



Spatiotemporal topological kriging of runoff time series

Jon Olav Skøien^{1,2} and Günter Blöschl¹

Received 27 November 2006; revised 24 May 2007; accepted 8 June 2007; published 27 September 2007.

[1] This paper proposes a geostatistical method for estimating runoff time series in ungauged catchments. The method conceptualizes catchments as space-time filters and exploits the space-time correlations of runoff along the stream network topology. We hence term the method topological kriging or top kriging. It accounts for hydrodynamic and geomorphologic dispersion as well as routing and estimates runoff as a weighted average of the observed runoff in neighboring catchments. Top kriging is tested by cross validation on 10 years of hourly runoff data from 376 catchments in Austria and separately for a subset of these data, the Innviertel region. The median Nash-Sutcliffe efficiency of hourly runoff in the Innviertel region is 0.87 but decreases to 0.75 for the entire data set. For a subset of 208 catchments, the median efficiency of daily runoff estimated by top kriging is 0.87 as compared to 0.67 for estimates of a deterministic runoff model that uses regionalized model parameters. The much better performance of top kriging is because it avoids rainfall data errors and avoids the parameter identifiability issues of traditional runoff models. The analyses indicate that the kriging variance can be used for identifying catchments with potentially poor estimates. The Innviertel region is used to examine the kriging weights for nested and nonnested catchments and to compare various variants of top kriging. The spatial kriging variant generally performs better than the more complex spatiotemporal kriging and spatiotemporal cokriging variants. It is argued that top kriging may be preferable to deterministic runoff models for estimating runoff time series in ungauged catchments, provided stream gauge density is high and there is no need to account for causal rainfall-runoff processes. Potential applications include the estimation of flow duration curves in a region and near-real time mapping of runoff.

Citation: Skøien, J. O., and G. Blöschl (2007), Spatiotemporal topological kriging of runoff time series, *Water Resour. Res.*, 43, W09419, doi:10.1029/2006WR005760.

1. Introduction

[2] The estimation of runoff related variables at locations where no measurements are available is a key problem in hydrology which is generally termed predictions in ungauged basins (PUB) [Sivapalan *et al.*, 2003]. A range of methods addressing this problem have been proposed in the literature [Blöschl, 2005]. The traditional approach is to use a deterministic rainfall-runoff model with parameters inferred from neighboring, gauged catchments. The advantage of this method is that it explicitly represents causal processes such as precipitation and runoff generation. However, as Blöschl [2005] noted, there is significant uncertainty in the estimates, mainly due to a lack of representativeness of catchment data, rainfall data as well as identifiability problems of the runoff model parameters [Montanari, 2005b]. For some applications it may not be necessary to invoke causal relationships directly. For example, when assessing

the hydropower potential one is interested in the flow duration curves of ungauged sites [e.g., Castellarin *et al.*, 2004] and there is usually no need to vary rainfall characteristics. Another application is environmental flows where one is interested in the runoff dynamics and seasonality at ungauged sites without the need to examine scenarios [Gippel, 2005]. For these applications, an alternative to the traditional rainfall-runoff modeling approach may be to directly estimate the runoff time series from observed time series of neighboring catchments without recourse to rainfall data. This would be of particular interest in those countries where a rather dense stream gauge network exists.

[3] An obvious choice for estimating runoff time series directly from observed runoff of neighboring catchments are geostatistical methods. These methods assume that the estimates at locations without observations can be found as weighted averages of the neighboring observations, and consist of finding suitable weights from the spatial correlations. Although geostatistical methods such as kriging are best linear unbiased estimators (BLUE) [Journel and Huijbregts, 1978, p. 304] they have not been used much in catchment hydrology. This is because geostatistical methods have evolved in the mining industry where a typical objective is to estimate the expected ore grade of a

¹Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Vienna, Austria.

²Now at Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, Netherlands.

cuboid block from point samples. Similar methods have been developed in meteorology [Gandin, 1963] where a typical objective is to estimate meteorological variables on a square grid. The geostatistical estimation procedures make use of the variogram, which is the spatial correlation of pairs of points of the variable of interest plotted against their Euclidian distance. The problem in catchment hydrology is quite different in that, unlike mining blocks, catchments are organized into subcatchments and the organization is defined by the stream network. Upstream and downstream catchments would have to be treated differently from neighboring catchments that do not share a subcatchment. The Euclidian distance is hence not the natural way of measuring the spatial distance of catchments. Rather a topology needs to be used that is based on the stream network.

[4] Gottschalk [1993] was probably the first to develop a method for calculating covariance along a stream network. This method was extended by Sauquet *et al.* [2000] to map annual runoff along the stream network using water balance constraints in the estimation procedure and by Gottschalk *et al.* [2006] to map the coefficient of variation of runoff along the stream network. Skøien *et al.* [2006] extended their method and estimated the 100 year floods at ungauged locations as well as their uncertainty. Their method is based on the concept of Woods and Sivapalan [1999] that assumes that local runoff generation or rainfall excess can be defined at each point in space and can be integrated over a catchment to form catchment runoff. Skøien *et al.* [2006] termed their method topological kriging or top kriging as it exploits the topology of nested catchments in addition to the spatial correlation of runoff. This paper is concerned with estimating time series of runoff for ungauged locations rather than with estimating a single quantity as was the case with Sauquet *et al.* [2000] and Skøien *et al.* [2006]. It is hence necessary to extend the spatial estimation procedure of the earlier work to a space-time estimation procedure, taking into account both spatial and temporal correlations of runoff.

[5] Spatiotemporal geostatistical models have been used in a number of disciplines [e.g., Snepvangers *et al.*, 2003; Jost *et al.*, 2005]. In a review of spatiotemporal methods, Kyriakidis and Journel [1999] note that there are three options of representing the random variable—full space-time models, simplified representations as vectors of temporally correlated spatial random fields, and simplified representations as vectors of spatially correlated time series. The latter reduces to a spatial estimation problem and is of interest for variables with observations that are rich in time but poor in space as is the case of runoff [Rouhani and Wackernagel, 1990]. Full spatiotemporal kriging is more complicated than the two simplifications as the kriging system needs to be solved simultaneously for both spatial and temporal kriging weights [Kyriakidis and Journel, 1999]. As still another alternative, spatiotemporal cokriging has been suggested [Rouhani and Wackernagel, 1990], where information from different time steps are treated as covariates. Spatiotemporal cokriging includes more unbiasedness conditions, while spatiotemporal kriging is strictly only valid when the mean does not change with time [Bogaert, 1996].

[6] The objective of this paper is to propose a method of spatiotemporal top kriging that is able to estimate runoff

time series at ungauged locations. The paper goes beyond the work of Sauquet *et al.* [2000] and Skøien *et al.* [2006] by accounting for space-time correlations rather than spatial correlations and by including routing effects. The new method is tested by cross validation to assess its accuracy for the case of ungauged locations on the basis of an Austria data set. The characteristics of the method in terms of representing the dynamics of the hydrograph are examined and its predictive power is compared to that of traditional regionalization on the basis of deterministic rainfall-runoff models.

2. Data

[7] The data used in this paper stem from a comprehensive hydrographic data set of Austria. Austria has a varied climate with mean annual precipitation ranging from 500 mm in the eastern lowland region up to about 3000 mm in the western alpine region. Runoff depths range from less than 50 mm per year in the eastern part of the country to about 2000 mm per year in the Alps. Potential evapotranspiration is on the order of 600–900 mm per year. Austria has a dense stream gauge network. Hourly runoff data over the period 1 August 1990 to 31 July 2000 are used in this paper. The raw runoff data were screened to exclude catchments with significant anthropogenic effects, karst and strong lake effects. The remaining data set consisted of 376 stream gauges with catchment areas ranging from 10 to 10,000 km² (Figure 1).

[8] To analyze the dynamic characteristics of the estimation method, the Innviertel region in northwestern Austria is examined in more detail (Figure 2). The region covers an area of approximately 1500 km² and contains 19 stream gauges. Mean annual runoff is rather uniform in the region, ranging from 350 to 510 mm/year. However, there are differences in the runoff dynamics (Table 1). This is because of a rather complex hydrogeology consisting of a mixture of moraine, clay and marl with local gravel fillings of the valleys [Schubert *et al.*, 2003]. The temporal variance of runoff in the most dynamic catchment (Osternach) is more than ten times that of the slowest responding catchment (Weghof). The latter contains significant gravel deposits in the valleys while the former does not.

3. Method

3.1. Concepts of Top Kriging

[9] There are two main groups of variables that control streamflow (Figure 3). The first group consists of variables that are continuous in space, which are related to local runoff generation. These variables include rainfall, evapotranspiration and soil characteristics. In this context, local runoff generation is conceptualized as a point process; that is, locally generated runoff is assumed to exist at any point in the landscape. This concept is discussed by Woods and Sivapalan [1999]. In a similar way, other streamflow-related variables can be conceptualized as continuous point processes on the local scale. For characterizing these variables, Euclidian distances are appropriate. The spatial statistical characteristics of the point variables can be represented by the variogram [Skøien *et al.*, 2003].

[10] The second group of variables is related to routing in the stream network. These variables are affected by the

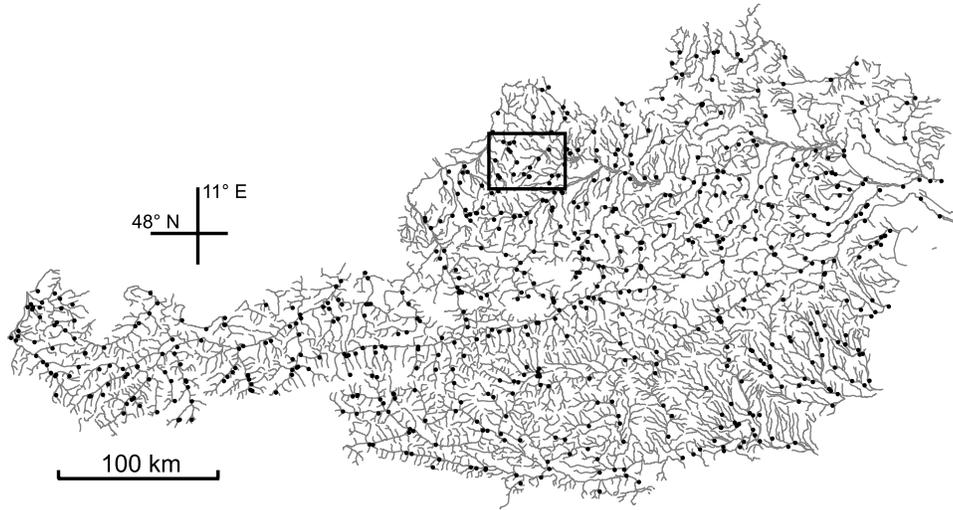


Figure 1. Stream gauges (circles) in Austria used in this paper. The square represents the Innviertel region.

catchment organization of nested catchments where runoff accumulates along the stream network. Variables of this type include stream flow, streamflow statistics, concentrations, turbidity and stream temperature. These variables are only defined for points on the stream network. Correlations between observations cannot be represented by Euclidian distances, as in ordinary kriging. Rather they need to be represented by methods that reflect the tree structure of the stream network.

[11] The method top kriging, presented by *Skøien et al.* [2006], combines these two groups of variables in a geostatistical framework. The continuous process in space defined for point variables is represented by a variogram. The channel network structure and the similarity between upstream and downstream neighbors are represented by the catchment area that drains to a particular location on the stream network. The catchment areas are defined by their boundaries in space. Extensions to the top kriging method for instantaneous runoff measurements are given below.

3.2. Spatiotemporal Estimation of Time Series

[12] We assume that specific runoff $q(\mathbf{x}_i, t_\omega)$ at location \mathbf{x}_i on the stream network and time t_ω can be represented as a spatiotemporal random field that is related to runoff Q by

$$q(\mathbf{x}_i, t_\omega) = Q(\mathbf{x}_i, t_\omega)/A_i \quad (1)$$

where A_i is the catchment area. Further arguments for this assumption is given in section 3.3. In this paper we propose three variants of top kriging which we term, for simplicity in terminology, spatial kriging, spatiotemporal kriging and spatiotemporal cokriging.

3.2.1. Spatial Kriging

[13] In the first variant we note that runoff time series represent spatiotemporal runoff with a much higher resolution in time than in space. As a simplification we

hence consider runoff to consist of a set of spatially correlated time series [*Kyriakidis and Journal*, 1999]. For each time step t_ω , specific runoff $\hat{q}(\mathbf{x}_i, t_\omega)$ of an ungauged target catchment defined by location \mathbf{x}_i is estimated from observed specific runoff $q(\mathbf{x}_j, t_\alpha)$ at the same point in time $t_\alpha = t_\omega$ of neighboring gauged catchments located at \mathbf{x}_j as

$$\hat{q}(\mathbf{x}_i, t_\omega) = \sum_{j=1}^n \lambda_j q(\mathbf{x}_j, t_\alpha) \quad (2)$$

where λ_j is the weight given to the runoff from each gauged catchment and n is the total number of stream gauges used. This means that each time step is treated

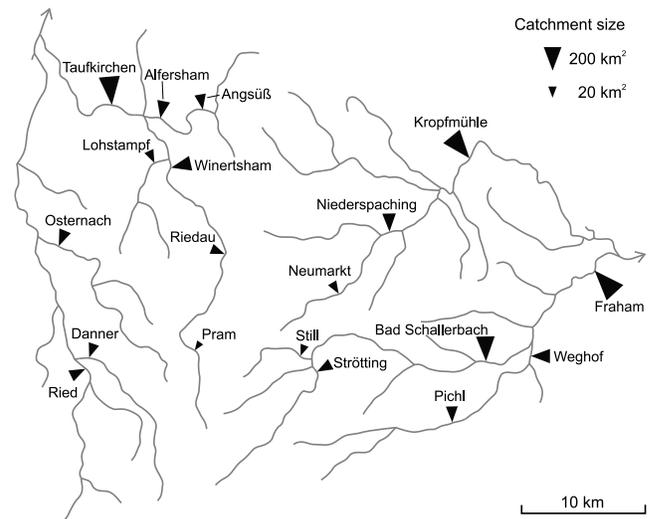


Figure 2. Innviertel region with stream gauges (triangles). The triangles are scaled according to catchment area. The catchments are listed in Table 1.

