

Spatial prediction on river networks: comparison of top-kriging with regional regression

G. Laaha,^{1*} J.O. Skøien² and G. Blöschl³

¹ Institute of Applied Statistics and Computing (IASC), University of Natural Resources and Life Sciences, BOKU Vienna, Austria

² Institute for Environment and Sustainability, Joint Research Centre of the European Commission, Ispra, Italy

³ Institute for Hydraulic and Water Resources Engineering (IWI), Vienna University of Technology, Austria

Abstract:

Top-kriging is a method for estimating stream flow-related variables on a river network. Top-kriging treats these variables as emerging from a two-dimensional spatially continuous process in the landscape. The top-kriging weights are estimated by regularising the point variogram over the catchment area (kriging support), which accounts for the nested nature of the catchments. We test the top-kriging method for a comprehensive Austrian data set of low stream flows. We compare it with the regional regression approach where linear regression models between low stream flow and catchment characteristics are fitted independently for sub-regions of the study area that are deemed to be homogeneous in terms of flow processes. Leave-one-out cross-validation results indicate that top-kriging outperforms the regional regression on average over the entire study domain. The coefficients of determination (cross-validation) of specific low stream flows are 0.75 and 0.68 for the top-kriging and regional regression methods, respectively. For locations without upstream data points, the performances of the two methods are similar. For locations with upstream data points, top-kriging performs much better than regional regression as it exploits the low flow information of the neighbouring locations. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS change of support; geostatistics; spatial interpolation; prediction in ungauged basins (PUB); stream distance; low flows and droughts

Received 20 March 2012; Accepted 1 October 2012

INTRODUCTION

The estimation of stream flow-related variables is usually based on regression methods between the stream flow variable and physiographic catchment characteristics. If the study domain is large or very heterogeneous in terms of the low flow processes, a number of authors have suggested to split the domain into regions and to apply a regression relationship to each of the regions independently (e.g. Nathan and McMahon, 1990; Gustard *et al.*, 1992; Aschwanden and Kan, 1999; Laaha and Blöschl, 2006b). This is termed the regional regression approach. Recently, geostatistical methods have been proposed for stream networks. One approach is to treat the estimation on a river network as a one-dimensional problem. Gottschalk (1993a) was probably the first to develop a method for calculating covariance along a stream network based on river distance. Ver Hoef *et al.* (2006) and Cressie *et al.* (2006) proposed a moving average approach for interpolating on the one-dimensional river system, where the estimation either proceeds upstream or downstream. Peterson and Ver Hoef (2010) proposed a combination of these two methods. Garreta *et al.* (2009) evaluated the three model types in the context of summer stream temperature and nitrate concentration and found the combined model to be superior to each of the individual models.

The second approach treats the river network as a two-dimensional problem. In this approach, the runoff process is conceptualised as a spatially continuous process which exists at any point in the landscape, and stream flow is the integral of local runoff over the catchment. Sauquet *et al.* (2000) and Gottschalk *et al.* (2006) proposed a block-kriging method where spatial dependence of catchments with different, non-zero support is modelled by a regularised covariogram. Skøien *et al.* (2006) extend the work of Sauquet *et al.* (2000) to account for the stronger spatial correlation between nested basins than between un-nested basins. They showed that their method, known as topological kriging or top-kriging, can be used, in an approximate way, for a range of stream flow-related variables including variables that do not aggregate linearly and are non-stationary. Also, they apply a kriging with uncertain data (KUD) estimator (de Marsily, 1986 p. 300; Merz and Blöschl, 2005) to account for local uncertainties of the observations.

The aim of this paper is twofold: (1) to test top-kriging in the context of low stream flows, and (2) to compare it with the regional regression method which constitutes the actual benchmark for low stream flow regionalisation. The analyses will be performed on a comprehensive Austrian data set.

TOP-KRIGING METHOD

There are two main groups of processes that control stream flow. The first group consists of runoff generating

*Correspondence to: G. Laaha, Institute of Applied Statistics and Computing (IASC), University of Natural Resources and Life Sciences, BOKU Vienna, Austria
E-mail: gregor.laaha@boku.ac.at

processes acting over the catchment area. These processes are continuous in space. The second group is related to runoff aggregation and routing. These processes are related to the river network topology. The main idea of top-kriging is to combine the two groups of processes in a geostatistical framework. For this, runoff generation is conceptualised as a spatially continuous process which exists at any point in the landscape. Instead of observing a pointwise realisation of this process, data $z(A_1), z(A_2), \dots, z(A_n)$ are observed at stream gauges, where

$$z(A_i) = \frac{1}{|A_i|} \int_{A_i} z(\mathbf{x}) d\mathbf{x} \quad (1)$$

and A_i denotes the spatial support of $z(A_i)$. For stream flow variables, A_i is the catchment which drains into a river location \mathbf{x}_i , and $|A_i|$ is its surface area. If we follow the stream from the source to the mouth, the support increases as the catchment area increases. In this context, the transfer of information between river locations adds up to the area-to-area change of support problem in geostatistics (e.g. Gotway and Young, 2002), based on the classical Euclidean distance metric. For spatial prediction on a river location \mathbf{x}_0 with catchment area A_0 from non-point samples $z(A_1), z(A_2), \dots, z(A_n)$, the linear block-kriging predictor given by

$$\hat{z}(A_0) = \sum_{i=1}^n \lambda_i z(A_i) \quad (2)$$

is used. Assuming runoff generation as an intrinsic stationary random process, the optimal weights λ_i can be found by solving the kriging system:

$$\begin{aligned} \sum_{j=1}^n \lambda_j \bar{\gamma}_{ij} - \lambda_j \sigma_i^2 + \mu &= \bar{\gamma}_{0i} & i = 1, \dots, n \\ \sum_{i=1}^n \lambda_i &= 1 \end{aligned} \quad (3)$$

The $\bar{\gamma}_{ij}$ refers to the expected semivariance between two observations i and j with non-zero support (see below), μ is the Lagrange multiplier from the constraint minimisation, and σ_i^2 represents the measurement error or uncertainty of observation i . The kriging equations (Equation (3)) result from minimising prediction mean squared error subject to unbiasedness constraints. The use of measurement errors in the kriging equations is termed KUD (de Marsily, 1986 p. 300).

