Extension of *k*-most similar neighbours methods by local polynomial regression

Sebastian Schoneberg
Northwest German Forest Research Station,
Dept. Ecoinformatics, Biometrics and Forest Growth
Georg-August University of Göttingen

07.11.2013

Introduction

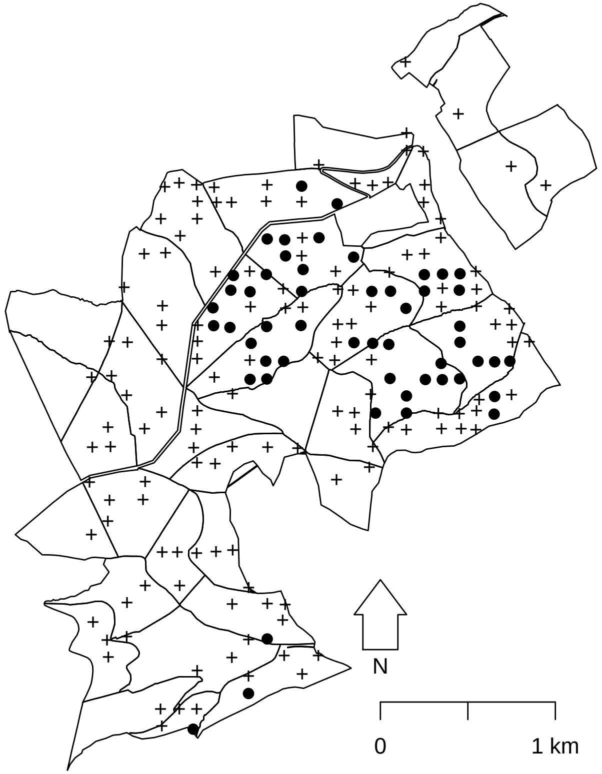


Figure: Spatial distribution of RFI-plots. Dots indicate plots with beech of DBH ≥ 60 cm.

- most dendrometrical variables are required for single stands
- only a few or even zero sample plots are located in a stand
- remote sensing bridge the gap between the RFI and the need for precise data for each of the small stands

Data

- the study site is a part of the Krofdorfer Wald, located in Hessen, Germany
- RFI was conducted as double sampling for stratification
- vegetation height calculated by using CIR images, spatial resolution of the CIR images is 20 cm × 20 cm

Table: Descriptive statistics of the forests within the study area. Based upon RFI.

		total	beech ($\mathrm{DBH} \geq 60\,\mathrm{cm}$)
	number of trees	1746	70
DBH [cm]	mean	32.65	66.04
age [years]	mean	84.43	149.74
	number of heights	448	50
height [m]	mean	25.07	37.34

Methods – target and auxiliary variables

- target variable:
  volume per hectare based on RFI data: $V_i = \sum_{j=1}^{m_i} \frac{10000}{\pi r_{ij}^2} v_{ij}$
- auxiliary variable:
  the study area was divided into 25 m × 25 m subareas,
  for each descriptive statistics were calculated
  by using the vegetation height and the $NDVI = \frac{NIR - red}{NIR + red}$

Methods – *k*-MSN regression

- *k*-NN estimate: $\hat{f}(x) = \underset{i \in N(x)}{\text{ave}} \; y_i$
- identify nearest neighbors: $d_i = \sqrt{(x_i - x)^T \Gamma \Lambda^2 \Gamma^T (x_i - x)}$ (Moeur & Stage, 1995)
  takes into account correlations and explanatory power among the auxiliary variables through canonical correlations analysis

Methods – LPR, general approach

- introduce distance dependent weights for the considered neighbours
- any smooth function can be locally approximated by a polynomial of certain order using a Taylor series

$$\begin{aligned} f(z_i) &\approx f(z) + (z_i - z) \cdot f^{'}(z) + (z_i - z)^2 \cdot \frac{f^{''}(z)}{2} + \cdots + (z_i - z)^l \cdot \frac{f^{(l)}(z)}{l!} \\ &= \gamma_0 + (z_i - z) \cdot \gamma_1 + (z_i - z)^2 \cdot \gamma_2 + \cdots + (z_i - z)^l \cdot \gamma_l \end{aligned}$$

- the general LPR approach can be transferred to our spatial prediction problem:

$$\begin{aligned} y_i &= f(x_i) + \varepsilon_i \\ &\approx \gamma_0 + d_i \cdot \gamma_1 + d_i^2 \cdot \gamma_2 + \cdots + d_i^l \cdot \gamma_l + \varepsilon_i . \end{aligned}$$

Methods – LPR, least square

- γ_0 stands for the target variable $y(x)$ at a subarea associated with the vector of auxiliary variables x.
- lead to a weighted least square criterion: $\sum_{i=1}^{n}\left(y_i - \sum_{m=0}^{l}\gamma_m d_{ij}^m\right)^2 w_\lambda(d_{ij})$
- $w_\lambda(d_{ij})$ becomes larger if d_{ij} becomes smaller, by using kernel functions: $w_\lambda(d_{ij}) = K\left(\frac{d_{ij}}{\lambda}\right)$

