

Using udometric network data to estimate an environmental covariate

A. Manuela Gonçalves^{*,†} and Marco Costa^{**,‡}

^{*}*Departamento de Matemática e Aplicações, Universidade do Minho, Portugal*

[†]*CMAT - Centro de Matemática da Universidade do Minho*

^{**}*ESTGA - Escola Superior de Tecnologia e Gestão de Águeda, Universidade de Aveiro, Portugal*

[‡]*CMAF - Centro de Matemática e Aplicações Fundamentais da Universidade de Lisboa*

Abstract. Many hydrological and ecological studies recognize the importance of characterizing the temporal and spatial variability of precipitation. In this study, geostatistical methodologies were developed in order to estimate a hydro-meteorological factor by (re)building the space-time distribution of the precipitation associated to monthly averages in a certain hydrological river basin that will be used in the modelling of surface water quality. A hydro-meteorological factor is constructed for each water quality monitoring site (WQMS), based on the analysis of the space-time behaviour of the precipitation observed in an udometric network located in a Portuguese river basin.

Keywords: hydrological basin, water quality, area rainfall estimation, Kriging, hydro-meteorological factor

PACS: 05.45.Tp, 02.50.Ey

Introduction

The problem of area precipitation measurement is presented in order to estimate a hydro-meteorological factor that will be used in the modelling of surface water quality of river basins.

A hydro-meteorological factor is constructed for each WQMS (totalling 20 sites) based on the analysis of the temporal and spatial behaviour of the precipitation (monthly total) observed in a udometric network consisting of a total of 19 meteorological stations located in the geographical area of the River Ave basin, located in the Northwest of Portugal, between 1931-2009. This covariate will integrate a hydro-meteorological component that is crucial in any water quality modelling process ([1] and [3]).

The proposed methodology has in consideration that the hydro-meteorological factor is obtained by respecting the proportionality of the influence area of each WQMS: each influence area overlaps the area of each of the 20 hydrological sub-basins (linked to the 20 WQMS) within the basin area. In this context, the influence regions of each WQMS were defined by the INAG (Portuguese Institute of Water) technicians and they are corroborated by the region's topography and by the land's drainage dynamics. We defined a neighbourhood around each WQMS in order to estimate monthly measurements of rainfall averages in area.

Figure 1 and Figure 2 show the region corresponding to the River Ave hydrological basin (approximately 1400 km^2) with the spatial distribution of the 20 WQMS and the influence areas delineated for each of its 20 associated sub-basins, as well as the entire basin discretized in 368 points (each point corresponding to cell centers of 2 $km \times 2 km$), and the 19 meteorological stations located in this basin: Bran, VChão, Guil, Gont, PLan, Font, MRei, Same, Fafe, Arad, Lour, Taip, Vila, Escu, Cast, Lord, VTod, Para, and VChã. Although the study of the 20 sub-basins of the WQMS has been performed, we only present the sub-basin of River Ave associated to the WQMS of Caniços.

The data is available under Cartesian (X, Y) co-ordinates (X - distance to meridian (km) and Y - distance to the perpendicular (km)) coincident with those of the military maps under a 1/25000 scale. The River Ave hydrological basin is situated between the co-ordinates $X = 147 km$ to East, $X = 209 km$ to West, $Y = 475 km$ to South and $Y = 521 km$ to North. The adopted measuring unit is the millimeter (mm).

