Geostatistische Modelle für Fließgewässer

Gregor Laaha (gregor.laaha@boku.ac.at)
Institute of Applied Statistics and Computing,
BOKU Vienna, Austria

ROeS Seminar, 6. November 2014, BOKU Wien.

Reference: Laaha G, Skøien J, Blöschl G 2012. Comparing geostatistical models for river networks. In: Abrahemsen P, Hauge R, Kolbjørnsen O (eds.), Geostatistics Oslo 2012, pp. 543–553, Springer.

Introduction: Spatial Interpolation

- Estimation at a certain location

e.g. Air pollutant concentrations were measured at different locations.

- What is the concentration at location X_o?

Introduction: Grid-based Estimation

- Estimation of a value for each cell of a grid

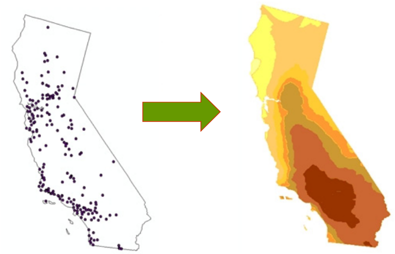
13	14	16	20	23
14	14	16	19	24
18	16	16	18	22
24	22	19	19	21
30	27	23	20	20

Introduction: Grid-based Estimation

- Presentation in form of a map (Mapping)

Example: Air pollutant concentrations were measured at different locations.

- Air pollution map



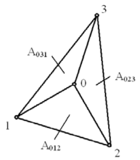
Simple Interpolation Methods

Example: Triangulation

Plain through the next three data points (1,2,3)

Calculation: Delaunay method

- linear combination
- weightings of opposing area



L $Z^* = \Sigma \lambda^\alpha Z_\alpha = \lambda^1 Z_1 + \lambda^2 Z_2 + \lambda^3 Z_3$

U $\Sigma \lambda^\alpha = 1$

Weighting: $\lambda^1 = \frac{A_{023}}{A_{123}}$, $\lambda^2 = \frac{A_{031}}{A_{123}}$, $\lambda^3 = \frac{A_{012}}{A_{123}}$

From Simple Interpolation to Geostatistics

In a nutshell:

- Estimation by linear combination of data points
- Unbiased estimator
- Empirical chosen weights (not optimal!)

Optimal weights?

- Estimation variance as a quality criteria

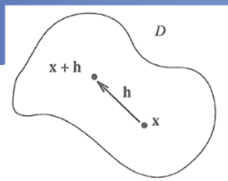
$$\sigma_E^2 = \text{var}[Z^* - Z] = E[(Z^* - Z)^2]$$

- Optimal estimation – minimum estimation variance

$$\sigma_E^2 = \text{var}[Z^* - Z] = \min!$$

➢ Geostatistical interpolation (Kriging)
Optimal weights on the basis of *spatial correlation*

Spatial correlation



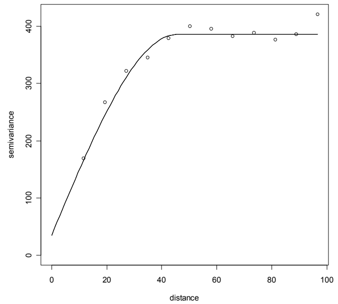
Variogram (Semivariogram)

… dissimilarity vs. distance (h)

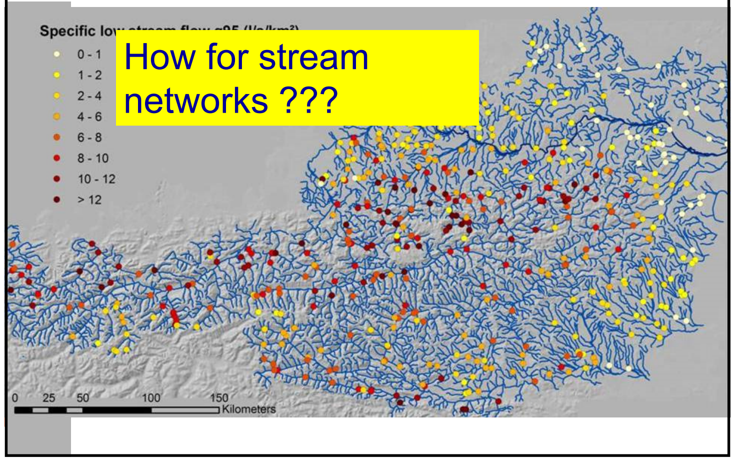
… for pairs of data points

$$\gamma^*_{\alpha\beta} = \frac{(z_\alpha - z_\beta)^2}{2} = \frac{(z(x+h) - z(x))^2}{2}$$

- In case of stationary/intrinsic random field the variogram is only a function of distance h.



