Which sampling design to monitor saturated hydraulic conductivity?

S.K. HASSLER [a,b], R.M. LARK [c], B. ZIMMERMANN [b, *] & H. ELSENBEER [b,d]

[a]*Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Germany,*

[b]*Institute of Earth and Environmental Sciences, University of Potsdam, Germany,*

[c]*British Geological Survey, Keyworth, Nottingham, U.K., and*

[d]*Smithsonian Tropical Research Institute, Balboa, Ancón, Panama*

Correspondence: S. K. Hassler. E-mail: sibylle.hassler@gfz-potsdam.de.

Running title: Efficiency of rotational stratified sampling

Key words: rotational stratified sampling, efficiency, reforestation, tropics, soil permeability

[*]*Present address: Research Institute for Post-Mining Landscapes (FIB e.V.), Finsterwalde, Germany.*

Summary

Soil in a changing world is subject to both anthropogenic and environmental stressors. Soil monitoring is essential to assess the magnitude of changes in soil variables and how they affect ecosystem processes and human livelihoods. However, we cannot always be sure of which sampling design is best for a given monitoring task.

We employed a rotational stratified simple random sampling (rotStRS) for the estimation of temporal changes in the spatial mean of saturated hydraulic conductivity (K_s) at three sites in central Panama in 2009, 2010 and 2011. To assess this design's efficiency we compared the resulting estimates of the spatial mean and variance for 2009 to those gained from stratified simple random sampling (StRS) which was effectively the data obtained on the first sampling time, and to an equivalent unexecuted simple random sampling (SRS).

The poor performance of geometrical stratification and the weak predictive relationship between measurements of successive years yielded no advantage of sampling designs more complex than SRS. The failure of stratification may be attributed to the small large-scale variability of K_s. Re-visiting previously sampled locations was not beneficial because of the large small-scale variability in combination with destructive sampling, resulting in poor consistency between re-visited samples. We conclude that for our K_s monitoring scheme, repeated SRS is equally effective as rotStRS. Some problems of small-scale variability might be overcome by collecting several samples at close range to reduce the effect of fine-scale variation. Finally, we give recommendations on the key factors to consider when deciding whether to use stratification and rotation in a soil monitoring scheme.

Introduction

Soil in a changing world is subject to both anthropogenic and environmental stresses. Yet soil provides the basis for food production and various ecosystem services. Changes in soil properties, their magnitude, rate and associated processes, are thus becoming increasingly important for management of natural resources and human livelihoods. For example, in many regions undergoing land-use change, soil is increasingly susceptible to erosion, leading to a decrease in fertility of agricultural areas and larger sediment loads in rivers (for example Giertz *et al.*, 2005; Huth *et al.*, 2012). In order to assess changes in soil properties on relevant spatial and temporal scales, soil monitoring studies, the repeated measurement of soil properties, are essential (Arrouays *et al.*, 2009).

When designing soil sampling schemes for monitoring purposes the first decision usually is whether to use a model-based or design-based approach (Brus & de Gruijter, 1993; Papritz & Webster, 1995; Brus & de Gruijter, 1997). A model-based approach is based on the assumption that the values of a soil variable in the study area can be modelled as a stochastic process. Because the model is the source of randomness in the subsequent data analysis the sampling need not be randomized and is commonly performed on a grid, which distributes the samples regularly over the study area and is especially suited for constructing maps of the soil variable. Inferences from these data are based on the model. However, if the assumptions of the model are not met, statistical inference from this design is invalid (Brus & de Gruijter, 1997; Arrouays *et al.*, 2012). Design-based methods, in contrast, do not assume an underlying model of the soil variable and base statistical inference solely on the inclusion probabilities of the sampling locations which are determined by the applied sampling design. They are often reported to be more suitable than model-based approaches for the determination of the spatial mean of an area and when only a small sample size is feasible (Brus & de Gruijter, 1993; 1997; Lark, 2009).

If the aim of a soil sampling scheme is to assess the spatial mean of a soil variable and having selected a design-based approach, the next step is to decide on the details of the sampling design. Two widely used designs are simple random sampling (SRS) and stratified simple sandom sampling (StRS), described in depth by de Gruijter *et al.* (2006). Whereas SRS uses the whole study area to select random samples, in StRS the study area is first sub-divided into strata before sampling randomly within the strata. Stratification can be based on previous knowledge of underlying processes influencing the target soil variable or simply by dividing up the study area into compact strata. To determine the average status and change of a soil variable over large regions, stratified designs have been shown to be more efficient than SRS in various studies (Papritz & Webster, 1995; Black *et al.,* 2008; Arrouays *et al.,* 2012). However, an increase in efficiency depends on a substantial proportion of the variation of the soil variable being accounted for by the stratification, resulting in smaller within-stratum variances compared to the overall variance.

The aims of sampling and the options for design are more complex in the case of monitoring. One key design decision is whether or not to re-visit some or all previously sampled locations in order to form a set of direct observations of change between the two sampling times. This approach is generally most efficient if the primary objective is to estimate change (Lark, 2009). However, if we are also interested in the spatial means for each sampling time, as in the present study, it may be advantageous to use a sampling design in which only a proportion of the sampling locations is re-visited and some additional locations are included in the second sampling time to increase the spatial extent of the sample (de Gruijter *et al*., 2006). This is termed a rotational design. The best sampling strategy depends, among other factors, on logistical constraints (maximum sample size, the challenges of re-locating sample sites and costs of repeated sampling campaigns) along with the spatio-temporal characteristics of the soil variable.

