On the Application of Nonparametric Regression Methods to Geodetic Data

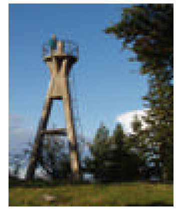
Beat FISCHER

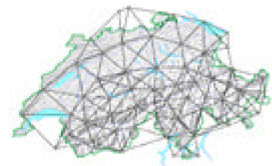
Philipp BRÜHLMANN

Paper 0460

Two Sets of coordinates

LV03 (1903)
conventional measurements
(triangulation)





1st order points 1903

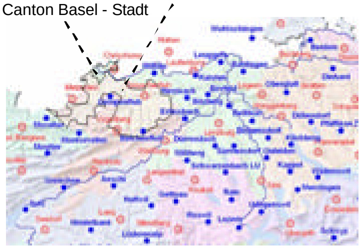
Two Sets of coordinates

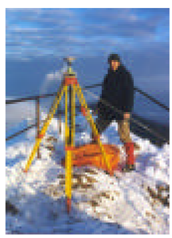
main points

points densifying the net

Canton Basel - Landschaft

Canton Basel - Stadt





LV95 (1995)
GPS based

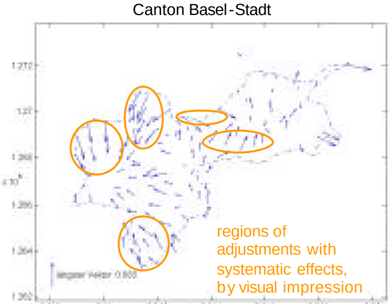
Transformation LV03 ? LV95

Problem: LV03 - coordinates show systematic distortions relative to the more homogeneous LV95 - coordinates. LV03 is the actual cadastral surveyreference frame. GPS measurements referencing to LV95 shouldbe usable.

swisstopo: The Cantons are orderedto triangulate their territories such that in each triangle the transformation LV03 ? LV95 can be performed with satisfying accuracy. Recommended are defor-mations less than 100ppm (1cm per 100 m). Starting in 2007, both reference frames are at disposal.

swisstopo April 2004:
http://www.cadastre.ch/pub/down/projet/rdlv95/ks_04_02_Beilage_de.pdf

Adjustments of Helmert-Transformation LV03 ? LV95
Canton Basel-Stadt



Problem:
How to choose the transformation triangles?

Semiparametric Regression (Collocation and Filtering)

Helmert transformation $l(t) = A(t)x$ with $\begin{cases} t = \text{target system coord.} \\ l = \\ x = \text{transformationparameters} \end{cases}$

and with least squares solution x of $l = A \cdot x - v, \; v^T P v = \min$

Extension of model $l(t) = A(t)x + s(t) + n$

noise

nonparametric signal

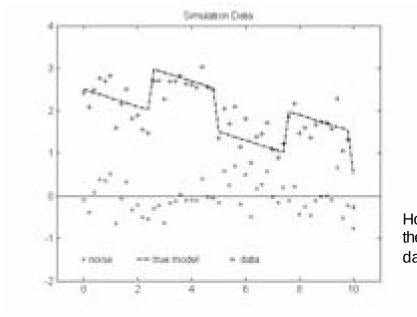
regularization parameter

regularizer

$a \cdot s^T R s + n^T P n = \min$

How to describe the signal function $s(t)$ and the regularizer R ?
How to determine the regularization parameter a ?

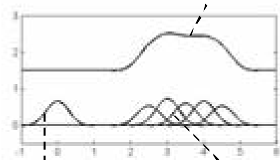
1-Dimensional Simulation Data



How to estimate the model from data?