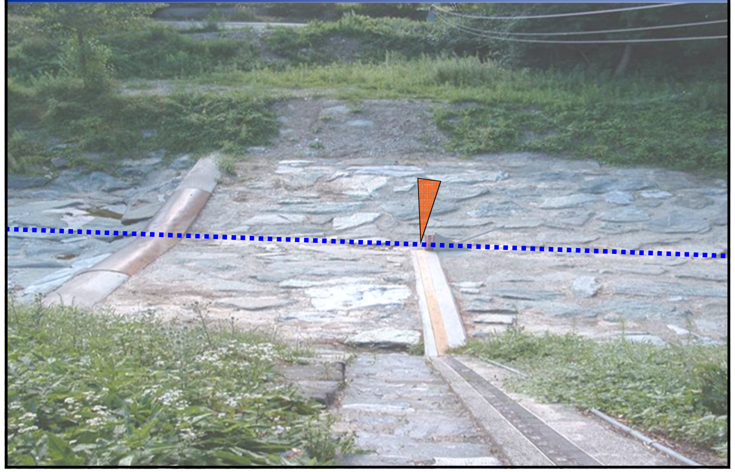
Geostatistics…



Outline

- Introduction
  … The river network problem
- Geostatistical Models for river networks
  … 1D and 2D conceptualisations
- Comparison of concepts
  … OK, 1D, 2D
- Review of case studies
  … Environmental variables, low flows, temperature
- Conclusions

Introduction: *The river network problem…*



Introduction: *The river network problem…*

- Estimation of streamflow and related variables
  … fundamental problem in WRM
- Gauged sites
  … summary statistics of observed time series
- Ungauged sites ?
  … regional transfer of observed information

Introduction: *The river network problem…*

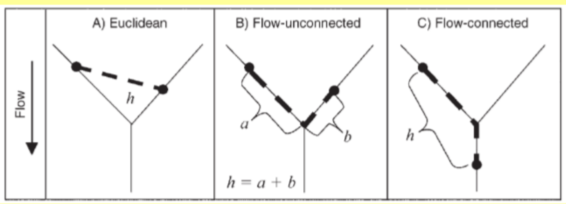
- Focus on geostatistical regionalisation methods
  … spatial average, weights according to spatial covariance
  … rarely used in practice
- Challenge: Tree structure of river network
  – catchments related to points of the river network are organised into subcatchments (i.e. they are *nested*)
  – they need to be treated differently from flow-unconnected neighbours which do not share a catchment
- Kriging on river networks – two concepts discussed:
  – 1D models, 2D models
  – Compared to OK-Euclid

1D Models

- Treat river network as 1D problem
- Support = river location
- Ordinary point-kriging predictor
- requires meaningful distance metric & valid Cov-Function

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i)$$

Euklidean Stream distance



From: Peterson EE, Ver Hoef JM. 2010:
A mixed-model moving-average approach to geostatistical modelling in stream networks.
Ecology **91**(3):644-651.