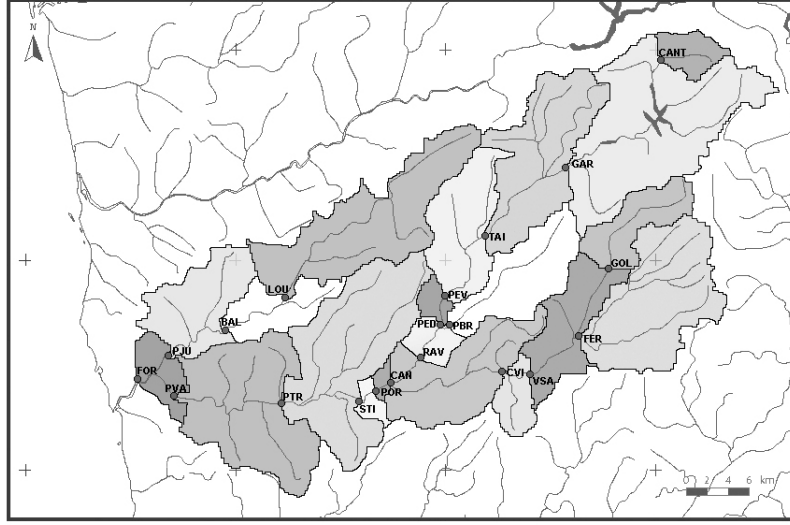


FIGURE 1. The River Ave hydrological basin: its sub-basins as defined by the 20 WQMS (left side), as well as the spatial distribution of 19 meteorological monitoring sites (right side), and its discretization in 368 points.

Methodologies

A geostatistical approach, namely the ordinary Kriging method ([2]) was chosen with the main goal of identifying models which estimate monthly average rainfall in a sub-basin associated to a given WQMS where there are no observed values.

We considered that the values of precipitation recorded throughout time are approximate replicas of the same process, which is valid if the temporal correlation is weak. For this reason, in our study, the observed values of the precipitation throughout the years were separated according to the twelve months of the year.

If we would like to have monthly measurements, $Z_t(s)$, at any given location s in a certain area A (in particular, in our case, a sub-basin area A_i , $i = 1, \dots, 20$ associated to the 20 WQMS), we could obtain the monthly mean area precipitation through the integral $Z_t(A) = \frac{1}{|A|} \int_A Z_t(s) ds$, where the integral is in the mean square sense and $|A|$ stands for the area of A .

The Kriging estimator evaluates this integral as a linear combination of the measurements of the 19 available meteorological monitoring sites, $\hat{Z}_t^*(A) = \sum_{j=1}^{19} \lambda_{A,j} Z_t^*(s_j)$ with the coefficients vector satisfying the relation $\sum_{j=1}^{19} \lambda_{A,j} = 1$. The Caniços sub-basin of the A_i region was regularly discretized by a grid of 76 points. The method used for the calculation of the empirical semivariogram is the method of moments ([5]), modified for a random space-time process $\{Z(s,t) : s \in \mathbb{R}^2, t = 1, \dots, T\}$. The process, in our case, will be a random spatial process $\{Z(s) : s \in \mathbb{R}^2\}$ which is isotropic and second-order stationary.

The sampling processes are $\{Z_t(s_i) : i = 1, \dots, n, t = 1, \dots, T\}$, for T equally spaced time instants (T months), data collected at n point locations on A ($n = 19$ meteorological monitoring points) in the region considered in \mathbb{R}^2 .

The semivariogram is obtained by averaging $\hat{\gamma}_Z(h,t)$ over the T months, a procedure that corresponds to considering the space-time process as a collection of independent T temporal replicates of a purely spatial process $\{Z(s) : s \in A\}$, in which case the purely spatial semivariogram $\hat{\gamma}_Z(h)$ characterizes all the space-time variability.

Under such hypothesis, two realizations corresponding to two different but close time instants may differ substantially, since they are independent, although their spatial variability pattern remains the same ([4]). Henceforth, the final estimator, the empirical semivariogram, is given by

$$\hat{\gamma}_Z(h|l) = \frac{1}{2T|N(h|l)|} \sum_{t=1}^T \sum_{(i,j) \in N(h|l)} [(Z_t(s_i) - Z_t(s_j))^2] \quad (1)$$

TABLE 1. Area precipitation estimates (in *mm*) in the neighbourhood of the WQMS of Caniços during the month of June

1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
150.68	35.54	23.50	18.37	6.68	37.89	48.88	14.38	28.13	114.48	28.22

with $N(h|l) = \{(i, j) : \|s_i - s_j\| - \|h\| \leq l; 1 \leq i \leq j \leq n\}$ and $|N(h|l)| = \#N(h|l)$.

