International Conference on Hydroinformatics

2014

Precipitation Sensor Network Optimal Design Using Time-Space Varying Correlation Structure

Juan Carlos Chacon-Hurtado

Leonardo Alfonso

Dimitri P. Solomatine

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

PRECIPITATION SENSOR NETWORK OPTIMAL DESIGN USING TIME-SPACE VARYING CORRELATION STRUCTURE

JUAN C. CHACON-HURTADO (1,2), LEONARDO ALFONSO (1), DIMITRI P. SOLOMATINE (1,2)

(1): Integrated Water Systems and Governance, UNESCO-IHE, Institute for Water Education, Westvest 7, Delft, 2611AX, the Netherlands.

(2): Water Resources Section, Delft University of Technology, Stevinweg 1, Delft, 2628CN, the Netherlands.

ABSTRACT

Design of optimal precipitation sensor networks is a common topic in hydrological literature, however this is still an open problem due to lack of understanding of some spatially variable processes, and assumptions that often cannot be verified. Among these assumptions lies the homoscedasticity of precipitation fields, common in hydrological practice. To overcome this, it is proposed a local intensity-variant covariance structure, which in the broad extent, provides a fully updated correlation structure as long as new data are coming into the system. These considerations of intensity-variant correlation structure will be tested in the design of a precipitation sensor network for a case study, improving the estimation of precipitation fields, and thus, reducing the input uncertainty in hydrological models, especially in the scope of rainfall-runoff models.

INTRODUCTION

Most of precipitation sensor network studies based on interpolation uncertainty reduction [1],[2],[3] rely on a long term spatial covariance structure. This assumption can be considered adequate for long term monitoring strategies however, for operative flow forecasting systems, this consideration may be not adequate since, in theory, every precipitation event will have a different correlation structure [4], which is even more prominent in short time steps.

Considering this fact, the selection of the design covariance structure for long term monitoring should be selected accordingly to the monitoring objectives. In this sense, the precipitation sensor network should provide data to minimise the uncertainty, especially when critical situations appear. Due to this, it is proposed an intensity-variant correlation structure.

For this purpose, an intensity-dependent semi-variogram methodology is being proposed to adjust each of the sensors in the system, in a Kriging-like [5] approach to estimate the precipitation field. This intensity dependent correlation structure is dynamically constructed based on local measurements of precipitation for each of the sensor, from which the spatial correlation structure will be established.

With this information, a conditional performance function can be established, based on the definition of optimality for the operational flow forecasting system. This function can assign weights to the different precipitation conditions, allowing for the establishment of an optimal network depending on the operational targets of the system and states.

STUDY CASE DESCRIPTION

This study was developed for the Brue catchment, U.K. due to the ample database obtained from the HYREX project [6]. In this catchment the data of 53 rain gauge stations from 1993 to 1999 was considered. The catchment size is about 135km², for which it has a concentration time between 10 to 12 hours and an average precipitation of 867 mm/year. The slopes in the region are mostly mild and the soil use is mostly rural.

MODELLING OF INTENSITY VARIANT CORRELATION STRUCTURES

Some authors have proposed the use of space-time varying correlation structures [2],[4],[7] considering the fact that different precipitation events are consequence of different atmospheric processes. Additionally, due to the spatial distribution of precipitation events, the selection of the semi-variogram is only valid for regionalised variables, (rigorously) leaving out the possibility of infinitely stretching this property to the whole catchment domain in despite of its extension.

To overcome this limitation, several approaches have been proposed such as event specific, average, assumed linear [4], Climatological [2], Fuzzy [8], space time [9], etc. Each of these approaches consider that the correlation structure of the precipitation is dynamic in time, however none of this approaches considers the fact that the spatial distribution of precipitation is not independent from the intensity.

In this respect, intensity dependent correlation structure is proposed for which a local semi-variogram is fitted based on pre-defined subsets of the recorded precipitation intensity for each of the sensors. This approach is justified by the experimental evidence found in the covariance analysis of different precipitation patterns, in which it was found (an apparently) straight conclusion: The recorded precipitation and the spatial correlation are proportional (Figure 1).

