

Spatial Prediction of Stream Temperatures Using Top-Kriging with an External Drift

Gregor Laaha · Jon O. Skøien · Franz Nobilis · Günter Blöschl

Received: 20 May 2011 / Accepted: 5 May 2013 / Published online: 2 June 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Top-kriging is a method for estimating stream flow and 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 a family of variogram models (regularisations) for different catchment areas (kriging support), which accounts for the different scales and the nested nature of the catchments. This assures that kriging weights are distributed to both hydrologically connected and unconnected sites of the stream network according to the data situation: top-kriging gives most weight to close-by sites at the same river system, but when the next hydrologically connected site is far away, more weight is given to a close-by site at an adjacent river system. The distribution of weights is in contrast to ordinary kriging and stream distance-based kriging which does not account for both spatial proximity and network connectivity. We extend the top-kriging method by incorporating an external

drift function to account for the deterministic patterns of the spatial variable. We test the method for a comprehensive Austrian stream temperature dataset. The drift is modelled by exponential regression with catchment altitude. Top-kriging is then applied to the regression residuals. The variogram used in top-kriging is fitted by a semiautomatic optimisation procedure. A leave-one-out cross-validation analysis shows that the model performs well for the study domain. The residual mean squared error (cross-validation) decreases by 20 % when using top-kriging in addition to the regression model. For regions where the observed stream temperatures deviate from the expected value of the drift model, top-kriging corrects these regional biases. By exploiting the topological information of the stream network, top-kriging is able to improve the local adjustment of the drift model for the main streams and the tributaries.

Keywords Block kriging · Change of support · Geostatistics · Regionalisation · Spatial interpolation · Stream distance

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

G. Laaha · G. Blöschl
Institute for Hydraulic and Water Resources Engineering (IWI),
Vienna University of Technology, Vienna, Austria

J. O. Skøien
Department of Physical Geography, University of Utrecht,
Utrecht, the Netherlands

J. O. Skøien
Institute for Environment and Sustainability, Joint Research Centre
of the European Commission, Ispra, Italy

F. Nobilis
Institute of Meteorology and Geophysics, University of Vienna,
Vienna, Austria

1 Introduction

Stream temperature is an important physical property of flowing waters. This is because of its enormous significance for all freshwater organisms, its influence on other aspects of water quality and pollution, its economic importance, and its sensitivity to both natural and human factors. Temperature has, traditionally, received less attention than other facets of water quality. However, recent years have witnessed a renewed interest in the thermal behaviour of flowing waters [3, 31].

The thermal regime of a river is usually characterised by variables such as mean annual, seasonal or monthly stream temperature. These variables are best estimated from observed stream temperature data, but for sites where these

data are unavailable, interpolation techniques can be used to infer them from other sites where stream temperature data have been collected. Statistical interpolation is usually based on some kind of correlation of the stream flow variable of interest. Multivariate methods exploit the information of auxiliary variables representing the physical or geographical properties of river locations, e.g. [17–19], whereas geostatistical methods make use of the spatial autocorrelation of the stream variable [11, 12]. A recent review of stream temperature modelling showed that for the larger (weekly to annual) timescale, regressions with air temperature or catchment altitude are the state of the art [1].

Surprisingly, little effort has been undertaken to apply geostatistics to stream temperature. This is most likely because of the tree-like structure of river networks, which poses specific challenges for geostatistical regionalisation [20]. Catchments related to points of the river network are organised into sub-catchments, i.e. they are nested. It is therefore clear that upstream and downstream catchments would have to be treated differently from neighbouring catchments that do not share a sub-catchment. The estimation of variables on stream networks needs to use a topology that is different from the Euclidian topology of geostatistical standard techniques. Most applications of geostatistics to river networks, so far, have indeed used Euclidian distance between catchments, usually measured as the Euclidian distance between the gauges or the catchment centroids (e.g. [32]). Gardner et al. [8] assessed the performance of various geostatistical metrics for predicting stream temperature on a river network. The study showed that kriging based on a one-dimensional conceptualisation of the river network which uses distances between sampling locations calculated along the stream network provided better predictions than kriging based on Euclidean distance between sampling locations. This is in line with the basic assumptions of the one-dimensional river network models of Gottschalk [11, 12], Ver Hoef et al. [29], Cressie et al. [6], and the more recent works of Peterson et al. [23], Garreta et al. [9], and Peterson and Ver Hoef [24] which apply kriging based on stream distance for interpolation. Ver Hoef et al. [26] developed valid covariance functions for river networks which either take upstream or downstream neighbours into account. Garreta et al. [9] evaluated the two model types in the context of summer stream temperature and found a combined model to be superior to each of the individual models. Wehrly et al. [32], however, pointed out that kriging based on stream distance provides no or only little kriging weight to neighbouring sites at an adjacent but hydrologically unconnected stream system (such as the situation shown in Fig. 1), even though these sites may also contain relevant information. In particular, kriging based on stream distance is not well suited for headwater catchments (small catchments at the source region) in sparsely gauged stream

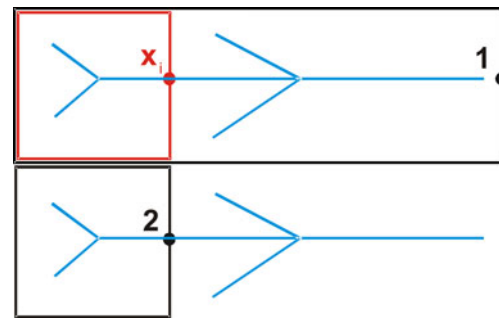


Fig. 1 Connected (1) and unconnected sites (2) of the stream network. x_i is the target site (red point) where an estimation is sought together with its catchment (red box); black points indicate the stream gauges together with their catchments (black boxes)

networks. Wehrly et al. [32] therefore have chosen to use ordinary kriging based on Euclidean distances between sampling locations rather than on stream distances. They found, however, that the performance of the model was still relatively modest and did not differ much from other simpler interpolation methods.

A two-dimensional geostatistical river network model might be more appropriate. This class of models assumes stream flow to be generated by a spatially continuous process which exists at any point in the landscape, and stream flow measurements represent an aggregate (*integral*) of point runoff over the catchment. In geostatistical jargon, the catchment is called the *support* of the measurement, which is not point-wise like for most other environmental variables, but an irregularly shaped area. As a consequence of runoff aggregation, the variogram (i.e. dissimilarity of the variable between catchments) decreases as the catchment area increases. This means that the point variogram used in standard kriging interpolation is not valid for catchments because of their different scales. It needs to be aggregated over the catchment, just the same as the variable of interest. The calculation of the aggregated variogram for catchments (or, more general, for areal support) from a point variogram is termed *regularisation*. The resulting kriging model for variables with areal support based on a regularised variogram is commonly called *block-kriging*.

Sauquet et al. [25] and Gottschalk et al. [13] were probably the first who proposed a block-kriging method where the spatial dependence of catchments with different, non-zero support is modelled by a regularised covariogram. Skøien et al. [28] extends the work of Sauquet et al. [25] to account for the stronger spatial correlation between nested basins than between un-nested basins. They suggested 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, as a function of drainage area, and are non-stationary. Top-kriging is based on the assumption of intrinsic stationarity, meaning that the dissimilarities

of the variable between catchments are only a function of distance and remain the same for any site of the study area. This assumption will not hold for stream temperature since the heat budget of a stream is strongly correlated with catchment altitude. Thus, some adaptation of the geostatistical model is needed in order to apply top-kriging to mixed random and deterministic fields. The approach proposed in this paper extends top-kriging with a deterministic drift model which exploits the information content of auxiliary variables.