The target monitoring variable of this study is the saturated hydraulic conductivity (K_s) of the soil, a critical parameter in the water cycle. In the humid tropics, K_s changes mainly because of shifts in land use. Conversion of tropical forest to pasture has been widely shown to affect top-soil soil hydrological properties including K_s (Alegre & Cassel, 1996; Martinez & Zinck, 2004). A consequence of this process can be the increased frequency of occurrence of overland flow and risk of top-soil erosion as the vertical water flow path is increasingly hindered by reduced K_s (Bonell & Gilmour, 1978; Hanson *et al.*, 2004; Germer *et al.*, 2010). In the last two decades, a different trend in land-use change has been observed; pastures and fields are being actively replanted with timber species or recolonized by secondary succession. With one exception (Zimmermann *et al.*, 2010a), the consequences of this reforestation for soil hydraulic properties have all been examined with space-for-time approaches which assume that soils at different sites under varying stages of reforestation can be regarded as examples of the temporal trend in soil properties under reforestation at a fixed location (Zimmermann *et al.*, 2008; Hassler *et al.*, 2011b; Nyberg *et al.*, 2012; Peng *et al.*, 2012). However, the space-for-time approach relies on various assumptions which have been criticised (Tye *et al*., 2013). In particular, this will not work if the likelihood of reforestation happening in a particular part of the landscape is not independent of the soil properties at that location. In order to provide definitive information on how hydraulic properties change under reforestation, unconfounded with possible spatial variation and space-time interactions, it is essential to monitor variables such as K_s at reforestation sites: however, it is not obvious which particular sampling design should be used for this task.

The aim of this study was twofold. (i) To employ a rotational stratified simple random sampling design (rotStRS) for the estimation of the temporal change of the spatial mean in K_s at three reforestation sites in Central Panama. In this design a proportion of sampling locations are re-visited at consecutive sampling times while new locations are also added. Furthermore, the random sampling is done within strata. (ii) To assess the efficiency of the employed design by comparing estimates of spatial mean and variance of the first sampling time to those of a StRS design, which effectively represents the first-year sampling of rotStRS. Additionally, we calculate the equivalent variance if the sample had been obtained from a SRS.

Sampling designs

This section gives an overview of the sampling designs that we considered, in addition to the rotStRS design used for sampling, and lists the equations to calculate their means and variances (adapted from de Gruijter *et al.*, 2006). A schematic to visualize the differences between the three designs is shown in Figure 1.

Simple random sampling (SRS)

Sampling points are selected at random within the study area. Equations are adapted from de Gruijter *et al.* (2006), page 82ff. In this presentation the *i*th observation of the target variable is denoted by z_i.

With sample size n the estimated spatial mean for SRS across the study area is calculated by

$$\hat{\bar{z}}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} z_i. \quad (1)$$

The sampling variance of the estimated spatial mean is given by

$$\hat{V}(\hat{\bar{z}}_{SRS}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(z_i - \hat{\bar{z}}_{SRS})^2, \quad (2)$$

and the spatial variance is estimated by:

$$\widehat{S^2}_{SRS}(z) = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \hat{\bar{z}}_{SRS})^2. \quad (3)$$

Stratified simple random sampling (StRS) with compact geographical stratification

The study area is divided into strata of equal size, random sampling is then done within the strata. Equations are adapted from de Gruijter *et al.* (2006).

For StRS the spatial mean can be estimated by

$$\hat{\bar{z}}_{StRS} = \sum_{h=1}^{H} a_h \, \hat{\bar{z}}_h, \tag{4}$$

where $\hat{\bar{z}}_h$ is the sample mean in stratum h, H is the number of strata and a_h is the relative area of stratum h.

The variance of $\hat{\bar{z}}_{StRS}$ can be estimated by

$$\hat{V}(\hat{\bar{z}}_{StRS}) = \sum_{h=1}^{H} a_h^2 \, \hat{V}(\hat{\bar{z}}_h), \tag{5}$$

with $\hat{V}(\hat{\bar{z}}_h)$ being the estimated variance of $\hat{\bar{z}}_h$ calculated as follows:

$$\hat{V}(\hat{\bar{z}}_h) = \frac{1}{n_h(n_h-1)} \sum_{i=1}^{n_h} (z_{hi} - \hat{\bar{z}}_h)^2. \tag{6}$$

Here $n_h$is the sample size in stratum h.

The spatial variance, that is to say the variance of the variable across the sampled area. can be estimated by

$$\widehat{S^2{}_{StRS}}(z) = \widehat{\overline{z^2}}_{StRS} - (\hat{\bar{z}}_{StRS})^2 + \hat{V}(\hat{\bar{z}}_{StRS}), \tag{7}$$

where $\widehat{\overline{z^2}}_{StRS}$ is the estimated mean of the target variable squared. It is calculated in the same way as $\hat{\bar{z}}_{StRS}$, but using squared values of the target variable.

For comparisons between sampling designs we can calculate the variance of the sample mean that we would obtain if we would sample applying SRS with the same total sample size, n , as StRS; $n = \sum_{h=1}^{H} n_h$,

$$\breve{V}(\hat{\bar{z}}_{SRS}) = \frac{\widehat{S^2}_{\mathrm{StRS}}(z)}{n}, \tag{8}$$

where the breve accent on $\breve{V}$ indicates that this variance is based on the estimate of the sample mean, and is not itself a design-based variance for a mean from a simple random sample.

Rotational stratified random sampling (rotStRS)

Rotational sampling is applied for soil monitoring, for example, if the spatial mean of a target variable is estimated at multiple sampling times. It includes the re-visiting of some sampling locations at consecutive sampling times, called the matched sample. If these observations are correlated, the efficiency of estimation of the spatial mean at the second sampling time can be increased by including a regression estimator gained from the matched sample. Not all sampling locations are re-visited, and at each subsequent sampling time, additional locations are established. The locations that are not re-visited and that are unique to one sampling time are called the unmatched sample. When the rotational design is based on stratified sampling in space, some of the points within each stratum are kept, and new ones are additionally established for consecutive sampling times. Equations are adapted from de Gruijter *et al.* (2006), page 226ff.