Table 1. Stream Gauges in the Innviertel Region With Catchment Area, Mean Annual Runoff, and Temporal Runoff Variance

| Stream Gauge | Stream | Area, km ² | Average Runoff, × 10 ⁻² m ³ s ⁻¹ km ⁻² | Temporal Variance, × 10 ⁻⁴ m ⁶ s ⁻² km ⁻⁴ |
|------------------|---------------|-----------------------|--|---|
| Ried | Rieder Bach | 69.3 | 1.53 | 7.79 |
| Danner | Antiesen | 55.6 | 1.45 | 4.14 |
| Osternach | Osternach | 68.6 | 1.42 | 10.97 |
| Pram | Pram | 14.2 | 1.60 | 9.36 |
| Riedau | Pram | 59.5 | 1.44 | 7.52 |
| Winertsham | Pram | 128.1 | 1.33 | 6.48 |
| Taufkirchen | Pram | 303.3 | 1.38 | 4.65 |
| Angsüß | Pfudabach | 64.1 | 1.59 | 3.80 |
| Alfersham | Pfudabach | 81.3 | 1.50 | 2.38 |
| Lohstampf | Messenbach | 39.3 | 1.31 | 7.04 |
| Still | Stillbach | 19.4 | 1.33 | 6.32 |
| Strötting | Trattnach | 52.0 | 1.44 | 3.06 |
| Bad Schallerbach | Trattnach | 183.8 | 1.23 | 3.19 |
| Pichl | Innbach | 66.2 | 1.23 | 1.32 |
| Weghof | Innbach | 116.5 | 1.09 | 0.79 |
| Fraham | Innbach | 361.8 | 1.14 | 1.62 |
| Neumarkt | Dürre Aschach | 29.7 | 1.22 | 6.78 |
| Niederspaching | Aschach | 104.0 | 1.30 | 7.17 |
| Kropfmühle | Aschach | 312.5 | 1.37 | 4.98 |
| Average | | 112.1 | 1.36 | 5.23 |
| Median | | 68.6 | 1.37 | 4.98 |

independently in the estimation procedure. The weights λ_j are found by solving the kriging system:

$$\sum_{k=1}^n \lambda_k \gamma_{jk} - \lambda_j \sigma_j^2 + \mu = \gamma_{ij} \quad j = 1, \dots, n$$

$$\sum_{j=1}^n \lambda_j = 1 \tag{3}$$

where the gamma values γ_{ij} and γ_{ik} are the expected semivariances (or variogram values) between the target catchment i and the neighbors j used for estimation, and between two neighbors j and k , respectively. The gamma values are obtained by regularizing a theoretical point variogram (see section 3.3). The term σ_j^2 represents the local uncertainty of runoff which may be due to measurement error and small-scale variability. The use of local uncertainty in the kriging equations is termed kriging with uncertain data (KUD) [de Marsily, 1986, p. 300; Merz and Blöschl, 2005]. Finally, μ is the Lagrange parameter. We limited the number of neighbors n to five to minimize potential numerical problems with the regularization procedure. The limitation of the number of neighbors also implies a local kriging approach, which reduces the effect of heterogeneous observations to a local region. Also, in some cases the kriging weights were adjusted as described in Appendix A before they were used in equation (2).

[14] It is assumed that the same variogram is applicable to all time steps and that the same stream gauges j can be used as neighbors for all time steps. The kriging equation has hence only to be solved once for each target catchment i , and the same weights are used for all time steps. As noted above, only runoff data for the time step of interest are used in the estimation.

3.2.2. Spatiotemporal Kriging

[15] Spatiotemporal kriging on the other hand does take into account information from different time steps [Kyriakidis and Journel, 1999]:

$$\hat{q}(\mathbf{x}_i, t_\omega) = \sum_{j=1}^n \sum_{\alpha=1}^p \lambda_{j\alpha} q(\mathbf{x}_j, t_\alpha) \tag{4}$$

where p is the number of time steps used for the estimation. Extending the kriging system of equation (3) gives

$$\sum_{k=1}^n \sum_{\beta=1}^p \lambda_{j\beta} \gamma_{jk\alpha\beta} - \lambda_{j\alpha} \sigma_j^2 + \mu = \gamma_{ij\alpha\omega}$$

for $j = 1, \dots, n \quad \alpha = 1, \dots, p$

$$\sum_{j=1}^n \sum_{\beta=1}^p \lambda_{j\beta} = 1 \tag{5}$$

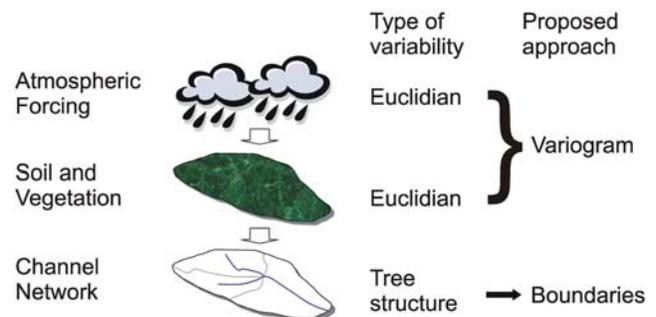


Figure 3. Atmospheric forcing and soil and vegetation contributions to the runoff generation process locally, which can be represented by point variograms. The channel network organizes runoff into streams, which can be represented by the catchment boundaries.

which consists of $n \cdot p + 1$ linear equations. $\gamma_{jk\alpha\beta}$ and $\gamma_{ij\alpha\omega}$ are, again, found by regularizing a theoretical point variogram (see section 3.3). Again, for numerical robustness, $n = 5$. We also limited the number of time steps p and considered two cases. In the first case, we used five time steps ($t_\alpha - t_\omega = -5, -2, 0, 2, 5$ hours) and in the second case we used nine time steps ($t_\alpha - t_\omega = -20, -10, -5, -2, 0, 2, 5, 10, 20$ hours).

3.2.3. Spatiotemporal Cokriging

[16] An alternative method, termed spatiotemporal cokriging, uses measurements from other time steps as covariates. Equation (4) is used here as well but the difference from spatiotemporal kriging is that the sum of weights for each time step is set equal to zero, except for the weights of the time step of interest, t_ω , which sum up to one. Cokriging hence involves more constraints than spatiotemporal kriging. The kriging weights $\lambda_{j\alpha}$ can be found by solving the following linear system [Rouhani and Wackernagel, 1990]:

$$\begin{aligned} \sum_{k=1}^n \sum_{\beta=1}^p \lambda_{j\beta} \gamma_{jk\alpha\beta} - \lambda_{j\alpha} \sigma_j^2 + \mu_j &= \gamma_{ij\alpha\omega} \\ \text{for } j &= 1, \dots, n \quad \alpha = 1, \dots, p \\ \sum_{j=1}^n \lambda_{j\omega} &= 1 \\ \sum_{j=1}^n \lambda_{j\beta} &= 0 \quad \forall \beta \neq \omega \end{aligned} \quad (6)$$

where $\gamma_{jk\alpha\beta}$ is the cross variogram between $q(\mathbf{x}_j, t_\alpha)$ and $q(\mathbf{x}_k, t_\beta)$. There exist different definitions of the cross variogram in the literature, and we adopt the definition of Clark *et al.* [1987], which has been recommended by Cressie [1991]. With information from different time steps as the covariates, the cross variogram is defined as

$$\gamma_{jk\alpha\beta} = \frac{1}{2} \text{var}[q(\mathbf{x}_j, t_\alpha) - q(\mathbf{x}_k, t_\beta)] \quad (7)$$

which is equal to a spatiotemporal variogram [Skøien and Blöschl, 2006]. As for spatiotemporal kriging, we used $n = 5$, and $p = 5$ and 9.

3.3. Catchments as Space-Time Filters

[17] The observed runoff at the catchment outlet is not only a result of the spatial aggregation described by Skøien *et al.* [2006] and in 3.1 but also a result of a complicated nonlinear averaging process within the catchment over a period of time. Skøien and Blöschl [2006] described this averaging process as a space-time filter that operates on local runoff. We adopt this concept in this paper to find the gamma values to be used in equations (3), (5) and (6).

[18] Following Woods and Sivapalan [1999] and Skøien and Blöschl [2006] we conceptualize the locally generated runoff or rainfall excess $R(\mathbf{r}, \tau)$ as a continuous point process in space \mathbf{r} and time t . To account for routing on the hillslopes and in the channels within the catchment, a

weighting function $u(\mathbf{r}, \tau)$ is introduced, which, assuming linear convolution, allows to combine locally generated runoff into runoff at the catchment outlet, Q_i :

$$Q_i(t) = \int_{A_i} \int_{t-T_i}^t R(\mathbf{r}, \tau) \cdot u(\mathbf{r}, t-\tau) d\tau d\mathbf{r} \quad (8)$$

where A_i is the catchment area and T_i is the time interval that influences the output. The weighting function $u(\mathbf{r}, \tau)$ can, for a given point \mathbf{r} in space, be seen as equivalent to a unit hydrograph for the runoff generated at this location. As an approximation, we assume that, for a given catchment, these weighting functions are constant within the integration limits both in space and time, i.e., $u(\mathbf{r}, \tau) = u_i = 1/T$. For a weighting function constant in space and time, equation (8) becomes a linear filter or a convolution integral. This can be seen as the equivalent of using the instantaneous unit hydrograph in a lumped model. It is important to notice, however, that this conceptualization is done for finding the statistical properties of the catchments only, not for estimation in the time domain. In time, the weighting function is the equivalent of a set of unit hydrographs that are constant between 0 and T_i and zero for other time steps. In space, the weighting function is constant within the catchment area and zero elsewhere. The specific runoff at the catchment outlet then becomes

$$q_i(t) = \frac{1}{A_i T_i} \int_{A_i} \int_{t-T_i}^t R(\mathbf{r}, \tau) d\tau d\mathbf{r} \quad (9)$$

[19] In geostatistical terminology, A_i and T_i are the spatial and temporal supports, respectively. Again following Skøien and Blöschl [2006], we assume that the temporal support is related to catchment area by

$$T_i = \mu \cdot A_i^\kappa \quad (10)$$

where μ and κ are parameters to be estimated from the data. For $\kappa > 0$ the temporal support (or time base of the unit hydrograph) increases with catchment area.

[20] In a geostatistical framework, the linear aggregation of equation (9) is performed on the second moments. We assume that the local runoff generation process, as the result of atmospheric forcing and filtering on the ground by soil and vegetation, can be described as a spatiotemporal random process. Because of elevation dependencies of precipitation, we cannot assume homogeneity of the runoff generation process, but we assume that the intrinsic hypothesis is applicable for the runoff generation process. Hence a point variogram of runoff represents the second moment of locally generated runoff, or local instantaneous runoff. From this point variogram with zero support in space and zero support in time, we can use equation (9) to estimate variograms that are valid for finite support areas and finite support times. This procedure is usually referred to as regularization [Journal and Huijbregts, 1978]. Following Cressie [1991, p. 66], and Skøien and Blöschl [2006], we find gamma value $\gamma_{ij\alpha\beta}$ between two catchments i and j , and

two different time steps, t_α and t_β by regularizing a spatiotemporal point variogram γ_{st} :

$$\begin{aligned} \gamma_{ij\alpha\beta} = & \frac{1}{A_i A_j T_i T_j} \int_{A_i} \int_{A_j} \int_0^{T_i} \int_0^{T_j} \gamma_{st}(|\mathbf{r}_1 - \mathbf{r}_2|, |\tau_1 + h_t - \tau_2|) \\ & \cdot d\tau_1 d\tau_2 d\mathbf{r}_1 d\mathbf{r}_2 \\ & - 0.5 * \left[\frac{1}{A_i^2 T_i^2} \int_{A_i} \int_{A_i} \int_0^{T_i} \int_0^{T_i} \gamma_{st}(|\mathbf{r}_1 - \mathbf{r}_2|, |\tau_1 - \tau_2|) \right. \\ & \cdot d\tau_1 d\tau_2 d\mathbf{r}_1 d\mathbf{r}_2 \\ & + \frac{1}{A_j^2 T_j^2} \int_{A_j} \int_{A_j} \int_0^{T_j} \int_0^{T_j} \gamma_{st} \\ & \left. \cdot (|\mathbf{r}_1 - \mathbf{r}_2|, |\tau_1 - \tau_2|) d\tau_1 d\tau_2 d\mathbf{r}_1 d\mathbf{r}_2 \right] \end{aligned} \quad (11)$$

where $h_t = |t_\alpha - t_\beta|$, \mathbf{r}_1 and \mathbf{r}_2 are spatial integration vectors within the two catchments, and τ_1 and τ_2 are the temporal integration variables. For the integration in equation (11) over A_i and A_j the catchment boundaries from the digital database were used for each catchment. For spatial kriging, $h_t = 0$ in equation (11).

[21] On the basis of a comparison of different variogram models of Skøien and Blöschl [2006] we chose their exponential model, because the number of parameters is limited and some physical interpretation of the parameters is possible:

$$\gamma_{st}(h_s, h_t) = a \left(1 - \exp\left(-((ch_t + h_s)/d)^b\right) \right) + a_s h_s^{b_s} + a_t h_t^{b_t} \quad (12)$$

[22] The first term of this variogram represents the stationary part. a gives the variance of the (stationary) process, c relates space and time, d is the combined correlation length, and b gives the slope of the variogram. The second and the third terms give the nonstationary parts of the variogram in space and time, respectively.

3.4. Simple Routing Model

[23] The space-time filter accounts for the effects of routing within the catchment on the spatial and temporal variance. However, it does not explicitly represent time lags between catchments, as a result of the water flowing from upstream to downstream catchment. catchments as a result of the water flowing from an upstream to a downstream catchment. We have simplified the correlation structure by the use of temporally symmetric variograms, i.e., assuming that past and future time steps will have the same correlation to the target time step. To account for time lags as a result of in-stream routing in the estimation procedure we use a routing model that consists of applying a time lag that is constant with time. To estimate runoff at time step t_ω of catchment i , we use $q(z_j, t_\alpha^*)$ at time step $t_\alpha^* = t_\alpha + t'_\alpha$ instead of t_α in equations (2) and (4). Depending on the relative position on the stream network of catchments i and j , the time lag t'_α can be positive or negative.