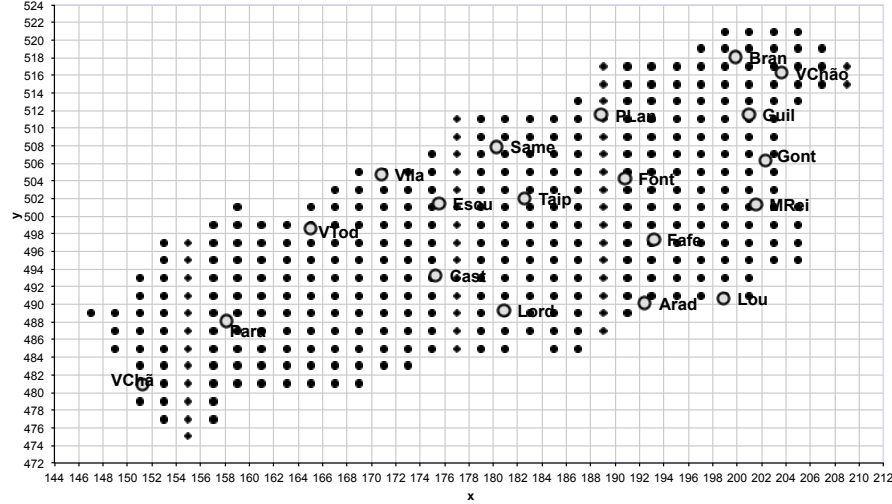


FIGURE 2. The River Ave hydrological basin discretization in 368 points.

The empirical semivariograms of the 12 months (a model for each of the 12 months of the year) was obtained through the estimator defined in (1). For each month, the empirical semivariograms were calculated for the tolerance l in order to define the sets $N(h)$. The least squares adjustments to several stationary models have been performed with an additional condition, enforcing the adjusted variances to be equal to the empirical variance, since this is known to be the best estimator.

The semivariogram model that has best performed has been the Gaussian with a nugget effect, for each of the 12 months. In order to assess the quality of the semivariogram fitting, a cross-validation procedure was carried out, and finally we evaluated the corresponding residuals (differences between estimated and true points values of $Z_t(s_0)$).

Estimation of the Hydro-meteorological factor

As mentioned before, the area A_i of the sub-basin associated to Caniços corresponds to the grid of 76 points of region A_i , which is included in the total area A of the River Ave hydrological basin. The value of ordinary Kriging estimation error variance is given by $\sigma_{ok}^2 = 397.891$. Table 1 only presents the mean area precipitation $\hat{Z}_t(A_i)$ estimates between 1999-2009 for the month of June in the sub-basin of Ave river associated to the WQMS of Caniços.

Thus, for each WQMS, the monthly mean area precipitation was computed in its influence region based on the average point prediction by using Kriging stochastic methodology. The precipitation amount of the sub-basin A_i associated to the WQMS i in month t is estimated by $p_t^{*(i)} = \hat{Z}_t(A_i) \times a_i$, where a_i is the sub-basin's area in Km^2 . For re-scale purposes, it is considered the proportional value $p_t^{(i)} = p_t^{*(i)} \times 10^{-3}$.

In order to construct a hydro-meteorological factor with a stronger linear correlation to the water quality response variable, after an exploratory analysis the hydro-meteorological factor $h_t^{(i)}$ is defined as the logarithm of a linear combination of the covariates $p_{t-1}^{(i)}$ and $p_{t-2}^{(i)}$, as follows $h_t^{(i)} = \log(0.7 p_{t-1}^{(i)} + 0.3 p_{t-2}^{(i)})$.

Concluding remarks

This estimate represents the approximate measure of the average volume of water flow in a given hydrological sub-basin and it can also be used as a hydro-meteorological factor for space-time modelling of water quality. We have also asked the opinion of INAG experts concerning the developed methodology, and they have considered it "realistic" and well-achieved for the resolution of our problem.

REFERENCES

1. M. Costa, A. M. Gonçalves, Clustering and forecasting of dissolved oxygen concentration on a river basin, *Journal of Stochastic Environmental Research & Risk Assessment* **25**, 151–163 (2010).
2. N. A. C. Cressie, *Statistics for Spatial Data*, John Wiley & Sons, New York, 1993.
3. A. M. Gonçalves, T. Alpuim, Water quality monitoring using cluster analysis and linear models, *Environmetrics*, DOI:10.1002/env.1112 (2011).
4. P. C. Kyriakidis, A. C. Journel, Geostatistical space-time models: a review, *Mathematical Geology* **31(6)**, 651–684 (1999).
5. G. Matheron, Principles of geostatistics, *Economic Geology* **58**, 1246–1266 (1963).