The spatial mean for the second sampling time is estimated by the composite estimator

$$\hat{\bar{z}}_{2c} = \widehat{w}_1\, \hat{\bar{z}}_{2gr}^{(m)} + \widehat{w}_2\, \hat{\bar{z}}_{2\pi}^{(u)} . \tag{9}$$

The second component of this estimator, $\hat{\bar{z}}_{2\pi}^{(u)}$, is the π -estimator for the mean of z_2 estimated only from the unmatched sample, according to the stratification (Equation 4). The first component, $\hat{\bar{z}}_{2gr}^{(m)}$ is a regression estimator of the spatial mean of z_2. This is calculated by

$$\hat{\bar{z}}_{2gr}^{(m)} = \hat{\bar{z}}_{2\pi}^{(m)} + b(\hat{\bar{z}}_{1\pi} - \hat{\bar{z}}_{1\pi}^{(m)}), \tag{10}$$

where $\hat{\bar{z}}_{2\pi}^{(m)}$ is the π-estimator for the mean of z_2 estimated from the stratified matched sample and b is the regression coefficient from the regression of the matched sample from the second sampling time on the matched sample from the first sampling time. The estimate $\hat{\bar{z}}_{1\pi}$ is the mean of the stratified entire sample at the first sampling time, and $\hat{\bar{z}}_{1\pi}^{(m)}$ is the mean of the stratified matched sample only at the first sampling time. The two separate estimates of the spatial mean at the second sampling time are combined in Equation (9) by weights that sum to one. These weights are calculated by

$$\hat{w}_1 = 1 - \hat{w}_2 = \frac{\hat{V}\left(\hat{\bar{z}}_{2\pi}^{(u)}\right)}{\hat{V}\left(\hat{\bar{z}}_{2gr}^{(m)}\right) + \hat{V}\left(\hat{\bar{z}}_{2\pi}^{(u)}\right)} \quad (11)$$

where $\hat{V}\left(\hat{\bar{z}}_{2\pi}^{(u)}\right)$ is the estimated variance of the π-estimator for the mean of z_2 from the stratified unmatched sample, calculated according to Equation (6) and $\hat{V}\left(\hat{\bar{z}}_{2gr}^{(m)}\right)$ is the estimated variance of the regression estimator. The variance of the regression estimator is given by

$$\hat{V}\left(\hat{\bar{z}}_{2gr}^{(m)}\right) = \frac{\widehat{S^2}(e)}{m} + \frac{\widehat{S^2}(z_2) - \widehat{S^2}(e)}{n}, \quad (12)$$

where $\widehat{S^2}(e)$ is the estimated variance of the regression residuals (from the matched sample, ignoring stratification) and $\widehat{S^2}(z_2)$ is the estimated spatial variance of the stratified whole sample at the second sampling time, calculated according to Equation (7).

Finally, the variance of the composite estimator is estimated by

$$\hat{V}\left(\hat{\bar{z}}_{2c}\right) = \frac{1 + 4\,\hat{w}_1 \hat{w}_2 \left(\frac{1}{m-1} + \frac{1}{n-m-1}\right)}{\frac{1}{\hat{V}\left(\hat{\bar{z}}_{2gr}^{(m)}\right)} + \frac{1}{\hat{V}\left(\hat{\bar{z}}_{2\pi}^{(u)}\right)}}. \quad (13)$$

Similarly, the spatial mean of the first sampling time can be estimated with these equations by incorporating the appropriate information gained from the second-year data-set and the regression estimator based on the regression of the first-year matched sample on the second-year matched sample.

Materials and methods

Study site

The study was conducted in central Panama in the watersheds of Río Agua Salud and Río Mendoza, which drain into the Panama Canal, partly covering the project area of the Agua

Salud Project (Figure 1A). The study area is characterised by a strongly dissected pre-Tertiary basalt plateau at an elevation between 53 and 331 m above mean sea level, with narrow interfluves, linear slopes averaging 42% and narrow or no valley floors. Top-soil textures in the area vary from silty clay to clay, pH values (in water) range from 4.4 to 5.8 (J.S. Hall *et al*., unpublished data).

The climate of the study area is tropical with a distinct dry season from mid-December to April. According to long-term records from nearby Barro Colorado Island, annual rainfall averages 2641 ±485 mm (mean ±1 standard deviation, n = 82, data from 1929 to 2010, by courtesy of the Environmental Science Program, Smithsonian Tropical Research Institute, Republic of Panama), and mean daily temperature varies little throughout the year, averaging 27 °C (Dietrich *et al.*, 1996).

Land use in the area varies over short spatial and temporal scales and includes pastures, timber plantations and secondary forests of different ages. This study was focussed on three catchments undergoing reforestation. The first site was a 34-ha plantation with native species, established in 2008. Formerly the catchment had been actively used as pasture, but included some larger trees. The second site was also a small former pasture catchment, covered by 5.7 ha of 3-year old secondary succession. The third catchment holds a 10.9-ha teak plantation planted in 2008, formerly covered by a mixed land use which was partly pasture and partly shrub-land.

Sampling design

Each site was sampled to determine the spatial mean of K_s in the years 2009, 2010 and 2011 in a rotStRS design with compact geographical stratification. We first divided each of our catchments into twenty compact strata of equal area (19 in the case of the secondary-succession catchment) with a k-means clustering algorithm (Brus *et al.*, 1999) from the R package SPCOSA (Walvoort *et al.*, 2010). Within each stratum, we randomly selected two sampling locations in the first year (2009) and marked them after sampling. In the following year we kept one of these two points per stratum, discarded the other and randomly chose a new sampling point. For the third-year campaign the sampling points in the matched sample for 2009 and 2010 were discarded, the unmatched sample points from 2010 were retained (now the matched set for 2010/2011) and a new sample point was randomly selected within each stratum to constitute the unmatched sample set for 2011; see Figure 2 for an example. The re-sampling of points in the matched set in any year took place within a maximum distance of one metre from the previous year's sampling point. Note that this initial sampling design is not appropriate for rotStRS because there is only one matched and one unmatched sample point per stratum in any year (which does not permit the calculation of a within-stratum variance). For this reason we merged adjacent strata so that each of the new strata contained (ideally) two matched and two unmatched points in any one year. In the case of the secondary-succession catchment (19 initial strata) one cluster of 3 strata were merged, and the remainder were merged in pairs.