The aim of this paper is fourfold: (1) to present top-kriging with external drift as a suitable method to perform spatial prediction on a river network, (2) to test top-kriging with external drift in the context of stream temperature and (3) assess its performance relative to a simple nonlinear regression model, and (4) to illustrate the differences in kriging weights among the various kriging methods. The analyses will be performed on a comprehensive Austrian dataset.

2 Top-Kriging Method

2.1 Concept of Top-Kriging

There are two main groups of forcings that control stream flow. The first group consists of variables that are continuous in space such as rainfall, evapotranspiration and soil characteristics, which are related to local runoff generation. The second group is related to runoff aggregation and routing along the stream network. The main idea of top-kriging is to combine the two groups of forcings in a geostatistical framework that takes the river network topology into account. For this, runoff generation is conceptualised as a spatially continuous process which exists at any point in the landscape. Instead of observing a point-wise realisation of this process, data $z(A_1)$, $z(A_2)$, ..., $z(A_n)$ are collected, 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. In this context, the interpolation between river locations adds up to the classical change of support problem in geostatistics (e.g. [14]). This is illustrated in Fig. 2 which shows a schematic of two river locations, i (upstream) and j (downstream), with catchment area, A_i , nested in catchment, A_j . The same principle applies for the case of river locations at unconnected rivers (i.e. non-nested catchments). The sole difference is

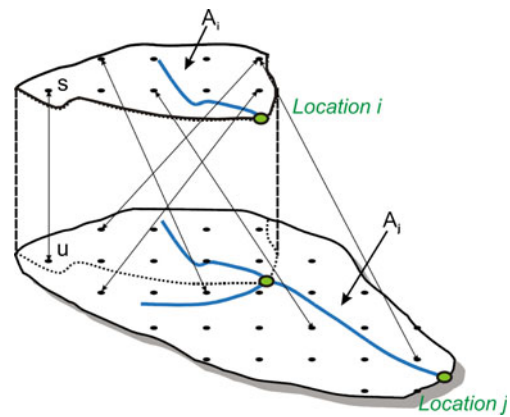


Fig. 2 Schematic of two nested catchments and their discretisation. Redrawn from [28]

that the distance between the connected catchments is generally smaller than for unconnected catchments due to the overlapping support of connected river locations.

For spatial prediction on a river location, \mathbf{x}_0 , with catchment area, A_0 , from non-point samples $z(A_1)$, $z(A_2)$, ..., $z(A_n)$, 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

$$\sum_{j=1}^n \lambda_j \bar{\gamma}_{ij} - \lambda_i \sigma_i^2 + \mu = \bar{\gamma}_{0i} \quad i = 1, \dots, n$$

$$\sum_{i=1}^n \lambda_i = 1 \quad (3)$$

These equations result from minimizing prediction mean squared error subject to unbiasedness constraints. $\bar{\gamma}_{ij}$ refers to the variogram model value for two measurements, i and j , with non-zero support (see below), $\bar{\gamma}_{0i}$ is the variogram model value between estimation point 0 and measurement i , μ is the Lagrange multiplier from the constraint minimisation, and σ_i^2 represents the measurement error or uncertainty of measurement i . The use of measurement errors in the kriging equations is termed kriging with uncertain data [7] (p. 300).

Since the measurements have a non-zero support A , the variogram model value $\bar{\gamma}_{ij}$ for the measurements needs to be obtained by regularisation [4] (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 variogram γ_p , the variogram model

value $\bar{\gamma}_{ij}$ between two measurements with catchment areas A_i and A_j , respectively, is

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

where \mathbf{s} and \mathbf{u} are position vectors within each catchment used for the integration. In top-kriging, the integrals in Eq. 4 are computed by discretizing the catchment area into a grid of points, according to common geostatistical practice [15]. This is illustrated in Fig. 2 for the case of two nested catchments, where non-bold arrows symbolize the separation vectors between discretization points of catchments i and j . The first part of Eq. 4 integrates all the variance between the grid points of the two catchments, whilst the second part subtracts the variance between the grid points within the catchments. The resulting $\bar{\gamma}_{ij}$ of the two catchments increase with increasing distance (general property of the variogram) and decrease with catchment area (because of the smoothing effect of the support). Consequently, $\bar{\gamma}_{ij}$ will be lowest for close-by locations at the same river as they are separated by a smaller distance than the same-sized catchments at different rivers.

2.2 Extending Top-Kriging with an External Drift Function

Top-kriging assumes that the spatial variable is the outcome of a translation-invariant random process. There are, however, cases where the observed patterns of the variable reveal systematic deviations from a purely stochastic field due to the influence of some deterministic process acting over the domain. Stream temperature is a typical example since the spatial variation of climatic parameters, such as elevation, affects the local values. Hence, some adaptation of the model is in order, and we propose an approach similar to external drift kriging (e.g. [30]) to model non-stationarity. We term the extended model top-kriging with external drift because of this analogy.

Top-kriging with external drift uses a decomposition of the spatial variable $Z(\mathbf{x})$ into deterministic and stochastic components to account for systematic patterns of the spatial variable $Z(\mathbf{x})$.

$$Z(\mathbf{x}) = m(\mathbf{x}) + Y(\mathbf{x}) \quad (5)$$

The deterministic component $m(\mathbf{x})$ is called the drift. It is modelled by an external drift function which we estimate by regression analysis. $Y(\mathbf{x})$ are the residuals of the external drift function, which are assumed an intrinsic stationary stochastic variable. They are estimated by the top-kriging

model based on the residual variogram. The superposition of the estimates of both components according to Eq. 5 yields the prediction of $Z(\mathbf{x})$.

3 Application to Stream Temperature

3.1 Dataset

The dataset used in this study consists of about 8,000 catchments in Austria. For all these catchments, the stream network and catchment boundaries were available. For a subset of 214 catchments, temperature measurements for the period 1976–2000 were provided by the Austrian Hydrographical Service [10]. The target variable of this study is the mean annual stream temperature, T (in degree Celsius). Figure 3 shows the spatial distribution of the observed T in Austria. It is subject to strong spatial variability. The values range from 2.2 °C in the western High Alps to 13.6 °C in the eastern lowlands; the average temperature is 8.0 °C and the temperature standard deviation is 1.9 °C. The observed patterns are apparently correlated with catchment altitude.