Since the observations have a non-zero support A , the expected semivariances $\bar{\gamma}_{ij}$ between the observations need to be obtained by regularisation (Cressie, 1993, p. 66). This means that instead of one variogram model, a family of variogram models for different catchment areas (kriging support) is used, which accounts for the different scales and the nested nature of the catchments. Assuming the existence of a point semivariogram γ_p , the expected semivariance $\bar{\gamma}_{ij}$ between two observations with catchment areas A_i and A_j , respectively, is:

$$\begin{aligned} \bar{\gamma}_{ij} &= 0.5 * Var(z(A_i) - z(A_j)) \\ &= \frac{1}{|A_i||A_j|} \int_{A_j} \int_{A_i} \gamma_p(\mathbf{s} - \mathbf{u}) ds du \\ &\quad - 0.5 * \left[\frac{1}{|A_i|^2} \int_{A_i} \int_{A_i} \gamma_p(\mathbf{s} - \mathbf{u}) ds du \right. \\ &\quad \left. + \frac{1}{|A_j|^2} \int_{A_j} \int_{A_j} \gamma_p(\mathbf{s} - \mathbf{u}) ds du \right] \end{aligned} \quad (4)$$

\mathbf{s} and \mathbf{u} are position vectors within each catchment used for the integration. The first part of this expression integrates all the variance between the two catchments, while the second part subtracts the variance within the catchments. Consequently, the $\bar{\gamma}_{ij}$ will be lowest for close-by locations at the same river, due to the overlapping support. In top-kriging, the integrals in Equation (4) are computed by discretising the catchment area into a grid of points, according to common geostatistical practice. Figure 1 (from Skjøien *et al.*, 2006) shows a schematic of two nested catchments, their discretisation by a square grid, and the distances between the discretised points within the catchments.

EXAMPLE OF LOW STREAM FLOWS

Data set

We test the performance of top-kriging for the example of low stream flows. The data set used in this study consists of about 8000 catchments in Austria. For 490 of these catchments, daily streamflow measurements Q are available for the period 1977–1996 (Laaha and Blöschl, 2006a). Q was standardised by the catchment area to transfer stream

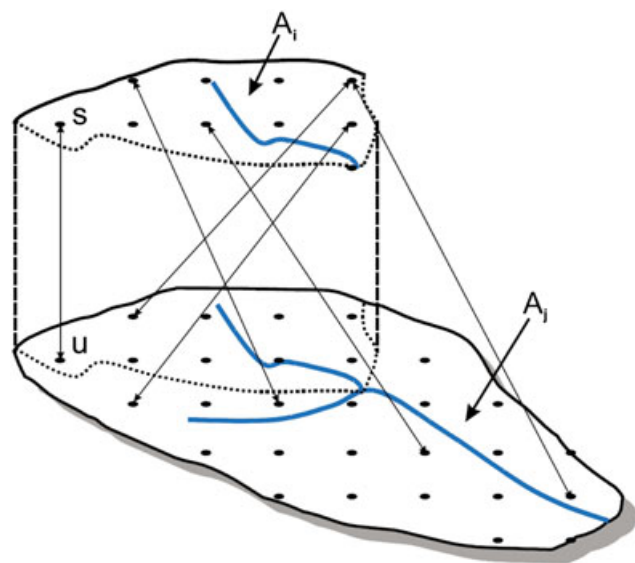


Figure 1. Calculation of regularised variogram for two nested catchments (schematic, from Skjøien *et al.*, 2006). Catchment i is nested within catchment j , A_i and A_j are the respective catchment areas (support), and \mathbf{s} and \mathbf{u} are the discretisation points within each catchment

flow into a continuous spatial variable q with point support. From these data, the specific discharge q_{95} ($l s^{-1} km^{-2}$) that is exceeded on 95% of the time, $Pr(q > q_{95}) = 0.95$, was calculated. q_{95} was used as the target variable in this study (see Figure 2). The measurement errors of q_{95} were taken from Laaha and Blöschl (2005 and 2007). From this assessment, the observed low flow characteristics exhibit errors between 3% and 24% of q_{95} for 20- to 5-year stream flow records. The so-obtained data set of low flow characteristics and their measurement errors was used in top-kriging estimation.

The top-kriging approach was compared with the regional regression model presented in Laaha and Blöschl (2006b). The regression model was fitted to low flow data of 325 Austrian catchments from a 20-year period, and 31 physiographic catchment characteristics were used as potential predictor variables, including sub-catchment area (A), topographic elevation (H), topographic slope (S), precipitation (P), geological classes (G), land use classes (L), and stream network density (D), see Laaha and Blöschl (2006a). The comparison of top-kriging and regional regression performances was carried out on 300 catchments which constitute the intersection of both data sets.

Estimation of variogram

For applying top-kriging, a variogram model is needed, and following Skøien *et al.* (2003), a point variogram with a nugget effect of the following shape was used:

$$\gamma_p(h) = ah^b \left(1 - e^{-(h/c)^d} \right) + C_{0p} \quad (5)$$

a , b , c , and d are parameters. a is related to the sill of the variogram, c is a correlation length, b and d define the long and short distance slope of the variogram in a log-log plot, and C_{0p} is the point nugget effect. Despite it is also possible to use many of the classical variogram models in top-kriging (e.g. exponential, Gaussian, or spherical variogram), we have chosen to use the model according to Equation (5) as it proved well suited in an earlier flood regionalisation study (Skøien *et al.*, 2003). To ensure a valid kriging model, the

point variogram generally needs to be estimated and modelled with a valid conditionally negative-definite function based on the data. We proved the validity of the model for parameters $a, b, c, d > 0$ and $2b + d < 1$ by showing that for all $z > 0$, $e^{-zy^{(h)}}$ is positive definite (Cressie, 1993, p.86ff). The proof was given through the sufficient condition that a decreasing real function which is even and convex on $(0; \infty)$ is positive definite (e.g. Berg and Forst, 1975, Theorem 5.4).

For stream flow and related variables the point variogram cannot be directly fitted to the sample variogram because of the different support (catchment area) of the measurements. Rather, it can be determined by the back-calculation approach of Kyriakidis (2004) and Mockus (1998). In this approach, a number of theoretical point variograms with different parameter sets are assumed. Then, each point variogram model is regularised for all pairs of observations using Equation (4), in order to find the point variogram whose regularisations fit best to sample variogram values. According to Cressie (1985), the sum of weighted squared residuals between regularised theoretical semivariances and observed semivariances was used as the optimality criterion. The observed semivariances were estimated using a classified variogram estimator which is similar to Matheron's (1965) traditional estimator. The classes were defined along distance between centroids (h), area of the smaller catchment A_i , and area of the bigger catchment A_j .