Field sampling of saturated hydraulic conductivity (K_s)

The saturated hydraulic conductivity (K_s) was measured on undisturbed soil cores. Two soil cores of 8.9 cm diameter were simultaneously extracted at depths 0–6 cm and 6–12 cm on levelled ground using a standard coring device (Soilmoisture Equipment Corporation, Santa Barbara, USA). Core ends were cut flat with a sharp knife and the samples were slowly saturated upside down over a period of 64 hours to prevent air entrapment. We measured K_s by applying a constant water head and following a simplified version of the methodology of Reynolds *et al.* (2002). After establishing a constant flow rate, K_s can be calculated according to Darcy's Equation for saturated conditions:

$$q = -K_s \, \mathrm{d}h/\mathrm{d}s, \tag{14}$$

where q is the flux density [m s^{-1}], K_s is the saturated hydraulic conductivity [m s^{-1}] and $\mathrm{d}h/\mathrm{d}s$ is the hydraulic gradient. The flux density can be expressed as $q = Q/A$ with Q being the water flux [m^3 s^{-1}] and A the cross-section of the sample [m^2].

Data analysis

Our data exhibited the well-known skewness for K_s. To obtain normally distributed data-sets for the analysis of the different sampling design estimates, we performed a Box-Cox transformation (Box & Cox, 1964). A common Box-Cox exponent was estimated for all data-sets (grouped by site, year and depth), and the BOXCOX procedure from the MASS package in R (Venables & Ripley, 2002) was used for estimation by maximum likelihood. The estimated value of the exponent was 0.16. Thus, all analyses were carried out with the transformed K_s values as follows:

$$z_{\mathrm{BC}} = \frac{(z^{0.16}-1)}{0.16}. \tag{15}$$

After estimating means and variances of the means, we calculated 95% confidence intervals around the means. The back-transformation of the means and confidence interval limits was done by

$$z = (z_{\mathrm{BC}} \times 0.16 + 1)^{1/0.16}. \tag{16}$$

Because the transformation is non-linear the simple back-transformation of the sample mean yields a value which is a biased estimate of the mean on the original scale of measurement. However, assuming normality of the transformed data, the back-transformed mean can be regarded as an estimate of the median on the original scale of measurement (Pawlowsky-Glahn & Olea, 2004) since the mean and median of a normal variable are coincident, and order statistics can be back-transformed simply for a monotonic transformation such as ours. Because of this monotonic property the upper and lower confidence interval limits can also be back-transformed directly.

The data analysis was split into three parts: First, we examined the temporal change in spatial mean of K_s according to the employed sampling design rotStRS, by plotting the means and confidence intervals for the different years, catchments and depths.

Second, we assessed the efficiency of the three different sampling designs rotStRS, StRS and SRS for estimating the sample mean of the first sampling time, 2009, by comparing the width of the respective confidence intervals. Calculations were done according to the equations cited in the Sampling Design section. We can do this design comparison because the first-year sampling considered in isolation can be analysed as a StRS, and Equation (8) provides the means to calculate the variance of the sample mean for a notional SRS with the same sample size as the StRS. Analysis of the rotStRS requires merged strata without missing samples, whereas the calculations according to a StRS could also be based on the original stratification without merging strata and hence, on a larger sample size. In order to have the same sample size for the comparison of sampling design efficiency we used a reduced data-set for each catchment and depth which satisfied the conditions for both rotStRS and StRS.

Third, we examined the benefits of stratification by comparing the spatial variance of the first-year StRS according to Equation (7) with a pooled within-stratum variance based on Equation (6), calculated as follows:

$$\hat{V}_{\text{pooled}} = \frac{1}{N-H}\sum_{h=1}^{H}\sum_{i=1}^{n_h}(z_{hi} - \hat{\bar{z}}_h)^2, \quad (17)$$

where N is the total sample size and H is the number of strata. We then assessed the benefits of including the regression estimator by evaluating the consistency between re-visited sampling locations with the regression of the matched sample of the second sampling time on the matched sample of the first sampling time.

All statistical analyses were carried out in the language and environment R (R Development Core Team, 2009).

Results

Change in K_s in the three catchments from 2009 to 2011

The estimated means in the native-species catchment suggested a decline in K_s from 2009 to 2011 at both depths (Figure 3), with the largest change from 2009 to 2010. The differences were particularly pronounced for the 6–12-cm depth where the confidence intervals for the 2009 and 2010/2011 estimates did not overlap. For the teak catchment at both depths any differences were small relative to the confidence intervals for the spatial mean in any one year. The secondary-succession catchment, however, showed an increase in K_s at the 0–6-cm depth which was large relative to the confidence intervals for 2009 and 2011. K_s at the 6–12-cm depth also showed an increase, but this was smaller.

Comparison of the different sampling designs

We assessed the efficiency of the different sampling designs by comparing the resulting estimated spatial means of K_s and their confidence intervals after back-transformation for the common data-set from 2009. The confidence interval limits of StRS, SRS and rotStRS (Figure 4) showed only negligible differences, therefore there was no general increase in efficiency. A possible exception could be seen for the teak catchment at the 0–6-cm soil depth as the confidence intervals for SRS and rotStRS were slightly wider than for StRS.

Analysis of spatial and within-stratum variance components and of the relationship between matched samples

In StRS, an increase in efficiency would be expected if the within-stratum variance was smaller than the spatial variance of the variable across the whole area. We assessed this by comparing the spatial variance of StRS with a pooled within-stratum variance (Table 1). The results showed that these variances values were within the same range, in two cases the pooled within-stratum variance was even larger than the spatial variance, thus hinting at only a very small or no increase in efficiency caused by stratification.

The rotational sampling is dependent on the regression of matched samples in consecutive years. Exemplary scatterplots of the matched samples of 2010 on 2009 for the three catchments and two soil depths are illustrated in Figure 5. The plots showed that there was no strong relationship between the matched samples of the two years. Similar weak relationships could be seen for the matched samples for the years 2011 on 2010 (plots not shown).

Discussion

Change of K_s in the three catchments

The observed decrease in K_s at both depths in the native-species catchment (Figure 3) could result from the consequences of rapid land cover change. In 2008, this catchment was an extensively managed pasture with some large trees, which were removed for reforestation with native species. During the felling and removal of the tree stumps, the soil was probably loosened to some extent, leading to an initial increase in K_s in 2009. The subsequent decrease back to values close to the baseline data might suggest a settling of the soil after the initial disturbance.

In the teak catchment any differences were small relative to the confidence intervals (Figure 3). The variation between the three years was probably also attributable to the rapid transformation of land cover, as here the formerly shrubby and diverse vegetation was removed for the teak plantation.