3.2 External Drift Model

Godina et al. [10] investigated the thermal regime of streams in Austria and found a strong, nonlinear relationship between mean annual stream temperature, T (in degree Celsius), and altitude of the stream gauge, H (in metres above sea level). On this theoretical basis, we fitted an exponential regression model to the dataset. The model exhibits the following prediction equation:

$$m^*(\mathbf{x}) = 11.487e^{-0.0008H} \quad (6)$$

Figure 4 shows that the model fits well to the observed relationship between T and H . Hence, it was used as the external drift function in subsequent analysis. The residuals of the model are

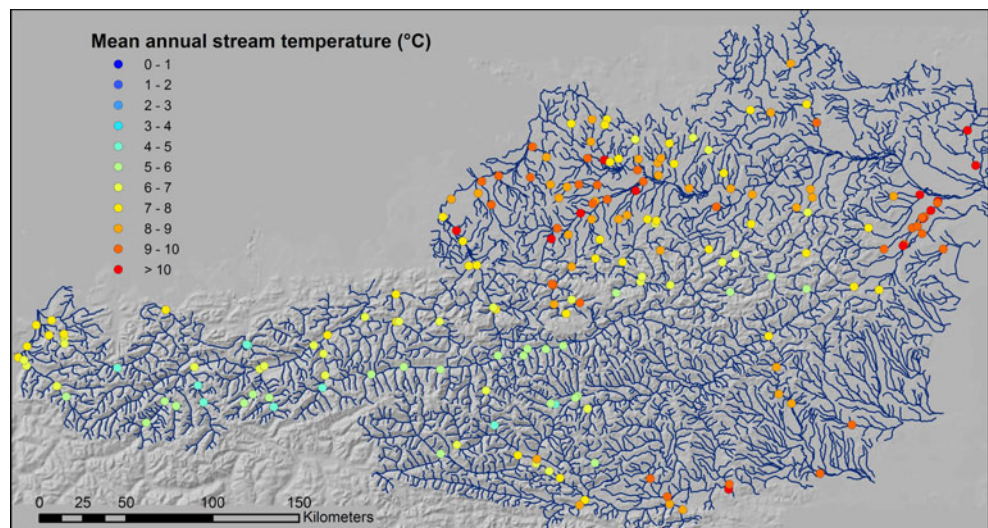
$$y^*(\mathbf{x}) = z(\mathbf{x}) - m^*(\mathbf{x}) \quad (7)$$

The spatial structure of residuals is exploited by the stochastic residual model in the second step of the analysis.

3.3 Variogram of the Residuals

For the stochastic residual model, the point variogram of the regression residuals $y^*(\mathbf{x})$ is needed. This variogram cannot be directly fitted to the regression residuals because of their different scale. Rather, it can be determined by the back-calculation approach of Kyriakidis [16] and Mockus [22]. In

Fig. 3 Situation of stream gauges used in this study. Colours indicate the magnitude of mean annual stream temperature calculated from the observed stream flow records ($0-1$ denotes $0 \leq T < 1$). The area shown is Austria, which is 600 km across



this approach, a number of theoretical point variograms with different parameter sets are assumed. Each point variogram model is regularised for all pairs of observations using Eq. 4 in order to find the point variogram with fits best to sample variogram values. Following Skøien et al. [27], we used a point variogram model with a nugget effect of the following shape:

$$\gamma_p(h) = ah^b \left(1 - e^{-\left(h/c \right)^d} \right) + C_{0p} \quad (8)$$

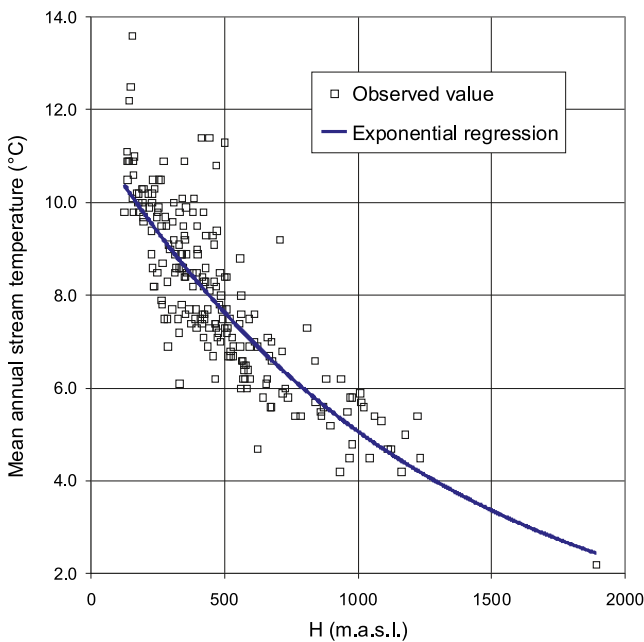


Fig. 4 Exponential regression model between the mean annual stream temperature (in degree Celsius) and altitude of the gauge (in metres above sea level) used as external drift function, fitted to 214 catchments

h is the separation distance between two measurements; a, b, c and d are parameters. a is related to the sill of the variogram, c is a correlation length, whilst b and d define the long and short distance slope of the variogram in a log-log plot, respectively, and C_{0p} is the point nugget effect. γ_p was optimised to give the best fit to the sample variogram [15, 21]. According to Cressie [4, 5], the sum of the weighted squared residuals between the regularised theoretical semi-variances and the observed semi-variances was used as an optimality criterion. We estimated a sample variogram with the pairs grouped into bins, similar to Matheron's [21] traditional estimator, but with three dimensions: the distance between centroids (h), the area of the smaller catchment (A_i) and the area of the bigger catchment (A_j). From the automatic fitting procedure, we obtained a preliminary point variogram model with parameters $a=0.339, b=0.001, c=16.359, d=1.698$ and $C_{0p}=0.100$.

Parameter optimisation is frequently complicated by a non-monotonic optimality function which leads to a multi-solution problem, where the local optimum depends on the initial parameter setting. We therefore suggest in this paper a semiautomatic fitting procedure where we use a lumped variogram (i.e. a conventional variogram where all catchments areas are in one class) as guidance for variogram fitting. Although the lumped variogram does not account for the different scales of catchments and, therefore, will not be genuinely suitable to describe the spatial structure of observations with different support, we expect that it will reflect most closely the structure of some average, or most frequent, support size as these observations will dominate the experimental variogram. The similarity between the regularisation for the dominant support size and the lumped variogram is therefore a useful indicator for a suitable model.

We tested this semiautomatic variogram fitting procedure for the residuals of the stream temperature regression model. The lumped variogram of the residuals $\gamma^*(\mathbf{x})$ is shown in Fig. 5a. Its shape should be most representative for the median catchment size of about 425 km². By varying the initial parameter settings, we obtained a theoretical variogram model whose regularisation for the median support size (represented by the 300 vs. 300 km² line) fits well to the structure of the lumped variogram (Fig. 5b). The parameter estimates of the model are $a=30.087$, $b=0.00009$, $c=15.125$, $d=0.011$ and $C_{op}=38.947$. A preliminary cross-validation analysis yielded that the model based on analogy to the lumped variogram was indeed superior to the automatically fitted model; hence, it was used in further analysis.

3.4 Spatial Estimates and Error Standard Deviations

The combination of the estimates of the drift model (Eq. 6) and the top-kriging of residuals according to Eq. 5 yields the spatial estimates of average stream temperature T^* (in degree Celsius). The model was applied to predict the 8,000 catchments related to the nodes of the river network. Assuming homogeneity of the sub-catchments in terms of stream temperature, the estimated values are representative for the river segment upstream of the node, and the estimates can thus be plotted as vector maps of the stream network.

The spatial estimates are presented in Fig. 6a. The patterns mainly reflect relief and climate. Alpine catchments ($T^*=1-4$ °C) exhibit lower stream temperature than lowland

catchments ($T^*=8-11$ °C). Catchments in the East and South exhibit somewhat higher values than in the remaining part of the study area. This reflects the influence of the warmer Pannonic climate on these regions. Overall, the estimates correspond well with the observed spatial patterns (Fig. 3).

