

Added-value from a multi-criteria selection of donor catchments in the prediction of continuous streamflow series at ungauged pollution control-sites

G. Drogue¹, W. Ben-Khediri¹, C. Conan²

5 ¹Laboratoire LOTERR, Université de Lorraine, Metz, 57045 Cedex 1, France

²Agende de l'Eau Rhin-Meuse, Moulins Lès Metz, BP 30019 57161, France

Correspondence to: G. Drogue (gilles.drogue@univ-lorraine.fr)

Abstract. We explore the potential of a multi-criteria selection of donor catchments in the prediction of continuous streamflow series by the spatial proximity method. Three criteria have been used: (1) spatial proximity; (2) physical similarity; (3) stream gauging network topology. An extensive assessment of our spatial proximity method variant is made on a 149 catchment-data set located in the Rhine-Meuse catchment. The competitiveness of the method is evaluated against spatial interpolation of catchment model parameters with ordinary kriging. We found that the spatial proximity approach is more efficient than ordinary kriging. When distance to upstream/downstream stream gauge stations is considered as a second order criterion in the selection of donor catchments, an unprecedented level of efficiency is reached for nested catchments. Nevertheless, the spatial proximity method does not take advantage from physical similarity between donor catchments and receiver catchments due to the low overlapping between hydrological cousins and physical cousins of receiver catchments.

1 Introduction

Because quantitative hydrological information contribute to the understanding of short and medium-term fluctuations in pollutants concentrations recorded in river flows (e.g. Burt *et al.*, 2010), there is a strong demand to reconstruct continuous flow time series at ungauged pollution-control sites for short (*i.e.* up to 15 days) and medium (*i.e.* up to 180 days) lead times, before and after the river water sampling has occurred. The best way to handle this problem would be to set up a coordinated monitoring network in order to get continuous streamflow series at the point where water quality data are collected. Unfortunately, different brakes like financial costs, logistics and most often diverging interest between water flow managers prevent the design of such a coordinated data-acquisition network. Therefore, rainfall-runoff modelling strategies can be viewed as surrogate models for reconstructing and simulating continuous flow time series at ungauged pollution control-sites. In the French part of the Rhine-Meuse catchment, we made different attempts to produce daily streamflow series through regionalisation of catchment model parameters (Drogue and Plasse, 2014; Plasse *et al.*, 2014). In the worst case, where no hydrological information is available at the point of interest, we came to the conclusion that, in our temperate non

Mediterranean area, the spatial proximity method is the best approach to solve the regionalisation problem of hydrograph prediction provided for the hydrological network is sufficiently dense (*i.e.* 1 station per 250 km²). But our results have also shown that there is still a considerable room for progress in reducing the prediction error in such hydrological information at ungauged catchments. For that reason and also because there is an important part of nested catchments in our catchment-data set, we test a variant of the spatial proximity approach where the donor catchments are not only selected according to their spatial proximity but also according to their physical similarity and gauging network connectivity from the receiver catchment.

1 Setting the scene: study area and datasets

1.1 Study area

10 The investigated territory corresponds to the French part of the Rhine-Meuse catchment (North-Eastern part of France). The presence of the Vosges Mountains induces climatic gradients among highest of France. Considering the weak influence of snow on the hydrological regime of the upstream mountain rivers, the snow component is not taken into account in the regionalisation study reported in this paper.

1.2 Datasets

15 We split the entire period of observations (1990-2002) into two periods: from 1 January 1990 to 31 December 1995 and from 1 January 1996 to 31 December 2002. Warm up periods of one year has been used in both cases. For the efficiency estimation of the catchment model regionalisation we used the second sub-period as a confirmatory period.

1.2.1 Catchment-data set

20 We made an extensive assessment of our regionalisation scheme on a dense stream gauging network (approximately 1 gauging station for 250 km²) comprising 149 reliable stations gauging non-regulated rivers and providing daily streamflow values validated on the target period 1990-2002. More than half of these stations are also dedicated to water quality monitoring. The drainage areas lie between 5 and 11500 km². The catchment set includes 40 % of nested catchments (*i.e.* 60 % of adjacent catchments). The territory covered by our catchment sample is approximately 38,000 km². This accounts for 15% of the total drainage area of the Rhine-Meuse catchment ($\approx 200,000$ km² at the outlet).

