left corner. In this area, E_H are explored with the USV at different times, with a bigger number of samples. It is a way to highlight the importance of the measuring process dynamics. The time and the changes on the parameters could be shown according to the variance. However, this variance is minimum for the non explored area in the bottom right corner. It is mainly due to the step based on neighborhood propagation. This is the main drawback of this approach.

At the opposite, as expected, the variance of measurements from the KMG is very much smaller. The variance associated to the non explored E_H in the bottom right corner is higher providing implicitly the biggest uncertainties on these measurement.

4.5 Discussion

The map extraction from field measurements which consists in detecting the explored area and to discretize it into small entities is very generic. It can be applied to areas with different shapes. Besides, it supports varying the shape of the entity used for discretizing the explored area. However, convex covering shapes, such as triangles, squares or hexagons are preferably used. These shapes share common edges and can be organized to cover the whole area without leaving blanks. Overlapping between entities is possible, though the usefulness of such overlaps is yet to be investigated. In the case of maps generated to

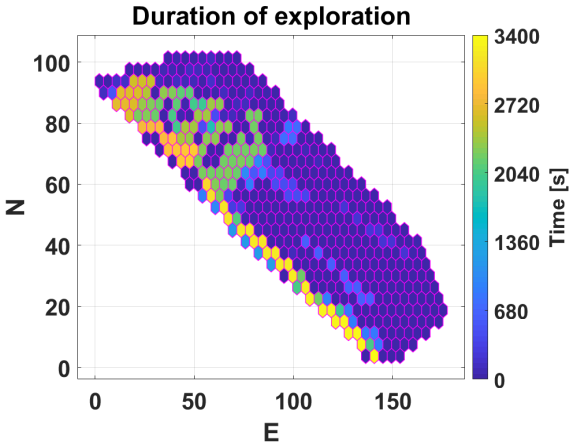


Fig. 14. Duration of exploring each hexagonal entity.

create simulation environments for robots, overlapping entities is *a priori* not required. Additionally, the size of each entity can be tuned according to the required accuracy required for robot navigation.

The Kriging approach leads to accurate map with a small variance, even if some estimations seem inaccurate. However, this approach is not suitable to take into account dynamics in the measuring process. In this case spatio-temporal Kriging is required.

5. CONCLUSION AND PERSPECTIVES

This paper targets environmental missions to assess water quality by means of aquatic robots. Experiments we have conducted expose critical points related to data collection that can be improved by the study and the implementation of control strategies for Unmanned Surface Vehicles. To this end, a limnological map of the region of interest is necessary to conduct simulations with different scenarios before the field implementation. In this context, and based on validated real data collection, we have provided an estimated parameters map of a region of interest. The proposed map was computed using a known data interpolation method, namely Kriging. Kriging shows some limitations in this kind of non-studied scenario, mainly due to the gap of time between measurements.

To the best of our knowledge, there exist no study about the application of this approximation method to scenarios similar to ours: water characteristics are dynamic and evolve over time while measurements are asynchronous. This is why we have introduced a novel method for data approximation, based on polyhedral sets. The potential benefits of this novel strategy were illustrated by a detailed analysis and simulation results.

In future works, in order to improve the parameter estimation of both methodologies, the strategies need to be adapted to a more appropriate dynamic context. We need to consider the time when the measures were taken as part of the estimation. In the case of Kriging, the use of spatio-temporal Kriging (Montero et al., 2015) must be assessed. In addition, the generated map will be utilised in simulation scenarios to test predictive control strategies for USV. The goal is to reduce the gap of time between measurements by appropriate optimal strategies that involves a trade-off between the size of the exploration area and the necessary amounts of measurements to create a reliable map. The general objective is to improve the data collection of water quality parameters in different sophisticated environments scenarios.

ACKNOWLEDGMENT

Authors want to thanks the company Bathy drone Solutions (BDS) for its participation in the experiments, and the Department of Economic Transformation, Industry, Knowledge and Universities of the Andalusian Government (PAIDI 2020) [Ampliación Aquacollect, ref. P18-HO-4713]. The Region Hauts de France and the French government are warmly acknowledged for the co-funding of the CPER ECRIN and the VERDEAU project.

REFERENCES

Brus, D.J. and Heuvelink, G.B. (2007). Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, 138(1), 86–95. doi:https://doi.org/10.1016/j.geoderma.2006.10.016.