From the automatic fitting procedure, we obtained a point variogram model with parameters $a = 112$, $b = 0.001$, $c = 4000$, $d = 0.100$, and $C_{0p} = 0.580$. Figure 3 shows the point variogram together with a number of regularised variograms for different catchment areas. In all cases, a square catchment shape was assumed. As the catchment area increases, the gamma values decrease because of the smoothing effect of regularisation. Due to their different supports, catchments of different size will always have a semivariance > 0 , also in the hypothetical limiting case when the centre-to-centre distance is zero. This is the reason why all variograms between catchments of different size

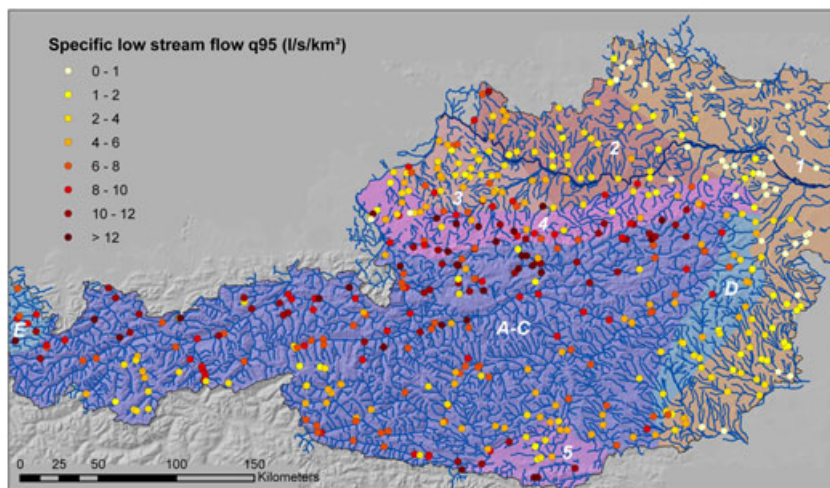


Figure 2. Situation of stream gauges used in this study. Point symbol colours indicate magnitude of specific low flows q_{95} calculated from observed stream flow records. Background shading refers to seasonality types (A–E winter seasonality, 1–5 summer seasonality). The area shown is Austria which is 600 km across

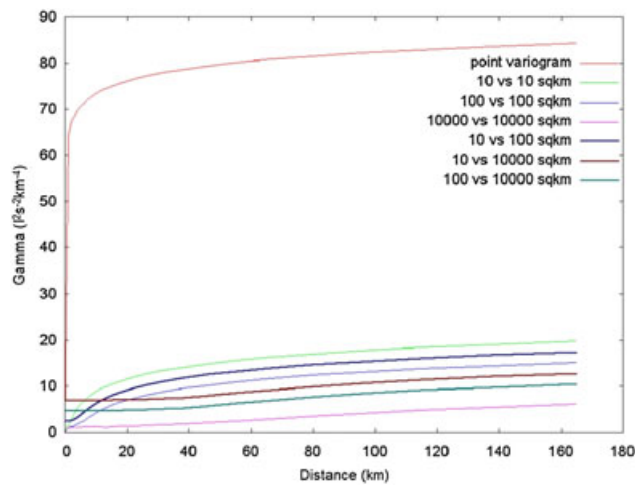


Figure 3. Variogram model of q95 top-kriging

start with an apparent nugget effect. A scatter plot between the theoretical variogram values γ_{mod} and observed variogram values γ_{obs} shows that the variogram fits well to observations, although the model has a tendency of underestimating some of the larger gamma values (Figure 4).

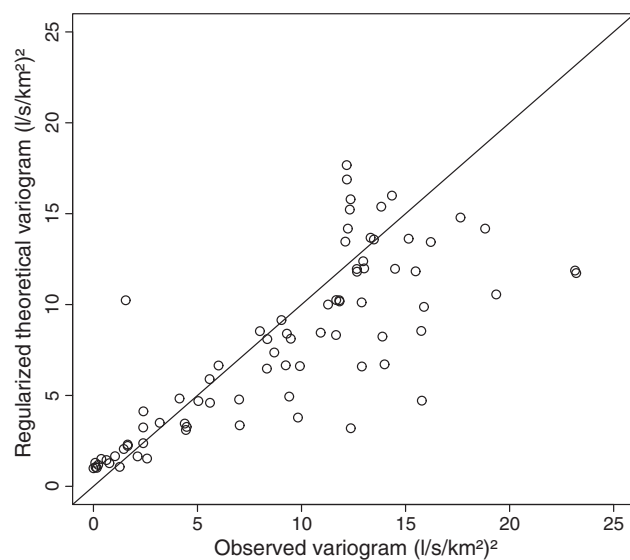


Figure 4. Goodness-of-fit between theoretical and observed variogram

For the purpose of demonstrating the characteristics of top-kriging, however, the fit was considered acceptable.

Regression model

For application of the regression model, the study domain was subdivided into eight geographically contiguous regions of similar low flow seasonality (Figure 2). For each region, a multiple regression relationship between low flow characteristic q95 and catchment characteristics was fitted to data. Nested catchments were split into sub-catchments between subsequent stream gauges. To minimise inter-correlations and multicollinearity, a stepwise regression approach was adopted using Mallow’s Cp (Weisberg, 1985, p. 216) as the criterion of optimality. The method was made more robust by an interactive outlier detection based on Cook’s distance criterion.

The resulting regional regression models are shown in Table I. The equations suggest that *precipitation* is one of the most important controls of low flows in Austria. It determines the water fluxes into the catchment which are available during the recession periods. Hence, it has a positive effect on low flows in both, Alpine and lowland regions. The positive effect on winter low flows in the Alps may be related to a tendency of precipitation periods to be generally warmer than dry winter periods. Catchment *topography* is represented in all regional models, generally by one altitude parameter or by one slope parameter. The topography further has a strong influence on precipitation, and appears as an equally important control of low flows in Austria as precipitation. Altitude has a positive effect on summer low flows (less evaporation) and a negative effect on winter low flows (lower temperature), which is most pronounced in the highest alpine ranges. *Slope* generally has a positive effect on low flows; it is possibly correlated with storage volume in high mountains. Catchment *geology* is represented in many regional models. Low flows increase with the porosity of the formations, so that sediments and limestone formations have a positive effect on low flows, and crystalline rocks have a negative effect on low flows. *Land use* seems to play a subordinate role. Apart from forest areas, the proportion of rocks and the proportion of wetlands appear in the regression models of the alpine regions, but

Table I. Components of the regional regression model. R² denotes the goodness-of-fit coefficient of determination; \hat{q}_{95} predicted low flow q95 [l s⁻¹ km⁻²]

Group	Region	R ²	Model
A–C	Alps	51%	$\hat{q}_{95} = 0.67 + 0.40 \cdot P + 0.17 \cdot G_Q - 0.01 \cdot G_C + 6.43 \cdot L_{WE} + 0.14 \cdot S_M - 0.04 \cdot L_R - 0.20 \cdot H_0$
1	Flatland and hilly terrain	71%	$\hat{q}_{95} = -0.12 + 0.11 \cdot S_M + 0.05 \cdot G_{GS} + 0.02 \cdot G_C$
2	Bohemian Massif	64%	$\hat{q}_{95} = -3.31 + 1.96 \cdot P_W$
3	Foothills of Alps (Upper Austria)	68%	$\hat{q}_{95} = -10.04 - 0.76 \cdot D + 3.27 \cdot P - 2.22 \cdot H_0$
4	Flyschzone	63%	$\hat{q}_{95} = -6.17 + 0.06 \cdot G_L + 2.07 \cdot P_S - 0.06 \cdot L_F$
5	Lower Carinthia	83%	$\hat{q}_{95} = -17.48 + 3.56 \cdot D + 20.06 \cdot L_{WE}$
D	Pre-Alps (Styria)	89%	$\hat{q}_{95} = -7.99 + 1.08 \cdot P + 0.04 \cdot L_F$
E	Pre-Alps (Vorarlberg)	60%	$\hat{q}_{95} = 18.20 - 0.18 \cdot S_{MO}$

Symbols: P, P_W, P_S mean annual / winter / summer precipitation [10² mm], H₀ altitude of the stream gauge [10² m], S_M mean slope [%], and S_{SL} area percentages of moderate slope [%]; G_Q, G_L, G_C, G_{GS} area percentages of Quaternary sediments / Limestone / Crystalline rock / areas with shallow groundwater table [%]; L_F, L_R, L_{WE} area percentages of forest / rocks / wetland [%]; D stream network density [10² m/km²].

they are rather associated with high mountains or hydrological conditions than with classical land uses.