25 1.2.2 Climate forcings and rainfall-runoff model

In addition to streamflow data, we also collected daily precipitation and potential evapo-transpiration from the SAFRAN gridded climatology data. The daily lumped GR4J rainfall-runoff model (Perrin *et al.*, 2003) has been calibrated over the

training period (1990-1995) by using the Broyden–Fletcher–Goldfarb–Shanno algorithm (hill climbing optimization technique; see Byrd *et al.*, 1995). Parameters have been optimized with the Nash-Sutcliffe coefficient computed on the square root of the daily streamflow series ($NS_{\text{sqrt}Q}$). This power transformation makes possible to define a parameter set representative of the catchment behaviour on all the streamflow range. The repetition of the split-sample test for all the catchment-data set allows setting up a regional library containing 149 vectors of four optimal parameters.

2 How strengthening the spatial proximity approach?

In a first simulation experiment, we optimized the performance of the spatial proximity approach allowing the output averaging option and four geographic neighbours (Plasse *et al.*, 2014). In this basic version of the spatial proximity approach, we filter out poorly modelled donor catchments, *i.e.* the ones having a $NS_{\text{sqrt}Q}$ below 0.7 in calibration mode. In this study we go one step further by selecting geographic neighbours according to a regional composite rank mixing elementary ranks related to spatial proximity, physical similarity and stream gauging network topology. The experiment design is described in Fig. 1.

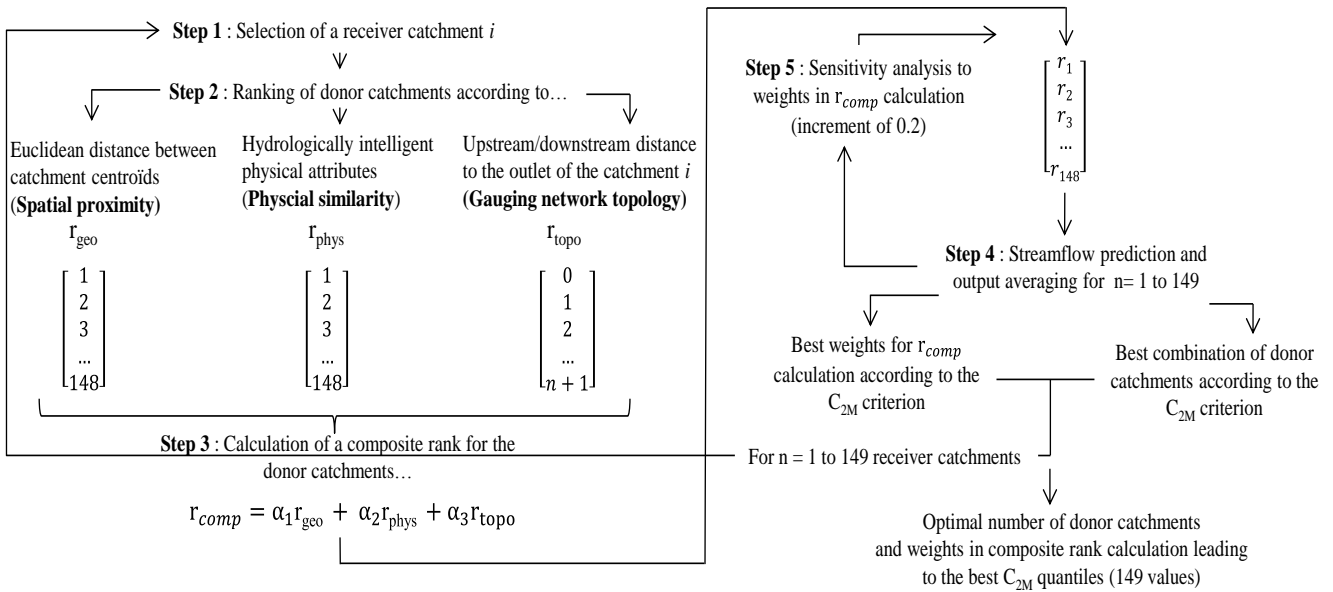


Figure 1. Flowchart of the regionalisation scheme applied to our catchment-data set.

15

As for a trully ungauged catchment, streamflow hydrograph is unkown, we apply a method that introduces hydrological catchment behaviour in the assessment of catchment physical similarity (Oudin *et al.*, 2010). We compute two catchment classifications (a hydrological one and a physical one) where similarity is defined as an Euclidean distance in the catchment property space: for the hydrological classification we use seven hydrological signatures calculated on the 1990-2002 period (runoff coefficient, lag time, Base Flow Index, slopes of the Flow Duration Curve for the high flow range, the low flow range and the medium flow range, rising limb density); for the physical classification we use 70 catchment attributes related