1-Dimensional Data: Description of Signal Functions $s(t)$ and Regularizer R

sum $1.5+\sum_{i=1}^{5}\frac{k_i}{a}b(t-t_i)$, 1.5 is added only for the clarity of the figure

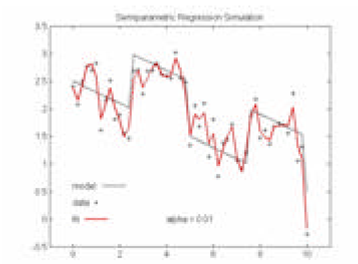


B-spline basis function $b(t)$

shifted and scaled B-Spline basis functions $\frac{k_i}{a}b(t-t_i),\ i=1,\ldots,5$

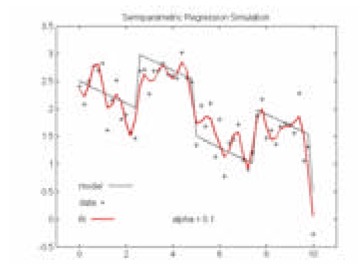
regularizer R given by $R_{ij}^{-1}=b(|t_i-t_j|)$

Dependence of model + signal estimate on *a*



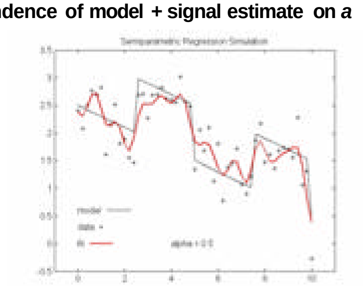
***a* = 0.01**

Dependence of model + signal estimate on *a*



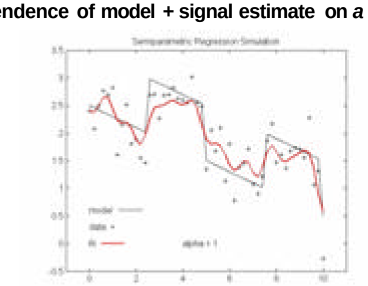
***a* = 0.1**

Dependence of model + signal estimate on *a*



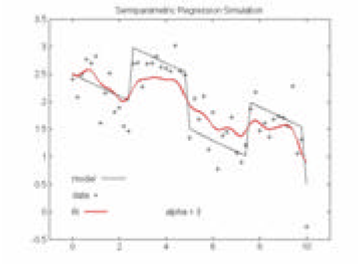
***a* = 0.5**

Dependence of model + signal estimate on *a*



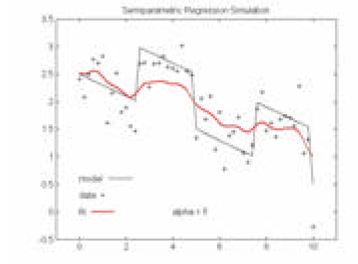
***a* = 1**

Dependence of model + signal estimate on *a*



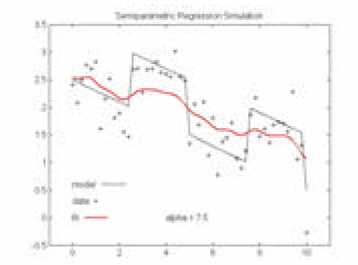
a **= 3**

Dependence of model + signal estimate on *a*



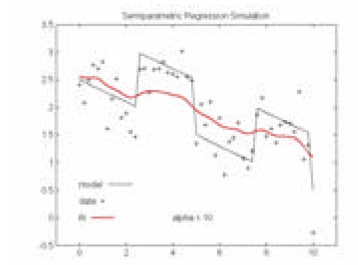
a **= 5**

Dependence of model + signal estimate on *a*



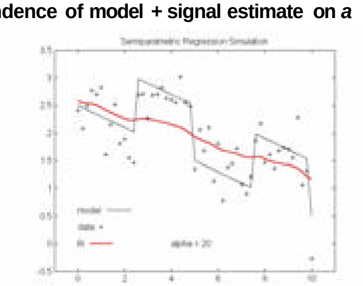
a **= 7.5**

Dependence of model + signal estimate on *a*



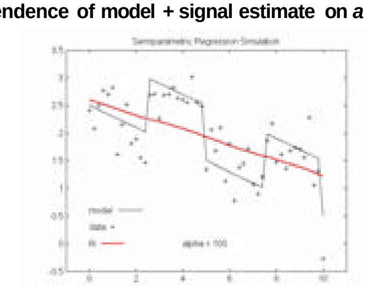
a **= 10**

Dependence of model + signal estimate on *a*



a **= 20**

Dependence of model + signal estimate on *a*



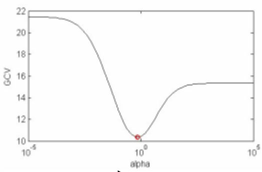
a **= 100**

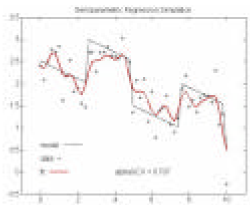
How to choose *a* ?

Choosing *a* by GCV (Generalized Cross Validation)

a is chosen in order to minimize the average predictive squared error PSE of an additional observation. GCV is an approximate estimate of