[24] The lag time was defined differently for nested catchments and nonnested catchments. For nested catch-

ments, we inferred traveltimes from cross variograms for catchment pairs estimated from the runoff data (equation (17) below). The minimum variance of the cross variograms typically occurred for a temporal separation larger than 0. Dividing the spatial distance between the stream gauges by that temporal separation gave a flow velocity for each pair of catchments. For the Innviertel region we found an average velocity of $v = 0.67$ m/s:

$$t'_\alpha = d_{ij}/v \quad \text{if } i \text{ and } j \text{ are nested} \quad (13)$$

where d_{ij} is the distance between the two stream gauges. d_{ij} is positive when j is a downstream neighbor and negative when j is an upstream neighbor of i . If d_{ij}/v does not correspond to an integer time step, a weighted average of q of the two time steps is used.

[25] A slightly different assumption was made about the time lags for nonnested catchments. Typically, the characteristic velocity of precipitation is much larger than that of flow routing processes in catchments [Skøien et al., 2003]. This means that the difference in the timing of a flood event of two catchments of different size that are close to each other will be mainly due to routing differences as the time difference due to the convective motion of the rainfall system will be much smaller. As an approximation, we hence relate the time lag to the size of the catchments. For a single catchment i the time lag T_{Li} of runoff relative to rainfall has been estimated as

$$T_{Li} = \xi \cdot A_i^\psi \quad (14)$$

following Melone et al. [2002]. For the Innviertel region $\xi = 1.5$ and $\psi = 0.35$ [Merz and Blöschl, 2003]. The time lag of two nonnested catchments is then assumed to be the difference of the individual time lags:

$$t'_\alpha = T_{Lj} - T_{Li} \quad \text{if } i \text{ and } j \text{ are nonnested} \quad (15)$$

[26] To analyze the effect of the routing model on runoff estimation we examined an alternative variant where we set all time lags to zero:

$$t'_\alpha = 0 \quad (16)$$

[27] The variant with a routing model is termed “routing all” below, whereas the variant with all time lags set equal to zero is termed “no routing.”

3.5. Hydrologic Interpretation of Top Kriging

[28] There are two main groups of processes that control runoff. The first group consists of variables that are continuous in space and include rainfall, evapotranspiration and soil characteristics. In top kriging their variability is represented by the point variogram that is based on Euclidian distances. The second group of processes is related to routing on the hillslope and in the stream network. Their effect cannot be represented by Euclidian distances. Top kriging represents these processes in three ways.

[29] 1. The channel network structure and the similarity between upstream and downstream neighbors are represented by the catchment area that drains to a particular

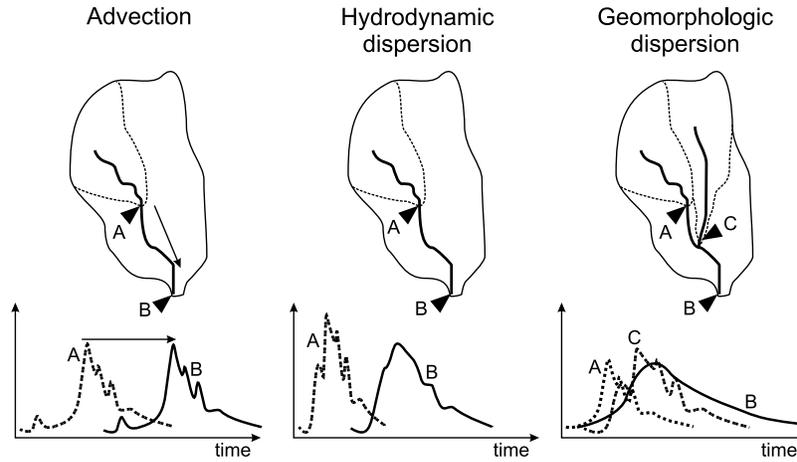


Figure 4. Schematic of runoff routing processes represented by top kriging. Hydrographs for locations A, B, and C are shown.

location on the stream network. The catchment areas are defined by their boundaries in space.

[30] 2. Advective runoff routing (Figure 4) is represented by a simple routing model (equations (13)–(16)). This model takes into account the traveltime between upstream and downstream neighbors and involves direction in both space and time. The routing is incorporated directly in the estimation procedure (equations (2) and (4)) by a time lag in estimating runoff as a weighted average of the runoff of neighboring catchments.

[31] 3. Dispersive routing is represented by the space-time filter (equation (11)). Dispersive effects include hillslope routing and what *Rinaldo et al.* [1991] refer to as hydrodynamic and geomorphologic dispersion (Figure 4). Hydrodynamic dispersion is caused by different traveltimes in the stream within individual reaches and is related to the pressure term in the St. Venant equation. Geomorphologic dispersion is related to the different lengths and junctions of the stream network and results in a superposition of runoff from the tributaries. The representation of dispersion in top kriging has an analogy in the unit hydrograph concept. Application of the unit hydrograph concept involves two steps-estimation of catchment rainfall (e.g., by areal reduction factors [Sivapalan and Blöschl, 1998]) and convolution of catchment rainfall with the unit hydrograph. The former step is a spatial filter, the latter step a temporal filter. Both are represented in the space-time filter of top kriging that is used to estimate the gamma values and hence the kriging weights.

3.6. Estimation of the Spatiotemporal Point Variogram

[32] For applying top kriging, a spatiotemporal point variogram is needed in the region of interest. We estimated the point variogram from the runoff data in the Innviertel region in the following way. In a first step, we estimated temporal cross variograms $\hat{\gamma}_{ij}(h_t)$ for all pairs of catchments i and j within the Innviertel region, which resulted in a family of temporal variograms, one for each pair of catchments:

$$\hat{\gamma}_{ij}(h_t) = \frac{1}{2n(h_t)} \sum_{i=1}^{n(h_t)} (q(\mathbf{x}_i, t_i + h_t) - q(\mathbf{x}_j, t_i))^2 \quad (17)$$

where $q(\mathbf{x}_i, t_i)$ is runoff at time t_i of stream gauge i with spatial location \mathbf{x}_i , h_t is the temporal lag, and $n(h_t)$ is the number of pairs of runoff measurements in time for the temporal separation bin associated with h_t .

[33] In a second step, we estimated theoretical gamma values $\gamma_{ij}(h_t)$ for pairs of catchments from a theoretical point variogram γ_{st} using the same regularization method as in equation (11). The parameters of the point variogram equation (12) ($a, b, c, d, a_s, a_t, b_s, b_t$) are initially unknown as are the parameters of equation (10) (μ, κ). To obtain these parameters, we fitted the theoretical gamma values $\gamma_{ij}(h_t)$ to the sample variograms $\hat{\gamma}_{ij}(h_t)$ by minimizing the objective function Φ using the shuffle complex evolution method [Duan et al., 1992]:

$$\Phi = \frac{2}{N(N+1)M} \sum_{i=1}^N \sum_{j=i}^N \sum_{m=1}^M \cdot \min \left\{ \left[\frac{\hat{\gamma}_{ij}(h_t)}{\gamma_{ij}(h_t)} - 1 \right]^2, \left[\frac{\gamma_{ij}(h_t)}{\hat{\gamma}_{ij}(h_t)} - 1 \right]^2 \right\} \quad (18)$$

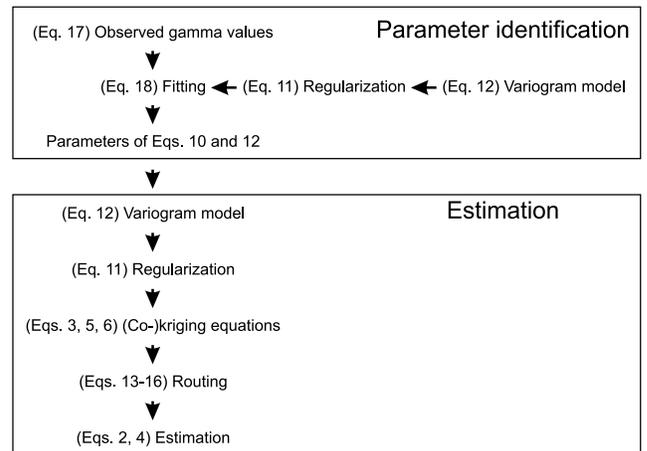


Figure 5. Flow scheme of the estimation process presented in this paper.

Table 2. Parameters of the Point Variogram (Equation (12)) and Equation (10) Estimated for the Innviertel Region

| Parameter | Value | Units |
|-----------|---------|-------------------|
| a | 0.00139 | $m^6 s^{-2} km^4$ |
| b | 0.445 | |
| c | 0.300 | $km hr^{-1}$ |
| d | 2.31 | km |
| a_s | 0.00003 | $m^6 s^{-2} km^4$ |
| a_t | 0.00009 | $m^6 s^{-2} km^4$ |
| b_s | 0.0247 | - |
| b_t | 0.186 | - |
| μ | 2.90 | hours |
| κ | 0.167 | - |

[34] N is the total number of catchments and M is the total number of temporal bins. Equation (18) is a modified version of the weighted least squares (WLS) method. The WLS method as introduced by Cressie [1985] only contains the first term of $\min\{\dots\}$ in equation (18). This gives asymmetrical errors, as errors of overestimation are limited to 1, while errors of underestimation can be very large and can mask the errors from other catchment pairs and temporal lags. The modified version used here limits all errors to the range from 0 to 1. Equation (18) is normalized to also give in the range from 0 to 1.

[35] Initial tests indicated that the type of routing method did not change the parameter values of the point variogram much. The parameters for all methods are hence estimated from cross variograms inferred without routing.

[36] As guidance for the reader, Figure 5 presents a flow scheme of the estimation process.

3.7. Evaluation of Methods

[37] To examine the predictive performance of the proposed method for ungauged catchments we performed a cross-validation analysis. We withheld one runoff record from the data set, estimated the runoff time series for that catchment from measurements of the neighboring stream gauges and finally compared the estimates with the runoff data of that catchment. As performance statistics we calculated the model efficiency (ME_i) according to Nash and Sutcliffe [1970] for each target catchment i :

$$ME_i = 1 - \frac{\sum_{\omega=1}^{\Omega} (q_i(\omega) - \hat{q}_i(\omega))^2}{\sum_{\omega=1}^{\Omega} (q_i(\omega) - \overline{q_i(\omega)})^2} \quad (19)$$

where Ω is the number of time steps estimated ($\Omega = 87672$ for the 10 years analyzed in this paper) and $\overline{q_i(\omega)}$ is the mean of the runoff data for the same time period. The efficiency is less equal unity, where $ME = 1$ indicates perfect estimation and $ME = 0$ means that the estimation method performs no better than the mean of the runoff data. In addition to the Nash-Sutcliffe efficiency, we examined the estimated hydrographs visually to understand the dynamics of the estimated runoff time series.

[38] A total of 10 estimation variants were tested; two routing models (no routing and routing model for all catchments), five methods of top kriging (spatial kriging,

and spatiotemporal kriging and spatiotemporal cokriging with five and nine time steps) and combinations thereof.

[39] The proposed method is an alternative to deterministic runoff models that use regionalized model parameters. To illustrate the relative merits of the two genres of methods we compared the top kriging estimates with simulations of a deterministic rainfall-runoff model taken from the study of Parajka *et al.* [2005]. The runoff model is a conceptual soil moisture accounting scheme that uses precipitation and air temperature data as inputs and runs on a daily time step. It consists of a snow routine, a soil moisture routine and a flow routing routine and involves 14 model parameters. Three of the parameters were preset in their study, leaving 11 parameters to be found by model calibration. Parajka *et al.* [2005] first calibrated the model to 320 catchments in Austria. They then regionalized the calibrated model parameters by different methods and examined the model performance for the ungauged catchment case by cross validation. We contrast their results (both locally calibrated and regionalized) with the results from top kriging. As their analysis is based on a daily time step, we averaged the hourly top kriging estimates to daily values and compared the daily runoff time series. A first comparison focuses on the Innviertel region and involves 17 stream gauges that are common to both studies. To provide context, a second comparison examines 208 stream gauges in Austria that are common to both studies. We used the results from their calibration period (1987–1997) which has 74% overlap with the period used in this study (1990–2000).

4. Results

4.1. Estimation of Point Variogram

[40] The estimated parameters of the point variogram are shown in Table 2. The parameters are similar to those obtained by Skoien and Blöschl [2006] for 488 catchments in Austria although differences exist. For example, the spatial correlation length of $d = 2.3$ km found here is somewhat larger than that of Skoien and Blöschl [2006] (1.0 km) which

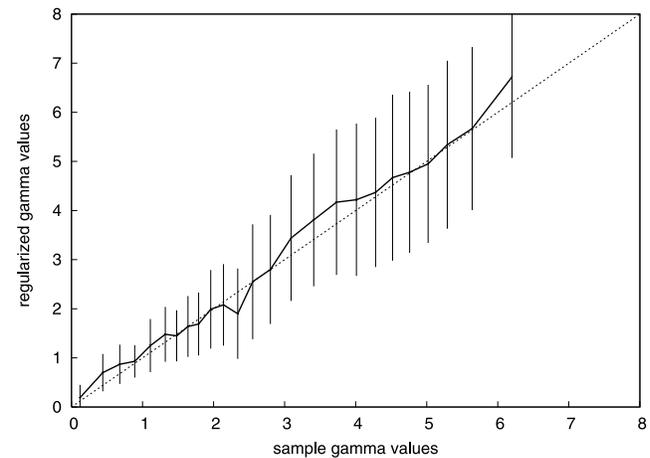


Figure 6. Comparison of sample gamma values (equation (17)) with gamma values obtained by regularizing the point variogram (equation (12) and Table 2) in terms of the mean (thick solid line) and standard deviation (error bars). Dashed line shows 1:1 line. Units are $m^6 s^{-2} km^{-4} \times 10^{-4}$.