1D Models … *valid covariance function*

- Gottschalk (1993a) first calculated covariance along stream network based on river distance
  … exponential Cov-Function well suited
  … added water balance constraints to kriging system
  to ensure predicted lateral inflow = difference b/w gauges
- Ver Hoef et al. (2006), Cressie et al. (2006)
  Spatial Cov-Function *C(h)*
  … derived by moving average (*kernel convolution*)
  … different kernel shapes -> relate to different Cov-Functions
  => classical Cov-Functions are valid for river networks
- Restriction: only unilateral kernels
  … downstream (Tail-down model) or
  … upstream (Tail-up model)

1D Models … *Pointkriging using stream distance*

Tail-down Tail-up

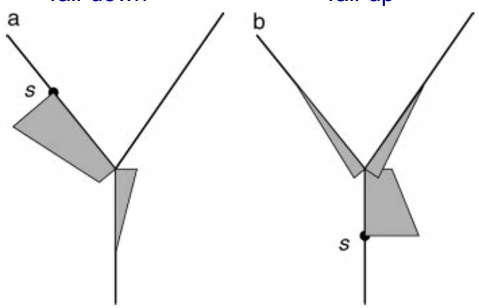


Figure 2. Kernel shapes for (a) *down* model and (b) *up* model

From: Ver Hoef JM, Peterson E, Theobald D. 2006: Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics* **13**: 449-464.

1D Models … *Pointkriging using stream distance*

Ver Hoef et al. (2006) Tail-up model performs better
… but needs auxiliary variables for weighting confluents

- catchment area (Ver Hoef, 2006);
- stream order (Cressie, 2006)
  … as surrogate of discharge (sic!)

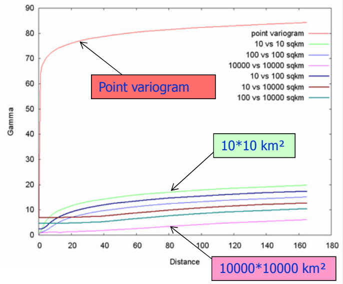
-> ct. Gottschalk (1993a,b): discharge constraints

2D Models

- Runoff generation = continuous spatial process
  …existing in any point of the landscape
- Discharge at river site = integral of point runoff over catchment
  $$z(A_i) = \frac{1}{|A_i|} \int_{A_i} z(\mathbf{x})\, d\mathbf{x}$$
- Support = catchment area
- Regional transfer ("prediction") = Change of support
  … Block-kriging
  … irregular support (!)
  … river network topology (!)
- Implementation not trivial, but consistent hydrological concepts of runoff generation

2D-Model Top-Kriging (Skøien et al. 2006)

- Regularised variogram
  … spatial correlation b/w
  - pairs of catchments
  - with different support (area)
- Variograms for pairs of catchments … a function of distance (h) and support (A_1,A_2)

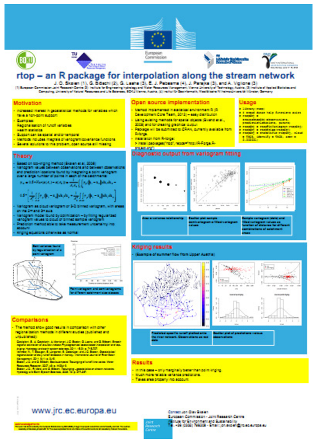


"extension variance" :

$$\bar{\gamma}_{12}(h) = \bar{\gamma}(h, A_1, A_2) - 1/2\left[\bar{\gamma}(h, A_1, A_1) + \bar{\gamma}(h, A_2, A_2)\right]$$

… is smaller for overlapping catchments
=> More weight to nested catchments

R-package rtop – see Poster P-033, J.O. Skøien et al.



Comparison of geostatistical models

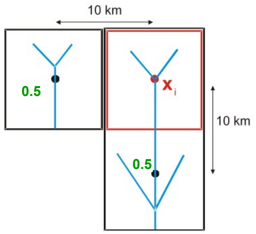
- Kriging = spatial weighted average

$\Rightarrow$ Methods are as good as kriging weights
… and how they are distributed in space

$\Rightarrow$ How are weights distributed b/w connected and unconnected neighbours?