$$PSE(a) \approx GCV(a) = \frac{1}{m}\cdot\frac{\sum_{i=1}^{m}\left(l_i - \hat{f}_a(t_i)\right)^2}{\left(1-\mathrm{tr}\,H(a)\right)^2} \quad \text{with} \quad Hl = A\hat{x} + \hat{s} = \hat{f}\,.$$





Simulation: a, GCV

Helmert Transformations

start system $t = \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x_i \\ y_i \end{bmatrix}$ target system $l = \begin{bmatrix} x \\ h \end{bmatrix}, \begin{bmatrix} x_i \\ h_i \end{bmatrix}$

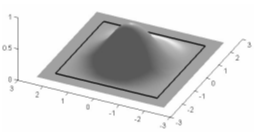
transformation + signal $\begin{bmatrix} x \\ h \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_x \\ a_y \end{bmatrix} + \begin{bmatrix} s_x(x,y) \\ s_h(x,y) \end{bmatrix}$

observation equations

$$\begin{bmatrix} x_1 \\ h_1 \\ x_2 \\ h_2 \\ \dots \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 \\ y_1 & -x_1 & 0 & 1 \\ x_2 & y_2 & 1 & 0 \\ y_2 & -x_2 & 0 & 1 \\ & \dots & & \dots \end{bmatrix}\begin{bmatrix} a \\ b \\ a_x \\ a_y \end{bmatrix} + \begin{bmatrix} s_{1x} \\ s_{1h} \\ s_{2x} \\ s_{2h} \\ \dots \end{bmatrix} + \begin{bmatrix} n_{1x} \\ n_{1h} \\ n_{2x} \\ n_{2h} \\ \dots \end{bmatrix}$$

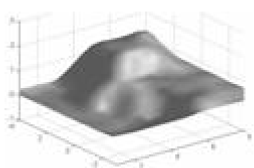
minimization requirement $a \cdot s^T R s + n^T P n = \min$

Signal Description for Coordinate Transformations



B-Spline basis functions

$b(|x - x_i|)\cdot b(|y - y_i|)$

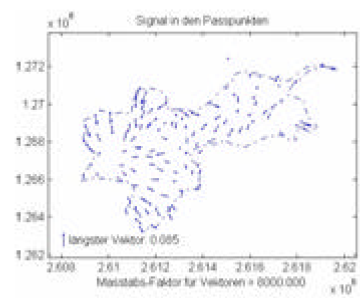


both target system coordinate signals are linear combinations

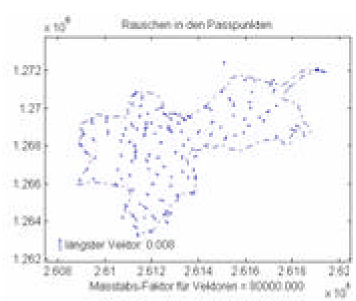
$$s(x,y) = \frac{1}{a}\cdot\sum_i k_i \cdot b(|x - x_i|)\cdot b(|y - y_i|)$$