Table 3. Model Efficiencies of Hourly Top Kriging Estimates of Runoff in the Innviertel Region for the Period 1 August 1990 to 31 July 2000^a

| Catchment | No Routing | | | | | Routing All Model | | | | |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------|------|-------------|-------------|
| | S | STK1 | STK2 | STC1 | STC2 | S | STK1 | STK2 | STC1 | STC2 |
| Ried | 0.76 | 0.77 | 0.74 | 0.76 | 0.76 | 0.75 | 0.77 | 0.74 | 0.76 | 0.76 |
| Danner | 0.76 | <i>0.78</i> | 0.76 | 0.76 | 0.76 | 0.79 | 0.80 | 0.78 | 0.78 | 0.78 |
| Osternach | 0.82 | 0.79 | 0.70 | 0.81 | 0.81 | 0.82 | 0.79 | 0.70 | 0.82 | 0.82 |
| Pram | 0.80 | 0.74 | 0.67 | 0.76 | 0.76 | 0.80 | 0.75 | 0.67 | 0.77 | 0.77 |
| Riedau | 0.84 | <i>0.88</i> | 0.85 | 0.87 | 0.87 | 0.87 | 0.90 | 0.87 | 0.89 | 0.89 |
| Winertsham | 0.88 | <i>0.90</i> | 0.83 | 0.89 | 0.89 | 0.91 | 0.92 | 0.84 | 0.93 | 0.93 |
| Taufkirchen | <i>0.94</i> | <i>0.94</i> | 0.89 | 0.92 | 0.92 | 0.96 | 0.96 | 0.89 | 0.96 | 0.96 |
| Angsüß | 0.77 | <i>0.79</i> | 0.77 | 0.74 | 0.75 | 0.82 | 0.84 | 0.80 | 0.79 | 0.80 |
| Alfersham | 0.73 | 0.70 | <i>0.76</i> | 0.56 | 0.60 | 0.78 | 0.70 | 0.77 | 0.53 | 0.57 |
| Lohstampf | 0.84 | 0.83 | 0.73 | <i>0.86</i> | 0.85 | 0.88 | 0.83 | 0.73 | 0.86 | 0.86 |
| Still | 0.89 | 0.90 | 0.81 | 0.91 | 0.91 | 0.86 | <i>0.87</i> | 0.82 | <i>0.87</i> | <i>0.87</i> |
| Strötting | <i>0.84</i> | 0.83 | 0.76 | 0.82 | 0.81 | 0.87 | 0.82 | 0.77 | 0.83 | 0.83 |
| Bad Schallerbach | <i>0.92</i> | 0.90 | 0.83 | 0.91 | 0.91 | 0.93 | 0.91 | 0.84 | 0.91 | 0.91 |
| Pichl | <i>0.80</i> | 0.70 | 0.59 | 0.66 | 0.62 | 0.82 | 0.72 | 0.58 | 0.67 | 0.63 |
| Weghof | <i>0.83</i> | 0.43 | 0.44 | 0.27 | 0.22 | 0.85 | 0.45 | 0.44 | 0.29 | 0.24 |
| Fraham | <i>0.89</i> | 0.84 | 0.85 | 0.81 | 0.81 | 0.92 | 0.90 | 0.86 | 0.90 | 0.89 |
| Neumarkt | 0.89 | 0.85 | 0.76 | 0.87 | 0.87 | 0.88 | 0.86 | 0.77 | 0.89 | 0.89 |
| Niederspaching | 0.92 | 0.83 | 0.75 | 0.86 | 0.87 | 0.92 | 0.86 | 0.75 | 0.89 | 0.89 |
| Kropfmühle | <i>0.90</i> | 0.88 | 0.83 | 0.88 | 0.88 | 0.91 | 0.88 | 0.83 | 0.87 | 0.86 |
| Average | <i>0.84</i> | 0.80 | 0.75 | 0.79 | 0.78 | 0.86 | 0.82 | 0.76 | 0.80 | 0.80 |
| Median | <i>0.84</i> | 0.83 | 0.76 | 0.82 | 0.81 | 0.87 | 0.84 | 0.77 | 0.86 | 0.86 |

^aS refers to spatial kriging; STK1 and STK2 refer to spatiotemporal kriging with five and nine time steps, respectively; and STC1 and STC2 refer to spatiotemporal cokriging with five and nine time steps, respectively. The largest model efficiency (ME) for each catchment is shown in bold, and the largest ME for each catchment and the routing/no routing model is shown in italics.

may be due to the somewhat more homogeneous runoff in the Innviertel region as compared to the rest of Austria. However, it must be noted that catchments are correlated also on distances longer than the here indicated 2.3 km, as a result of the regularization of the point variogram. The exponent κ relating temporal and spatial supports in (equation (10)) is smaller ($\kappa = 0.17$ instead of 0.4) and the scale μ is similar (2.9 as compared to 1.9) indicating that the catchments in the Innviertel respond somewhat faster than the average of the catchments in Austria. For example, the temporal support of a 100 km² catchment found in this study is $T_i = 6$ hours, as compared to 12 hours found by Skøien and Blöschl [2006]. For consistency across the two scales (Innviertel and Austria) we however used the parameters of Table 2 for all analyses in this paper.

[41] Figure 6 shows a comparison of all the sample gamma values for different catchment pairs and different temporal lags found from equation (17) and the gamma values obtained by regularizing the point variogram (equation (12), Table 2). The sample gamma values have been grouped into bins, and the mean and standard deviation of the corresponding gamma values obtained from regularizing the point variogram are shown as a line and the error bars, respectively. The regularized gamma values exhibit some scatter as indicated by the error bars, but they are very close to unbiased. This suggests that the point variogram parameters of Table 2 describe a realistic representation of the space time variability of runoff.

[42] As estimates of the local uncertainty in equations (3), (5) and (6) were unavailable in this study, we assumed a value of $\sigma_j^2 = 0.000005 \text{ m}^6 \text{ s}^{-2} \text{ km}^{-4}$ on the basis of test

simulations. This value is about 1% of the average temporal variance of runoff in the Innviertel region (Table 1).

4.2. Estimation Performance

[43] Table 3 shows model efficiencies (ME) of the hourly runoff time series estimated by top kriging in the Innviertel region. The largest model efficiency (ME) for each catchment is shown in bold. The largest ME for each catchment for the routing model (no routing or routing all) that did not have the highest ME is shown in italics.

[44] Considering spatial kriging first, ME ranges between 0.73–0.96 with an average of 0.84 for no routing and 0.86 for the routing all model. The routing all model gives the highest ME for 13 out of the 19 catchments and the no routing model gives the highest ME for three catchments indicating that a routing model does improve the runoff estimates over the variant where no routing is used. When we include the results from spatiotemporal kriging and spatiotemporal cokriging, the advantage of the routing model is similar, with the routing all model giving the highest ME for 15 catchments, whereas the variant without routing performs best only for one catchment.

[45] A comparison of the efficiency of spatial kriging with that of spatiotemporal kriging and cokriging for the routing all model indicates that spatial kriging outperforms or is as good as the other variants of kriging for 12 catchments. In the remaining seven catchments the efficiency of spatiotemporal kriging or cokriging is in most cases only marginally higher than that of spatial kriging, whereas one or more models of spatiotemporal kriging and cokriging perform considerably poorer than spatial kriging for almost all catchments. The median and the average show a similar

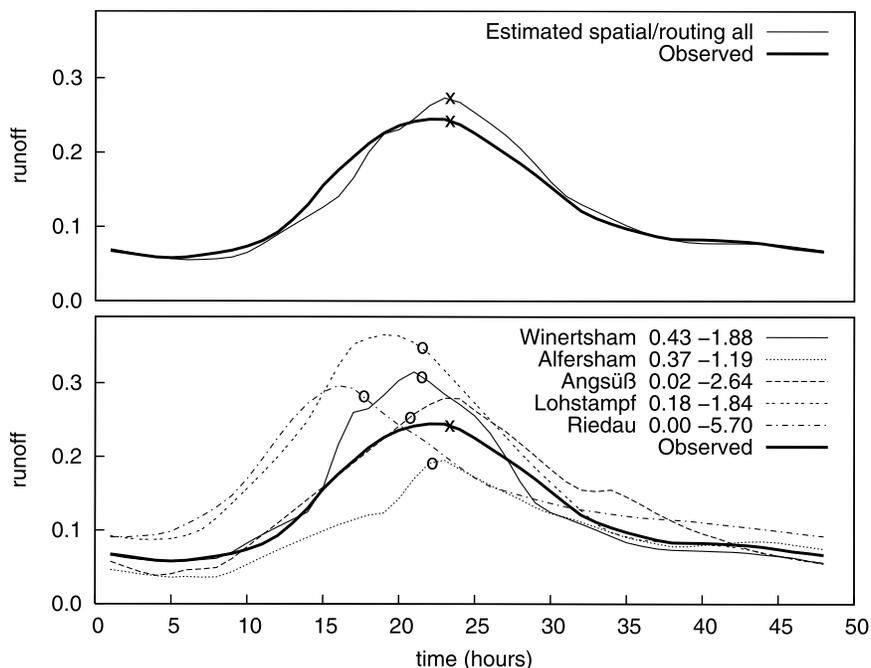


Figure 7a. Observed and estimated hydrographs for (top) Taufkirchen (303 km²) and (bottom) the neighbors for the period 17–18 March 1993. The first number after the neighbor name represents the kriging weight of the neighbor, and the second number represents the lag of the routing model in hours. Estimate is from spatial kriging, routing all model. Units of runoff are m³ s⁻¹ km⁻². Points illustrate example in text.

tendency for spatial kriging to outperform spatiotemporal kriging and spatiotemporal cokriging.

[46] The best runoff estimates are obtained for those catchments that have one or more close upstream neighbors. The highest ME is found for Taufkirchen (ME around 0.96, depending on the choice of routing model). Upstream gauges cover 248 km² of the 303 km² of the catchment. It is therefore not surprising that a weighted average of the upstream neighbors gives an estimated hydrograph that is similar to the observed hydrograph. The more of the catchment area that is shared with neighboring catchments the larger the efficiency. ME is around 0.9 for most of the catchments with several upstream neighbors, or which are a large part of a downstream catchment (e.g., Winertsham, Kropfmühle, Fraham, Bad Schallerbach). The catchments without such neighbors generally have ME on the order of 0.8 or lower (e.g., Ried, Osternach, Danner, Pram, Pichl).

4.3. Runoff Dynamics

[47] To analyze the ability of top kriging to reproduce the runoff dynamics we examined a large number of event hydrographs visually. A few examples are given below to illustrate the main characteristics of top kriging. The illustration begins with estimates of the spatial kriging method. Taufkirchen is the catchment with the highest model efficiency (ME) according to Table 3. One event is presented in Figure 7a. Although there is a large variety in the runoff from the neighboring catchments (bottom plot) and the peaks occur at different times, top kriging with routing for all catchments is able to estimate both the magnitude and the timing of the peak with high accuracy (top plot).

[48] The bottom plot of Figure 7a gives the kriging weights (first number) and the time lags (second number)

of the neighbors used. A positive time lag refers to a downstream or larger catchment (use of future measurements) while a negative number refers to an upstream or smaller catchment (use of earlier measurements). Winertsham (128.1 km²) is the largest tributary (Figure 2), and is associated with the largest weight of 0.43. Alfershams (81.3 km²) is smaller and has a weight of 0.37, while its upstream neighbor, Angsüß, is associated with a weight of

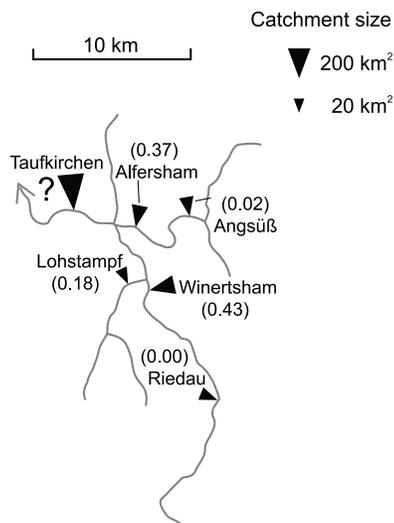


Figure 7b. Weights of neighboring catchments for estimating runoff at Taufkirchen from Figure 7a.

Table 4. Example Estimation of Runoff at Taufkirchen (Time Step 23 Hours) From Runoff at the Neighboring Catchments as in Figure 7

| Neighbor | Weight | Time Lag, hours | Time, hours | Observation, $\text{m}^3 \text{s}^{-1} \text{km}^{-2}$ | Weighted Observation, $\text{m}^3 \text{s}^{-1} \text{km}^{-2}$ |
|--------------|--------|-----------------|-------------|--|---|
| Winertsham | 0.43 | -1.88 | 21.12 | 0.31 | 0.133 |
| Alfersham | 0.37 | -1.19 | 21.81 | 0.19 | 0.070 |
| Angsüß | 0.02 | -2.64 | 20.36 | 0.26 | 0.005 |
| Lohstampf | 0.18 | -1.84 | 21.16 | 0.35 | 0.063 |
| Riedau | 0.00 | -5.70 | 17.30 | 0.28 | 0.000 |
| Sum=Estimate | | | | | 0.271 |

0.02. This very low weight is a result of the large correlation between Alfersham and Angsüß. The sum of weights from this tributary is then 0.39. The smallest tributary is Lohstampf (39.3 km^2) with a weight of 0.17. Riedau has a weight of -0.001 . Angsüß and Riedau are upstream neighbors of Alfersham and Winertsham respectively, and each pair of nested catchments can be regarded as a clustered sample, reducing the weight for the upstream neighbor in this case. Figure 7b illustrates the weights given to the different neighboring catchments.

[49] The time lags are all negative for Taufkirchen. This is because all catchments are upstream neighbors, and a peak flow will reach these catchments before it reaches Taufkirchen. The time lag is largest for Riedau, which is furthest away (13.7 km), with 5.70 hours. The smallest time lag is for Alfersham (2.9 km away) with 1.19 hours.

[50] To illustrate the method, one target time step (23 hours) has been marked by a cross in the top plot of Figure 7a. The observation at this time step is $0.245 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$ but it is assumed not to be known. Rather the runoff is to be estimated from the neighboring catchments. The observations of the neighboring catchments (circles in Figure 7a) that are used to obtain this estimate are shifted by a time lag as represented by the routing model. The weighted average of the observations (circles) then gives the final estimate for

Taufkirchen ($0.271 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$), indicated by a cross in Figure 7a. The calculation is illustrated in Table 4. For this time step, top kriging slightly overestimates runoff.

[51] Although top kriging works best with several upstream neighbors, the estimates are also satisfying for the smaller catchments. Figure 8 shows the observed and the estimated hydrographs of an event in August 1991 for the 52 km^2 Strötting catchment. This is a catchment without upstream neighbors, but with two downstream neighbors, Bad Schallerbach (184 km^2) and Fraham (362 km^2). As Bad Schallerbach is the first downstream neighbor, it gets the largest weight with 0.61. This is partly compensated by a negative weight of -0.10 for Fraham which is the downstream neighbor of Bad Schallerbach, so that the total weight of the downstream measurements is 0.51. These catchments are clustered, and top kriging gives a negative weight to the one least correlated with Strötting. The non-nested neighboring catchment Pichl (66 km^2) gets a weight of 0.42, partly compensated by the weight of -0.15 of its downstream neighbor, Weghof (117 km^2). Danner (56 km^2) is slightly smaller than Pichl and gets a weight of 0.22.