The catchment under secondary succession did not suffer from these severe changes; cattle grazing stopped here in the summer of 2006, after which secondary succession took over. The data exhibited a pronounced increase at the 0–6-cm depth and a weak increase at the 6–12-cm depth when compared with the baseline data. They showed the recovery of mainly top-soil K_s after abandoning pasture use and were consistent with other studies conducted in the same area (Hassler *et al.*, 2011b) and in other regions in the humid tropics (Zimmermann *et al.*, 2008; 2010a; Peng *et al.*, 2012).

Efficiency of stratification for better estimation of the variance

Confidence interval widths as calculated for StRS and SRS were very similar (with the possible exception of the teak catchment for the 0-6-cm soil depth). The stratification did not increase efficiency because of the very small difference between the spatial variance and pooled within-stratum variance (see Table 1). As noted above, the benefits of stratification are seen when the strata are internally uniform with regard to the target soil variable and most of the variation is seen between the strata. For K_s these differences could result from land cover or marked changes in soil type. In our catchments, however, land cover and soil type were relatively uniform; consequently, we divided the catchment into compact geographical strata. This type of stratification may nonetheless be beneficial, but only if the target soil variable exhibits spatial structure at larger scales, when the range of spatial autocorrelation is large (de Gruijter *et al.*, 2006; Walvoort *et al.*, 2010; Zimmermann *et al.*, 2010b). However, K_s frequently fails to exhibit large-scale structure, it is often characterized by substantial small-scale variability, partly because of the biotic influences acting on this scale which determine soil structure and partly an artificial effect when K_s is sampled with limited support such as small soil cores (Bouma, 1983; Mallants *et al.*, 1997; Sobieraj *et al.*, 2004; Hassler *et al.*, 2011a).

Efficiency of the rotational design

Rotational designs increase efficiency by incorporating knowledge about change of a soil property *via* the regression of matched samples. In our study, the regression estimator obviously did not improve the variance estimate substantially, as the confidence intervals of the rotational design rotStRS were similar or wider than for the non-rotational case StRS (Figure 4).

The reason for this became clear when examining the scatterplots of the matched samples for the different catchments, soil depths, comparing the years 2009 with 2010 and 2010 with 2011. The plots showed a weak relationship between the matched samples of all data-sets. The regression estimator will have advantages over alternatives estimating the spatial mean from single year data when there is a strong regression of the matched samples. With only a weak relationship the regression estimator may perform poorly because of the substantial uncertainty in the regression coefficient. We think that the reason why was there such little temporal consistency of samples taken at the same location was that we undertook destructive sampling by using soil cores. When re-sampling a sampling location, the core in the second year could be taken from no less than 50 cm from the previous year's sampling location in order to sample undisturbed conditions. Sometimes it was necessary to sample further from the original location if the nearer sites were affected by compression or large roots and so could not be sampled. Consequently, in some cases matched samples were located about one metre apart, and because of the small sample support and large small-scale variability they may not have qualified as matched samples (Goidts *et al.*, 2009). To overcome this problem, taking several samples at the same location to account for K_s small-scale variability could be a suitable approach. There are some examples in the soil organic carbon literature that detected significant temporal changes by expanding their support from locations to larger areas (Arrouays *et al.*, 2012).

Conclusions

The rotational stratified simple random sampling design that we used for our K_s monitoring studies did not yield the expected improvement in efficiency over simpler designs such as simple random sampling. The reasons for this were the small-scale variability and lack of large-scale structure in K_s: hence the strata were no less internally variable than the study site as a whole. Including a regression estimator of the spatial mean in the rotational design also did not yield benefits because of the poor consistency of the matched samples. The lack of consistency is probably because of the large short-range variability of K_s. Thus, when destructively sampling K_s using soil cores, matched samples, albeit close in space, might have very different K_s values.

For K_s studies, taking more samples at the same location to better incorporate small-scale variability and reduce the uncertainty of estimates for the spatial mean might overcome some of these problems. Generally, we recommend that the spatial structure and temporal consistency of the target variable are given careful thought when designing monitoring schemes. Appropriate information on the spatial variability of the variable and the potential to make consistent estimates at matched sites can be collected in a pilot study.

We summarize these conclusions in some practical recommendations for designing an efficient sampling scheme for soil monitoring:

- In a design-based approach, stratification is a good way to spread out the samples across the study site. However, in terms of improving the variance estimate, stratification is only useful either if the strata show marked differences in factors influencing the target variable, or, in the case of compact geographical strata, if the target variable exhibits large-scale spatial structure. A pilot study can give insights into the spatial structure of the variable or potential strata.
- Rotational designs are helpful in estimating a temporal trend of a target variable in other circumstances. In order to take advantage of the regression estimator, there must be a strong consistency between repeated observations at the same location.

An assessment of the ‘best possible’ consistency between re-visited samples could be done in a pilot study: if a set of exploratory sampling locations were sampled and then re-sampled, this would indicate how consistent matched observations can be in the absence of temporal change. If the consistency is poor then it would be clear that a rotational design has no advantages. Additionally, sampling more points at close range and thus increasing the support of the sample can be beneficial if sampling is destructive and therefore cannot target the same soil volume at consecutive sampling times.

- Judging how the target variable complies with the abovementioned conditions is the paramount step in deciding whether to include stratification or a rotational approach. If, as was the case for our K_s sampling, the conditions are not fully met, choosing SRS over more complicated designs will barely affect the efficiency of the estimates of the means and variances. In some cases SRS might even improve the estimates. Thus, if other considerations such as potential difficulties in re-visiting the exact same sampling points for rotational sampling or straightforwardness of data analysis play a role, repeatedly applying SRS poses a very suitable design option.

Acknowledgements

The Agua Salud Project (ASP), in which this study was undertaken, is a collaboration among the Panama Canal Authority (Autoridad del Canal de Panamá, ACP), the National Environmental Authority of Panama (Autoridad Nacional del Ambiente, ANAM), and the Smithsonian Tropical Research Institute (STRI). It is financially supported by the HSBC climate partnership (HCP), with additional funding from STRI, ACP, the Frank Levinson family foundation, and the Motta family foundation. RML's contribution to this work is published with the permission of the Director of the British Geological Survey (NERC). The field work of this study would not have been possible without the help of a number of interns and Bachelor students from the University of Potsdam, as well as local support by several ASP staff and people from the village El Giral.