to climate, geology, land cover, hydrology and morphology and a weighted version of the Euclidean distance for estimating the catchment similarity. Then, we sought physical attributes allowing the best matching between both classifications according to the Adjusted Rand Index -ARI- (Hubert and Arabie, 1985). We also incorporate the hydraulic connectivity of nested catchments in the composite rank calculation (Fig. 1). We proceed in a very simple way: for a target point having upstream/downstream neighbour(s), we assign a topological rank to neighbour(s) according to the stream network distance between the target point and its neighbours. The closest is the gauging station the smallest is the rank. For adjacent stations, we add one to the r_{topo} value of the furthest upstream/downstream neighbour(s) rank value. For an adjacent receiver catchment, the topological rank is set to 0. We examine the efficiency of our regionalisation method by jackknife cross validation (Fig. 1). As the catchment-data set is quite large, a [-1; 1] bounded version of the NS_{sqrtQ} criterion (called C_{2M}) is calculated for efficiency estimation (Mathevet *et al.*, 2006). For sake of robustness evaluation, the GR4J model parameters are also regionalised by ordinary kriging (OK). The variogram properties of GR4J model parameters are summarized in Table 1. As the X2 parameter has no spatial autocorrelation, the median value is used.

model parameter	Definition	Variogram model
X1 (mm)	maximal capacity of the production reservoir	spherical; range : 100 km; sill : 250 000 mm ²
X3 (mm)	capacity of the non linear routing reservoir	spherical; range : 37 km; sill : 3100 mm ²
X4 (day)	unit hydrograph time base	spherical; range : 35 km; sill : 0.25 day ²

Table 1. Variogram models used to interpolate the X1, X3 and X4 GR4J model parameters in the study area.

3 Results and discussion

3.1 Which physical attribute(s) of catchment could be used as proxie(s) for hydrological similarity?

The catchment classification based on hydrological signatures leads to nine clusters. The two physical attributes optimizing the ARI index are the catchment perimeter and the proportion of Strahler stream order 5 (Table 2). Nevertheless, the intermediate value of ARI (0.545) shows that physical catchment characteristics are poor proxies for describing hydrological behaviour similarity patterns.

Iterations	Physical attributes	Number of clusters	Weights in Euclidean distance metric	Adjusted Rand Index (-)
1	Catchment perimeter	13	1	0.333
2	Catchment perimeter (CP) + proportion of Strahler stream order 5 (SSO5)	13	CP : 0.5 SSO5 : 0.5	0.545

Table 2. Physical attributes maximizing the Adjusted Rand Index (ARI) computed between the classification based on physical attributes and the classification based on proxies of hydrological behaviour.

In order to orient future researches, a focus should be made on hydrological signatures that discriminate the most hydrological cousins and physical cousins (Oudin *et al.*, 2010).

3.2 Overall performance of the multi-criteria spatial proximity approach

Looking at Fig. 2 we can see that the spatial proximity method produces less prediction error than OK over the confirmatory period. An improvement of model hydrograph prediction at ungauged sites could be obtained when using the multi-criteria selection approach with four donor catchments, especially for well regionalised catchments (Fig. 2). Optimal weights in composite rank leading to that result are respectively $\alpha_1 = 0.8$ for r_{geo} (spatial proximity), $\alpha_2 = 0.1$ for r_{phys} (physical similarity) and $\alpha_3 = 0.1$ for r_{topo} (distance between upstream/downstream gauging stations). It means that for the considered monitoring stream gauge network, physical similarity and network topology have a second order effect in the selection of pertinent donor catchments in comparison to Euclidean distance between catchment centroids.