As one merit, geostatistical models provide estimates of the model uncertainty, given by the kriging error. Since top-kriging with external drift estimates are obtained by the superposition of two independent component models, the error standard deviation of top-kriging with external drift can be estimated by

$$e = \sqrt{e_{ED}^2 + e_{TK}^2} \quad (9)$$

where e_{ED}^2 denotes the prediction error variance of regression and e_{TK}^2 denotes the kriging error variance of top-kriging.

The error standard deviation of top-kriging with external drift estimates is presented in Fig. 6b. From the patterns, predictive uncertainty is mainly related to the density of the gauging network. Close to a gauge and along the main rivers, top-kriging with external drift yields low predictive uncertainties, with error standard deviations between 0 and 2 °C. In some of the smaller catchments which are notably distant from an observation point, higher error standard deviations of 2–5 °C occur. The higher uncertainty for the smallest catchments in ungauged regions is very realistic because of the higher predictive uncertainty in the sparsely gauged regions and also because these catchments are much smaller than the sampling scale. Hence, their prediction will be subject

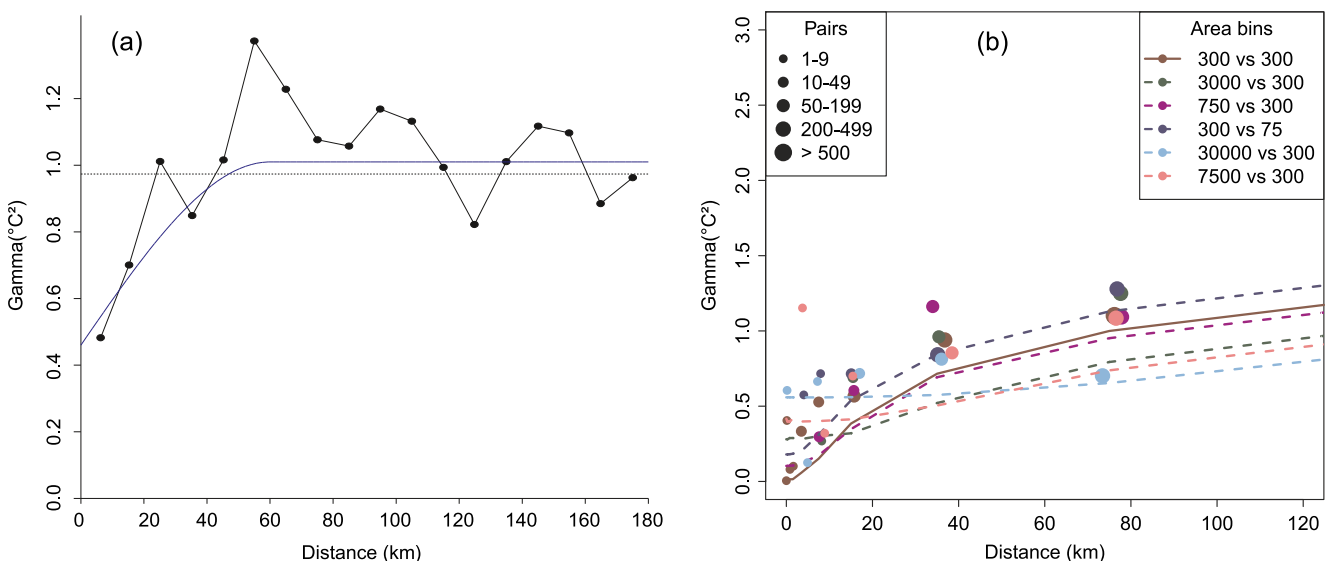
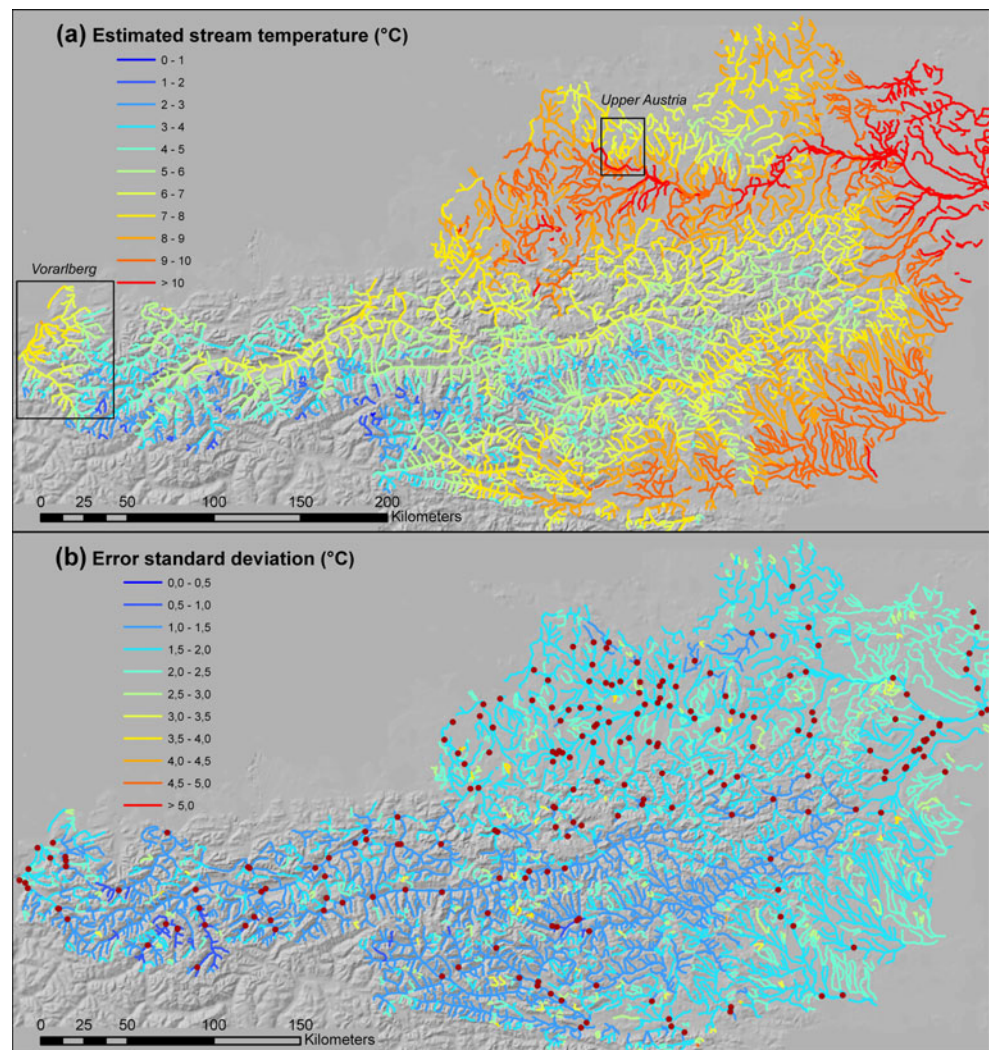


Fig. 5 Variogram model ($\gamma=\gamma$) of regression residuals of annual stream temperature (in degree Celsius) fitted to 214 catchments. **a** Lumped variogram. **b** Regularised variograms

Fig. 6 Estimate and estimation error of the mean annual stream temperature (in degree Celsius) for top-kriging with external drift



to additional downscaling errors. Because of the tree structure of river networks, scale issues constitute a typical problem in river network modelling [2, 26]. For top-kriging, scaling is implicit to the model since it is based on geostatistical regularisation. Top-kriging is therefore the most natural way of predicting stream flow-related variables on a river network. The external drift model additionally accounts for the systematic impact of catchment altitude on the regionalised variable.

