Endorheic waterbodies delineation from remote sensing as a tool for immersed surface topography

Carole Delenne, Jean-Stéphane Bailly, Antoine Rousseau, Renaud Hostache, Olivier Boutron

Endorheic waterbodies delineation from remote sensing as a tool for immersed surface topography.

Carole Delenne[1], Jean-Stephane Bailly[2], Antoine Rousseau[3], Renaud Hostache[4], and Olivier Boutron[5]

[1]Inria, team LEMON / HSM, Univ. Montpellier, CNRS, IRD, France.
[2]LISAH, Univ. Montpellier, INRAE, IRD, Institut Agro, AgroParisTech, France
[3]Inria, team LEMON / IMAG, Univ. Montpellier, CNRS, France.
[4]Luxembourg Institute of Science and Technology, Belvaux, Luxembourg
[5]Tour du Valat Research Institute, Arles, France

May 3, 2023

Abstract

Since several decades, it becomes possible to delineate waterbodies and their dynamics from optical or radar images, that are now available at high spatial and temporal resolutions. We present here an interpolation approach that takes benefits from this waterbodies delineation in endorheic areas. It consist in computing isovalue contourlines to improve topography estimates classically obtained from measurement points only. The approach, based on a minimisation problem, uses Thin Plate Splin interpolation functions, whose coefficient are determined along with the unknown water level of each curve. Results obtained on a generated topography show that this approach, applied with three contour-line curves, yields a lower root mean square error using only one measurement point compared to the one obtained with nine points and the classical approach.

1 Introduction

Endorheic and temporarily submerged natural or man-made areas (wetlands, deltas, reservoirs...) are a vital resource for ecosystems, agrosystems and populations. Hydraulic models may be used for their study and management (see e.g. [5, 11]); but if bidimensional models are powerful tools for shallow water flows representation, they require precise knowledge of the topography of the modelled domain. The extent, heterogeneity and the generally very little marked or evolving topography of such submerged area, make practical models application difficult. Proven interpolation methods are most often applied, using point surveys carried out in the field. Ideally, higher resolution LIDAR data could be available but their acquisition and processing cost remains prohibitive.

These interpolations are therefore inaccurate and can quickly turn out to be obsolete when annual floods change the system's geometry. Indeed, during a flood period, the sediments carried by surface water modify the geometry of natural rivers and become blocked in water reservoirs, leading to a reduction in storage capacity [1]. Reservoirs geometry is also an essential factor in the modelling of their hydrological functioning [4]. Bathymetric monitoring is the standard method for monitoring reservoir geometries but it remains an expensive method, and given the tidal range, especially in semi-arid areas, alternatives are possible [18, 24].

Remote sensing data from optical or radar images are now widely accessible and available at high spatial resolution and high temporal repeatability. It becomes possible to extract the water limits and their dynamics (*e.g.* floods limits [17], seasonal tidal ranges [23, 18], or long-term changes [25]) and use them for calibrating or validating hydraulic/hydrologic models (e.g. [17, 13, 21]) or to retrieve water elevation [12]. In the event of scarce field observation, it appears interesting to take advantage of these dynamics of water surfaces, which in endorheic areas form isovalue curves, in order to improve the estimation of the topography. Recent papers have been published [25, 9] to derive bathymetry for dynamic waterbodies using information on land elevation and long temporal series of water surfaces provided by satellite data. However, these studies are applied on very large areas only, since they uses high resolution images combined to radar altimeters. The bathymetry is recovered with an overall accuracy of several meters, which may be unsuitable for smaller waterbodies, while these represent the majority of the stored volume in European regions [16]. Moreover, if remote sensing make the geolocation of the isovalue curves always possible, it generally does not allow to directly assign them a value, *i.e.* an elevation, except for areas under altimeter tracks, which represents a very small part of continental areas.

The approach proposed in this paper aims to find an interpolated field which takes into account several measurement points and the information from a few number of unvalued isovalue curves. It is based on the optimisation of an objective function in two parts: a "classical part", representing the distance of the interpolation to the measurement points, and an additional part integrating the information from the contour lines. A difficulty here is that the elevation itself of each curve is not known. The elevation of the contour lines is consequently computed in the proposed optimisation algorithm and is then defined as a parameters of the interpolation functions. The following section presents the proposed methodology applied first in a controlled environment to an artificially generated elevation field and next in a real-world experiment to the Rascaillan pond (Rhône delta, France), presented in section 3. Section 4 analyses the accuracy of the proposed approach with respect to the parameter settings for the artificial test case and provides numerical results for both the artificial and real application test cases.