of shifted and scaled basis functions

Canton Basel-Stadt: Signal estimate in fiducial points



Canton Basel-Stadt: noise in fiducial points



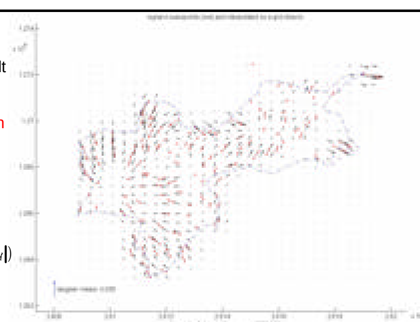
A randomly varying noise is a necessary condition for a good fit.

Canton Basel-Stadt

signal estimates on fiducial points

interpolated signal

$s(x,y) = \sum_i \frac{k_i}{a}\cdot b(|x - x_i|)\cdot b(|y - y_i|)$



The interpolated signal can be used to detect regions of similar net distortion.

Definition:
Floe = Region of similar signal or similar net distortion (= "Scholle" in German)

Methods for floe identification:

? Inspecting the **signal variation**, given by the signal derivative $s'(t)$, if the signal is differentiable. Floe boundaries are marked by large $|s'(t)|$.

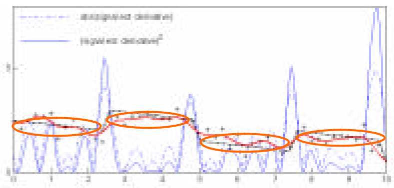
For the derivative of target coordinate signal functions vector $\begin{bmatrix}\hat{s}_x(x,y)\\ \hat{s}_h(x,y)\end{bmatrix}$

we take $\sqrt{|\text{grad}\ \hat{s}_x(x,y)|^2 + |\text{grad}\ \hat{s}_h(x,y)|^2}$.

? Cluster Analysis (not discussed here)

1-dimensional data: Detection of floes

floes

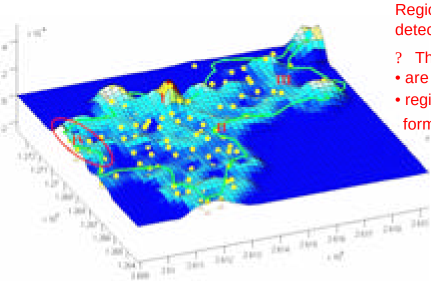


The peaks of the derivative of signal estimate allows the identification of floe boundaries. The square of the signal derivative enhances the peaks.

Canton Basel - Stadt: Derivative of signal estimate.

? fiducial points

— boundary

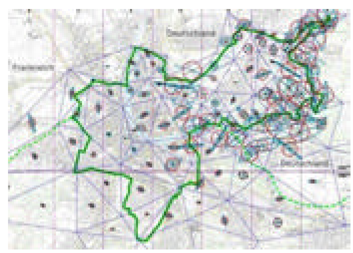


I – IV are regions of strong signal variation.

Region IV was newly detected.

? These regions
- are floe boundaries or
- regions of small transformation triangles

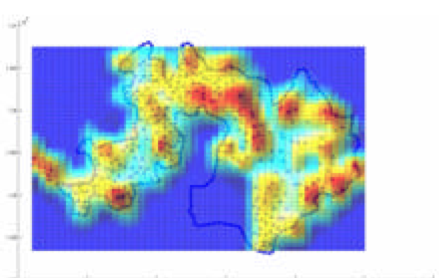
Canton Basel - Stadt: Triangles for Transformation LV03 ? LV95 with deformation ellipses.