The coefficients of determination of Table I show that the regional regression model based on seasonality regions performs well in most regions, with coefficients of determination ranging from 60 to 80%. The exception is the Alpine, winter low flow-dominated region, where the coefficient of determination is only 51%. This low coefficient of determination may be related to lumping three types of seasonality (A, B, C) that do not form contiguous regions into a single contiguous region. The regression model for the Pre-Alps of Styria exhibits a larger coefficient of determination of 89%. Overall, the results show that the seasonality characteristics seem to contain a lot of information highly relevant to low flow regionalisation.

In addition to average errors represented by the coefficient of determination, it is important to have an understanding of the errors committed when predicting at an individual, ungauged site. The regression standard error of predicting individual observations represents the uncertainty of regression estimates and is given by:

$$e = \sqrt{s^2 + se^2(\hat{q}95_0)} \tag{6}$$

where s is the standard deviation of the residuals of multiple regression, and $se(\hat{q}95_0)$ is the standard error of the predicted mean value $\hat{q}95_0$ (Draper and Smith, 1998, p. 130). The resulting standard errors range between 0.4 and $2.8 \text{ ls}^{-1} \text{ km}^{-2}$, depending on the region.

Spatial estimates and error standard deviations

The fitted models were applied to estimate specific low flows $q95$ for 8000 nodes of the river network. Assuming sub-catchments as homogeneous units with constant $q95$, these estimates are representative for the river segment upstream of the node and can thus be plotted as vector maps of the stream network.

The spatial estimates obtained by top-kriging and regional regression are presented in Figure 5. Although estimated patterns of both models are similar on a larger scale, there are clear differences in terms of small-scale variability, and top-kriging yields more heterogeneous patterns than regression. The small-scale variability of estimates corresponds with the weight a model gives to local information relative to regional information. Regression is fitted to a regional data set without local weighting of data points. Thus, the variability of estimates solely originates from the spatial variability of predictors reflecting topography and

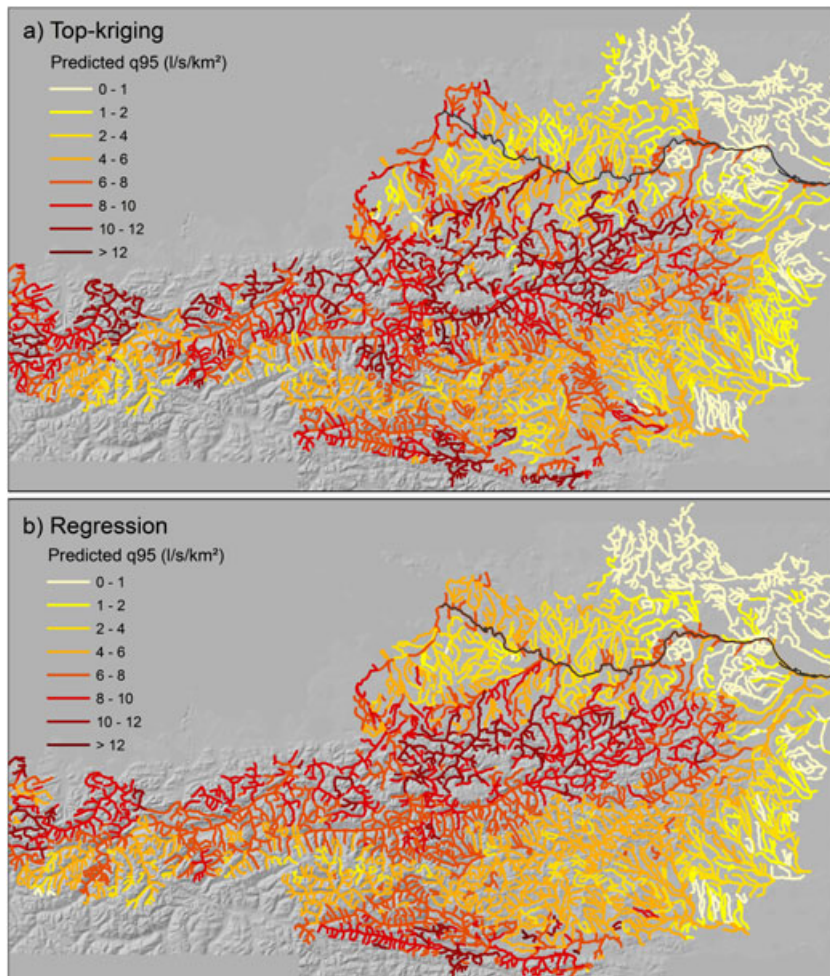


Figure 5. Estimated specific low stream flows $q95 \text{ (ls}^{-1} \text{ km}^{-2}\text{)}$

climate of the study area. Top-kriging, however, distributes weights according to proximity and topology of catchments, depending on the chosen variogram model and the size of the local neighbourhood. For the setting used in this study, much weight is given to local data, and the patterns of top-kriging estimates are therefore more heterogeneous than regression estimates.

The uncertainties estimated by top-kriging and regression are very different for most of the stream network. The prediction errors of regression (Figure 6b) are nearly constant over large areas which, to some degree, correspond to the geographic regions described in section 3.3. They basically reflect the estimation variance of each regional regression model. The prediction errors of top-kriging (Figure 6a) are more heterogeneous. Top-kriging gives relatively small uncertainties on the main river with error standard deviations between 0 and $1 \text{ l s}^{-1} \text{ km}^{-2}$. This is only slightly larger than the error standard deviations of the observations. On the other hand, the uncertainties of some of the tributaries are considerably larger, in particular of those tributaries where no observations are available. It is interesting that the uncertainty also increases substantially with decreasing catchment area. For some of the smallest catchments, error standard deviations of more than

$8 \text{ l s}^{-1} \text{ km}^{-2}$ are estimated. The higher uncertainty for the smallest catchments is very realistic, because catchments of such size are underrepresented in the sample. Hence, their prediction will be subject to additional downscaling errors. Because of the tree structure of river networks, scale issues constitute a typical problem in river network modelling (Blöschl and Sivapalan, 1995; Skøien and Blöschl, 2006). Top-kriging explicitly takes scaling into account, since it is actually one solution to the traditional change of support problem in geostatistics (Gotway and Young, 2002). Top-kriging is therefore the natural way of predicting stream flow-related variables on a river network. Regression, however, ignores the differences of the support along the river network.