2 Methodology

2.1 Input data

Let a domain $\mathscr{D}$ containing N punctual records, defined by their coordinate vectors:

$$\mathbf{s}_k = \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix}, \, k = 1, \ldots, N \tag{1}$$

and L contour-lines (*i.e.* L sets of points with the same elevation):

$$\mathscr{C}_l = \left\{ \mathbf{s}^l = \left(x^l, y^l, z^l\right)^{\mathrm{T}} \mid z^l = a^l \right\}, \; l = 1, \ldots, L \tag{2}$$

where the points $\left(x^l, \, y^l\right) = \mathbf{X}^l$ are known since the shape of the contour-line is supposed to be extracted from remote sensing data, but the actual elevation a^l of the contour-line is not known.

2.2 Approximation function

The approximated elevation $Z(\mathbf{X})$ for each point $\mathbf{X}$ of the domain $\mathscr{D}$ is defined as a sum of radial functions:

$$\forall \, \mathbf{X} \in \mathscr{D}, \;\; Z(\mathbf{X}) = \sum_{c=1}^{N_\varphi} \lambda_c \varphi_c \left(\| \mathbf{X} - \mathbf{X_c} \| \right) \tag{3}$$

where $\mathbf{X_c}$ is the centre of function φ_c and N_φ is the number of functions φ chosen.

The interpolation function chosen in this study is a sum of Thin Plate Splin (TPS), introduced by Duchon in [8]. TPS has a natural representation in terms of radial basis functions, and can be written on its normalized form as:

$$\varphi_c(r) = \frac{r^2}{R^2} \ln \left(\frac{r^2}{R^2} \right) \tag{4}$$

where r is the Euclidian distance from the spline node (or center) c. According to [20], we compute the parameter R as follows:

$$R = \frac{1}{2} \max_k \left(\| \mathbf{X}_c - \mathbf{G} \| \right) \tag{5}$$

where $\mathbf{G}$ is the barycenter of all the center points $\mathbf{X_c}$.

The main advantage of TPS on other radial functions such as Gaussian or inverse multiquadratic functions is that it does not depend on the scale of the topographic map of interest and has also good properties concerning the minimization of the overall curvature of the generated surface (see e.g. [3, 15]) as well as the speed of convergence towards the desired surface [6, 19]. However, any other interpolation function may be used in the proposed approach. For example, the classical inverse-distance function was tested and lead to similar results.

2.3 Minimisation problem

The problem can be expressed as follows: determine an approximation function $Z(\mathbf{X}; p)$, where $\mathbf{X} = (x, y)$ stands for any points of the domain and p stands for the possible parameters of the function Z.This would classically be defined as a minimisation problem with the following cost function (Mean Square Error):

$$J_1(p) = \frac{1}{N} \sum_{k=1}^{N} \left[Z(\mathbf{X}_k; p) - z_k \right]^2 \tag{6}$$

The purpose of this study is to introduce another information based on the contour-lines that can be visible on aerial or satellite images, via the following cost function:

$$J_2(p) = \frac{1}{l} \sum_{l=1}^{L} \left(\frac{1}{N_l} \sum_{i=1}^{N_l} \left[Z(\mathbf{X}_i^l; p) - \tilde{a}^l \right]^2 \right) \tag{7}$$

where N_l is the number of points $\mathbf{X}_i^l$ chosen on each contour-line $\mathscr{C}_l$ and $\tilde{a}^l$ the estimation of the contour-line level a^l. Indeed, the particularity of this approach is that the elevation a^l of each contour-line is not known and will be assessed along with the parameters p during the minimisation process.

The two following approaches will thus be compared in this paper:

- Approach 1 (A1): minimisation based only on the measurement points using cost function $J_1(p)$ defined by equation (6).
- Approach 2 (A2): minimisation based on measurement points and contour-lines with unknown levels using cost function $J_{12}(p) = J_1(p) + J_2(p)$, defined by equations (6) and (7).