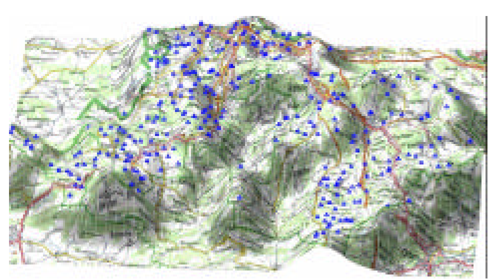
The choice of triangles based on the derivative of the signal estimate showed to very efficient. The same procedure was applied to the town districts with denser sets of fiducial points.

Canton Basel – Landschaft: Derivative of signal estimate

? fiducial points



Canton Basel – Landschaft:
Derivative of signal estimate as landscape on a map
? fiducial points



Discussion

- ? Semiparametric regression is a valuable tool for detection and diagnosis of effects with a nonparametric part. Examples of application:
  - Performing the layout of LV03 ? LV95 transformation triangles
  - Actual running project: Modelling the locally different shifts of a sliding slope

$$\textit{new location} = \textit{old location} + \textit{shift} + \textit{signal} + \textit{noise}$$

- ? Semiparametric regression is an **ill posed method** in the sense that the estimated model and signal are not uniquely determined. The choice of basis functions or method for determining *a* can be varied. Emphasis lies more on reproducability of the results than on objectivity.

- ? Semiparametric Regression versus Collocation and Filtering:

  The signal is a function and not the realization of a stochastic process. Thus prediction reduces to simple interpolation.

- ? Modern nonparametric methods may also be applied to other problems, e.g the fit of a parametrised curve to kinematic data.

REFERENCES

Bruehlmann, P.: Koordinatentransformationen: Nichtparametrische Regression und Schollendetektierung. Diploma thesis, dept. of geodesy, FHBB, 2004 (notpublished)

Fischer, B. and Hegland, M.: Collocation, Filtering and Nonparametric Regression. ZfV 124(1), S. 17-24, 124(2), S. 46-52.

Fischer, B., Schwaninger, M. and Sievers, B.: Nichtparametrische Regression am Beispiel der Helmerttransformation. Geomatik Schweiz 6/2003, S. 346-47

Green, J.P. and Silverman, B.W.: Nonparametric regression and generalized linear models. Chapman and Hall, London 1994.

Hansen, P. and O'Leary, D.P.: The use of the L-curve in the regularization of discrete ill-posed problems. SIAM (Society of Industrial and Applied Mathematics) J. Sci. Comp. 14(6): 1487-1503, 1993.

Hastie, T.J. and Tibshirani, R.J.: Generalized additive models. Chapman and Hall, London 1990.

Jansa, J. and Augustin, L.: Ein Vorschlagfuer die Homogenisierung des oesterreichischen Fixpunktnetzes. Vermessung & Geoinformation 2/ 2004, p. 59– 67.

Liechti, H.U. and Haffner, P.: Dreiecksvermaschung Kanton Basel-Stadt, Technischer Bericht, Basel 2006 (notpublished).

Moritz, H.: Advanced physical geodesy, 2nd ed.. Wichmann, Karlsruhe 1989.

Ruppert, D., Wand, M.P. and Carroll, R.J.: Semiparametric Regression, Cambridge University Press 2003.

Schwaninger, M.: Nichtparametrische Regression: Softwareentwicklung. Diploma thesis, dept. of geodesy, FHBB, 2002 (not published)

swisstopo: Transformation LV03? LV95, Bereitstellung der Transformationsgrundlagen für die AV, Weisungen betreffend die Erstellung der kantonalen Dreiecksvermaschung, 2000.

swisstopo: Transformation LV03? LV95, Bereitstellung der Transformationsgrundlagen für die AV, Empfehlungen bez üglich der Qualitätsanforderungen bei der Erstellung der Dreiecksvermaschung durch die Kantone, April 2004.

swisstopo: Transformation LV03? LV95, Bereitstellung der Transformationsgrundlagen für die AV, Technische Hinweise zur Erstellung der Dreiecksvermaschung, Mai 2004.