Cross-validation

In order to examine the relative performances of models more quantitatively, we performed leave-one-out cross-validation where one withholds the stream flow observation of a particular data point, makes an estimate for that location, and then compares the estimate with the stream flow observation, repeating the procedure for all data points. This procedure emulates the case of estimating at sites without stream flow observation. For top-kriging, we used the

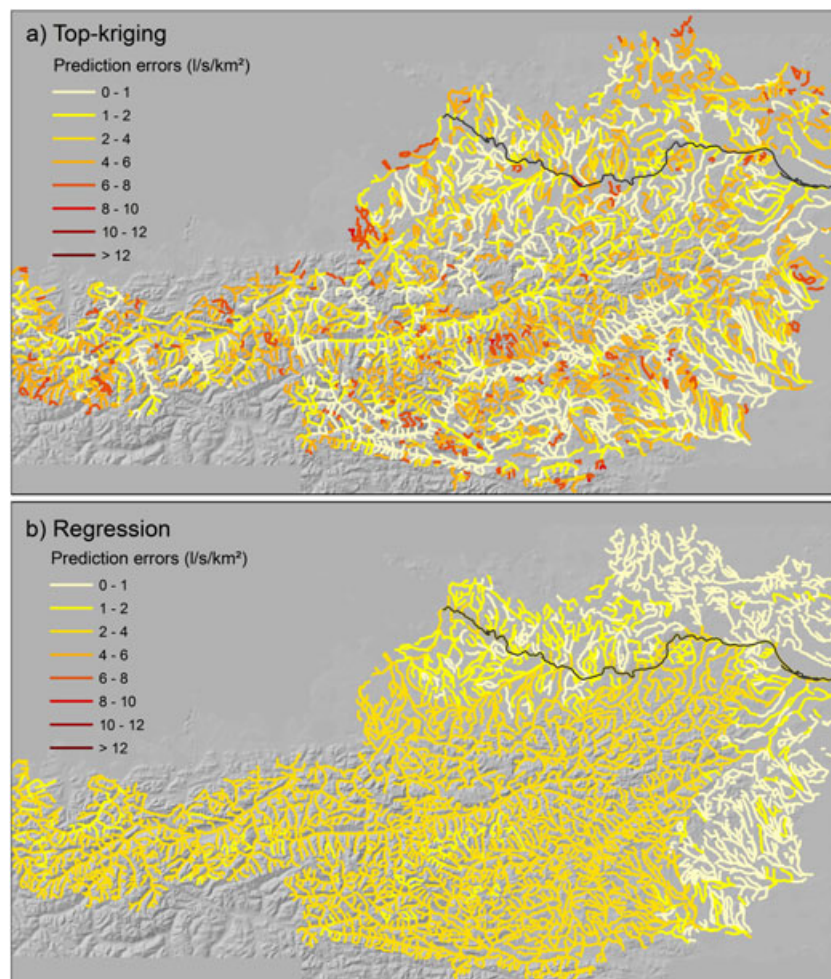


Figure 6. Error standard deviation of spatial estimates of low stream flows q_{95} ($\text{l s}^{-1} \text{ km}^{-2}$)

variogram model obtained from all n observations to predict by kriging each point in turn using the remaining $n-1$ observations (e.g. see Cressie, 1993, p102). The approach therefore assumes a known variogram model. For regional regression, the model consists of individual regressions for contiguous sub-regions (section 3.3). Because of the large number of observations, the boundaries will not change much when leaving out single data points, so we did not change the regions in the cross-validation. For each point in turn, we updated the regression equation for remaining $n-1$ data points, and the so-obtained regression model was used to predict low flows at the site of interest. From the resulting vector of cv -errors of each model, the cross-validation error variance V_{cv} , the root mean squared error $rmse_{cv}$, the coefficient of determination R_{cv}^2 , and the bias of cross-validation estimates $bias_{cv}$ were estimated:

$$V_{cv} = \frac{1}{n} \sum_{i=1}^n (e_{cv,i})^2 \tag{7}$$

$$rmse_{cv} = \sqrt{V_{cv}} \tag{8}$$

$$R_{cv}^2 = \frac{V_q - V_{cv}}{V_q} \tag{9}$$

$$bias_{cv} = \frac{1}{n} \sum_{i=1}^n (e_{cv,i}) \tag{10}$$

In these equations, $e_{cv,i}$ is the cv -residual of a model for catchment i , and V_q is the spatial variance of the observed specific low flow discharges q_{95} . The root mean squared error $rmse_{cv}$ and the coefficient of determination R_{cv}^2 provide composite measures of systematic and random errors, whereas the of cross-validation estimates $bias_{cv}$ provides a measure of systematic errors only. In addition, relative error measures $rmse_{cv} = rmse_{cv} / \text{mean}(q_{95})$ and $rbias_{cv} = bias_{cv} / \text{mean}(q_{95})$ were calculated.

The results of the cross-validation are shown in Figure 7 in terms of the error distribution for the set of 300 catchments, and residual statistics are presented in Table II. All residual statistics (R_{cv}^2 , $rmse_{cv}$, $bias_{cv}$, etc.) were estimated without using the 5% outliers for robustness. These outliers

can be explained by karst effects or seepage and are hence not genuinely attributable to regionalisation errors. Figure 7, however, does include these outliers. The results indicate that top-kriging ($R_{cv}^2 = 0.75$, $rmse_{cv} = 1.781 \text{ ls}^{-1} \text{ km}^{-2}$) clearly outperforms regression, which gives R_{cv}^2 of 0.68 and $rmse_{cv}$ of $1.999 \text{ ls}^{-1} \text{ km}^{-2}$. The regression estimates are nearly unbiased ($bias_{cv} = -0.032 \text{ ls}^{-1} \text{ km}^{-2}$), while top-kriging slightly overestimates low stream flows ($bias_{cv} = 0.294 \text{ ls}^{-1} \text{ km}^{-2}$, $rbias_{cv} = 0.05$). The possible reasons will be assessed through regional examples in the next section.