2.4 Resolution

The optimisation function depends on $N_\phi + L$ parameters, *i.e.* the chosen number of interpolation functions plus the number of contour-lines whose elevation is unknown. The approximation $Z(X)$ given by equation (3) is computed for each measurement point s_k and each point on the contour-lines s_i^l. Starting from a random set of initial parameters values, a quasi-Newton approach [22] is used to find parameters minimising the cost function given by equations (6-7).

3 Data

3.1 Academic test case

The proposed methodology is first applied on an academic test case, for which all input values are controlled, for a better parametric analysis. A random Gaussian field is generated to simulate a smooth and realistic topography used as reference for the validation test case. The domain extent is 1000 m×1000 m with a 10 m spatial resolution; the model used to generate this $2D$ field is a

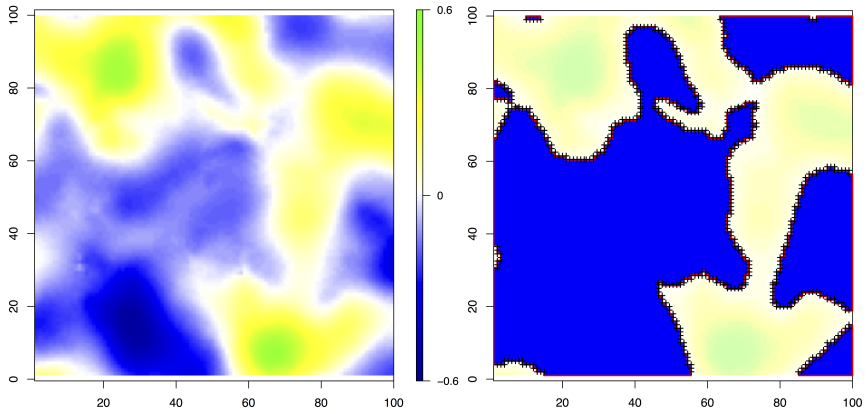

Figure 1: Generated topography for the validation test case (left) and example of the 0 m contour-line (right)

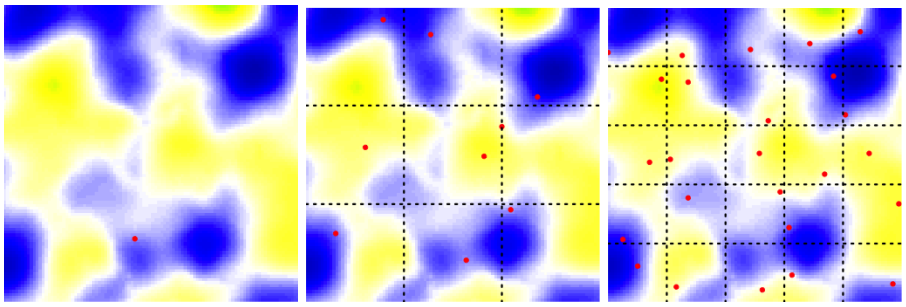

Figure 2: Measurement points selection for $d = 1$, 3 or 5.

Gaussian model with a 150 m practical range, a 0 m nugget and a 1/36 sill in order to get a smooth field with elevation values ranging from -0.5 m to 0.5 m (see Fig. 1).

Contour-lines, such as those that could be obtained by remote sensing approaches, are extracted from this field (see an example Figure 1-right). All the elevation pixels falling on a contour-line are selected to constitute the $\mathscr{C}_l$ set of points defined by equation (2).

In order to better compare the two approaches, several numbers of punctual measurements are selected with the following random stratified method. The domain is divided into d^2 subdomains in which one point is randomly selected, with d varying from 1 to 5 hereafter (see Figure 2). The minimisation process is run 100 times for each number of points and for each of the two approaches.

The TPS radial functions φ used in this test case are defined on a regular grid which density is also studied.