[52] The most important stream gauge for the estimation of runoff at Strötting is Bad Schallerbach. The peak of this catchment is quite similar to the peak of Strötting, but slightly delayed. The other catchments east and south of

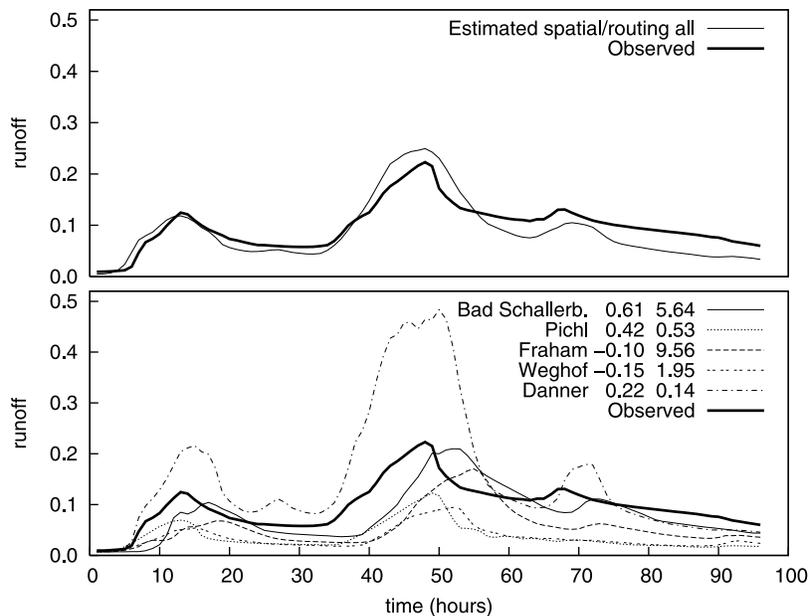


Figure 8. Observed and estimated hydrographs for (top) Strötting (52 km^2) and (bottom) the neighbors for the period 1–4 August 1991. The first number after the neighbor name represents the kriging weight of the neighbor, and the second number represents the lag of the routing model in hours. Estimate is from spatial kriging, routing all model. Units of runoff are $\text{m}^3 \text{ s}^{-1} \text{ km}^{-2}$.

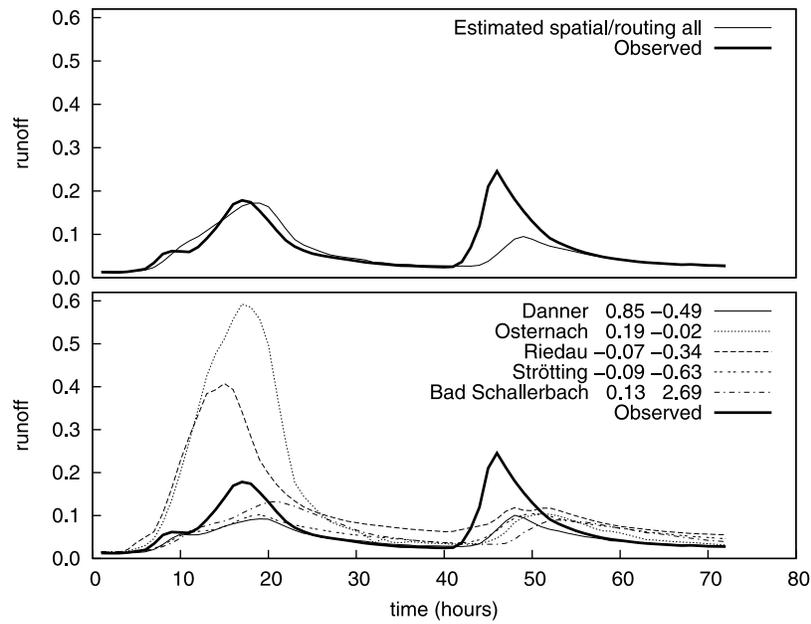


Figure 9. Observed and estimated hydrographs for (top) Ried (69 km²) and (bottom) the neighbors for the period 29 October to 1 November 1998. The first number after the neighbor name represents the kriging weight of the neighbor, and the second number represents the lag of the routing model in hours. Estimate is from spatial kriging, routing all model. Units of runoff are m³ s⁻¹ km⁻².

Strötting show considerably lower peaks, while the non-nested Danner catchment has a peak more than twice of that at Strötting. The weighted average gives a reasonably well-estimated peak. The close match is also a result of the time lags applied. A time lag of 5.64 hours was applied to Bad Schallerbach, which is also apparent in Figure 8. As Pichl,

Weghof and Danner are not nested with Strötting, their time lags have been estimated by equations (14)–(15). Although the time lags are considerably smaller than those of Bad Schallerbach and Weghof, Figure 8 does confirm that the time lags are reasonable.

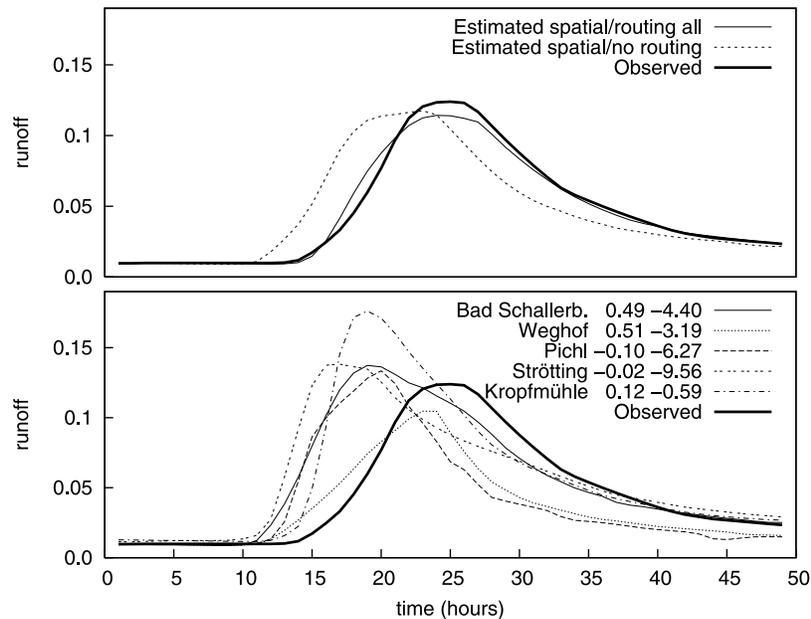


Figure 10. Observed and estimated hydrographs for (top) Fraham (362 km²) and (bottom) the neighbors for the period 23–25 October 1993. The first number after the neighbor name represents the kriging weight of the neighbor, and the second number represents the lag of the routing model in hours. Estimate is from spatial kriging, routing all model. Units of runoff are m³ s⁻¹ km⁻².

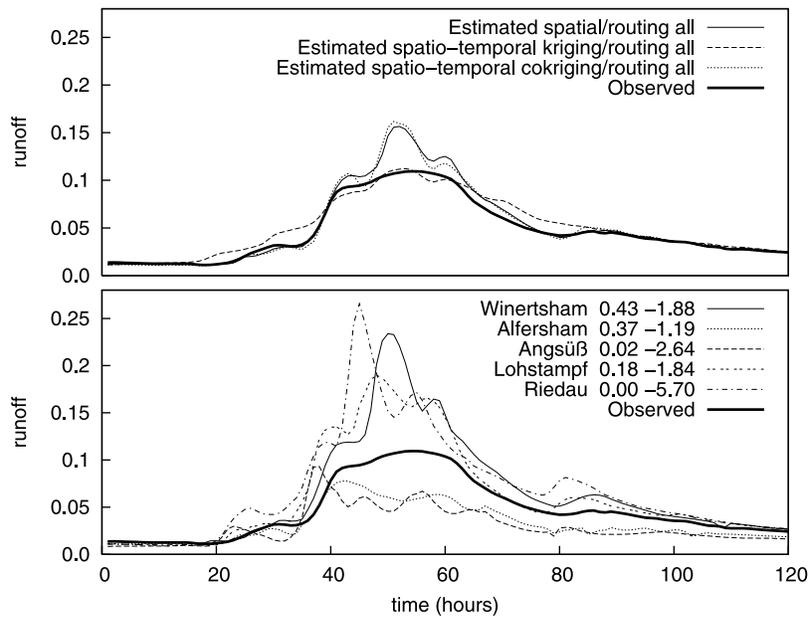


Figure 11. Observed and estimated hydrographs for (top) Taufkirchen (303 km²) and (bottom) the neighbors for the period 18–23 November 1990. The first number after the neighbor name represents the kriging weight of the neighbor, and the second number represents the lag of the routing model in hours. Estimate is from spatial kriging, spatiotemporal kriging (nine time steps), and spatiotemporal cokriging (nine time steps). Units of runoff are m³ s⁻¹ km⁻².

[53] There are also some events where the top kriging estimates differ significantly from the observed hydrographs. This typically happens for local events that do not appear in the hydrographs of the neighboring catchments. One example for Ried (69 km²) is presented in Figure 9. This catchment has neither nested upstream nor downstream neighbors. In the three day period shown two events occurred. The first event (shown at time 18 hours) was properly estimated. The second event (shown at time 48 hours) was not. This is because it was apparently the result of a local precipitation event. None of the surrounding catchments experienced more than a small change in runoff for this period. It is interesting to note that the local event was also missed by the rain gauges in the region, so rainfall-runoff models could not have simulated this event either.

[54] The merits of the routing model, in many cases, are obvious in the hydrographs. The top plot of Figure 10 shows the estimated hydrograph for Fraham (362 km²) with and without the use of a routing model. The time lags presented in the bottom plot show that all the neighbors have their peaks earlier than Fraham. The largest weights are given to the two large tributaries, Bad Schallerbach (184 km²) with 0.49 and Weghof (117 km²) with 0.51. The larger weight for Weghof is partly compensated by the negative weight of -0.10 of Pichl, giving 0.41 as the total weight for this tributary. The upstream neighbor of Bad Schallerbach, Strötting (52 km²), has a negative weight of -0.02. The nonnested neighbor of equal size, Kropfmühle (313 km²), has a weight of 0.12. Fraham is a large catchment, so the time lags are considerably larger than those of the earlier examples. Strötting is 23 km from Fraham and the time lag is 9.6 hours. The other nested stream gauges are closer, with time lags ranging from 3 to 6 hours.

[55] The model efficiencies presented in Table 3 indicate that spatiotemporal kriging and spatiotemporal cokriging

usually do not perform as well as spatial kriging. However, there are events where the spatiotemporal approach performs better than spatial kriging. One example is an event at Taufkirchen in November 1990, presented in Figure 11. The top plot shows that spatial kriging and spatiotemporal cokriging are quite similar and overestimate the peak. The hydrograph estimated by spatiotemporal kriging on the other hand is closer to the observed hydrograph, although it is slightly too flat at the beginning and the end of the event. There are two reasons for this effect. First, spatiotemporal kriging, in this case, involves weights for nine time steps ($t_\alpha - t_\omega$ ranging from -20 to 20 hours) as opposed to spatial kriging where all weights are associated with the same time step ($t_\alpha = t_\omega$). As the weights are distributed over more time steps, the estimation method more strongly smoothes peaks observed at the neighboring catchments. Second, the weights were rescaled (see appendix A). This tended to reduce the weights of the closest neighbors and increase the weights of the more remote neighbors. Summed over all time steps, the weights associated with the neighbors of Taufkirchen ranged from 0.15 to 0.26 instead of 0.00 to 0.43 for spatial kriging. These two effects consistently gave smoother hydrographs than spatial kriging. In many cases they were smoother than the observed hydrographs.

4.4. Temporal Average and Variance of the Time Series

[56] When estimating runoff time series of ungauged catchments it is important to represent the statistical moments of the temporal variability well. The left plot of Figure 12 presents the average runoff estimated by spatial kriging plotted against the observed averages for the Innviertel catchments. There is a good correspondence with a slight tendency for the small averages to be overestimated and the larger averages to be underestimated. There is no

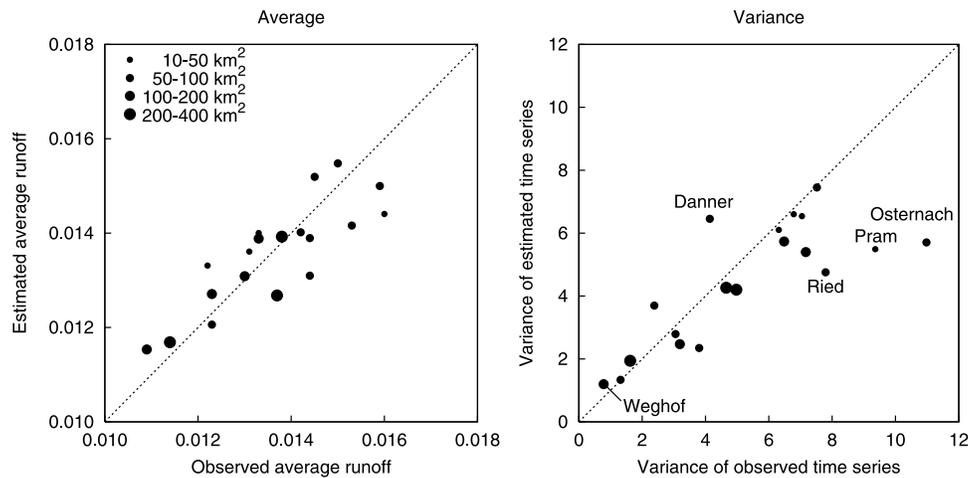


Figure 12. (left) Average runoff estimated by top kriging (spatial kriging, routing all model) plotted against average runoff of the observed time series for the 19 catchments in the Innviertel region. (right) Temporal variance of the estimated time series plotted against the variances of the observed time series. The size of the points indicates catchment area. Units are $\text{m}^3 \text{s}^{-1} \text{km}^{-2}$ for average runoff and $\text{m}^6 \text{s}^{-2} \text{km}^{-4} \times 10^{-4}$ for the variance.

obvious relationship between this tendency and catchment area.