$\Rightarrow$ Focus on limiting situations
(i) equally distant neighbours
(ii) more distant flow-connected neighbour

OK … *Point-kriging (Euclidean distance)*



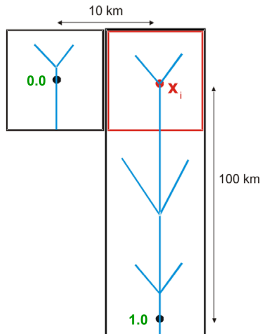
⇒ Distribute weights according to distance only
⇒ Topology not taken into account!!
⇒ **Too much weight according to distance in geographic space, and too little weight according to river network topology**

1D Models … *Point-kriging using stream distance*

Weights (upstream model)

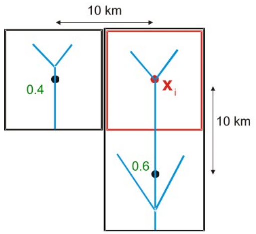
⇒All weight given to flow-connected neighbour and **no weight for flow-unconnected neighbour**

… Prediction of source area by river mouth, rather than by next source



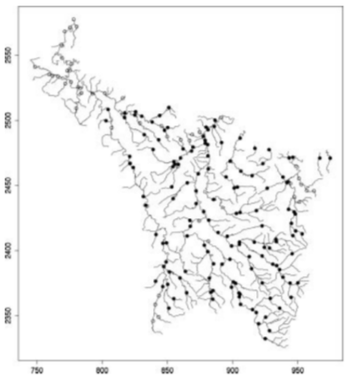
⇒ **Good results if most information at connected sites**
⇒ **Overall too much weight according to topology and too little weight according distance in geogr. space**

2D-Models … *Top-kriging*



=> Distribute weights according to distance and river network topology, depending on data situation

Case study 1: 1D-modelling of environmental variables



- Garreta et al. (2009)
- 141 nitrate and 187 temperature stations
- situated at the Meuse and Moselle basin in north-eastern France.

Reference: Garreta, V, Monestiez, P & Ver Hoef, J M 2009. Spatial modelling and prediction on river networks: up model, down model or hybrid? *Environmetrics* 439–456.

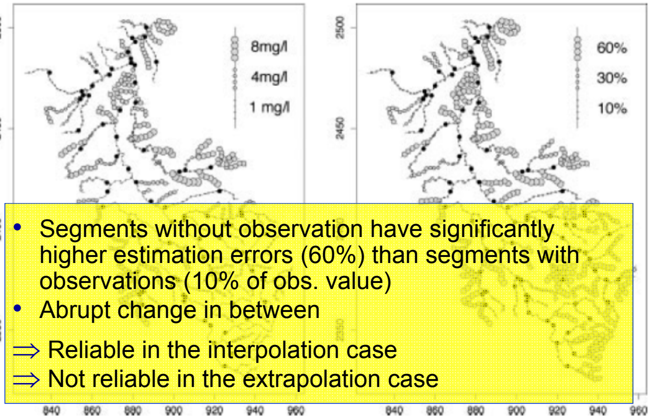
Case study 1: 1D-modelling of environmental variables

Results (Garreta et al. 2009)

- Summer temperature: the Tail-up model performed better
- Nitrate: the inverse was true
- A hybrid model which (= combination Tail-up & Tail-down) performed significantly better than each of the models separately.

Reference: Garreta, V, Monestiez, P & Ver Hoef, J M 2009. Spatial modelling and prediction on river networks: up model, down model or hybrid? *Environmetrics* 439–456.

Nitrate loads, Hybrid model: Prediction errors (left) and confidence interval (right)

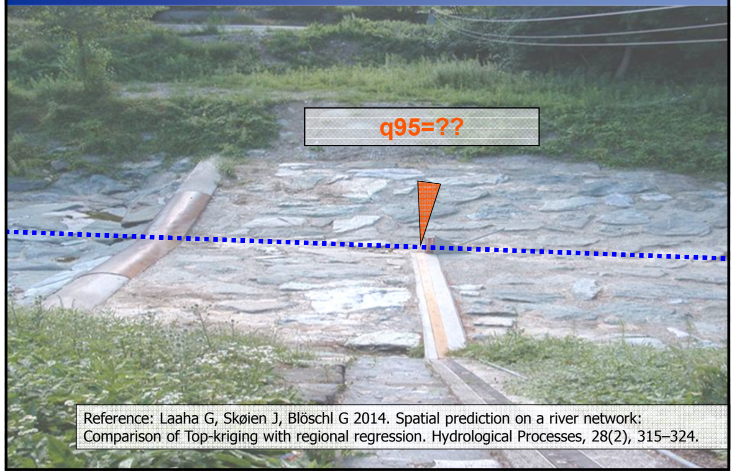


- Segments without observation have significantly higher estimation errors (60%) than segments with observations (10% of obs. value)
- Abrupt change in between