The results will be assessed via the classical Root Mean Square Error (RMSE) computed between the "real" topography and its estimation by the two ap-

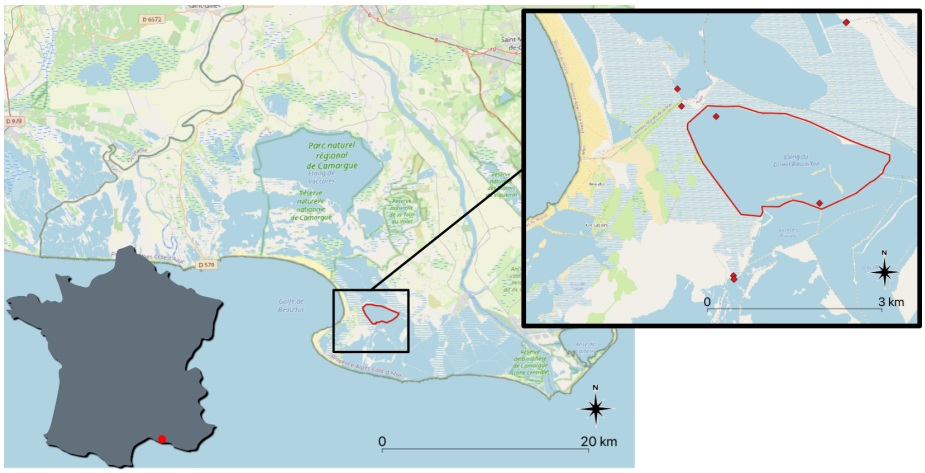

Figure 3: Situation of Rascaillan pond, south of France.

,

proaches on each pixel. The relative error defined by RMSE divided by the [min,max] interval of the real field (equal to 1.04 m here) is also given. Approach 2 can also be evaluated in terms of contour-lines level estimation.

3.2 Real test case: the Rascaillan pond

3.2.1 Study area and validation data

The Rascaillan is a pond located in the Rhone delta, south of France, as part of the UNESCO's Man and Biosphere reserve of Camargue (Figure 3). It comprises an area of c.a. 420 ha, and is a shallow pond (average water depth of 0.6 m at a water level of 0.5 m NGF (General Leveling of France). It is subject to very strong evaporation rates [5], which can result in a partial drying up. Figure 4 shows the real bathymetry of the pond obtained by a topo-bathymetric LIDAR survey carried out in September 2016 with the Optech Titan DW600 of Nantes and Rennes Universities platform and processed by D. Lague [14]. The spatial resolution is 0.5×0.5 m. The vertical resolution is 0.01 m, with a vertical accuracy of 0.05 m.

3.2.2 Remote-sensing data and water extent extraction

To extract varying waterbody extents of the Rascaillan pond, we processed a series of Synthetic Aperture Radar (SAR) images acquired by the Sentinel-1 satellite constellation. These images were acquired in Interferometric Wide swath mode with a pixel spacing of 10 m, corresponding to an average resolution of 20 m. The images were calibrated and re-sampled at 20 m pixel spacing. The waterbody extent was retrieved based on the method proposed in [10].

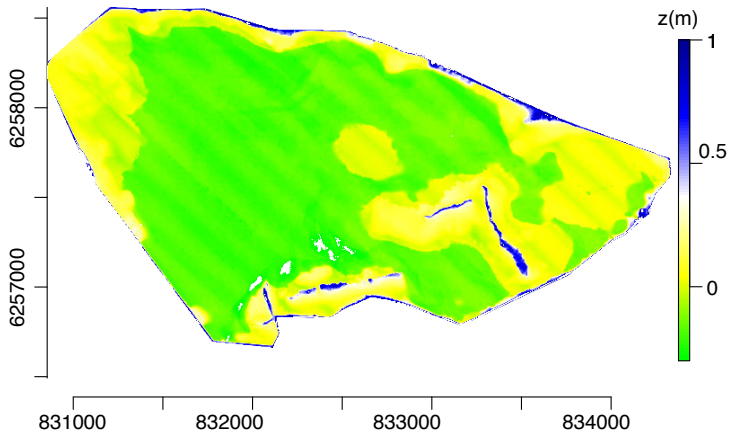

Figure 4: Real bathymetry of Rascaillan pond obtained from lidar survey

This algorithm is based on the Bayes theorem and assigns to any pixel of a SAR image a probability of being water using its backscatter values in two steps. It first estimates automatically the probability distribution functions of backscatter values associated with water and non-water pixels respectively, by decomposing the histogram of backscatter values. These distributions are next used to estimate for each pixel its probability to be of class water. The resulting probabilistic map is finally used to extract waterbody contours, assuming that pixels with probability greater than 0.5 correspond to water.