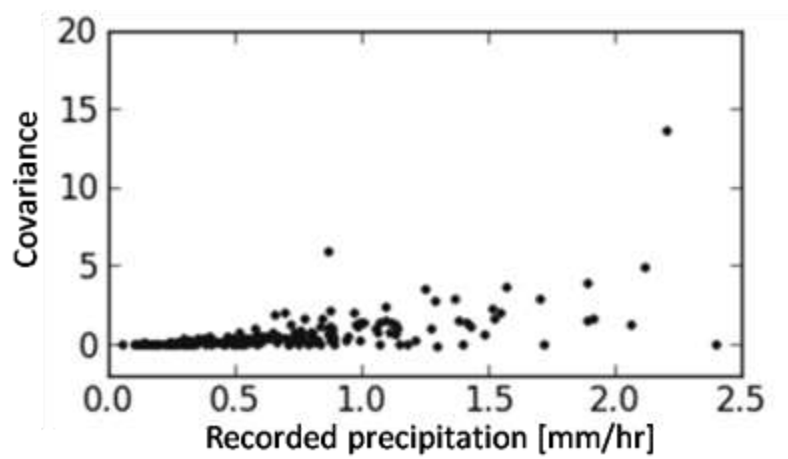

Figure 1 Covariance and recorded precipitation intensity for a sensor pair

If this property is neglected in the estimation of the autocorrelation structure, this will lead to an underestimation of the spatial autocorrelation for high precipitation rates. This is a consequence of the significantly greater percentage of events with low precipitation in a minimum error based theoretical semi-variogram fitting. When this is considered, a less uncertain precipitation fields, especially in high precipitation events, can be adjusted, not only for a reduction in the semi-variogram uncertainty, but also due to a higher correlations between measurement stations (more predictable precipitation field).

CONSTRUCTION OF INTENSITY VARIANT PRECIPITATION FIELDS

For the construction of the precipitation field, a dynamic assignation based on each sensor measurement is applied. In this sense the construction of the covariance matrix of the measurements can be assumed as the average bidirectional covariance between sensors. This situation occurs in despite of the assumption of isotropy, because different sensors will have different semi-variograms that must be merged, in this case, by means of arithmetic averaging. Once the correlation matrix between observations is constructed, the Kriging system can be solved.

This intensity variant precipitation fields will lead to a dynamic uncertainty in the Kriging interpolation. This happens due to the particularisation of the semi-variogram with respect to individual measurements per sensor. If the intensity variant system is selected from a global switch parameter (that affects the selection of the semi-variogram for all the sensors at the same time), the uncertainty in the field will become a scalar function between scenarios. This will lead to an equally optimal configuration of the sensor network in despite of the absolute value of the variance in the interpolation.

In a more realistic case, the performance function using as target the minimisation of the uncertainty estimation will change from an average minimum uncertainty reduction, to a weighted minimum uncertainty approach. This weighted is going to be relevant in the definition of the most optimal configuration of the sensor network, aimed to minimise the uncertainty in the estimation during most relevant conditions. The selection of these weights have to be determined experimentally on the data and will be case dependent, based on the targets of the operative flow forecasting system.

MINIMUM MODEL ERROR

The first approach in this respect was the definition of optimality of the sensor network as such that minimises the error of a hydrological model at the outlet of the catchment. This holds the intrinsic assumption that models are systematically adequate to represent the dynamics of the hydrological cycle in the catchment, as well as the analysis over a complete picture of the precipitation series will allow to determine the marginal effect on the error of the lack of information.

For this purpose, a HBV-96 model was calibrated in the catchment, considering all available information. Considering this as the ground truth, model performance was tested based on the limited availability of information coming from the sensors at potential locations within the catchment domain. The function to evaluate this performance is the Nash-Sutcliffe Efficiency.

MINIMUM OVERALL WEIGHTED APPROACH

The proposed approach to evaluate the overall optimality of the solution is based in the selection a set of weights depending on the importance of the uncertainty of the measured variable. These weights have to be defined prior, and should be consistent with the monitoring objectives of the network. For example, in operative flow forecasting systems for early warning flood, higher precipitation events should be of more relevance, in comparison with lower precipitation conditions.

In this context, the optimisation problem can be formulated as:

$$OF = \sum_{t=1}^{T} \sum_{i=1}^{I} \sigma_{i,t}^{2} w_t$$

(1)

Where t is the time step, T is the total length of the precipitation series, i is the discrete position of the interpolation target, I is the total number of targets to be interpolated and w corresponds to the weighting strategy.