Chen, Y.C., Yeh, H.C., and Wei, C. (2012). Estimation of river pollution index in a tidal stream using kriging analysis. *International journal of environmental research and public health*, 9(9), 3085–3100.

Chiles, J.P., Delfiner, P., et al. (1999). Modeling spatial uncertainty. *Geostatistics, Wiley series in probability and statistics*.

Edition, F. (2011). Guidelines for drinking-water quality. *WHO chronicle*, 38(4), 104–108.

George, P.L., Borouchaki, H., Frey, P.J., Laug, P., and Saltel, E. (2007). *Mesh Generation and Mesh Adaptivity: Theory and Techniques*, chapter 17. John Wiley Sons, Ltd. doi:https://doi.org/10.1002/9780470091357.ecm012.pub2. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470091357.ecm012.pub2`.

Herceg, M., Kvasnica, M., Jones, C.N., and Morari, M. (2013). Multi-parametric toolbox 3.0. In *2013 European control conference (ECC)*, 502–510. IEEE.

Hervagault, Y. (2019). *Design and Implementation of an Effective Communication and Coordination System for Unmanned Surface Vehicles (USV)*. Ph.D. thesis, Université Grenoble Alpes.

Ivanovsky, A., Belles, A., Criquet, J., Dumoulin, D., Noble, P., Alary, C., and Billon, G. (2018). Assessment of the treatment efficiency of an urban stormwater pond and its impact on the natural downstream watercourse. *Journal of Environmental Management*, 226, 120–130. doi:https://doi.org/10.1016/j.jenvman.2018.08.015. URL `https://www.sciencedirect.com/science/article/pii/S0301479718308867`.

Kannel, P.R., Lee, S., Lee, Y.S., Kanel, S.R., and Khan, S.P. (2007). Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment. *Environmental monitoring and assessment*, 132(1), 93–110.

Karimifard, S. and Moghaddam, M.R.A. (2018). Application of response surface methodology in physicochemical removal of dyes from wastewater: a critical review. *Science of the Total Environment*, 640, 772–797.

Kumar, P.S. and Thomas, J. (2019). Seasonal distribution and population dynamics of limnic microalgae and their association with physico-chemical parameters of river noyyal through multivariate statistical analysis. *Scientific reports*, 9(1), 1–14.

Madeo, D., Pozzebon, A., Mocenni, C., and Bertoni, D. (2020). A low-cost unmanned surface vehicle for pervasive water quality monitoring. *IEEE Transactions on Instrumentation and Measurement*, 69(4), 1433–1444.

Montero, J.M., Fernández-Avilés, G., and Mateu, J. (2015). *Spatial and spatio-temporal geostatistical modeling and kriging*, volume 998. John Wiley & Sons.

Mougin, J. (2021). Beyond high frequency monitoring: an optimised automatic sampling. In *EGU General Assembly Conference Abstracts*, EGU21–1969.

Shtiliyanova, A., Bellocchi, G., Borras, D., Eza, U., Martin, R., and Carrère, P. (2017). Kriging-based approach to predict missing air temperature data. *Computers and Electronics in Agriculture*, 142, 440–449. doi:https://doi.org/10.1016/j.compag.2017.09.033.

Sinisterra, A.J., Dhanak, M.R., and Von Ellenrieder, K. (2017). Stereovision-based target tracking system for usv operations. *Ocean engineering*, 133, 197–214.

Siyang, S. and Kerdcharoen, T. (2016). Development of unmanned surface vehicle for smart water quality inspector. In *2016 13th International conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON)*, 1–5. IEEE.

Snepvangers, J., Heuvelink, G., and Huisman, J. (2003). Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma*, 112(3), 253–271. doi:https://doi.org/10.1016/S0016-7061(02)00310-5. Pedometrics 2001.

Wang, Z., Yang, S., Xiang, X., Vasilijević, A., Mišković, N., and Na, . (2021). Cloud-based mission control of usv fleet: Architecture, implementation and experiments. *Control Engineering Practice*, 106, 104657.

Yang, D., Gu, C., Dong, Z., Jirutitijaroen, P., Chen, N., and Walsh, W.M. (2013). Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging. *Renewable Energy*, 60, 235–245. doi:https://doi.org/10.1016/j.renene.2013.05.030.