⇒ Reliable in the interpolation case
⇒ Not reliable in the extrapolation case

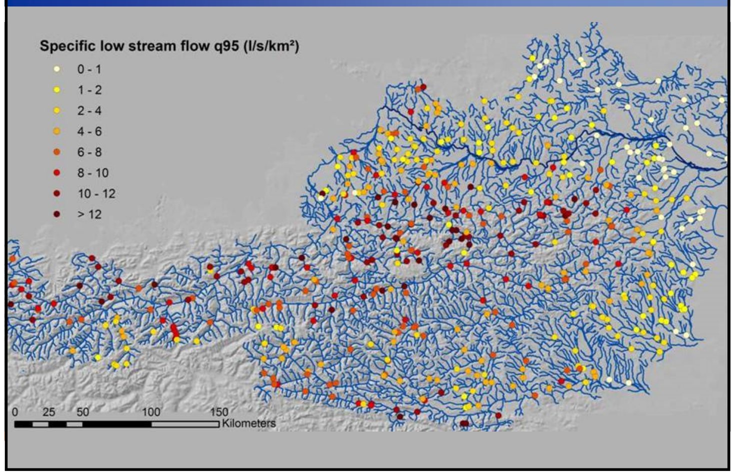
Reference: Garreta, V, Monestiez, P & Ver Hoef, J M 2009. *Environmetrics* 439–456.

Case study 2: 2D-modelling of low streamflows



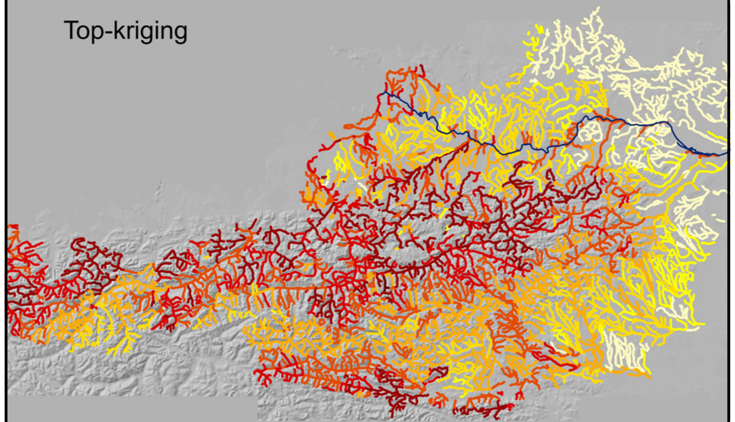
Reference: Laaha G, Skøien J, Blöschl G 2014. Spatial prediction on a river network: Comparison of Top-kriging with regional regression. Hydrological Processes, 28(2), 315–324.

Data: Austria, 491 gauges



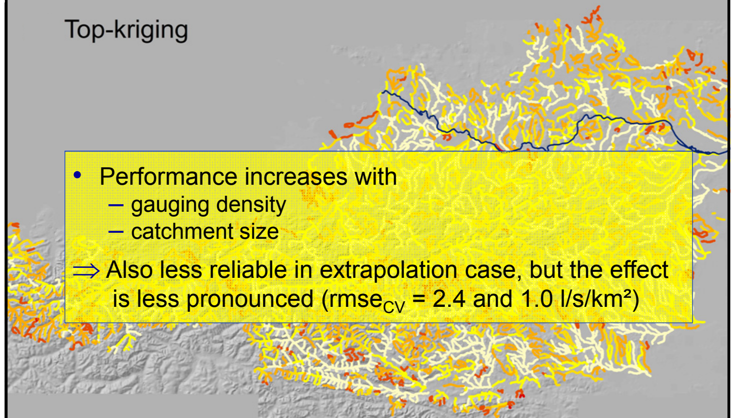
Prediction

Top-kriging



Uncertainty (Kriging - standard error)