USE OF OPTIMISATION TECHNIQUES TO FIND OPTIMAL POSITION OF SENSORS

It is clear that the selection of the position of gauges is a major combinatorial problem, for which an automatic solution has to be implemented. For the solution of this system, several algorithms have been studied, including metaheuristic search (NSGAII [10], Differential Evolution [11], Augmented Lagrangian Harmonic Search [12]), Particle Swarm Optimisation [13] and Simulated Annealing [14]. The selection of these algorithms is based on the characteristics of the problem, such as size of the search space, as well as the constant present of local minima, which may compromise the validity of the proposed solution.

RESULTS

The results shown that for this approach, the uncertainty in the estimation will strongly depend on the selection of the neighbours for the computation of the interpolation. This influence can be seen in Figure 2 and Figure 3 due to the step-like increase in the uncertainty, present due to the influence of measurements only in local neighbourhoods.

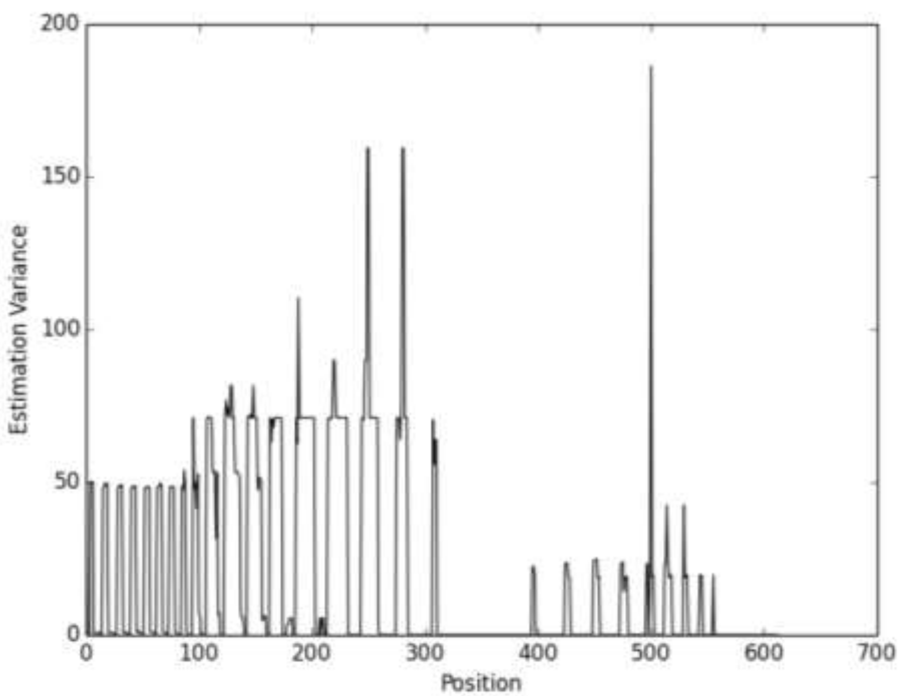

Figure 2 Estimation variance in interpolated precipitation field (event 1)

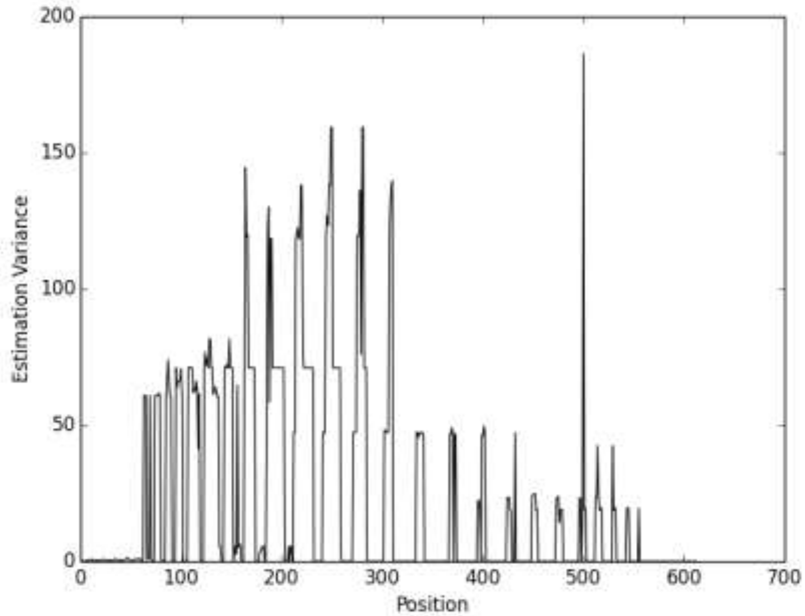

Figure 3 Estimation variance in interpolation precipitation field (event 2)