[57] The right plot of Figure 12 presents an analogous plot for the temporal variance of runoff. Overall, there is a relatively good correspondence between the estimated and observed variances although the largest variances are underestimated. This is particularly the case for the Ried, Osternach and Pram catchments. These are headwaters, and Osternach and Ried do not have nested neighbors at all. Danner is located between Osternach and Ried but has a considerably lower temporal variance which is, conversely, overestimated. Clearly, top kriging has difficulties with representing local hydrological effects well that are not reflected in the runoff data of neighboring catchments. When examining a hydrogeological map [Schubert *et al.*, 2003] there are apparent hydrological reasons for the differences in the temporal variances of these catchments. Danner has a large number of artesian springs while Osternach has none. The subsurface contribution to runoff in Danner is hence larger than that in Osternach. Some of the Danner catchment area is forested with the remaining land being agricultural while all of the Osternach catchment is agricultural which is likely another factor contributing to the differences in the dynamics. Out of the catchments in the Innviertel region, the Weghof catchment has the smallest temporal variance, which can be estimated well even though the low variance is due to local gravel deposits in the valleys of the catchment.

[58] The underestimation of the large temporal variances is more pronounced for spatiotemporal kriging and spatiotemporal cokriging than for spatial kriging as illustrated in Figure 13. In 14 out of the 19 catchments, spatiotemporal kriging and cokriging underestimate the variance, which is clearly not a desirable property. Figure 13 indicates that spatiotemporal cokriging retains more of the observed variance than spatiotemporal kriging. It also indicates that the spatiotemporal estimation methods with 5 time steps retain more of the variance than those using 9 time steps. It is particularly the catchments with large observed variances

where the estimated time series have less variance than the observed time series. For catchments with small variances, the estimates tend to be slightly too large and there are only minor differences between the different spatiotemporal estimation methods.

4.5. Comparison With a Deterministic Rainfall-Runoff Model

[59] The runoff time series estimated by top kriging were finally compared with the simulation results of a deterministic rainfall-runoff model obtained by Parajka *et al.* [2005]. For comparison, the top kriging estimates were aggregated to daily values so there is almost no difference between the routing models. The results are hence only presented for spatial kriging with the routing all model. Out of the 19 catchments in the Innviertel, 17 were part of the data set of Parajka *et al.* [2005].

[60] The model efficiencies (ME) for top kriging and the deterministic model are shown in Table 5. The first column gives ME for the deterministic model calibrated at site. This is the case where the parameters have been obtained by fitting the runoff simulations directly to the runoff data; that is, it is the gauged catchment case. The second column gives the ungauged catchment case of the deterministic model using parameters that have been regionalized from the neighboring catchments by kriging. This is the regionalization method of Parajka *et al.* [2005] that performed best; see Table 2, line 6, and Figure 9, top left green line, of Parajka *et al.* [2005]. The at site calibration efficiency is in the range of 0.58–0.79, with an average of 0.71. When using regionalized parameters, the efficiency decreases to a range of 0.43–0.66, with an average of 0.58. This is quite a significant decrease which is a result of the uncertainty associated with the ungauged catchment problem. The third column of Table 5 shows the efficiencies for top kriging which, again, is an ungauged catchment case. Note that the top kriging efficiencies in Table 5 are considerably larger than those in Table 3. This is because errors in estimating the time of peak and the dynamics of the runoff hydrograph

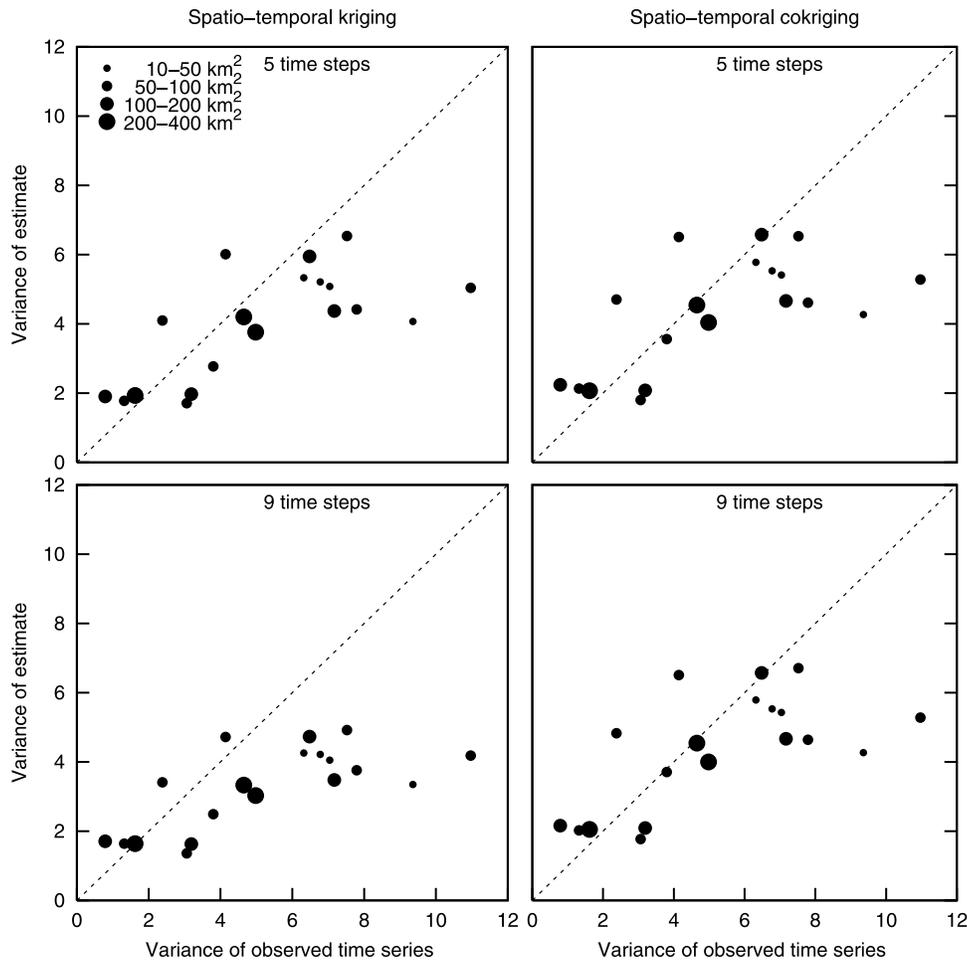


Figure 13. Temporal variance of the runoff time series estimated by various top kriging variants plotted against the variances of the observed time series. (left) Spatiotemporal kriging with (top) five and (bottom) nine time steps. (right) Spatiotemporal cokriging with (top) five and (bottom) nine time steps. Units are $\text{m}^6 \text{s}^{-2} \text{km}^{-4} \times 10^{-4}$.

will be less important for the daily averages of Table 5 than for the hourly values of Table 3. Top kriging gives efficiencies in the range of 0.87–0.98 with an average of 0.93. This is remarkably better than the corresponding results of the deterministic model (average of 0.58). In fact, for every single catchment the top kriging estimates are much better than the corresponding deterministic estimates of regionalization. This is not surprising for, say, Taufkirchen where 80% of the catchment area is covered by upstream gauges. However, the much better performance is true for all catchments including those without upstream or downstream neighbors.

[61] To illustrate the differences of the two approaches in terms of representing the runoff dynamics Figure 14 shows an example of an event for the Kropfmühle catchment (313 km^2). The solid line is the runoff time series estimated by top kriging. The dashed and dotted lines are the deterministic model results based on at site and regionalized parameters, respectively. The observed runoff is shown as points. The top kriging estimates are very close to the observed hydrograph for the entire period including the peak of the events and the recession. The deterministic model results tend to underestimate the peak. Interestingly,

the first event is better captured by the regionalized model parameters while the second event is better captured by the locally calibrated parameters. It is also of interest that top kriging tends to estimate the low flow period better than either of the two deterministic model setups.

[62] Although the hydrogeology of the Innviertel region is complex, average annual runoff is rather uniform (Table 1). It is therefore of interest to put the results of the Innviertel into the context of all of Austria, which is hydrologically much more diverse. For the entire Austrian data set of 367 catchments, the median model efficiency of top kriging in terms of estimating hourly runoff time series is 0.75 (Table 6). This is lower than the median of 0.87 for the Innviertel region where the stream gauge density is higher. It is also of interest to compare the results to the deterministic results of *Parajka et al.* [2005]. There were a total of 208 catchments that were common to *Parajka et al.* [2005] and this paper. For these catchments, the regionalized deterministic model gives a median efficiency is 0.67 while it is 0.87 for top kriging, both on the basis of daily runoff time series for the ungauged catchment case (Table 6). This means that top kriging performs much better than the deterministic model. To illustrate the between-catchment

Table 5. Model Efficiencies of Daily Runoff Estimated by Top Kriging (Spatial Kriging, Routing All Model) and the Deterministic Runoff Model in the Innviertel Region^a

| Stream Gauge | Deterministic Model at Site | Deterministic Model Regionalized (PUB) | Top Kriging (PUB) |
|------------------|-----------------------------|--|-------------------|
| Ried | 0.71 | 0.56 | 0.87 |
| Osternach | 0.69 | 0.60 | 0.88 |
| Pram | 0.74 | 0.50 | 0.90 |
| Riedau | 0.77 | 0.64 | 0.94 |
| Winertsham | 0.70 | 0.55 | 0.96 |
| Taufkirchen | 0.73 | 0.62 | 0.98 |
| Angsüß | 0.73 | 0.63 | 0.92 |
| Alfersham | 0.71 | 0.66 | 0.90 |
| Lohstampf | 0.67 | 0.59 | 0.92 |
| Still | 0.74 | 0.57 | 0.94 |
| Strötting | 0.75 | 0.62 | 0.91 |
| Bad Schallerbach | 0.72 | 0.64 | 0.96 |
| Pichl | 0.58 | 0.43 | 0.89 |
| Weghof | 0.66 | 0.52 | 0.89 |
| Fraham | 0.71 | 0.52 | 0.97 |
| Niederspaching | 0.72 | 0.59 | 0.96 |
| Kropfmühle | 0.79 | 0.61 | 0.94 |
| Average | 0.71 | 0.58 | 0.93 |
| Median | 0.72 | 0.59 | 0.92 |

^aPUB indicates the ungauged catchment case as assessed by cross validation.

variability of model performance, Figure 15 shows the cumulative distribution function (cdf) of the model efficiencies for the deterministic model (both at site calibration and regionalized) and top kriging. Figure 15, again, indicates that top kriging indeed outperforms the regionalized deterministic model substantially. There are a small number of catchments where top kriging gives efficiencies less than zero. A closer examination of these catchments suggested that a number of factors contribute to

the poor performance. The majority of these catchments were close to the border of Austria and/or spatially isolated with a large distance to the nearest stream gauge. Some of the catchments, however, did have close neighbors but the hydrogeology differed significantly from that of the neighboring catchments. A typical example in Austria are gravel deposits in some valleys that increase the groundwater component of runoff and produce much slower response as compared to neighboring catchments without such gravel deposits.

[63] While the top kriging cdf in Figure 15 is always above that of the regionalized deterministic model, the efficiencies do not necessarily relate to the same catchments. In Figure 16, the top kriging efficiencies have been plotted against those of the regionalized deterministic model. The comparison indicates that there are indeed catchments where top kriging performs more poorly than the deterministic model. As indicated above, these are mostly catchments that do not have a stream gauge in their vicinity. However, this is the case for only 15% of the catchments. For the remaining catchments top kriging performs better.

[64] To identify locations with potentially poor estimates, geostatistical analysis generally makes use of the kriging variance. This is also possible for the top kriging approach [Skøien *et al.*, 2006]. For spatial kriging (equations (2)–(3)), the kriging variance is

$$\sigma_{Ri}^2 = \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu \tag{20}$$

[65] In this paper the kriging variance σ_{Ri}^2 was estimated in a cross-validation model; that is, in solving the kriging system to obtain the set of weights λ_j for a given catchment

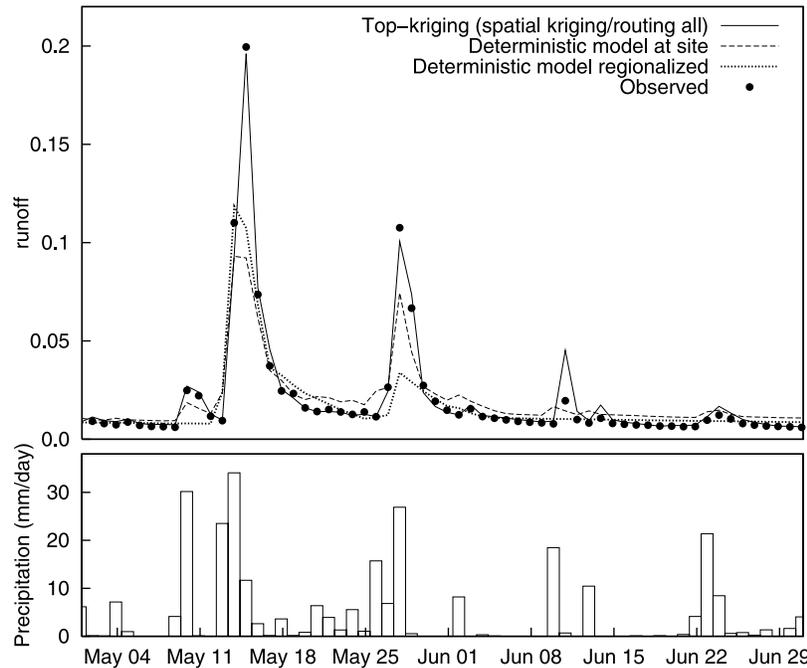


Figure 14. Comparison of runoff time series estimated by top kriging and the deterministic model for the period of 1 May to 30 June 1996 at Kropfmühle (313 km²). The top kriging weights of the neighbors are Niederspaching, 0.53; Fraham, 0.39; Taufkirchen, 0.31; Bad Schallerbach, -0.13; and Winertsham, -0.09. Units of runoff are m³ km⁻²s⁻¹.