References

ACP & ANAM. 2006. *Monitoreo de la Cuenca Hidrográfica del Canal de Panamá*. Programa de Vigilancia de la Cobertura Vegetal. URL: http://www.micanaldepanama.com/wp-content/uploads/2012/06/cobertura-vegetal.pdf (accessed 06/06/14), ACP & ANAM, Panama City, Panama

Alegre, J.C. & Cassel, D.K. 1996. Dynamics of soil physical properties under alternative systems to slash-and-burn. *Agriculture Ecosystems & Environment*, **58**, 39–48.

Arrouays, D., Bellamy, P.H. & Paustian, K. 2009. Soil inventory and monitoring. Current issues and gaps. *European Journal of Soil Science*, **60**, 721–722.

Arrouays, D., Marchant, B.P., Saby, N.P.A., Meersmans, J., Orton, T.G., Martin, M.P. *et al*. 2012. Generic Issues on Broad-Scale Soil Monitoring Schemes: A Review. *Pedosphere*, **22**, 456–469.

Black, H., Bellamy, P., Creamer, R., Elston, D., Emmett, B., Frogbrook, Z. *et al.* 2008. *Design and Operation of a UK Soil Monitoring Network*. Environment Agency. Bristol, UK.

Bonell, M. & Gilmour, D.A. 1978. Development of overland flow in a tropical rainforest catchment. *Journal of Hydrology*, **39**, 365–382.

Bouma, J. 1983. Use of soil survey data to select measurement techniques for hydraulic conductivity. *Agricultural Water Management*, **6**, 177–190.

Box, G.E.P. & Cox, D.R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, **2**, 211–252.

Brus, D.J. & de Gruijter, J.J. 1993. Design-based versus model-based estimates of spatial means - theory and application in environmental soil science. *Environmetrics*, **4**, 123–152.

Brus, D.J. & de Gruijter, J.J. 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma*, **80**, 1–44.

Brus, D.J., Spätjens, L.E.E.M. & de Gruijter, J.J. 1999. A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. *Geoderma*, 89, 129–148.

de Gruijter, J.J., Brus D.J., Bierkens, M.F.P. & Knotters, M. 2006. *Sampling for Natural Resource Monitoring*. Springer-Verlag, Heidelberg, Germany.

Dietrich, W.E., Windsor, D.M. & Dunne, T. 1996. Geology, Climate, and Hydrology of Barro Colorado Island. In: *The Ecology of a Tropical Forest: Seasonal Rhythms and Long-Term Changes* (eds E.G. Leigh, S.A. Rand & D.M. Windsor), pp. 21-46. Smithsonian Institution, Washington DC, USA.

Germer, S., Neill, C., Krusche, A.V. & Elsenbeer, H. 2010. Influence of land-use change on near-surface hydrological processes: Undisturbed forest to pasture. *Journal of Hydrology*, **380**, 473–480.

Giertz, S., Junge, B. & Diekkruger, B. 2005. Assessing the effects of land use change on soil physical properties and hydrological processes in the sub-humid tropical environment of West Africa. *Physics & Chemistry of the Earth*, **30**, 485–496.

Goidts, E., van Wesemael, B. & Crucifix, M. 2009. Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales. *European Journal of Soil Science*, **60**, 723–739.

Hanson, D.L., Steenhuis, T.S., Walter, M.F. & Boll, J. 2004. Effects of soil degradation and management practices on the surface water dynamics in the Talgua River Watershed in Honduras. *Land Degradation & Development*, **15**, 367–381.

Hassler, S.K., Lark, R.M., Milne, A.E. & Elsenbeer, H. 2011a. Exploring the variation in soil saturated hydraulic conductivity under a tropical rainforest using the wavelet transform. *European Journal of Soil Science*, **62**, 891–901.

Hassler, S.K., Zimmermann, B., van Breugel, M., Hall, J.S. & Elsenbeer, H. 2011b. Recovery of saturated hydraulic conductivity under secondary succession on former pasture in the humid tropics. *Forest Ecology & Management*, **261**, 1634–1642.

Huth, T., Porder, S., Chaves, J. & Whiteside, J.H. 2012. Soil carbon and nutrient changes associated with deforestation for pasture in southern Costa Rica. *Biotropica*, **44**, 661–667.

Lark, R.M. 2009. Estimating the regional mean status and change of soil properties: two distinct objectives for soil survey. *European Journal of Soil Science*, **60**, 748–756.

Mallants, D., Mohanty, B.P., Vervoort, A. & Feyen, J. 1997. Spatial analysis of saturated hydraulic conductivity in a soil with macropores. *Soil Technology*, **10**, 115–131.

Martinez, L.J. & Zinck, J.A. 2004. Temporal variation of soil compaction and deterioration of soil quality in pasture areas of Colombian Amazonia. *Soil & Tillage Research*, **75**, 3–17.

Nyberg, G., Bargués Tobella, A., Kinyangi, J. & Ilstedt, U. 2012. Soil property changes over a 120-yr chronosequence from forest to agriculture in western Kenya. *Hydrology & Earth System Sciences*, **16**, 2085–2094.

Papritz, A. & Webster, R. 1995. Estimating temporal change in soil monitoring: II. Sampling from simulated fields. *European Journal of Soil Science*, **46**, 13–27.

Pawlowsky-Glahn, V. & Olea, R.A. 2004. *Geostatistical Analysis of Compositional Data*. Oxford University Press, New York, USA.

Peng, S.L., Wu, J. & You, W.H. 2012. Recovery of saturated hydraulic conductivity along a forest successional series from abandoned land to mature, evergreen broad-leaved forest in eastern China. *Soil Research*, **50**, 257–266.

Reynolds, W.D., Elrick, D.E., Youngs, E.G., Amoozegar, A., Booltink, H.W.G. & Bouma, J. 2002. Saturated and Field-Saturated Water Flow Parameters. In: *Methods of Soil Analysis, Part 4* (eds J.H. Dane, J.H. & G.C. Topp), pp. 797–878. Soil Science Society of America, Inc., Madison, USA.