4 Results

4.1 Academic test case

Figure 5 shows an example of application with three different contour-lines selected at -0.1 m, 0 m and 0.1 m, 9 punctual records (red points) and 100 interpolation functions (one each 10m). In this particular test case, the RMSE between the reconstructed field and the real one is equal to $e_1 = 0.32$ m for Approach 1 corresponding to a relative error of 30%. Approach 2 enables a drop of the error to $e_2 = 0.10$ m (9.87%). It is first noticeable that the use of the contour-lines in the optimisation process reduces the error by three. Moreover, the approach enables the estimation of the contour-lines level with a RMSE equal to 4.7cm only.

A parametric analysis is then performed with these generated data. Each of the following tests has been repeated 100 times with a random selection of measurement points.

The density of interpolation functions has first been tested on the classical approach with 3 contour-lines (selected at -0.1 m, 0 m and 0.1 m) and 9 punctual records. It can be seen (see Figure 6) that the distance between two φ

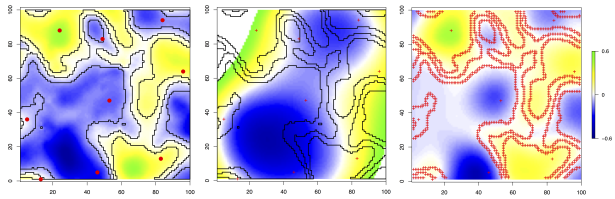

Figure 5: Example of application. a) Real field, with 3 contour-lines and 9 punctual records (red crosses). Results obtained using b) punctual records only (red crosses) and J_1 cost function, c) punctual records plus contour-lines (red crosses) and J_{12} cost function.

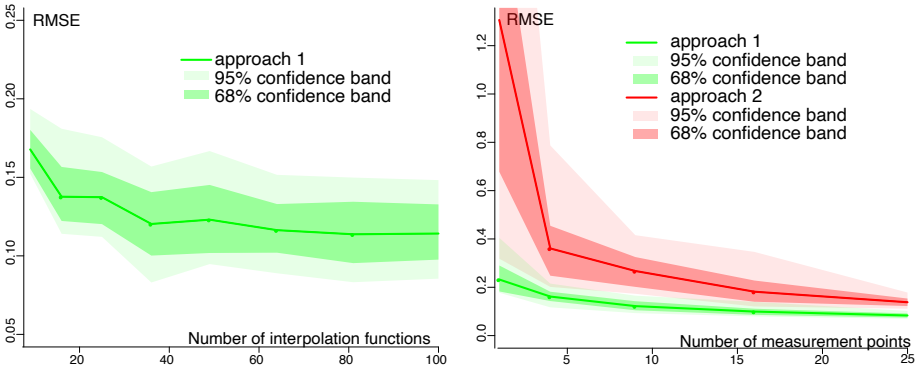

Figure 6: Left: RMSE (m) as a function of number of interpolation functions for the classical approach (A1). Right: RMSE as a function of number of measurement points for the two approaches. The red lines correspond to the mean values, the green envelop to the 66% confidence interval and the blue one to the 95% confidence interval.

centres should be of the same order of magnitude than the field range of spatial correlation. The following results are thus obtained with 7×7 functions (e.g. about 15 m apart from each other).

The influence of the number of measurement points have then been assessed, with the three same curves. As expected, it is first noticeable that for both approaches, the result accuracy increases with the number of interpolation points. What is more interesting is that the use of contour-lines extracted from remote sensing enables a very significant gain in accuracy with few measurement points (see Figure 6 right). Indeed the RMSE obtained by Approach 2 with only one measurement point and the three curves is between the ones obtained using 9 and 16 points in the classical approach.