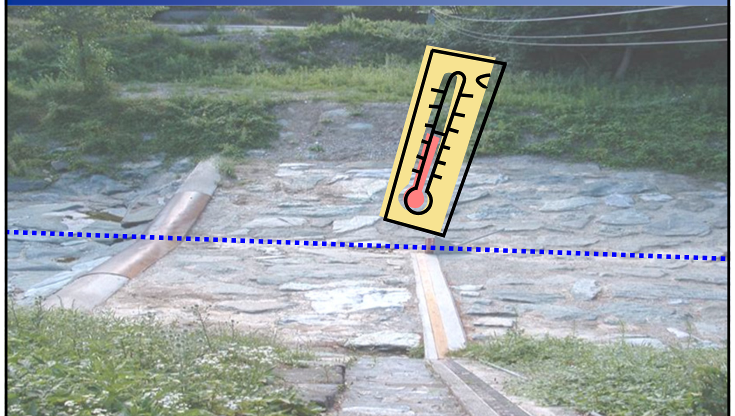
Top-kriging



- Performance increases with
  - gauging density
  - catchment size

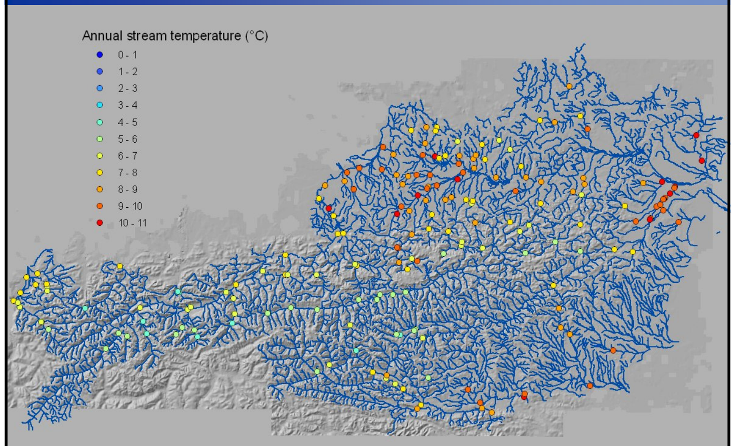
⇒ Also less reliable in extrapolation case, but the effect is less pronounced ($rmse_{CV}$ = 2.4 and 1.0 l/s/km²)

Case study 3: Annual stream temperature

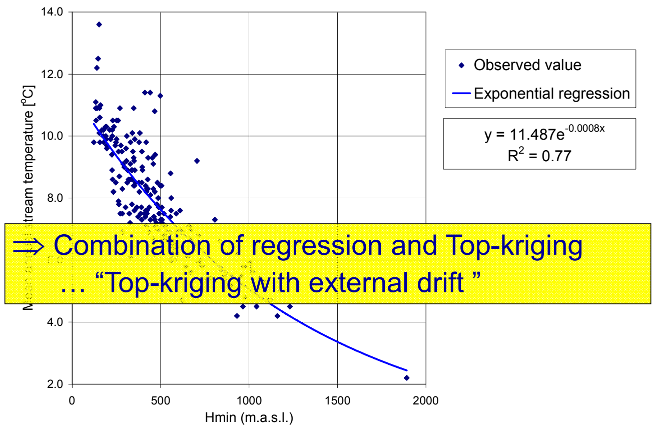


Reference: Laaha G, Skøien JO, Nobilis, F, Blöschl G 2013. Spatial prediction of stream temperatures using top-kriging with an external drift. Envir.Mod.Ass., 18(6), 671–683.

Data: 214 gauges, mean annual streamflow temperature [°C]