Sobieraj, J.A., Elsenbeer, H. & Cameron, G. 2004. Scale dependency in spatial patterns of saturated hydraulic conductivity. *Catena*, **55**, 49–77.

R Development Core Team. 2009. R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Tye, A.M., Robinson, D.A. & Lark, R.M. 2013. Gradual and anthropogenic soil-change for fertility and carbon on marginal sandy soils. *Geoderma*, **207-208**, 35–48.

Venables, W.N. & Ripley, B.D. 2002. *Modern Applied Statistics with S*. Springer, New York, USA.

Walvoort, D.J.J., Brus, D.J. & de Gruijter, J. J. 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. *Computers & Geosciences*, **36**, 1261–1267.

Zimmermann, B. & Elsenbeer, H. 2008. Spatial and temporal variability of soil saturated hydraulic conductivity in gradients of disturbance. *Journal of Hydrology*, **361**, 78–95.

Zimmermann, B., Papritz, A. & Elsenbeer, H. 2010a. Asymmetric response to disturbance and recovery: Changes of soil permeability under forest-pasture-forest transitions. *Geoderma*, **159**, 209–215.

Zimmermann, B., Zimmermann, A., Lark, R.M. & Elsenbeer, H. 2010b. Sampling procedures for throughfall monitoring: A simulation study. *Water Resources Research*, **46**.

FIGURE CAPTIONS

Figure 1 Schematic representation of the three sampling designs that we compare in this study. Abbreviations are: SRS for simple random sampling (A), StRS for stratified simple random Sampling (B), rotStRS for rotational stratified simple random sampling (C), SP for sampling points, Y1 and Y2 for Year 1 and Year 2.

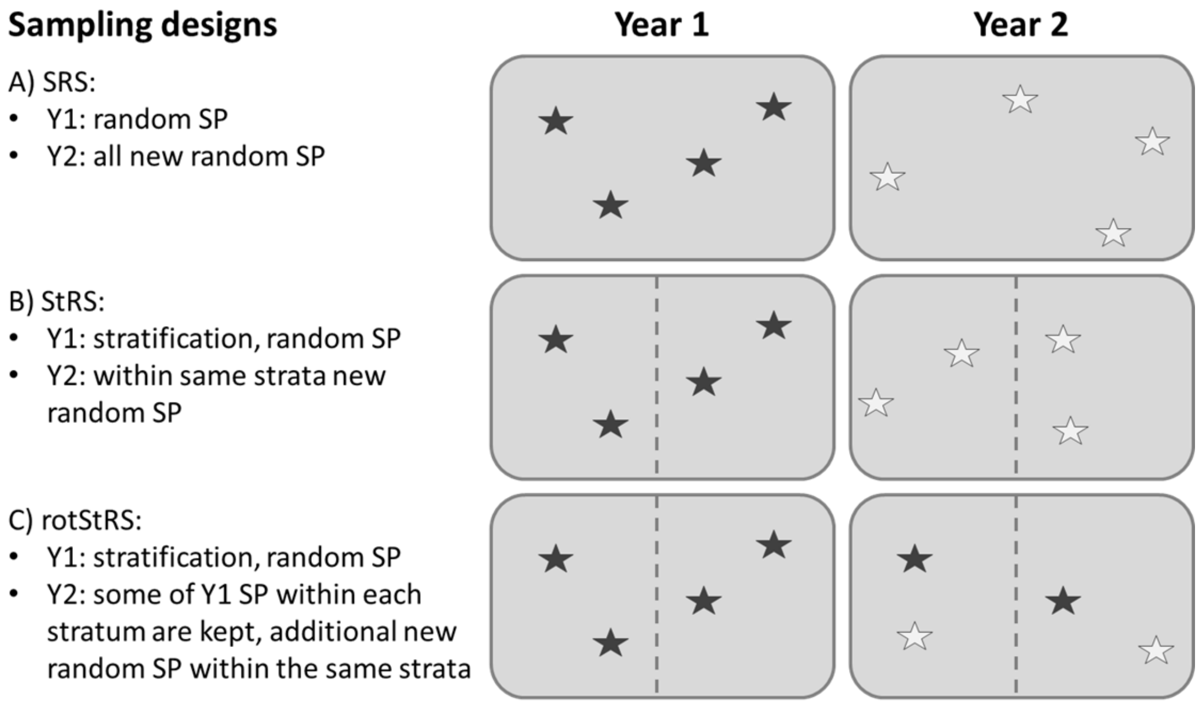
Figure 2 (A) Location of the study in Central Panama, (B) Map of the sampling design in the native-species catchment in 2010. Shown are the sampling points of 2010, the matched sample that was sampled in 2009 and re-sampled in 2010 and the stratification.

Figure 3 Means and confidence intervals of K_s calculated according to rotStRS (rotational stratified simple random sampling). Shown are the comparisons between the three years 2009, 2010 and 2011, the three study sites, covered by native species, teak and secondary succession, and the two depths 0–6 cm and 6–12 cm. The dashed lines within the plots show the baseline data before reforestation, sampled according to a StRS, however, the strata were different from those for the rotStRS monitoring design.