The last parameter of interest here is the number of contour-lines available, taken from 2 to 6. At each of the 100 repetitions, the nine measurement points are randomly selected as explained before and the contour-lines are also ran-

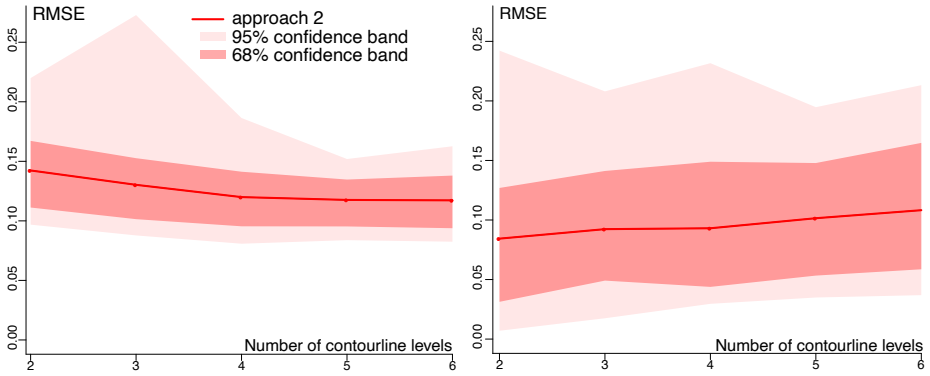

Figure 7: RMSE (m) as a function of number of curves for the proposed approach (A2). Left: computed on the whole field; Right: computed on contour-lines points only.

domly selected, with levels ranging between −0.4 m and 0.4 m. It can be seen (Figure 7) that the benefit of adding contour-lines rapidly drops (best results are obtained with four curves). Moreover, although adding curves slightly enhance the altitude assessment on the whole domain on average, local biases degrade the estimates of each contour-lines level. This result shows that we face a bias-variance tradeoff which has been also observed for other spatial regularization problems [2].

4.2 Real case application

Figure 8 shows two contourlines extracted from Sentinel 1 images (using approach presented in [10]) at dates with different water levels (2017-11-03 and 2016-09-16) and an example of 8 punctual records selected with stratified random procedure from Lidar data. As for the academic case, several tests were run with contourlines number varying from two to four and with punctual records number from 5 to 10. The results are assessed through the RMSE between the estimated and Lidar topographies. The standard approach based on records yields error ranging from 0.22 m to more than 0.6 m depending on the number and position of the randomly selected points. However, the proposed approach provides systematically better results with errors generally between 0.18 m and 0.22 m. Moreover, using more than two contourlines does not increase accuracy. This may be due to the uncertainty of the contourlines extracted from sentinel images as can be seen in Figure 8-right, that enable the comparison of the estimated contourlines to the real topography. Several factors may be in cause, such as image resolution, variation of topography between Lidar and images acquisition, or wind that often inclines the pond surface elevation in this area (several centimeters of difference from one side to the other).

Figure 9 shows an example of results obtained for the Rascaillan pond using the classical approach with five punctual records interpolation and the proposed

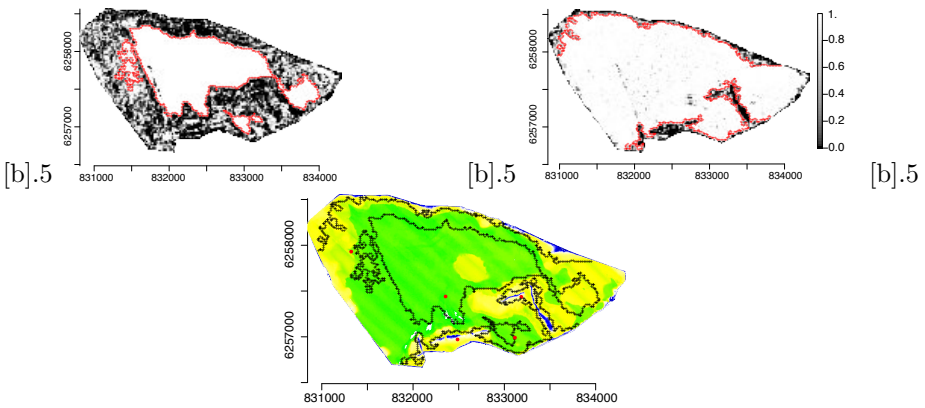

Figure 8: a) Example of two contourlines (red crosses) extracted from Sentinel 1 images (the probability for each pixel of beeing water is given in gray scale), and b) five punctual "records" selected from lidar data (red points) along with contourlines (black crosses) and lidar topograhy (in color).