Table 6. Median Model Efficiencies for Top Kriging as Compared to the Deterministic Runoff Model for the Ungauged Catchment Case^a

| Region | Number of Catchments | Top Kriging Hourly Runoff | Top Kriging Daily Runoff | Regionalized Deterministic Model, Daily Runoff |
|------------|----------------------|---------------------------|--------------------------|--|
| Innviertel | 19 | 0.87 | 0.92 | |
| Innviertel | 17 | 0.87 | 0.92 | 0.59 |
| Austria | 376 | 0.75 | 0.82 | |
| Austria | 320 | | | 0.68 |
| Austria | 208 | 0.82 | 0.87 | 0.67 |

^aFor 208 catchments, both Top kriging and deterministic model results were available; 17 of these are in the Innviertel region.

i it was assumed that it is ungauged and only runoff data for the neighboring catchments j were available. This means that σ_{Ri}^2 should be an indicator of the uncertainty of estimating runoff in ungauged catchments. To examine this idea, Figure 17 shows the top kriging efficiencies of 376 catchments plotted against the kriging variance. The model efficiencies are indeed well correlated with the kriging variance indicating that the kriging variance can be used as an indicator to identifying catchments with potentially poor estimates. For example, catchments with kriging variances less than $4 \times 10^{-5} \text{ m}^6 \text{ s}^{-2} \text{ km}^{-4}$ are associated with median model efficiencies of hourly time series of 0.9. If the kriging variance is more than $12 \times 10^{-5} \text{ m}^6 \text{ s}^{-2} \text{ km}^{-4}$, median model efficiencies of 0.5 can be expected.

5. Discussion and Conclusions

[66] In this paper we have extended the top kriging approach of Skoien *et al.* [2006] to account for hydro-

dynamic and geomorphologic dispersion as well as routing. The main appeal of the method is that it is a best linear unbiased estimator adapted for the case of stream networks. We hence believe it is the most natural way of statistically estimating runoff time series in ungauged catchments.

[67] Examination of the top kriging results in the Innviertel region suggests that the kriging weights are plausible. In particular, they take into account the upstream and downstream similarities of catchments as well as similarities with catchments that do not share a subcatchment. They also account for clustering effects, giving less weight to clusters of catchments that consist of catchments with highly correlated runoff records.

5.1. Top Kriging Performance

[68] The cross-validation analyses in the Innviertel region indicate that the Nash-Sutcliffe model efficiency of the spatial kriging variant of top kriging is on the order of 0.8–0.95 for hourly runoff time series and 0.85–0.98 for daily runoff time series. The efficiency is slightly lower for the rest of Austria with a median of 0.82 for hourly runoff

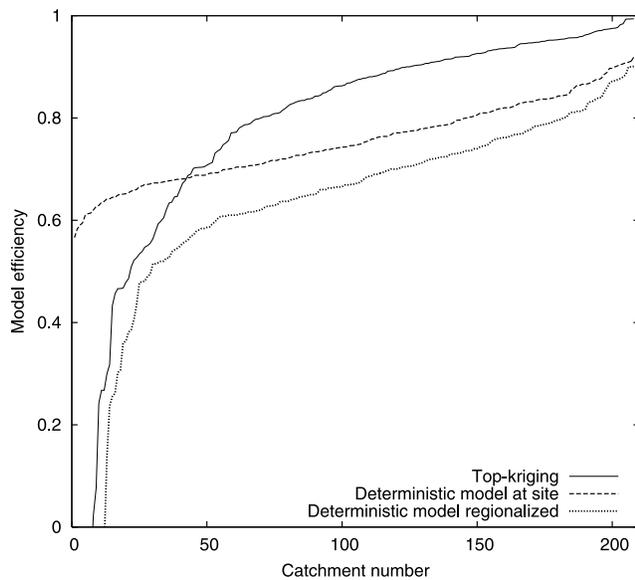


Figure 15. Cumulative distribution functions of model efficiencies (ME) of daily runoff estimated by top kriging (spatial kriging, routing all model) and the deterministic runoff model for 208 catchments in Austria. Top kriging and the regionalized deterministic model relate to the ungauged catchment case as assessed by cross validation.

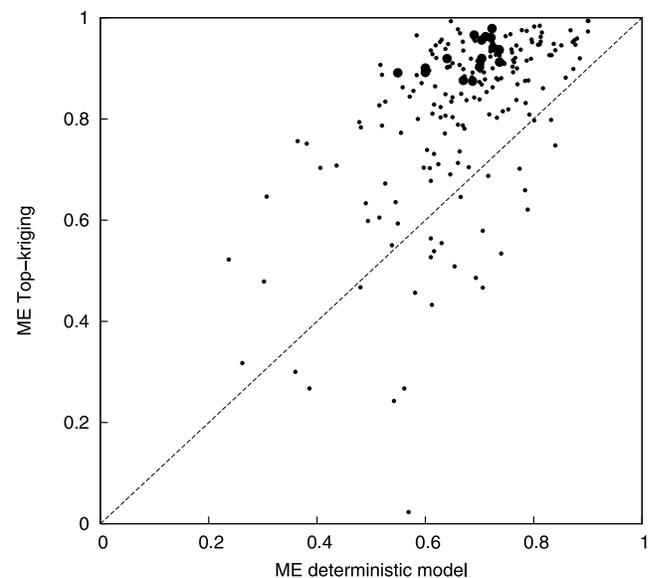


Figure 16. Model efficiencies (ME) of daily runoff estimated by top kriging (spatial kriging, routing all model) and the regionalized deterministic runoff model for 208 catchments in Austria. Large circles indicate the Innviertel catchments. All model efficiencies relate to the ungauged catchment case as assessed by cross validation.

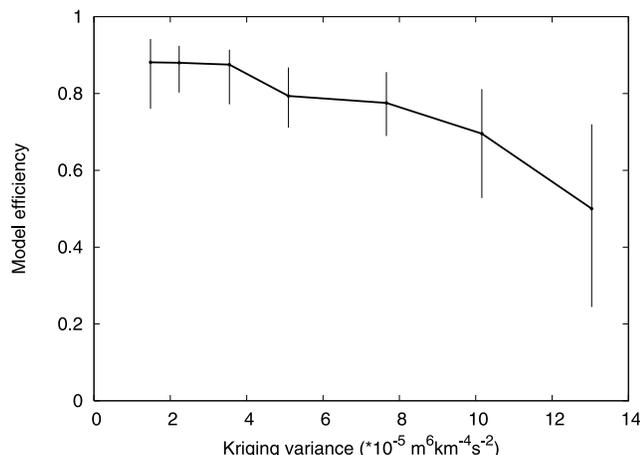


Figure 17. Model efficiencies (ME) of hourly runoff estimated by top kriging (spatial kriging, routing all model) plotted as a function of the kriging variance (equation (20)) for 376 catchments in Austria. The thick line represents the median ME, and the error bars represent the 25 and 75 percentiles. Both model efficiencies and the kriging variance relate to the ungauged catchment case obtained in a cross-validation mode.

and 0.87 for daily runoff for a set of 208 catchments, and 0.75 for hourly runoff and 0.82 for daily runoff for a set of 376 catchments. These efficiencies are excellent as compared to what can usually be obtained by deterministic rainfall-runoff models in ungauged catchments. For the same 208 catchments, the deterministic runoff model of *Parajka et al.* [2005] gave median efficiencies of daily runoff $\text{ME} = 0.67$ if the catchments were treated as ungauged. This means that the top kriging efficiency is 0.20 higher than the corresponding efficiency of a deterministic runoff model.

[69] These efficiencies also compare well with other studies around the world. A review of *Merz et al.* [2007] suggests that, typically, the runoff modeling efficiency of daily runoff in ungauged catchments ranges between 0.6 and 0.7, depending on data availability. If the focus is on simulating runoff with hourly temporal resolution, the efficiency is usually lower, mainly because of poorer rainfall information and errors in the timing of runoff dynamics. An example is the DMIP distributed model intercomparison project [*Reed et al.*, 2004]. In the DMIP project, 12 deterministic runoff models were used to simulate hourly runoff for 8 catchments (65 km^2 to 2484 km^2) in the USA for a period of 7 years (1994–2000). The models were run both in an *at site* mode where local runoff data were used to calibrate model parameters and in an uncalibrated mode that represents the ungauged catchment case. In the *at site* (calibrated parameter case) the median of the hourly model efficiencies over the eight catchments ranged between 0.30 and 0.73, depending on the model with a mean of 0.58. The efficiencies for the uncalibrated case was significantly lower with the median over the eight catchments ranging between -0.50 and 0.61 , depending on the model, with a mean of 0.22. The median efficiencies obtained by top kriging (0.82 and 0.75 for the 208 and 376 catchment data sets, respectively) is much higher than that of the best model

of that study. There may exist differences in the efficiency that are due to differences in the climate but the relatively low DMIP efficiencies are typical of the general difficulty of estimating runoff in ungauged catchments. What sets top kriging apart from deterministic models such as those of *Parajka et al.* [2005] and *Reed et al.* [2004] is that top kriging does use concurrent runoff data of neighboring catchments. This means that, for some of the applications envisaged by *Parajka et al.* [2005] and *Reed et al.* [2004], such as flood forecasting, top kriging is not applicable. However, for those applications where concurrent runoff data are available, the better performance of top kriging as compared to the traditional deterministic modeling approach can be a real advantage.

[70] Other indicators of model performance are the statistical moments of temporal runoff variability. Capturing the temporal variability well is of particular importance for assessing the hydropower potential of streams and for ecological studies. The analyses of the Innviertel region indicate that, overall, top kriging does represent the moments well with little bias. However, the temporal variances of the catchments with the largest observed temporal variances were underestimated and the temporal variance of the catchments with the smallest observed temporal variances were overestimated. This means that the spatial variance of the temporal variance was underestimated. This would be expected as kriging, generally, does not retain the spatial variance, being a best estimator. Rather, kriging introduces smoothing as it minimizes the error variance. Methods such as stochastic simulation do retain the spatial variance but the errors of such methods are larger [e.g., *Deutsch and Journel*, 1992]. A similar effect of spatial smoothing was found by *Gottschalk et al.* [2006, Figure 6] (1 hour) who estimated the first two moments of runoff directly from neighboring catchments by geostatistical techniques for 17 catchments in France.

5.2. Factors Controlling Performance

[71] For deterministic models, obviously, the uncertainty in rainfall is one of the main factors contributing to simulation uncertainty, particularly for hourly models [*Faures et al.*, 1995]. Other uncertainties are problems with parameter identifiability and the representativeness of catchment characteristics such as soils data [*Blöschl*, 2005]. The factors controlling the uncertainties of top kriging are quite different. The main limitations are the assumption of spatial homogeneity of runoff, the spatial distribution of stream gauges and the quality of the runoff measurements. Much of the spatial and temporal variability of the atmospheric forcing is filtered by the catchments. Both the runoff of the target catchment and the runoff at the neighboring stream gauges therefore represent aggregated values that have much larger spatial and temporal correlations than the atmospheric input [*Skøien et al.*, 2003; *Skøien and Blöschl*, 2006]. With this in mind, it is clear that the Innviertel region with a high density of stream gauges is a favorable case for the top kriging approach. For the deterministic model, the Innviertel region was on the other hand a region with poorer results than the average of Austria. This can be attributed to two reasons. First, the response times in this region are relatively small, which makes it more difficult for a daily runoff model to simulate

runoff well (see Figures 7–11). Second, although the mean runoff is relatively homogenous in the Innviertel region, the dynamics reflect a rather heterogeneous hydrogeology which may contribute to parameter identifiability issues and regionalization uncertainty. This is also indicated by the decrease in median model performance from 0.72 to 0.59 in the Innviertel region when moving from gauged to ungauged catchments which is substantial.

[72] In other parts of Austria, stream gauge density is lower and there are, in particular, a number of stream gauges located at the Austrian border with no close neighbors. If one moves from the Innviertel to all of Austria, the top kriging performance hence decreases from a median efficiency of 0.87 to 0.75 (hourly data). In addition to the lower stream gauge density, the variogram may contribute to the decrease in performance as it was estimated for the Innviertel and used for all of Austria for consistency. Also, the quality of the Innviertel data set may be above average. The data set of 208 catchments used by *Parajka et al.* [2005] gives a better top kriging performance (ME = 0.82 for hourly data) than the entire Austrian data set because *Parajka et al.* [2005] have been more restrictive in selecting high-quality runoff records.

[73] It was shown that the kriging variance (Figure 17) is an indicator of expected model performance. The kriging variance takes into account the catchment shape, stream network organization, routing characteristics and, most importantly, the location of the stream gauges including the effects of catchment nesting [*Skøien et al.*, 2006]. However, kriging variance does not take into account local particularities in the hydrogeology. Because of this, for some catchments, the efficiency was considerably poorer than what could be expected from the kriging variance. To capture these effects, maps of the hydrogeology would have to be consulted or, preferably, reconnaissance field trips undertaken to assess the runoff dynamics in ungauged catchments [*Blöschl*, 2005].

5.3. Top Kriging Variants

[74] We tested a number of variants of top kriging to examine what variant would produce the most realistic runoff hydrographs. The results indicated that inclusion of a routing model almost always improves the estimated runoff time series. Figure 10 shows a typical example. We also tested spatiotemporal kriging and cokriging variants of top kriging. From a conceptual point of view spatiotemporal kriging and cokriging are more complete than spatial kriging as they account for the temporal filtering (i.e., the unit hydrograph) effects of catchments in the estimation equation (equation (4)) while spatial kriging estimates runoff from concurrent time steps only. It should be noted that spatial kriging does account for the temporal filtering effects on the regularized variogram; that is, the unit hydrograph enters the estimation indirectly through the gamma values. The analyses indicate that spatiotemporal kriging and spatiotemporal cokriging, in some cases, improved the results over spatial kriging, but this was not generally the case. The median model efficiency of spatiotemporal cokriging was similar to that of spatial kriging but there was a clear trend of underestimating the temporal variance of runoff. Spatiotemporal kriging underestimated the temporal variance more and the underestimation increased with the number of time steps used in the estimation. Clearly, this is related

to the smoothing effect of kriging. The better performance of spatiotemporal cokriging as compared to spatiotemporal kriging is a result of the additional unbiasedness constraints. Given that spatiotemporal kriging and spatiotemporal cokriging increases the computational burden significantly over spatial kriging, the additional complexity does not seem to be warranted in the light of the results of this paper. This is in line with the results of *Goovaerts et al.* [2006] who suggested that, in their case, spatiotemporal kriging did not generally improve the estimation results over spatial kriging.