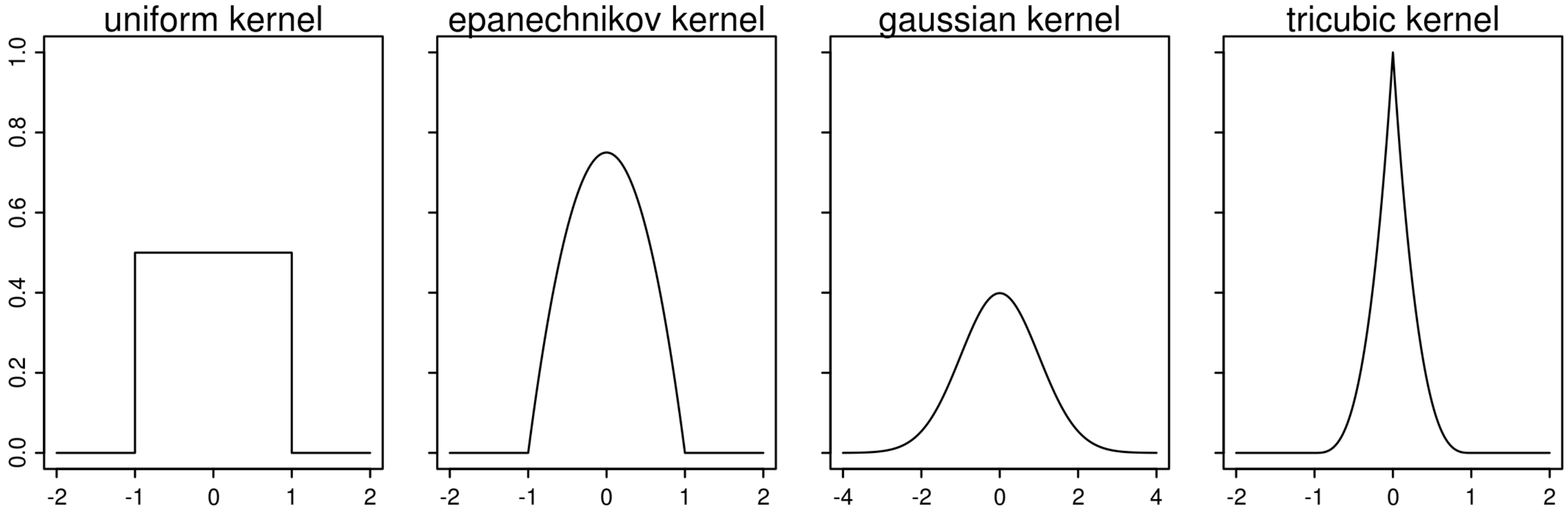


Figure: Used kernel functions.

Methods – LPR, estimation of $\hat{f}(x)$

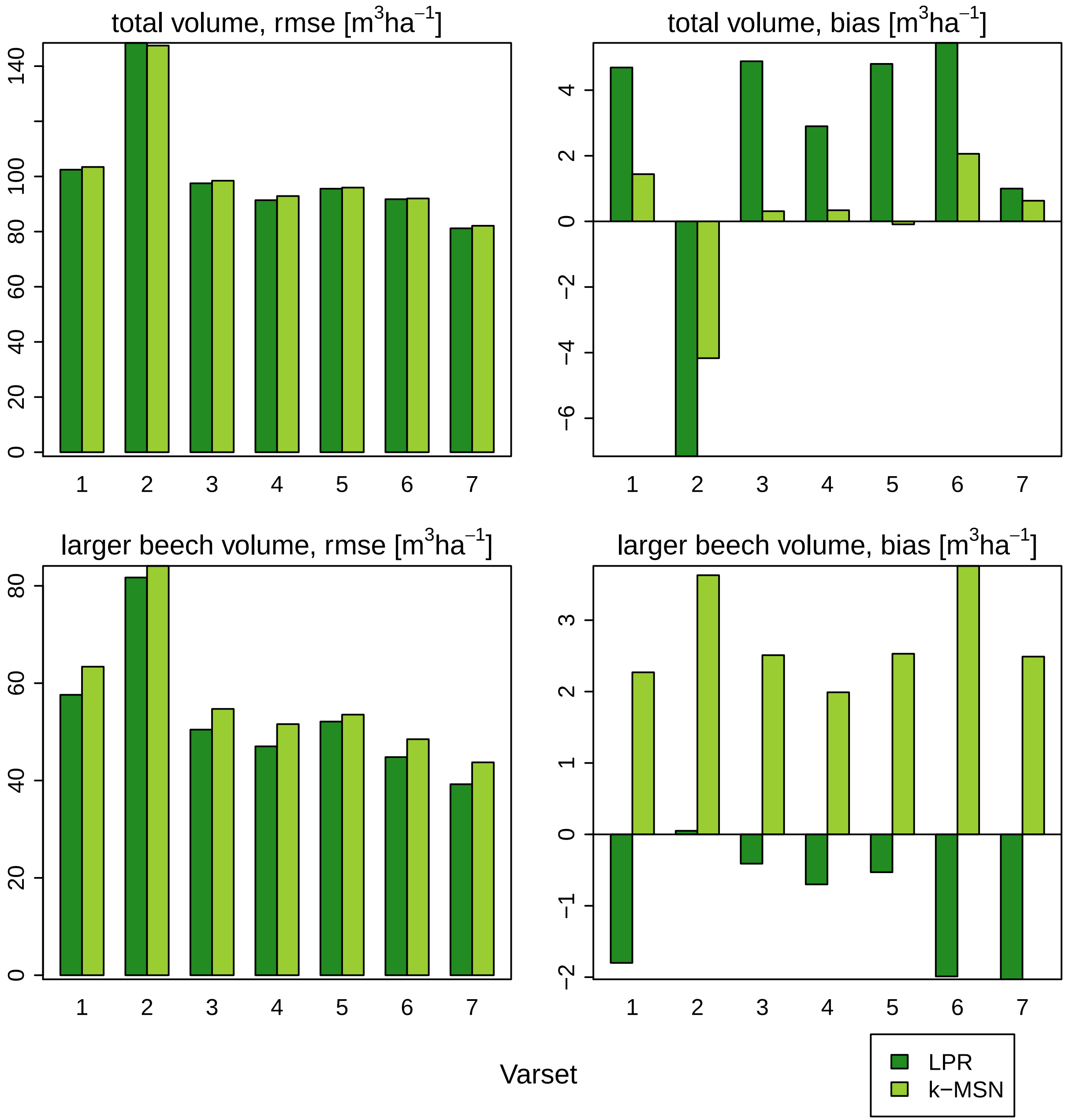
- weighted least square estimator:

$$\hat{f}(x) = \hat{\gamma}_0 = e^T (Z^T W Z)^{-1} Z^T W y$$

- with

$$Z = \begin{pmatrix} 1 & d_1^1 & \dots & d_1^l \\ \vdots & & & \vdots \\ 1 & d_n^1 & \dots & d_n^l \end{pmatrix}, \; W = \begin{pmatrix} w_\lambda(d_1) & & 0 \\ & \ddots & \\ 0 & & w_\lambda(d_n) \end{pmatrix}$$

Results


total volume, rmse [m³ha⁻¹]
0
20
40
60
80
100
140
total volume, bias [m³ha⁻¹]
−6
−4
−2
0
2
4
larger beech volume, rmse [m³ha⁻¹]
0
20
40
60
80
larger beech volume, bias [m³ha⁻¹]
−2
−1
0
1
2
3
1
2
3
4
5
6
7
Varset
LPR
k−MSN

Discussion and conclusions

- with regard to predictions of total volume LPR achieves a slightly higher precision than k-MSN
- for volume prediciton of large beech trees precision of LPR is clearly higher than with k-MSN
- the systematic error is mostly not larger than about 5 % of the according RMSE

Thank you

for your attention!

Literature

Anttila, P. (2002). Nonparametric estimation of stand volume using spectral and spatial features of aerial photographs and old inventory data. *Canadian Journal of Forest Research*, 32(10), 1849–1857.

Eskelson, B. N. I., Barrett, T. M., & Temesgen, H. (2009). Imputing mean annual change to estimate current forest attributes. *Silva Fennica*, 43(4), 649–658.

Fahrmeir, L., Kneib, T., & Lang, S. (2007). *Regression: Modelle, Methoden und Anwendungen ; mit 51 Tabellen*. Berlin; Heidelberg: Springer.

Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., & Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment*, 112(5), 2232–2245.

Latifi, H., Nothdurft, A., & Koch, B. (2010). Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry*, 83(4), 395–407.

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2004). *Remote sensing and image interpretation*. New York: Wiley.

Malinen, J. (2003). Locally adaptable non-parametric methods for estimating stand characteristics for wood procurement planning. *Silva Fennica*, 37(1), 109–120.

McRoberts, R. E. (2012). Estimating forest attribute parameters for small areas using nearest neighbors techniques. *Forest Ecology and Management*, 272, 3–12.

Mikhail, E. M., Bethel, J. S., & Ch., M. J. (2001). *Introduction to Modern Photogrammetry*. Wiley, New York.

Moeur, M. & Stage, A. R. (1995). Most similar neighbor: An improved sampling inference procedure for natural resource planning. *Forest Science*, 41(2), 337–359.

Nothdurft, A., Saborowski, J., & Breidenbach, J. (2009). Spatial prediction of forest stand variables. *European Journal of Forest Research*, 128(3), 241–251.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Saborowski, J., Marx, A., Nagel, J., & Böckmann, T. (2010). Double sampling for stratification in periodic inventories-infinite population approach. *Forest Ecology and Management*, 260(10), 1886–1895.

Straub, C., Weinacker, H., & Koch, B. (2010). A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. *European Journal of Forest Research*, 129(6), 1069–1080.