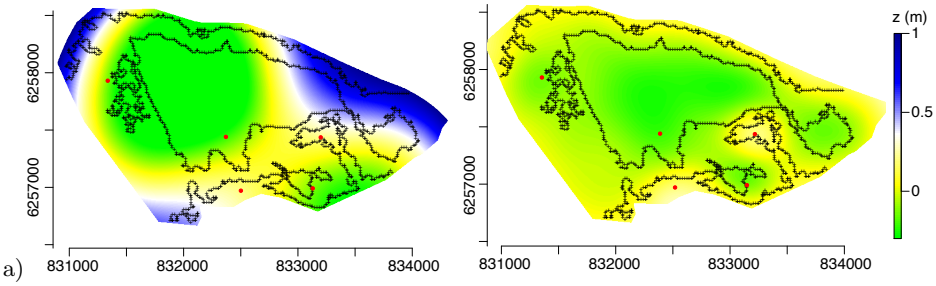

Figure 9: a) Results obtained with punctual records only and b) with two contourlines extracted from Sentinel 1 images.

approach that accounts for the two contourlines extracted from Sentinel 1 images provided in Figure 8. The RMSE obtained here for the first approach is 0.42 m whereas it drops to 0.17 m using the two contourlines.

5 Conclusion

In this study we evaluate the impact of adding topography contourlines to traditional point-wise measured altitudes in a global topography reconstruction. The location of the isolines is an available information from remote sensing data, whereas their altitude remains unknown. Indeed, the approach takes benefits of widely accessible optical or radar images and, since the contourlines altitudes are assessed by the optimisation process, it does not require more rarely available altimeters information. The application on generated data en-

abled a parametric study with many repetitions. The results show a substantial reduction of the RMSE error when adding only a few contourlines (up to 4), whereas additional isolines do not seem to provide any useful information. The real-case application to the Rascaillan pond shows that this approach can easily be implemented using widely available data such as Sentinel images, yielding a large decrease of the root mean square error between the interpolated topography and the reference LIDAR acquisition, even using only two satellite images. Depending on the study area and the available data, other methodologies may be used to extract waterbodies contourlines such as the ones presented in [26] or [7].

References

[1] J. Albergel, J. Collinet, Y. Pepin, P. Zante, S. Nasri, M. Boufaroua, A. Droubi, and A. Merzouk. Sediment budgets on hill reservoirs of small catchments in North Africa and the Middle East. *IAHS-AISH Publication*, 291:323–331, 01 2005.

[2] M. Belkin, A. Rakhlin, and A. B. Tsybakov. Does data interpolation contradict statistical optimality? In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1611–1619. PMLR, 2019.

[3] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Patterns Analysis and Machine Intelligence*, 11(6):567 – 585, 1989.

[4] M. Bouteffeha, C. Dagès, R. Bouhlila, and J. Molenat. A water balance approach for quantifying subsurface exchange fluxes and associated errors in hill reservoirs in semiarid regions. *Hydrological Processes*, 29, 03 2015.

[5] O. Boutron, O. Bertrand, A. Fiandrino, P. Höhener, A. Sandoz, Y. Chérain, E. Coulet, and P. Chauvelon. An Unstructured Numerical Model to Study Wind-Driven Circulation Patterns in a Managed Coastal Mediterranean Wetland: The Vaccarès Lagoon System. *Water*, 7(11):5986–6016, 2015.

[6] M. D. Buhmann. Multivariate cardinal interpolation functions. *Constr. Approx.*, 8:225 – 255, 1990.

[7] M. C. R. Cordeiro, J.-M. Martinez, and S. Pena-Luque. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. *REMOTE SENSING OF ENVIRONMENT*, 253, FEB 2021.

[8] J. Duchon. "splines minimizing rotation-invariant semi-norms in sobolev spaces". In W. Schempp and K. Zeller, editors, *Constructive Theory of Functions of Several Variables*, pages 85–100, Berlin, Heidelberg, 1977. Springer Berlin Heidelberg.

[9] A. Getirana, H.-C. Jung, and K.-H. Tseng. "deriving three dimensional reservoir bathymetry from multi-satellite datasets". *Remote Sensing of Environment*, 217:366 – 374, 2018.