5.4. Top Kriging Assumptions

[75] Top kriging as presented in this paper involves a number of assumptions. Some of them were made for clarity of presentation and can be easily relaxed if suitable data are available. In this paper it was assumed that the temporal support (equation (10)) is a function of catchment area only but it can be easily made a function of physiographic catchment characteristics. Similarly, the simple routing model based on a constant velocity used here (equations (13) and (15)) was mainly chosen for clarity. Also, equation (14) can be relaxed by introducing more complex relationships between lag and catchment characteristics. Specific runoff was assumed to be a random field. It would be possible to include a temporal trend model [*Montanari*, 2005a] or a spatial trend model and apply top kriging to the residuals of the trend model. A spatial trend model could, for example, be based on mean annual rainfall, to account for differences in runoff in ungauged catchments as a result of climate. In fact, one could also use a deterministic runoff model as a trend model and apply top kriging as an error model to improve runoff estimates over those of the deterministic runoff model [*Goovaerts et al.*, 2006].

[76] An assumption that would be more difficult to relax, is the assumption of the variogram not changing with time. It is likely that low flows are correlated over larger distances in both space and time than flood flows, so one would expect temporal differences in the space-time variogram of runoff. In this paper we have argued that the variograms of runoff time series will be dominated by the variability of high flows as the variogram is the second moment of runoff. Changing the variogram with time would have to involve additional assumptions about temporal aggregation. Catchments were conceptualized to operate as a linear space-time filter which is at the heart of top kriging. Introducing nonlinearity would in effect modify top kriging to become a spatially distributed runoff model.

[77] Relaxing assumptions will generally require additional data and add complexity, which may not be warranted as suggested by the comparison of the spatial and spatiotemporal kriging variants of top kriging. Also, the high model efficiencies obtained in this paper indicate that the assumptions on which the spatial variant of top kriging is based may be highly appropriate for practical purposes of estimating runoff time series in ungauged catchments.

5.5. Potential Applications of Top Kriging

[78] Top kriging uses concurrent runoff data in neighboring catchments and does not explicitly account for rainfall-runoff processes. This means that there are a number of applications, such as flood forecasting and assessing the effects of land use change on runoff, for which top kriging is not the right method and deterministic runoff models are

needed. However, there are a range of applications where top kriging could be applied to estimate runoff time series in ungauged catchments. The main applications in a simulation mode are estimating the flow duration curve for assessing the hydropower potential of river reaches, estimating temporal flow (and low flow) variability in the context of environmental flow requirements, and producing high-resolution maps of runoff variability. These maps could not only involve mean annual runoff, as in the traditional runoff mapping approaches, but also high-resolution characteristics such as temporal variance and runoff regimes. Another simulation application is the generation of hydrographs for ungauged catchments to which deterministic rainfall-runoff models are calibrated. This may reduce some of the problems with regionalizing model parameters [Parajka *et al.*, 2007]. There are also potential online applications. One of them is near-real time monitoring and visualization of runoff. In top kriging, much of the effort is in estimating the kriging weights. Once the weights are known, estimation of runoff for all river reaches in a region from runoff data is a straightforward linear operation, ideally suited for online applications.

[79] The relatively high performance of top kriging found in this paper is because errors of rainfall data and parameter identifiability issues of traditional runoff models are avoided. The model efficiencies obtained by top kriging are considerably higher than what can be typically achieved by using a regionalized deterministic model. This means that, for certain applications, top kriging is an appealing method for estimating runoff in ungauged catchments. However, a prerequisite of using top kriging is a relatively dense stream gauging network. In the Innviertel and Austrian examples, a total of 19 stream gauges over an area of 1500 km² and 376 stream gauges over an area of 80,000 km², respectively, were available. If the stream gauging network is less dense, the benefits of top kriging would probably be less obvious. When the distances between observations become too large, it is likely that deterministic models will perform better than top kriging. However, given that the improvement over deterministic models is quite significant, top kriging may still be the method of choice for regions with considerably fewer stream gauges than Austria.

Appendix A: Adjustment of Kriging Weights

[80] If measurements are close to each other, i.e., clustered in space, they tend to be correlated. One advantage of kriging is to account for this effect by reducing the kriging weights of clustered measurements. However, in the limit, if two measurements are located at the same point in space, the kriging system will only give the sum of their weights and it is impossible to identify the weights individually. Measurements that are close may hence lead to ill-conditioned kriging systems and numerical problems may arise. Some of the kriging weights obtained may then be much smaller than zero. This is not a desirable result as they can produce artifacts in the estimates. As catchments are, occasionally, very close, and runoff measurements are highly correlated in time, this needs to be dealt with in top kriging. The issue was addressed in two ways in this paper. First, nonzero local uncertainty of runoff, σ_j^2 , was used. The larger σ_j^2 , the fewer numerical problems will arise

as it effectively decreases the correlation between neighboring catchments. σ_j^2 has been set to a physically realistic value of 1% of the average temporal variance of runoff. However, occasional large negative weights still remained. These were adjusted by rescaling them.

[81] The method of rescaling weights usually recommended in the literature [see Yamamoto, 2000] is to remove all negative weights and to rescale the remaining weights to a sum of unity. However, as a large negative weight is usually the counterpart of a large positive weight, this method gives a loss of information. Instead, in this paper we use a method that keeps some of the negative weights and, in most instances, maintains the internal ranking of the weights. The starting point is that ill-conditioned kriging systems can be identified by large absolute sum of weights:

$$\Lambda = \sum_{j=1}^n \sum_{\beta=1}^p |\lambda_{j\beta}| \gg 1 \quad (\text{A1})$$

where $\lambda_{j\beta}$ are the kriging weights. We defined an upper limit for this sum and, after testing different limits, found $\Lambda_{\max} = 1.5$ to be a suitable choice. All weights were then scaled according to 1350

$$\lambda'_{j\beta} = \lambda_{j\beta} \Lambda_{\max} / \Lambda \quad (\text{A2})$$

[82] As the sum of the weights is no longer unity the difference to unity is added to all weights: 1354

$$\lambda''_{j\beta} = \lambda'_{j\beta} + \frac{1 - \sum_{j=1}^n \sum_{\beta=1}^p \lambda'_{j\beta}}{np} \quad (\text{A3})$$

[83] For spatial and spatiotemporal kriging all weights 1355 were adjusted in this way while for spatiotemporal cokriging only the weights for time step t_ω were adjusted by equations (A2) and (A3). Equation (A1) was used to recalculate Λ using the weights of equation (A3), typically, giving $\Lambda > \Lambda_{\max}$. The procedure ((A2), (A3), (A1)) was then repeated until $|\Lambda - \Lambda_{\max}| < \delta_\Lambda$ with $\delta_\Lambda = 0.05$. We suggest that the weights adjusted by this procedure respect the kriging system to a larger extent than the methods presented by Yamamoto [2000].

[84] **Acknowledgments.** The ideas of top kriging have evolved over the past 10 years. Early discussions with Murugesu Sivapalan and Rodger Grayson and more recent discussions with Ralf Merz, Lars Gottschalk, Etienne Leblois and Eric Sauquet have significantly contributed to our own ideas. The Austrian Academy of Sciences, project HÖ 18, and the Austrian Science Foundation, project P18993_N10, provided financial support; the Austrian Hydrographic Service (HZB) provided the hydrographic data, and Juraj Parajka provided the deterministic runoff model results. We are grateful for all of these contributions.

References

- Blöschl, G. (2005), Rainfall-runoff modelling of ungauged catchments, in *Encyclopedia of Hydrological Sciences*, vol. 3, edited by M. G. Anderson, article 133, pp. 2061–2080, John Wiley, Chichester, U. K.
- Bogaert, P. (1996), Comparison of kriging techniques in a space-time context, *Math. Geol.*, 28(1), 73–86.
- Castellarin, A., L. Brandimarte, A. Montanari, A. Brath, and G. Galeati (2004), Regional flow-duration curves: Reliability for ungauged basins, *Adv. Water Resour.*, 27, 953–965.

- Clark, I., K. L. Basinger, and W. V. Harper (1987), MUCK—A novel approach to co-kriging, in *Proceedings of Conference on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radio-nuclide Transport Modeling*, edited by B. E. Buston, pp. 473–493, Battelle, Columbus, Ohio.
- Cressie, N. (1985), Fitting variogram models by weighted least squares, *Math. Geol.*, 17(5), 563–586.
- Cressie, N. (1991), *Statistics for Spatial Data*, 900 pp., John Wiley, New York.
- de Marsily, G. (1986), *Quantitative Hydrogeology*, 440 pp., Academic, London.
- Deutsch, C. V., and A. G. Journel (1992), *GSLIB—Geostatistical Software Library and User's Guide*, 340 pp., Oxford Univ. Press, New York.
- Duan, Q., S. Sorooshian, and V. K. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28, 1015–1031.
- Faures, J.-M., D. C. Goodrich, D. A. Woolhiser, and S. Sorooshian (1995), Impact of small-scale spatial rainfall variability on runoff modeling, *J. Hydrol.*, 173, 309–326.
- Gandin, L. S. (1963), *Objective Analysis of Meteorological Fields* (in Russian), Isr. Program for Sci. Transl., Jerusalem.
- Gippel, C. J. (2005), Environmental flows: Managing hydrological environments, in *Encyclopedia of Hydrological Sciences*, vol. 5, edited by M. G. Anderson, article 191, pp. 2953–2971, John Wiley, Chichester, U. K.
- Goovaerts, P., A. Auchincloss, and A. V. Diez-Roux (2006), Performance comparison of spatial and space-time interpolation techniques for prediction of air pollutant concentrations in the Los Angeles area, paper presented at 6th International Congress, Int. Assoc. for Math. Geol., Liège, Belgium, 4–8, Sept.
- Gottschalk, L. (1993), Correlation and covariance of runoff, *Stochastic Hydrol. Hydraul.*, 7, 85–101.
- Gottschalk, L., I. Krasovskaia, E. Leblois, and E. Sauquet (2006), Mapping mean and variance of runoff in a river basin, *Hydrol. Earth Syst. Sci.*, 10, 469–484.
- Jost, G., G. B. M. Heuvelink, and A. Papritz (2005), Analysing the space-time distribution of soil water storage of a forest ecosystem using spatio-temporal kriging, *Geoderma*, 128(3–4), 258–273.
- Journel, A. G., and C. J. Huijbregts (1978), *Mining Geostatistics*, 600 pp., Academic, London.
- Kyriakidis, P. C., and A. G. Journel (1999), Geostatistical space-time models: A review, *Math. Geol.*, 31(6), 651–684.
- Melone, F., C. Corradini, and V. P. Singh (2002), Lag prediction in ungauged basins: An investigation through actual data of the upper Tiber River valley, *Hydrol. Processes*, 16, 1085–1094.
- Merz, R., and G. Blöschl (2003), A process typology of regional floods, *Water Resour. Res.*, 39(12), 1340, doi:10.1029/2002WR001952.
- Merz, R., and G. Blöschl (2005), Flood frequency regionalization—Spatial proximity vs. catchment attributes, *J. Hydrol.*, 302, 283–306.
- Merz, R., G. Blöschl, and J. Parajka (2007), Regionalisation methods in rainfall-runoff modelling using large catchment samples, in *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment—MOPEX Proceedings of the Paris (2004) and Foz de Iguacu (2005) Workshops*, edited by V. Andréassian et al., *IAHS Publ.*, 307, 117–125.
- Montanari, A. (2005a), Deseasonalisation of hydrological time series through the normal quantile transform, *J. Hydrol.*, 313, 274–282.
- Montanari, A. (2005b), Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, 41, W08406, doi:10.1029/2004WR003826.
- Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models. part I—A discussion of principles, *J. Hydrol.*, 10, 282–290.
- Parajka, J., R. Merz, and G. Blöschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrol. Earth Syst. Sci.*, 9, 157–171.
- Parajka, J., G. Blöschl, and R. Merz (2007), Regional calibration of catchment models: Potential for ungauged catchments, *Water Resour. Res.*, 43, W06406, doi:10.1029/2006WR005271.
- Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.-J. Seo, and DMIP participants (2004), Overall distributed model intercomparison project results, *J. Hydrol.*, 298, 27–60.
- Rinaldo, A., A. Marani, and R. Rigon (1991), Geomorphological dispersion, *Water Resour. Res.*, 27, 513–525.
- Rouhani, S., and H. Wackernagel (1990), Multivariate geostatistical approach to space-time data analysis, *Water Resour. Res.*, 26, 585–591.
- Sauquet, E., L. Gottschalk, and E. Leblois (2000), Mapping average annual runoff: A hierarchical approach applying a stochastic interpolation scheme, *Hydrol. Sci. J.*, 45(6), 799–815.
- Schubert, G., H. Lampl, W. Pavlik, G. Pestal, C. Rupp, S. Shadlau, and M. Wurm (2003), Hydrologie, *Map 6.2*, in *Hydrologischer Atlas Österreichs (Hydrological Atlas of Austria)*, Bundesminister. für Land- und Forstwirtschaft., Umwelt und Wasserwirtschaft., Vienna.
- Sivapalan, M., and G. Blöschl (1998), Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves, *J. Hydrol.*, 204, 150–167.
- Sivapalan, M., et al. (2003), IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydrol. Sci. J.*, 48(6), 857–880.
- Skøien, J. O., and G. Blöschl (2006), Catchments as space-time filters—A joint spatio-temporal geostatistical analysis of runoff and precipitation, *Hydrol. Earth Syst. Sci.*, 10, 645–662.
- Skøien, J. O., G. Blöschl, and A. W. Western (2003), Characteristic space scales and timescales in hydrology, *Water Resour. Res.*, 39(10), 1304, doi:10.1029/2002WR001736.
- Skøien, J. O., R. Merz, and G. Blöschl (2006), Top-kriging—Geostatistics on stream networks, *Hydrol. Earth Syst. Sci.*, 10, 277–287.
- Snepevangers, J. J. J. C., G. B. M. Heuvelink, and J. A. Huisman (2003), Soil water content interpolation using spatio-temporal kriging with external drift, *Geoderma*, 112, 253–271.
- Woods, R., and M. Sivapalan (1999), A synthesis of space-time variability in storm response: Rainfall, runoff generation, and routing, *Water Resour. Res.*, 35(8), 2469–2485.
- Yamamoto, J. K. (2000), An alternative measure of the reliability of ordinary kriging estimates, *Math. Geol.*, 32(4), 489–509.

G. Blöschl, Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Karlsplatz 13/222, A-1040 Vienna, Austria.

J. O. Skøien, Department of Physical Geography, Faculty of Geosciences, Utrecht University, P.O. Box 80.115, NL-3508 TC Utrecht, Netherlands. (j.skøien@geo.uu.nl)