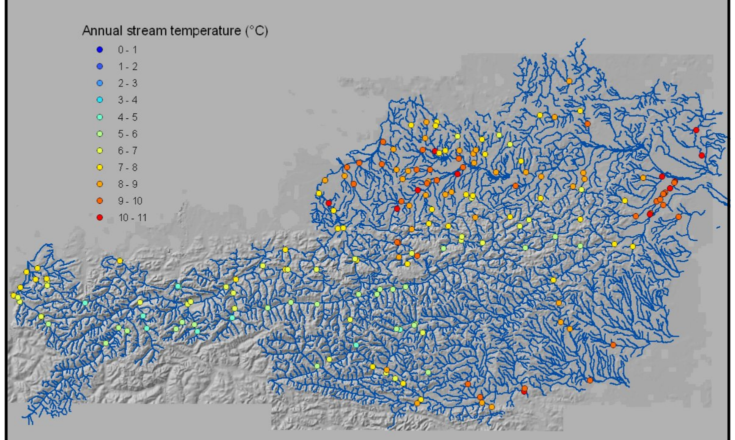


… using altitude as external drift



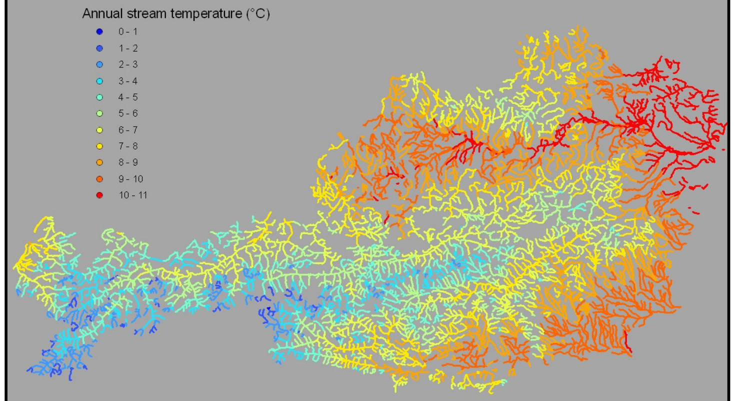
⇒ Combination of regression and Top-kriging … “Top-kriging with external drift ”

Data: 214 gauges

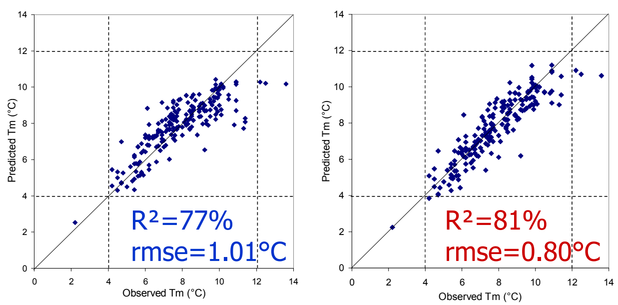


EDTK results: estimate

T**=T(Hmin)*-Resid*

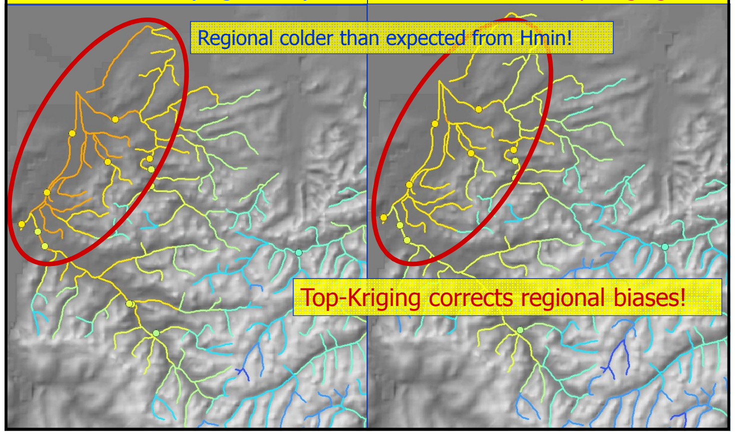


Cross-validation



Regional examples
(1) Vorarlberg (western Austria)

External drift (Regression)	External drift Top-Kriging



Conclusion

- We assessed geostatistical models for stream networks
- Ordinary-kriging (based on Euclidean distance) distribute weights according to distance only. Topology not taken into account!!
- 1D models give all weight to connected gauges at the same river, while close-by neighbors at unconnected rivers are not taken into account. Distribution of weights among tributaries is unsolved (Up-tail model).
- 2D models are more realistic; they distribute kriging weights according to spatial structure, distance and nestedness. They are consistent with hydrological concepts of runoff generation.
- Performance of 1D and 2D models was illustrated here in a meta-analysis of case studies. It would be interesting to perform a direct comparison on a common data set.

Thank you ...