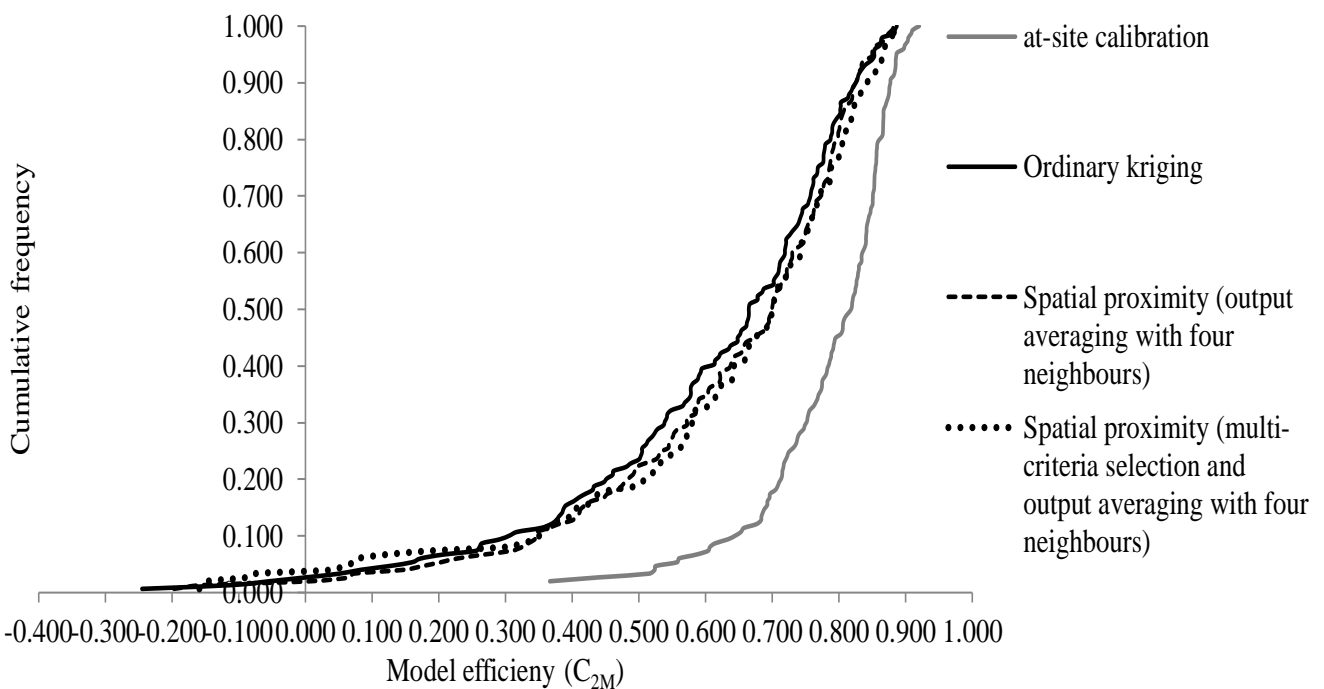


Figure 2. Cumulative distribution functions (CDFs) of GR4J model efficiencies using model parameters regionalised by three methods. CDF for at-site calibration efficiencies is also shown. Results are given in validation mode.

4 Conclusions

We tested a multi-criteria variant of the spatial proximity approach for daily hydrograph prediction at ungauged sites. Three criteria were used to choose the neighbours of a target catchment: spatial proximity, physical similarity (conditioned by catchment hydrological similarity patterns) and distances between upstream/downstream neighbours for nested catchments. No added value comes up from using physical similarity in the selection process of donor catchments. Poorly regionalised

catchments do not take advantage from the multi-criteria approach. For well modelled nested catchments, prediction error of hydrograph in ungauged conditions could be reduced by prioritizing upstream/downstream neighbours among the closest donor catchments. In light of these results, hydrograph prediction over ungauged catchments through catchment model regionalisation definitely appears as a learning process.

5

Acknowledgments

We would like to thank Météo France for providing the SAFRAN reanalysis data, the SCHAPI for providing access to the HYDRO streamflow archive, the German Federal Waterways and Shipping Administration as well as the German Federal Institute of Hydrology (BfG) for providing the German streamflow data, and the Service public de Wallonie, Direction générale opérationnelle “Mobilité et Voies hydrauliques”, Direction de la Gestion hydrologique intégrée, Service d’Etudes Hydrologiques (SETHY) for making the Belgian streamflow data freely available. We also would like to acknowledge the Rhine-Meuse French Water Agency for its financial support.

References

- 15 Burt T.P., Howden N.J.K., Worrall F., Whelan M.J. 2010. Long-term monitoring of river water nitrate: how much data do we need? *J. Environ. Monitor.*, 12, 71–79. doi:10.1039/b913003a.
- Byrd R.H., Lu P., Nocedal J., Zhu C. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16, 1190–1208. doi:10.1137/0916069.
- Drogue G, Plasse J. 2014. How can a few streamflow measurements help to predict daily hydrographs at almost ungauged
20 sites? *Hydrological Sciences Journal*, 59, 12. <http://dx.doi.org/10.1080/02626667.2013.86503>.
- Hubert L, Arabie P., 1985. Comparing partitions. *J. Classif.*, 2, 193–218.
- Mathevet T, Michel C, Andréassian V, Perrin C. 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. *IAHS Red Books Series*, 307, 211-219.
- Oudin L., Kay A.I., Andreassian V., Perrin C., 2010. Are seemingly physically similar catchments truly hydrologically
25 similar? *Water Resour. Res.*, vol.46, W11558.
- Perrin, C., Michel, C., and Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279, 275–289. doi:10.1016/S0022-1694(03)00225-7.
- Plasse J, Drogue G, François D, Battaglia P, Conan C. 2014. Using point flow measurements for guided reconstruction of daily streamflow time series through continuous flow simulation: a regional case study on the Rhine-Meuse district. *La
30 Houille Blanche, Revue Internationale de l'Eau*, 1, 45-52. <http://dx.doi.org/10.1051/lhb/2014007>.