ALTERRA WAGENINGEN UR

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, NO_3, NH_4) 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:

$$\bar{Z}(t_j) = \beta_1 + \beta_2 \cdot t_j + \eta(t_j) \;\; j = 1 \cdots r \qquad (1)$$

where $\eta(t_j)$ is the model residual (model error) of the spatial mean at time t_j. The slope parameter β_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 ϵ

$$\widehat{\bar{Z}}(t_j) = \beta_1 + \beta_2 \cdot t_j + \eta(t_j) + \epsilon(t_j) \;\; j = 1 \cdots r \qquad (2)$$

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.

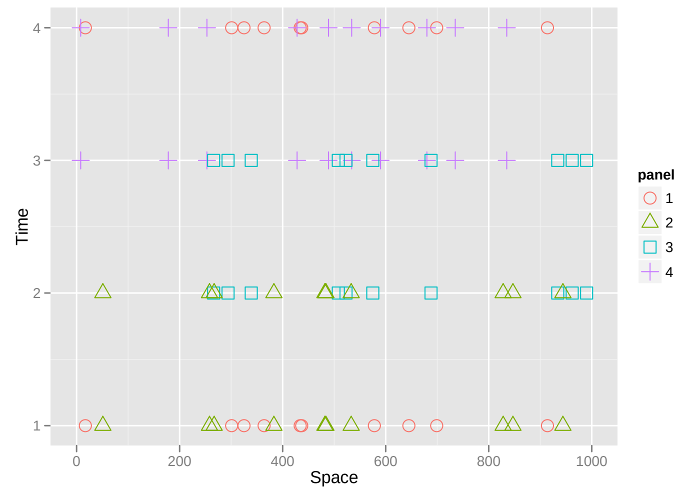

Figure 1: Rotating panel design with two panels of equal size (10 locations) per sampling time

4 Estimation of trend

The model parameters β_1 and β_2 were estimated by

$$\hat{\boldsymbol{\beta}} = (\mathbf{D}'\widehat{\mathbf{C}}_{\xi p}^{-1}\mathbf{D})^{-1}\mathbf{D}'\widehat{\mathbf{C}}_{\xi p}^{-1}\hat{\mathbf{Z}} \qquad (3)$$

with $\mathbf{D}$ the design-matrix with 1's in first column and the sampling times in the second column, $\hat{\mathbf{Z}}$ the vector with estimated spatial means (see hereafter), and $\hat{\mathbf{C}}_{\xi p}$ the sum of the estimated sampling variance-covariance matrix $\widehat{\mathbf{C}}_p$ and the estimated model variance-covariance matrix $\widehat{\mathbf{C}}_\xi$.

Estimation of spatial means and their sampling variances and covariances

The spatial means at the sampling times and their sampling variances and covariances were estimated by the usual design-based estimators, also referred to as the Horvitz-Thompson (HT) estimator. For simple random sampling the HT-estimator of the mean is the unweighted sample average. The HT estimator of the sampling variances is $\widehat{S^2}_j/n_j$ with $\widehat{S^2}_j$ the estimated spatial variance at time t_j, and n_j the sample size at time t_j. The HT-estimator of the sampling covariance of two estimated means is $\widehat{C}_{jk} = \frac{\widehat{S^2}_{jk} \cdot m_{jk}}{n_j \cdot n_k}$ with $\widehat{S^2}_{jk}$ the estimated spatial covariance of the observations at times t_j and t_k, and m_{jk} the number of points in the overlapping part of the samples at times t_j and t_k. The sampling covariance of the estimated means at times t_1 and t_3 is 0, as there is no overlap. The same holds for the sampling covariance of the means at times t_2 and t_4.

A more advanced estimation method for space–time designs with partial overlap such as the rotating panel design is to estimate the trend from the panel-specific estimates of the spatial means, see for details Brus and de Gruijter (2011, 2012). In Figure 1 we have two panels per sampling time, leading to eight estimated means.

Estimation of model variance-covariance matrix

We assumed that the spatial means were generated by a continuous AR(1) model, which has an exponential covariance function. The distance parameter a of this covariance function was estimated by fitting an exponential model to the pooled experimental correlograms at point locations. The variance parameter τ^2 was then fitted by REML, conditional on this estimated distance parameter

5 Results

$\hat{\beta}_2$: est. trend; se($\hat{\beta}_2$): standard error; $\hat{a}$: est. distance parameter (yr); $\hat{\tau}^2$: est. model variance. Last three columns: estimates obtained with panel-specific estimates of means

	$\hat{\beta}_2$	se($\hat{\beta}_2$)	$\hat{a}$	$\hat{\tau}^2$	$\hat{\beta}_2$	se($\hat{\beta}_2$)	$\hat{\tau}^2$
pHtop	0.028	0.028	6.9	0	0.033	0.068	0.060
NO_3top	0.011	0.32	3.6	0.81	0.033	0.27	0.54
NH_4top	-0.052	0.26	0.72	0.18	-0.040	0.26	0.19
pHmid	0.020	0.023	3.3	0	0.022	0.017	0
NO_3mid	0.025	0.091	0.76	0.028	0.0015	0.087	0.025
NH_4mid	-0.089	0.074	0.67	0.014	-0.079	0.065	0.0089
pHsub	*-0.044*	0.020	6.3	0	-	-	0.079
NO_3sub	0.020	0.067	1.7	0.015	-	-	0.024
NH_4sub	-0.030	0.018	1.5	0	-0.013	0.015	0

6 Conclusions

For all variables the estimated trend was not significant ($\alpha = 0.05$), except for pH in the subsoil. For NO_3 and NH_4 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

Brus, D. J., de Gruijter, J. J., 2011. Design-based Generalized Least Squares estimation of status and trend of soil properties from monitoring data. *Geoderma* **164**, 172–180.

Brus, D. J., de Gruijter, J. J., 2012. A hybrid design-based and model-based sampling approach to estimate the temporal trend of spatial means. *Geoderma* **173-174**, 241–248.

*corresponding author (✉ dick.brus@wur.nl)