3.5 Cross-Validation

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

model, we updated the regression equation for the remaining $n-1$ data points. The so obtained regression model was used for prediction at that location. For top-kriging, we used the residual variogram model obtained from all n observations to predict by kriging each point in turn using the remaining $n-1$ observations. The scatterplot of the predicted residuals (i.e. top-kriging estimates) vs. the observed residuals presented in Fig. 7 shows a positive correlation between the predicted and the observed residuals. However, the wide scatter and the tendency to underestimate the highest values point to a limited performance of the residual model alone. Finally, to obtain top-kriging with external drift estimates of stream temperature, the predictions of the residual model were superimposed with the predictions of the regression model according to Eq. 5.

From the resulting cross-validation (cv) predictions of top-kriging with external drift, the cross-validation error variance, V_{cv} , the root mean squared error, $rmse_{cv}$, and

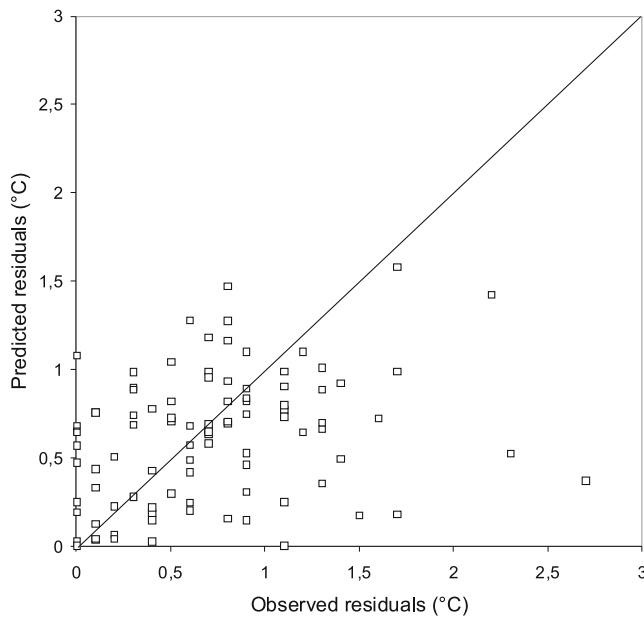


Fig. 7 Scatterplots of the predicted vs. the observed mean annual stream temperature (in degree Celsius) for top-kriging (without drift)

the coefficient of determination, R^2_{cv} , based on cross-validation were estimated

$$V_{cv} = \frac{1}{n} \sum_{i=1}^n (e_{cv,i})^2 \quad (10)$$

$$\text{rmse}_{cv} = \sqrt{V_{cv}} \quad (11)$$

$$R^2_{cv} = \frac{V_q - V_{cv}}{V_q} \quad (12)$$

In these equations, $e_{cv,i}$ is the cv residual (difference between cv predicted and observed values) of a model for catchment i and V_q is the variance of the observed T . In addition, the relative error $\text{rmse}_{cv} = \text{rmse}_{cv} / \text{mean}(T)$ was calculated based on the assumption that the freezing temperature 0°C constitutes the natural lower bound of stream temperature.

The results of the cross-validation are presented in Fig. 8 in terms of the error distribution for the set of 214 catchments. Figure 8a indicates that exponential regression fits rather well for the majority of catchments. The form of the scatter, however, shows a clustering related to differences in predictive accuracy. This suggests that some situations are better predicted than others. We will further explore this effect in Section 4.1. Top-kriging with external drift (Fig. 8b) exhibits a narrower scatter. The number of outliers is reduced, and they are better separated from the bunch of data.

The residual statistics show as well that top-kriging with external drift ($\text{rmse}_{cv} = 0.80^\circ\text{C}$, $\text{rmse}_{cv} = 0.10$, $R^2_{cv} = 0.81$) clearly outperforms exponential regression alone, which yields $\text{rmse}_{cv} = 1.01^\circ\text{C}$, $\text{rmse}_{cv} = 0.13$ and $R^2_{cv} = 0.77$. This corresponds to a reduction of predictive uncertainty by 20 % in terms of rmse_{cv} . The results demonstrate that top-kriging is well suited to exploit the spatial structure of the residuals of a deterministic model in a river network context.

4 Discussion

4.1 Regional Performances

The application study on the Austrian dataset yielded a clear improvement in predictive performance when using top-kriging to exploit the spatial structure of the residuals. The assessment of the estimated stream temperature by means of regional examples can be useful to better understand how top-kriging improves the estimates. We are interested here in the degree of realism of the estimated stream temperature patterns along the river network from both models. The first example is a region in Vorarlberg (western Austria). The example is typical for topologic effects in alpine catchments, where cold water from the mountains (southeast of the region) is transported to the valleys (northwest of the region). Figure 9 shows the estimates of regression alone (left panel) and top-kriging with external drift (right panels) plotted along the stream network. The measurements are shown as circles, using the same colour coding as for the estimates. The results of both models differ substantially. The red ellipses show a region where the observed stream temperature is colder than expected from the altitude of the gauge due to the fast transport of colder water from higher mountain ranges. This leads to systematic overestimation by the regression model. The estimates of top-kriging, however, correspond well with observations. This means that top-kriging of the residuals corrects regional biases detected by the drift model.

The second example is a region in Upper Austria (northern Austria). The example is a typical situation of a main stream (Danube, bold line) with tributaries. Figure 10 (left panel) reveals that the regression model does not fit well to none of the streams. It underestimates the temperature of the main stream and overestimates the temperature of the tributaries. Top-kriging, shown in the right panel, gives a more realistic estimate of the main stream and tributaries by exploiting local information of the stream network geometry and topology. Top-kriging of the residuals improves the local adjustment of the drift model, notably for dense gauging networks.