Regional performances

The first example is a typical situation of Tyrolean High-Alps. Catchments in this region are small, and many of the catchments are headwater catchment, i.e. catchments without upstream data point. Figure 8 shows the estimates (left panels) and error standard deviations (right panels) of top-kriging and regression plotted along the stream network. The observations are shown as circles, using the same colour coding as for the estimates. The results of both models differ substantially. Close to a gauge, top-kriging estimates fit very well to the observations, and kriging errors are of the same magnitude as measurement errors. The prediction errors of regression, however, are much larger than measurement errors, indicating a lower performance of regression in close proximity to gauges. The relative performances of models change when moving from the gauges to the headwater catchments (e.g. region indicated by the ellipses in Figure 8). Low flow discharges of Alpine catchments are expected to decrease with catchment altitude, as they are controlled by freezing processes which are more intense for higher altitudes. The expected low flow patterns are fairly well reproduced by regression, as altitude is parameterised in the regression model. Top-kriging, however, overestimates these values. The reason is that top-kriging distributes weights according to integral distances of catchment areas and therefore gives most weight to neighbouring gauges at the same river. Sources therefore correspond to a boundary area of top-kriging, and effects similar to the well-known boundary effects of ordinary-kriging occur. The boundary effects are reflected in top-kriging standard deviations (Figure 8c) which increase when moving from valleys to the

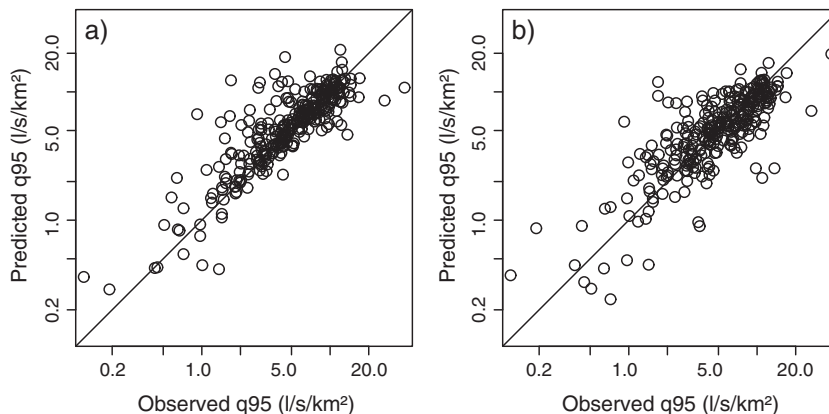


Figure 7. Scatter plots of predicted versus observed specific low stream flows q_{95} ($\text{ls}^{-1} \text{ km}^{-2}$) of (a) top-kriging and (b) regression

Table II. Predictive performance of top-kriging (TK) and regional regression (Reg) for different data situations (all catchments, headwater catchments, non-headwater catchments)

Situation	All catchments		Headwater		Non-headwater	
	TK	Reg	TK	Reg	TK	Reg
R_{cv}^2 (-)	0.75	0.68	0.59	0.56	0.91	0.82
$rmse_{cv}$ ($l s^{-1} km^{-2}$)	1.781	1.999	2.355	2.427	0.951	1.392
$bias_{cv}$ ($l s^{-1} km^{-2}$)	0.294	-0.032	0.558	0.119	0.036	-0.167
$rrmse_{cv}$ (-)	0.31	0.34	0.40	0.41	0.16	0.24
$rbias_{cv}$ (-)	0.05	-0.01	0.10	0.02	0.01	-0.03

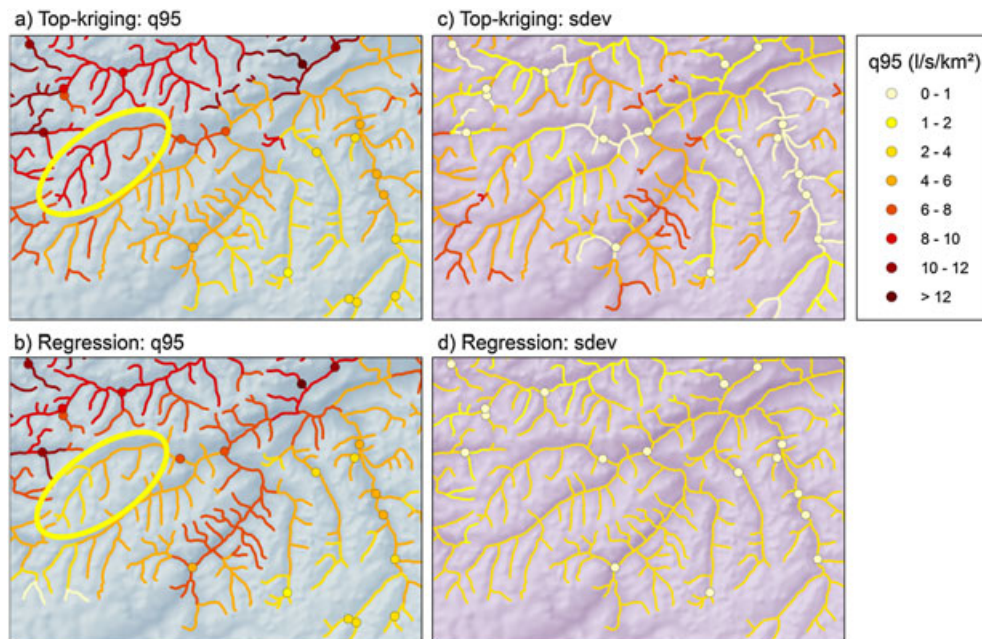


Figure 8. Regional example Tyrol: q95 low flow predictions in ungauged basins by top-kriging (a) and regression (b) for an Alpine region in Austria. (c) and (d) indicate error standard deviations. The area shown is 100 km across. Ellipses indicate a region where top-kriging extrapolates from gauges situated in the valleys to headwater catchments (without upstream data point) leading to systematic overestimation

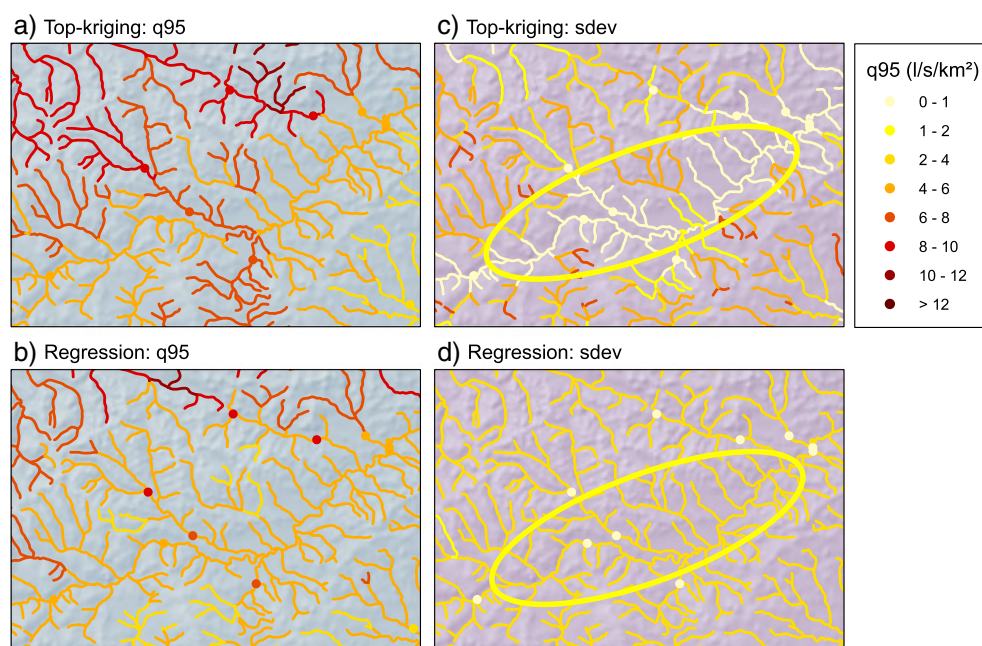


Figure 9. Regional example river Mur: q95 low flow predictions in ungauged basins by top-kriging (a) and regression (b) for a major river in Austria. (c) and (d) indicate error standard deviations. The area shown is 100 km across. For large rivers and high gauging density (see ellipsis) top-kriging outperforms regression

sources. In this situation, top-kriging extrapolates the higher values from gauges in the valleys to the headwaters.

