A hybrid design and model-based sampling approach for regional trend monitoring
Dick. J. Brus and Jaap J. de Gruijter

1 Introduction

This poster describes a novel statistical approach for sampling in space and time for estimating space–time parameters such as the space–time mean (total) or trend of the spatial mean (total). In this hybrid, design- and model-based approach sampling locations are selected by probability sampling. Due to this the underlying space–time process that generated the data need not be modeled fully. The spatial means can be estimated model-free by design-based inference, so that a model for the temporal variation of the spatial means suffices. The model contains two error terms, one for model inadequacy and another for sampling error in the estimated spatial means. Important advantages of the presented approach over the fully model-based approach are its simplicity and robustness to model assumptions. The hybrid approach is illustrated with the trend of the spatial mean of three soil-chemical variables (pH, NO₃, NH₄) measured at three depths in the soil profile.

2 Space–time parameter

In the hybrid, design and model-based approach the time-series of spatial means is described by a linear mixed model, for instance by:

\[ Z(t_j) = \beta_1 + \beta_2 \cdot t_j + \eta(t_j) \quad j = 1 \cdots r \]  

(1)

where \( \eta(t_j) \) is the model residual (model error) of the spatial mean at time \( t_j \). The slope parameter \( \beta_2 \) describes the trend of the spatial means (average change per time unit), and is the target parameter to be estimated.

In practice the spatial means are unknown and are estimated from a sample. Consequently, the model is extended with a sampling error \( \epsilon \):

\[ \tilde{Z}(t_j) = \beta_1 + \beta_2 \cdot t_j + \eta(t_j) + \epsilon(t_j) \quad j = 1 \cdots r \]  

(2)

3 Space–time design

Sample data were collected in 2004, 2005, 2006 and 2007 according to a rotational panel design. In each year 20 locations were selected by simple random sampling. 10 out of the 20 locations of 2004 locations were resampled in 2005, and 10 new locations were selected. These 10 new locations of 2005 were resampled in 2006 and so on.

4 Estimation of trend

The model parameters \( \beta_1 \) and \( \beta_2 \) were estimated by

\[ \hat{\beta} = (D' \tilde{C}_{tp}^{-1} D)^{-1} D' \tilde{C}_{tp}^{-1} \tilde{Z} \]  

(3)

with \( D \) the design-matrix with 1’s in first column and the sampling times in the second column, \( \tilde{Z} \) the vector with estimated spatial means (see hereafter), and \( \tilde{C}_{tp} \) the sum of the estimated sampling variance-covariance matrix \( C_p \) and the estimated model variance-covariance matrix \( C_t \).

5 Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>-0.05</td>
<td>NS</td>
</tr>
<tr>
<td>NO₃</td>
<td>0.02</td>
<td>NS</td>
</tr>
<tr>
<td>NH₄</td>
<td>0.01</td>
<td>NS</td>
</tr>
</tbody>
</table>

6 Conclusions

For all variables the estimated trend was not significant (\( \alpha = 0.05 \), except for pH in the subsoil. For NO₃ and NH₄ in the topsoil the estimated model variances were relatively large, leading to a strong contribution to the uncertainty about the trend. The sampling error can be manipulated by the space–time design (type of space–time design, type of spatial design, number of sampling locations, sampling frequency). Further research into the effect of the space–time design on the accuracy of the estimated regional trend is needed.

References


*corresponding author (D.J.brus@wur.nl)