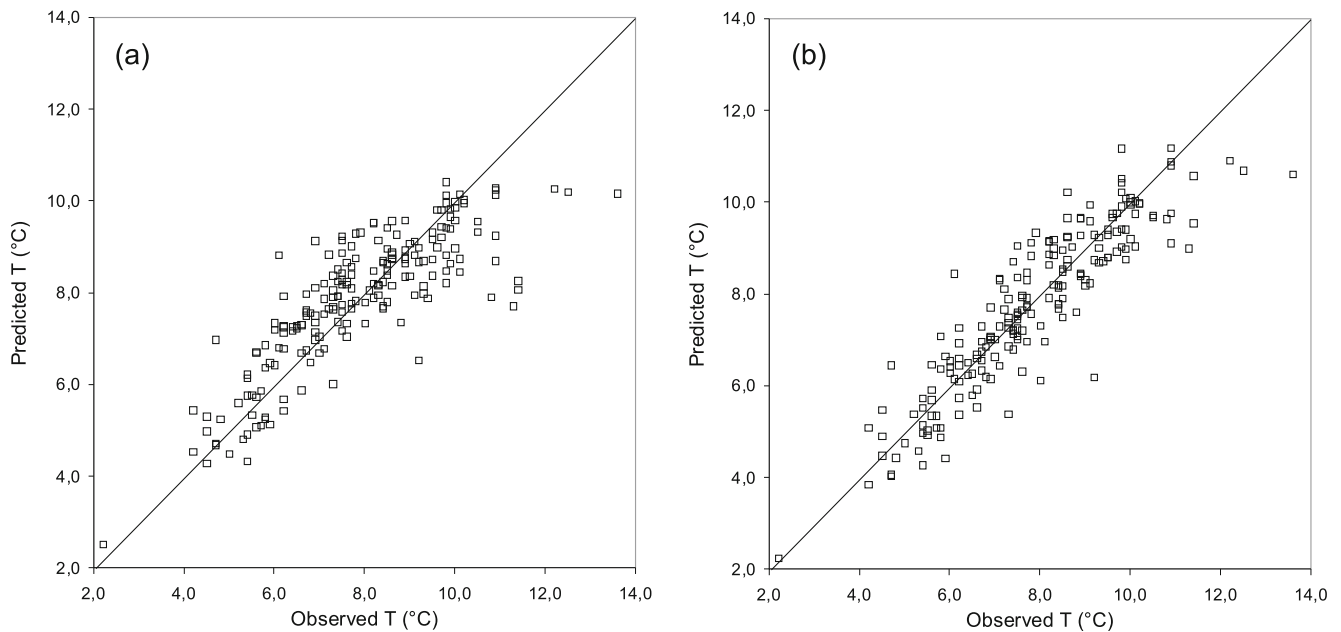


Fig. 8 Scatterplots of the predicted vs. the observed mean annual stream temperature (in degree Celsius). **a** Exponential regression ($R^2_{cv}=0.77$). **b** Top-kriging with external drift ($R^2_{cv}=0.81$)

4.2 Kriging Weights

We now focus on the properties of the spatial component model. The kriging estimates are weighted linear combinations of the observations; it is therefore interesting to shed light on how the kriging weights are distributed in top-kriging as compared to ordinary kriging based on the Euclidean (OK-Euclid) or on the stream distance metric (OK-Stream). In our examples, we have the minimal assumption of an intrinsic stationary and isotropic spatial variable which can be represented by an arbitrary valid variogram model (with exception of a pure nugget effect). For top-kriging, we further assume the variogram model of Fig. 5b. For ordinary kriging, we calculate distances between sampling locations (i.e. catchment outlet).

4.2.1 Interpolation

We first discuss the problem of interpolating between stream gauges, as illustrated in Fig. 11. The left picture shows the case of two connected neighbours which are situated in the same distance from the site of interest (10 km), and the catchments are assumed to be of rectangular shape. In OK-Euclid and OK-Stream, this would imply that the same weights, λ_i , are assigned to both neighbouring sites (0.5 in the case of two neighbours) as one would expect. Top-kriging, however, gives more weight to the downstream site ($\lambda_i=0.61$) and less weight to the upstream site ($\lambda_i=0.29$). The distribution of kriging weights of top-kriging seems surprising, but a closer look reveals that it reflects the different degree of nestedness between target site and neighbours: whilst the

Fig. 9 Regional example, Vorarlberg (western Austria). Stream temperature (in degree Celsius) estimated by exponential regression **(a)** and top-kriging with external drift **(b)**. *Ellipses* indicate a region which is colder than expected from H , leading to systematic overestimation by the regression model. Top-kriging corrects regional biases. The area shown is 60 km across; for location, see Fig. 6

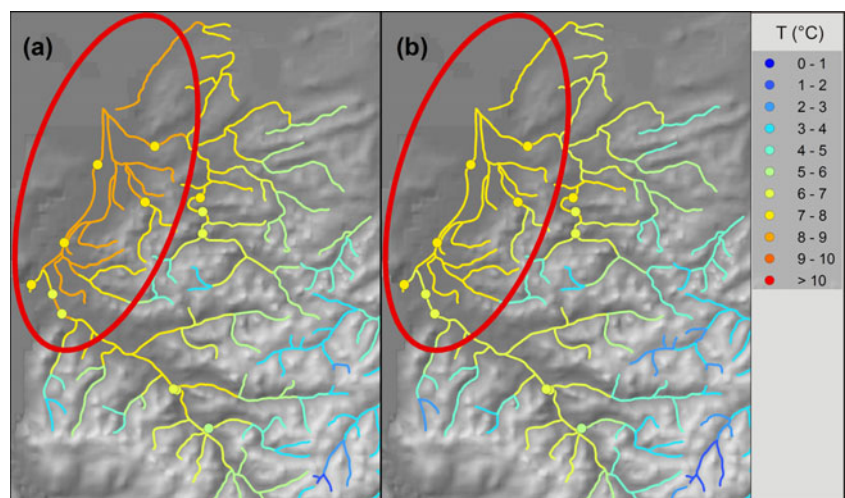
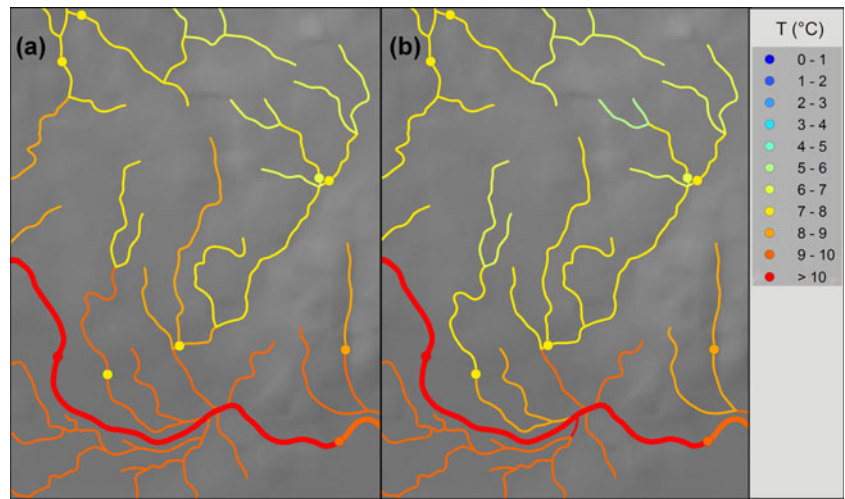


Fig. 10 Regional example, Upper Austria (northern Austria). Stream temperature (in degree Celsius) estimated by exponential regression (a) and top-kriging with external drift (b). Top-kriging improves local adjustment at main streams and tributaries. The area shown is 25 km across; for location, see Fig. 6



overlapping area with the upstream gauge amounts to 10 of 20 km² (i.e. 50 %), there is a larger overlap with the downstream gauge (20 of 30 km², or 67 %). Consequently, the discharge at the site of interest should be more similar to the downstream site. Top-kriging indeed gives more weight to the more similar downstream site and seems therefore better adapted than OK-Euclid and OK-Stream for interpolating along the stream network.