The second example is the river Mur and tributaries in low Alps of Styria, southern Austria. The river Mur is one of the major rivers in Austria, and river sites are related to large catchment areas (up to 4400 km²). Estimates and error standard deviations of top-kriging and regression are presented in Figure 9. Along the main river, top-kriging and regression estimates are consistent with observations. The estimates of top-kriging are also consistent with observations for the larger tributaries, but this is clearly not true for regression. For locations without upstream data point, however, top-kriging yields a rather arbitrary pattern, while regression estimates show the expected decrease with catchment altitude. The relative performances of both methods are manifested in the error standard deviations (Figure 9c,d). The errors of top-kriging are low for large rivers and a high gauging density, and the errors increase in the tributaries with decreasing catchment area. The errors of regression are nearly constant in this region. They do not seem to contain a lot of information about the regional performance of the model.

To generalise the findings from the regional examples to the whole study area, we conducted a stratified cross-validation analysis where catchments were separated in headwater catchments (without upstream gauge) and non-headwater catchments (with upstream gauge). For the Austrian setting, headwater catchments typically correspond to small catchments situated in Alpine areas (example 1), and non-headwater catchments correspond to larger catchments (example 2).

Table II indicates that for non-headwater catchments, top-kriging is nearly unbiased ($\text{bias}_{\text{cv}}=0.036 \text{ ls}^{-1} \text{ km}^{-2}$) and regression yields a slightly negative bias of $-0.167 \text{ ls}^{-1} \text{ km}^{-2}$. These biases are small compared to observed low flows in Austria, and systematic errors are a minor problem for the relatively large non-headwater catchments. For headwater catchments, the systematic error of regression is also low ($\text{bias}_{\text{cv}}=0.119 \text{ ls}^{-1} \text{ km}^{-2}$), but top-kriging exhibits somewhat larger systematic errors ($\text{bias}_{\text{cv}}=0.558 \text{ ls}^{-1} \text{ km}^{-2}$), corresponding to an average overestimation of 10%. As it has been discussed above, low flow discharges of Alpine catchments are expected to decrease with catchment altitude, and using spatial information from a downstream gauge is clearly a biased procedure. For top-kriging, the altitude gradient of low flows may result in non-stationary increments which violate the assumption of intrinsic stationarity. Hence, top-kriging estimates could be further improved by including the altitude gradient in the model, by an approach similar to external drift kriging (Wackernagel, 1995). This will likely reduce the bias of headwater catchments and increase the predictive performance of top-kriging over the regression model. Laaha *et al.* (2012) tested top-kriging with external drift to water temperatures in Austria. The approach assumes a deterministic relationship between water temperature and catchment altitude, through performing top-kriging on the residuals of an initial regression model. Their analysis showed that the model with

external drift was indeed free of regional biases. From the assumptions, the approach seems notably well suited when the auxiliary variable has a much lower measurement error than the target variable. Otherwise, a combined co-kriging and top-kriging approach would be better suited to include auxiliary information in the top-kriging model. This, however, would require a more complex algorithm which includes cross-variograms in all steps of the analysis.

We finally assess the total predictive performance of models. For headwater catchments, top-kriging ($R_{\text{cv}}^2=0.59$, $\text{rmse}_{\text{cv}}=2.355 \text{ ls}^{-1} \text{ km}^{-2}$) exhibits somewhat higher coefficient of determination and a somewhat lower root mean squared error than regression ($R_{\text{cv}}^2=0.56$, $\text{rmse}_{\text{cv}}=2.427 \text{ ls}^{-1} \text{ km}^{-2}$). For the non-headwater catchments, the predictive performances are generally higher than for headwater catchments. Here, top-kriging ($R_{\text{cv}}^2=0.91$, $\text{rmse}_{\text{cv}}=0.911 \text{ ls}^{-1} \text{ km}^{-2}$) clearly outperforms regression ($R_{\text{cv}}^2=0.82$, $\text{rmse}_{\text{cv}}=1.392 \text{ ls}^{-1} \text{ km}^{-2}$). The cross-validation analysis confirms the findings from the regional examples that top-kriging performs clearly better for the larger non-headwater catchments than regression. For smaller catchments in the tributaries and headwater catchments, the performance of top-kriging is indeed reduced because of the bias, but, nevertheless, still higher than the performance of the regression model. These findings correspond well with the results of Castiglioni *et al.* (2011) where top-kriging was compared with a smooth regional estimation method, known as geographical space-based interpolation (PSBI) (Castiglioni *et al.*, 2009) for the case of predicting low flows of 51 catchments in central Italy. Overall, top-kriging and PSBI have comparable performances (Nash–Sutcliffe efficiencies in cross-validation of 0.89 and 0.83, respectively), but top-kriging outperforms PSBI at larger river branches while PSBI performs better for headwater catchments. The main idea of PSBI is to apply standard point-kriging methods in a transformed space of catchment characteristics instead of kriging in geographical space. Consequently, PSBI and regression models exploit the same kind of information, i.e. the catchment characteristics, and this explains the similar features of both concepts. The correlations along the stream network are accounted for only indirectly through the catchment characteristics. However, top-kriging exploits the spatial correlations of low flows in an explicit way. Top-kriging is a best linear unbiased geostatistical estimation method which takes the specific structure of river networks into account.

CONCLUSIONS

We compared the performance of top-kriging relative to the regional regression model proposed by Laaha and Blöschl (2006b) by leave-one-out cross-validation. Regional regression is the standard regionalisation approach in low flow hydrology and constitutes a benchmark for any innovative model. On average, over the Austrian study area, top-kriging explains 75% of the variance of the low flows, while the regression models explain 68%. The performance of top-kriging mainly depends on the (intrinsic) homogeneity of

the observations and the density of the gauging network in the region. One would expect the performance of top-kriging to increase with increasing density of the network and increasing catchment size. The latter is because runoff is an integrating process, so low flows tend to vary more smoothly along the stream for large catchments. The analyses of the data set in this paper indicate that this is indeed the case. Table II suggests that top-kriging explains 91% and 59% of the variances in the non-headwater and headwater catchments, respectively. The median catchment sizes of these two groups of catchments are 252 and 62 km². The performance of the regression model, in contrast, hinges on the availability of meaningful catchment characteristics used in the regression model and the degree of correlation of these catchment characteristics with the low flows. Low flow generation is mainly governed by subsurface processes which are not accessible by area-wide data collection techniques, such as remote sensing. This makes it difficult to collect meaningful catchment characteristics. The choice between top-kriging and the regression model should therefore mainly depend on data availability and the characteristics of the observations, as the two methods use different sources of information. The data characteristics of this case study give some guidance on the choice of method: For top-kriging, the network density is relevant (490 observation points over 84 000 km²). For the regression models, the standard error is relevant (0.4 to 2.5 ls⁻¹ km⁻², depending on the region).