[10] L. Giustarini, R. Hostache, D. Kavetski, M. Chini, G. Corato, S. Schlaffer, and P. Matgen. Probabilistic flood mapping using synthetic aperture radar data. *IEEE Trans. Geosci. Remote Sens.*, 54(12):6958–6969, 2016.

[11] M.M. Haque, O. Seidou, A. Mohammadian, and A. Gado Djibo. "development of a time-varying modis/ 2d hydrodynamic model relationship between water levels and flooded areas in the inner niger delta, mali, west africa". *Journal of Hydrology: Regional Studies*, 30:100703, 2020.

[12] S.-H. Hong and S. Wdowinski. Multitemporal Multitrack Monitoring of Wetland Water Levels in the Florida Everglades Using ALOS PALSAR Data With Interferometric Processing. *IEEE Geoscience and Remote Sensing Letters*, 11(8):1355–1359, 2014.

[13] R. Hostache, X. Lai, J. Monnier, and C. Puech. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. *Journal of Hydrology*, 390(3):257 – 268, 2010.

[14] D. Lague and B. Feldmann. Chapter 2 - topo-bathymetric airborne lidar for fluvial-geomorphology analysis. In P. Tarolli and S. M. Mudd, editors, *Remote Sensing of Geomorphology*, volume 23 of *Developments in Earth Surface Processes*, pages 25–54. Elsevier, 2020.

[15] J. Meinguet. Surface spline interpolation: Basic theory and computational aspects. *Approximation Theory and Spline Functions, S.P. Singh et al. (eds.)*, pages 127 – 142, 1984. D. Reidel Publishing Company.

[16] M. L. Messager, B. Lehner, G. Grill, I. Nedeva, and O. Schmitt. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7(1):13603, 2016.

[17] A. Ogilvie, G. Belaud, C. Delenne, J.-C. Bader, A. Oleksiak, J. S. Bailly, L. Ferry, and D. Martin. Decadal monitoring of the Niger Inner Delta flood dynamics using MODIS optical data. *Journal of Hydrology*, 523:358–383, 2015.

[18] A. Ogilvie, G. Belaud, S. Massuel, M. Mulligan, P. Le Goulven, P.-O. Malaterre, and R. Calvez. Combining landsat observations with hydrological modelling for improved surface water monitoring of small lakes. *Journal of Hydrology*, 566:109 – 121, 2018.

[19] M. J. D. Powell. The Uniform Convergence of Thin Plate Spline Interpolation in Two Dimensions. *Numerische Mathematik*, 68(1):107–128, 1994.

[20] C. Rabut. B-splines polyharmoniques cardinales : interpolation, quasi-interpolation, filtrage, 1990. Thèse d'État, Université de Toulouse.

[21] G. Schumann, P. Matgen, L. Hoffmann, R. Hostache, F. Pappenberger, and L. Pfister. Deriving distributed roughness values from satellite radar data for flood inundation modelling. *Journal of Hydrology*, 344(1):96 – 111, 2007.

[22] D. F. Shanno. Conditioning of quasi-Newton methods for function minimization. *Mathematics of computation*, 24(111):647–656, 1970.

[23] V. Soti, A. Tran, J.-S. Bailly, C. Puech, D.L. Seen, and A. Bégué. Assessing optical earth observation systems for mapping and monitoring temporary ponds in arid areas. *International Journal of Applied Earth Observation and Geoinformation*, 11(5):344–351, 2009.

[24] K.-H. Tseng, C. K. Shum, J.-W. Kim, X. Wang, K. Zhu, and X. Cheng. Integrating Landsat Imageries and Digital Elevation Models to Infer Water Level Change in Hoover Dam. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(4):1696–1709, 2016.

[25] N. Xu, Y. Ma, W. Zhang, and X. H. Wang. Surface-Water-Level Changes During 2003-2019 in Australia Revealed by ICESat/ICESat-2 Altimetry and Landsat Imagery. *IEEE Geoscience and Remote Sensing Letters*, pages 1–5, 2020.

[26] X. Yang, Q. Qin, H. Yésou, T. Ledauphin, M. Koehl, Grussenmeyer P., and Z. Zhu. Monthly estimation of the surface water extent in france at a 10-m resolution using sentinel-2 data. *Remote Sensing of Environment*, 244:111803, 2020.