The right panel of Fig. 11 shows how the situation changes when an additional unconnected neighbour gauge is available. All three gauges are situated in the same distance from the site of interest. In top-kriging, the unconnected site receives no weight. Nestedness has an important effect on the similarity of discharges, and connected sites will mask out unconnected sites at an adjacent river when interpolating along the stream network. This is also valid for OK-Stream where the weights are still equally distributed between the upstream and

downstream gauges ($\lambda_i=0.5$). OK-Euclid, however, gives the same weight to all gauges ($\lambda_i=0.33$ in the case of the three gauges) regardless of the stream network topology.

4.2.2 Extrapolation

Figure 12 illustrates two typical extrapolation problems on a stream network. They correspond to the problem discussed by Wehrly et al. [32] presented in Section 1 of this paper. The left panel shows a situation of two neighbouring catchments having the same distance to the target catchment (10 km). One catchment is situated at the same river system, whilst the other catchment is situated at an adjacent, but unconnected river system. In OK-Euclid, this would imply that the same weights, λ_i , are assigned to all the neighbouring catchments (0.5 in the case of two neighbours). OK-Euclid again ignores the topology of the stream network. OK-Stream constitutes

Fig. 11 Interpolation: effect of stream network topology on the kriging weights, λ_i (green numbers), of top-kriging. x_i is the target site (red point) together with its catchment (red box); black points indicate the stream gauges together with their catchments (black boxes)

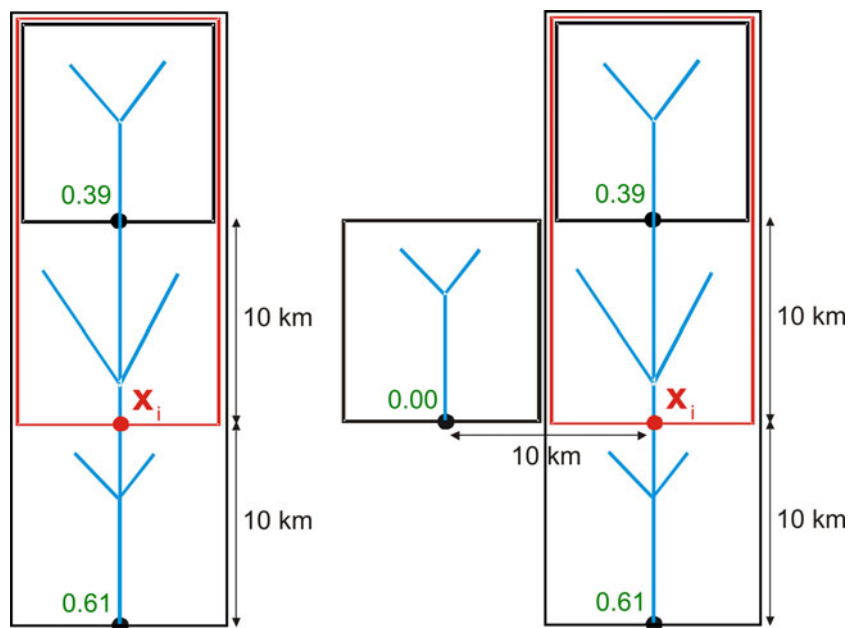
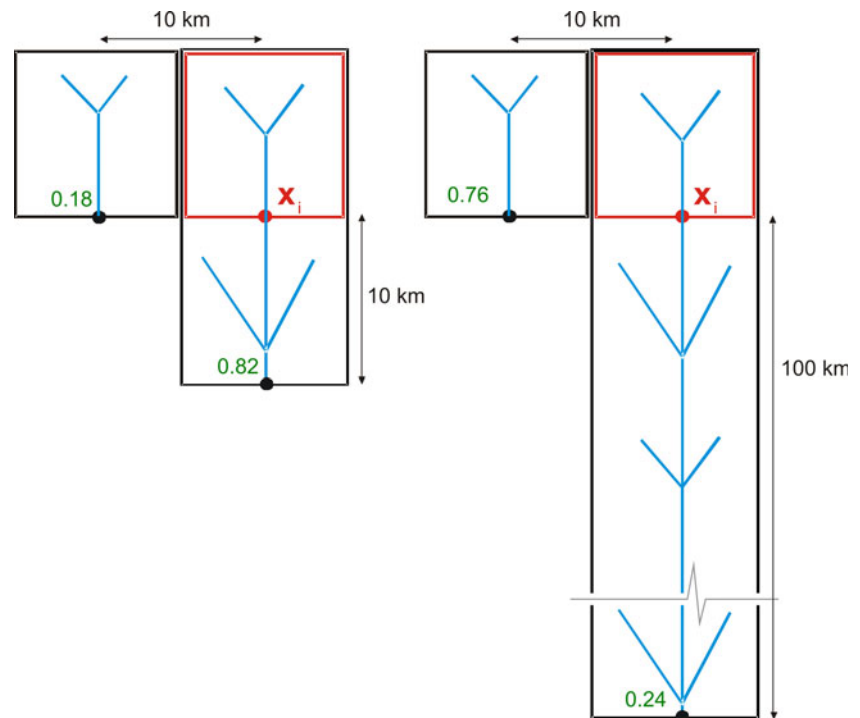


Fig. 12 Extrapolation: effect of stream network topology on the kriging weights, λ_i (green numbers), of top-kriging. x_i is the target site (red point) together with its catchment (red box); black points indicate the stream gauges together with their catchments (black boxes)



the opposite extreme since all weight ($\lambda_i=1$ in the case of two neighbours) is given to the connected site and no weight at all ($\lambda_i=0$) is given to the unconnected site. OK-Stream seems not optimally suited to the river network problem as it completely ignores the information in unconnected proximity.

The right panel of Fig. 12 shows a topologically similar situation, but the connected site has moved in far distance from the target site. Here, it is evident that the close-by catchment at the adjacent river system will, in most cases, be more similar to the target catchment than the distant catchment at the same river system. Hence, the unconnected site should receive most of the kriging weights. OK-Stream, however, still gives all weight to the distant connected site. In this case, even OK-Euclid, which is not adapted to stream networks, seems better suited since it always gives more weight to a close-by site, but regardless of the stream network topology.

In top-kriging, as opposed to OK-Euclid and OK-Stream, the kriging weights are distributed in a natural way, depending on the distance and stream network topology. When a close-by site at the same river system is available, top-kriging gives less weight to an adjacent river system. For the example shown in the left panel of Fig. 12, the connected site receives a weight $\lambda_i=0.82$, whereas the unconnected site receives a weight of only $\lambda_i=0.18$. However, when the next site at the same river system is far away, much weight is given to a close-by site at an adjacent river system. For the example shown in the right panel of Fig. 12, the distant connected site receives a weight of only $\lambda_i=0.24$, whereas the unconnected site receives a weight of $\lambda_i=0.76$. This distribution of kriging weights is very realistic and

corresponds well with physical understanding of stream flow process behaviour in space (c.f. [32]). The weights have been obtained by the variogram model shown in Fig. 5b. The relative effects of size and topology would be similar for other variograms.

5 Conclusions

We extend the top-kriging method with an external drift function to account for deterministic patterns of the spatial variable. We tested the method for a comprehensive Austrian stream temperature dataset. The drift was modelled by exponential regression with catchment altitude. Leave-one-out cross-validation analysis reveals that the model performs well for the study domain. When using top-kriging in addition to the regression model, the root mean squared error decreases by 20 %, from 1.01 to 0.80 °C. For regions where the observed stream temperatures deviate from the expected value of the drift model, top-kriging corrects regional biases. Top-kriging is able to improve the local adjustment of the drift model for the main streams and the tributaries by exploiting the topological information of the stream network. The additional performance increases with increasing density of the gauging network.