In the hydrological model based performance, it was found that even with the availability of few sensors, good results can be achieved. This is mainly due to the characteristics of hydrological processes to act as a low-pass filter of the disturbances of the system (precipitation-runoff). However, the solely use of this approach do not allows to provide enough consideration to suggest the network relocation or diminishing, since different objectives has to be taken into consideration due to the several objectives in the monitoring systems [15].

As a result of this exercise, the optimal location of 3, 4 and 5 sensors is determined. Results shows that absolute position of the sensors is not as relevant as relative position of these within the catchment, showing that the problem is an ill-posed one, yet to be refined within its conception. Also, position of the sensors shows a weakly dominant behaviour, making several combinations of solutions equally valid.

This optimisation problem was implemented in an Intel Core i-5-3320M CPU @ 2.6 GHz. The algorithmic implementations of the optimisation routines used in the development of this work were adapted from OpenOpt [16] and pyOpt [17].

REFERENCES

[1] Pardo-Igúzquiza E. Optimal selection of number and location of rainfall gauges for areal rainfall estimation using geostatistics and simulated annealing. *Journal of Hydrology*, (1998).
[2] Bastin G., Lorent B., Duque C. And Gevers M. Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. *Water Resources Research* (1984).
[3] Nowak W., de Barros F.P.J. And Rubin Y. Bayesian geostatistical design: Task-driven optimal site investigation when the geostatistical model is uncertain. *Water Resources Research* (2010).
[4] Verworn A. and Haberlandt, U. Spatial interpolation of hourly rainfall – effect of additional information, variogram inference and storm properties. *Hydrology and Earth Sciences System Sciences* (2011).

[5] Journel A. G., and Huijbregts C. J. *Mining Geostatistics.* Academic Press, London (1978).
[6] Moore R. J., Jones D. A., Cox D. R. And Isham V.S. Desing of the HYREX raingauge network. *Hydrology and Earth System Sciences* (2000).
[7] Rodriguez-Iturbe I. and Mejía J. M. The design of rainfall networks in time and space. *Water Resources Research* (1974).
[8] Bárdossy, A, I Bogardi, and W E Kelly. “Kriging with imprecise (Fuzzy) Variograms. I: Theory.” *Mathematical Geology* 22 (1990a): 63-79.
[9] Bogaert P. Comparison of Kriging techniques in space-time context. *Mathematical Geology* (1996)
[10] Deb, K., Mohan, M., and Mishra, S. A fast multi-objective Evolutionary Algorithm for Finding Well-Spread Pareto-Optimal Solutions. *Evolutionary Computation* (2005).
[11] Price, K. V. Differential evolution: A fast and simple numerical optimizer. 1996 Biennial conference of the North American Fuzzy Information Processing society, NAFIPS. (1996). Berkley, US.
[12] Geem, Z. W., Kim, J. H., and Loganathan, G. V. A New Heuristic Algorithm: Harmony Search. *Simulation*, 76(2), (2001). 60-68.
[13] Jarboui, B., Cheikh, M., Siarry, P., and Rebai, A. “Combinatorial particle swarm optimization (CPSO) for partitional clustering problem.” *Applied Mathematics and Computation*, 192(2). (2007).337-345.
[14] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. “Optimization by Simulated Annealing.” *Science*, 220 (4598), (1983). 671-680.
[15] Alfonso, L, and R Price. Coupling hydrodynamic models and value of information for designing stage monitoring networks. *Water Resources Research* 48 (2012): W08530.
[16] Kroshko, D. OpenOpt: Free scientific-engineering software for mathematical modelling and optimization. (2007).
[17] Perez, R. E., Jansen, P. W., Joaquim, R. R., and Martins, A. pyOpt: A Python-based object-oriented framework for nonlinear constrained optimization. *Structures and multidisciplinary optimization*. 45 (1), (2012). 101-118.