From the conceptualisation of stream flow generation processes, the distribution of kriging weights, the predictive performance, and the distribution of uncertainty along the stream network, top-kriging appears notably well adapted for the stream network problem. Top-kriging is easy to apply as it does not require additional auxiliary data. This is a major advantage over the classical regression approach if the auxiliary data are not well correlated with the variable of interest or unavailable. Also, top-kriging can be extended to combine the strengths of regional regression and top-kriging which is a topic we are currently exploring. Top-kriging also offers advantages over one-dimensional conceptualisations of correlations along the stream network, as there is no need for a decision whether to estimate the variable from upstream or downstream neighbours, or a combination thereof. We therefore suggest that top-kriging is the most natural method of spatial predictions on river networks.

ACKNOWLEDGEMENTS

The authors would like to thank the Austrian Science Foundation (FWF), project no. P18993-N10, and the Austrian Academy of Sciences 'Hydrological Predictability' project for financial support. The paper is a contribution to UNESCO's FRIEND-Water program. Comments from two anonymous reviewers contributed to the improvement of this paper.

REFERENCES

- Aschwanden H, Kan C. 1999. Die Abflussmenge Q347 - eine Standortbestimmung. *Hydrologische Mitteilungen* **27**. Landeshydrologie und -geologie, Bern.
- Berg C, Forst G. 1975. *Potential theory on locally compact Abelian groups*. Springer: Berlin.
- Blöschl G, Sivapalan M. 1995. Scale issues in hydrological modelling - a review. *Hydrological Processes* **9**: 251–290.
- Castiglioni S, Castellarin A, Montanari A. 2009. Prediction of low-flow indices in ungauged basin through physiographical space-based interpolation. *Journal of Hydrology* **378**: 272–280.
- Castiglioni S, Castellarin A, Montanari A, Sköien JO, Laaha G, Blöschl G. 2011. Smooth regional estimation of low-flow indices: physiographical space based interpolation and top-kriging. *Hydrology and Earth System Sciences* **15**: 715–727. DOI: 10.5194/hess-15-2011.
- Cressie N. 1985. Fitting variogram models by weighted least squares. *Mathematical Geology* **17**: 563–586.
- Cressie N. 1993. *Statistics for spatial data*, revised edn. Wiley: New York, NY.
- Cressie N, Frey J, Harch B, Smith M. 2006. Spatial prediction on a river network. *Journal of Agricultural, Biological, and Environmental Statistics* **11**: 127–150.
- Drajer NR, Smith H. 1998. *Applied Regression Analysis*, 3rd edn. Wiley: New York, NY.
- Garreta V, Monestiez P, Ver Hoef JM. 2009. Spatial modelling and prediction on river networks: up model, down model or hybrid? *Environmetrics* **21**: 439–456. DOI: 10.1002/env.995.
- Gottschalk L. 1993a. Correlation and covariance of runoff. *Stochastic Hydrology and Hydraulics* **7**: 85–101.
- Gottschalk L, Krasovskaia I, Leblois E, Sauquet E. 2006. Mapping mean and variance of runoff in a river basin. *Hydrology and Earth System Sciences* **10**: 469–484.
- Gotway CA, Young LJ. 2002. Combining Incompatible Spatial Data. *Journal of the American Statistical Association* **97**: 632–648.
- Gustard A, Bullock A, Dixon JM. 1992. *Low flow estimation in the United Kingdom*. Institute of Hydrology, Report No. 108, 88pp, append.
- Kyriakidis PC. 2004. A geostatistical framework for area-to-point spatial interpolation. *Geographical Analysis* **36**: 259–289.
- Laaha G, Blöschl G. 2005. Low flow estimates from short stream flow records—a comparison of methods. *Journal of Hydrology* **306**(1/4): 264–286.
- Laaha G, Blöschl G. 2006a. Seasonality indices for regionalizing low flows. *Hydrological Processes* **20**: 3851–3878.
- Laaha G, Blöschl G. 2006b. A comparison of low flow regionalisation methods—catchment grouping. *Journal of Hydrology* **323**(1/4): 193–214.
- Laaha G, Blöschl G. 2007. A national low flow estimation procedure for Austria. *Hydrological Sciences Journal* **52**(4): 625–644.
- Laaha G, Sköien JO, Blöschl G. 2012. External top-kriging of stream temperature in Austria. *Environmental Modeling and Assessment*, under review.
- de Marsily G. 1986. *Quantitative hydrogeology*. Academic Press: London, UK.
- Matheron G. 1965. *Les variables régionalisées et leur estimation*. Masson: Paris, France.
- Merz R, Blöschl G. 2005. Flood frequency regionalisation - Spatial proximity vs catchment attributes. *Journal of Hydrology* **302**: 283–306.
- Mockus A. 1998. Estimating dependencies from spatial averages. *Journal of Computational and Graphical Statistics* **7**: 501–513.
- Nathan RJ, McMahon TA. 1990. Identification of homogeneous regions for the purpose of regionalization. *Journal of Hydrology* **121**: 217–238.
- Peterson EE, Ver Hoef JM. 2010. A mixed-model moving-average approach to geostatistical modelling in stream networks. *Ecology* **91**(3): 644–651.
- Sauquet E, Gottschalk L, Leblois E. 2000. Mapping average annual runoff: a hierarchical approach applying a stochastic interpolation scheme. *Hydrological Sciences Journal* **45**: 799–815.
- Sköien JO, Blöschl G. 2006. Sampling scale effects in random fields and implications for environmental monitoring. *Environmental Monitoring and Assessment* **114**(1-3): 521–552.
- Sköien JO, Blöschl G, Western AW. 2003. Characteristic space-time scales in hydrology. *Water Resources Research* **39**: 1304–1323.
- Sköien JO, Merz R, Blöschl G. 2006. Top-kriging - geostatistics on stream networks. *Hydrology and Earth System Sciences* **10**: 277–287.
- Ver Hoef JM, Peterson E, Theobald D. 2006. Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics* **13**: 449–464.
- Wackernagel H. 1995. *Multivariate Geostatistics*. Springer: Berlin Heidelberg, Germany.
- Weisberg S. 1985. *Applied Linear Regression*, 2nd edn. Wiley: New York, NY.