We compared top-kriging with two alternative kriging approaches for two typical stream network estimation problems. The first problem is the standard case of interpolating at sites between stream gauges. Here, top-kriging gives more weight to the downstream site which shares a larger

proportion of catchment area with the site of interest and is therefore expected to be more similar. An additional site at an unconnected river will be masked out by the two connected neighbours as it does not share any catchment area with the site of interest and is therefore less similar than the connected sites. In contrast, ordinary kriging based on either Euclidean or stream distance gives equal weight to upstream and downstream sites and does not account for the nestedness of catchments.

The second estimation problem is the extrapolation to source areas; we analysed two typical situations (cf. [32]): in the first situation, a close-by site at the same river is available, whilst in the second situation a close-by site is situated at an adjacent, unconnected river system. In these situations, ordinary kriging based on Euclidean distance would completely ignore the topology of the stream network as it distributes the kriging weights solely based on geometry. Ordinary kriging based on stream distance seems not optimally suited to the river network problem as well as it completely ignores the information in catchments that are nearby but not hydrologically connected. Top-kriging, however, allocates the kriging weights to the connected and unconnected sites of the stream network according to the data situation. When a close-by site at the same river system is available, top-kriging gives less weight to an adjacent river system. On the other hand, when the next site at the same river system is far away, much weight is given to a close-by site at an adjacent river system. This distribution of kriging weights is very realistic and corresponds well with physical understanding of stream flow process behaviour in space (c.f. [32]).

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, we conclude that top-kriging is notably well adapted for the stream network problem. From all these advantageous properties, and also because it is the best linear unbiased estimator, we conclude that top-kriging is the most natural method to perform spatial prediction on river networks. The method is well suited for stream flow and many stream flow-related variables. For non-stationary situations, such as in the example of stream temperatures of this paper, top-kriging with external drift should be used to perform spatial predictions on a river network.

Acknowledgments Top-kriging was performed using the original FORTRAN code of [26] and the R package “Rtop” which can be freely downloaded from R-Forge platform at <http://R-Forge.R-project.org>. The authors would like to thank the Austrian Science Foundation (FWF, project no. P18993-N10), the Austrian Academy of Sciences “Hydrological Predictability” Project and the Austrian

Climate Research Program “CILFAD” Project for financial support. Many thanks to Franz Suppan (Institute of Surveying, Remote Sensing and Land Information, University BOKU Vienna) for providing specific GIS functionality for catchment aggregation and to two anonymous reviewers for their valuable comments on the manuscript.

References

1. Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T. B. M. J., & Bobee, B. (2007). A review of statistical water temperature models. *Canadian Water Resources Journal*, *31*, 179–192.
2. Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling—a review. *Hydrological Processes*, *9*, 251–290.
3. Caissie, D. (2006). The thermal regime of rivers: a review. *Freshwater Biology*, *51*(8), 1389–1406.
4. Cressie, N. (1985). Fitting variogram models by weighted least squares. *Mathematical Geology*, *17*, 563–586.
5. Cressie, N. (1991). *Statistics for spatial data*. New York, NY: Wiley.
6. 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.
7. de Marsily, G. (1986). *Quantitative hydrogeology*. London: Academic.
8. Gardner, B., Sullivan, P. J., & Lembo, A. J., Jr. (2003). Predicting stream temperatures: geostatistical model comparison using alternative distance metrics. *Canadian Journal of Fisheries and Aquatic Sciences*, *60*, 344–351.
9. Garreta, V., Monestiez, P., & Ver Hoef, J. M. (2010). Spatial modelling and prediction on river networks: up model, down model or hybrid? *Environmetrics*, *21*, 439–456. doi:10.1002/env.995.
10. Godina, R., Nobilis, F., Webb, B.W. (2005). Wassertemperaturen [Water temperature]. Karte 5.6. In Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft (Ed.), *Hydrological Atlas Austria*. Österreichischer Kunst- und Kulturverlag Wien. ISBN 3-85437-250-7.
11. Gottschalk, L. (1993). Correlation and covariance of runoff. *Stochastic Hydrology and Hydraulics*, *7*, 85–101.
12. Gottschalk, L. (1993). Interpolation of runoff applying objective methods. *Stochastic Hydrology and Hydraulics*, *7*, 269–281.
13. 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.
14. Gotway, C. A., & Young, L. J. (2002). Combining incompatible spatial data. *Journal of the American Statistical Association*, *97*, 632–648.
15. Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. New York, NY: Oxford University Press.
16. Kyriakidis, P. C. (2004). A geostatistical framework for area-to-point spatial interpolation. *Geographical Analysis*, *36*, 259–289.
17. Laaha, G., & Blöschl, G. (2006). Seasonality indices for regionalizing low flows. *Hydrological Processes*, *20*, 3851–3878.
18. Laaha, G., & Blöschl, G. (2006). A comparison of low flow regionalisation methods—catchment grouping. *Journal of Hydrology*, *323*(1/4), 193–214.
19. Laaha, G., & Blöschl, G. (2007). A national low flow estimation procedure for Austria. *Hydrological Sciences Journal*, *52*(4), 625–644.
20. Laaha, G., Skoien, J., Blöschl, G. (2013). Spatial prediction on river networks: comparison of top-kriging with regional regression. *Hydrological Processes*. doi:10.1002/hyp.9578.
21. Matheron, G. (1965). *Les variables régionalisées et leur estimation*. Paris, France: Masson.

22. Mockus, A. (1998). Estimating dependencies from spatial averages. *Journal of Computational and Graphical Statistics*, 7, 501–513.
23. Peterson, E. E., Theobald, D. M., & Ver Hoef, J. M. (2007). Geostatistical modelling on stream networks: developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology*, 52, 267–279.
24. Peterson, E. E., & Ver Hoef, J. M. (2010). A mixed-model moving-average approach to geostatistical modelling in stream networks. *Ecology*, 91(3), 644–651.
25. 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.
26. Skøien, J. O., & Blöschl, G. (2006). Sampling scale effects in random fields and implications for environmental monitoring. *Environmental Monitoring and Assessment*, 114(1–3), 521–552.
27. Skøien, J. O., Blöschl, G., & Western, A. W. (2003). Characteristic space–time scales in hydrology. *Water Resources Research*, 39, 1304–1323.
28. Skøien, J. O., Merz, R., & Blöschl, G. (2006). Top-kriging—geostatistics on stream networks. *Hydrology and Earth System Sciences*, 10, 277–287.
29. Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13, 449–464.
30. Wackernagel, H. (1995). *Multivariate geostatistics*. Berlin: Springer.
31. Webb, B. W., Hannah, D. M., Moore, R. D., Brown, L. E., & Nobilis, F. (2008). Recent advances in stream and river temperature research. *Hydrological Processes*, 22(7), 902–918.
32. Wehrly, K. E., Brenden, T. O., & Wang, L. (2009). A comparison of statistical approaches for predicting stream temperatures across heterogeneous landscapes. *Journal of the American Water Resources Association*, 45(4), 986–997.