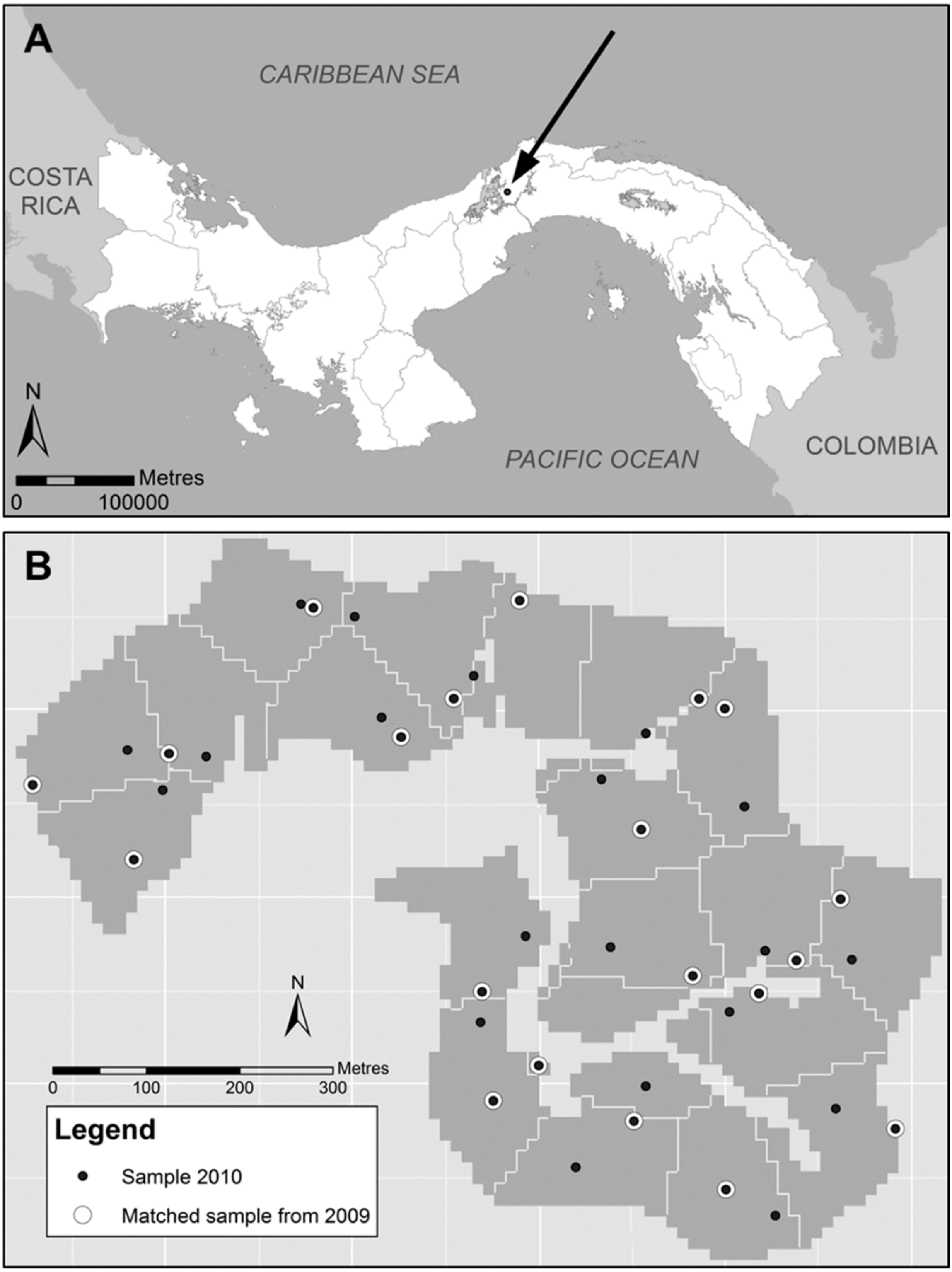
Figure 4 Means and confidence intervals of K_s for the year 2009, the three catchments and both depths, calculated according to StRS (stratified simple random sampling), SRS (simple random sampling) and rotStRS (rotational stratified simple random sampling).

Figure 5 Exemplary scatterplots for the matched samples of 2009 and 2010, for the three catchments and two depths. The transformed data are shown.

Figure 1

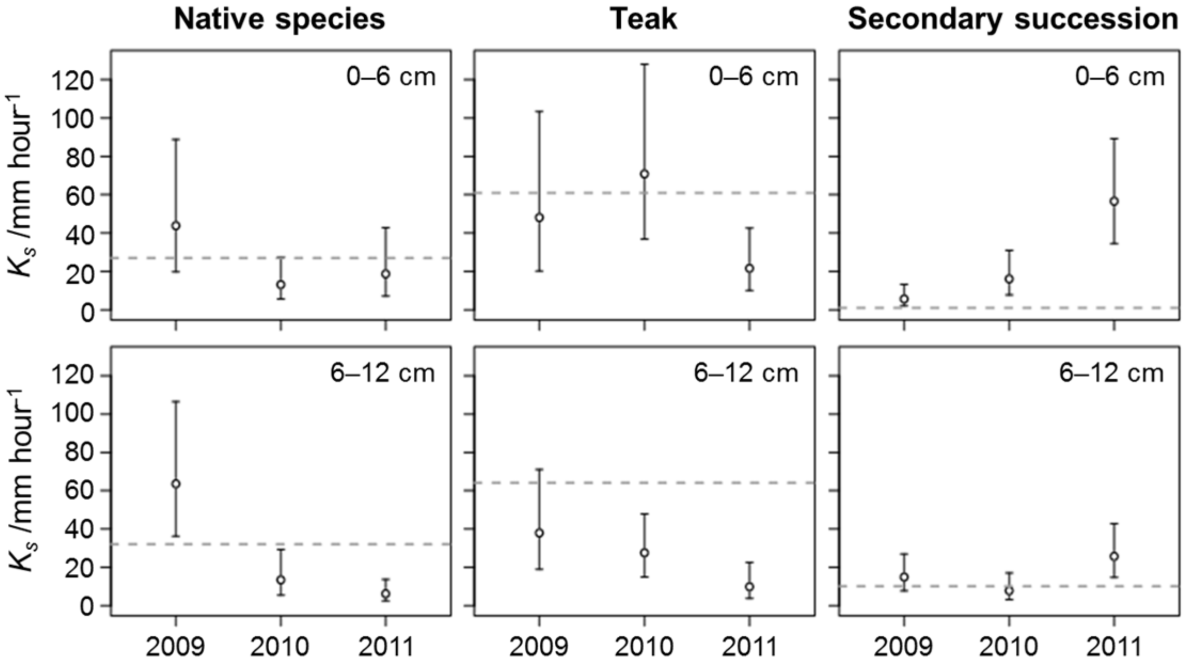

Sampling designs
Year 1
Year 2
A) SRS:
• Y1: random SP
• Y2: all new random SP
B) StRS:
• Y1: stratification, random SP
• Y2: within same strata new random SP
C) rotStRS:
• Y1: stratification, random SP
• Y2: some of Y1 SP within each stratum are kept, additional new random SP within the same strata

609 **Figure 2**



610

611 **Figure 3**



612

Figure 4

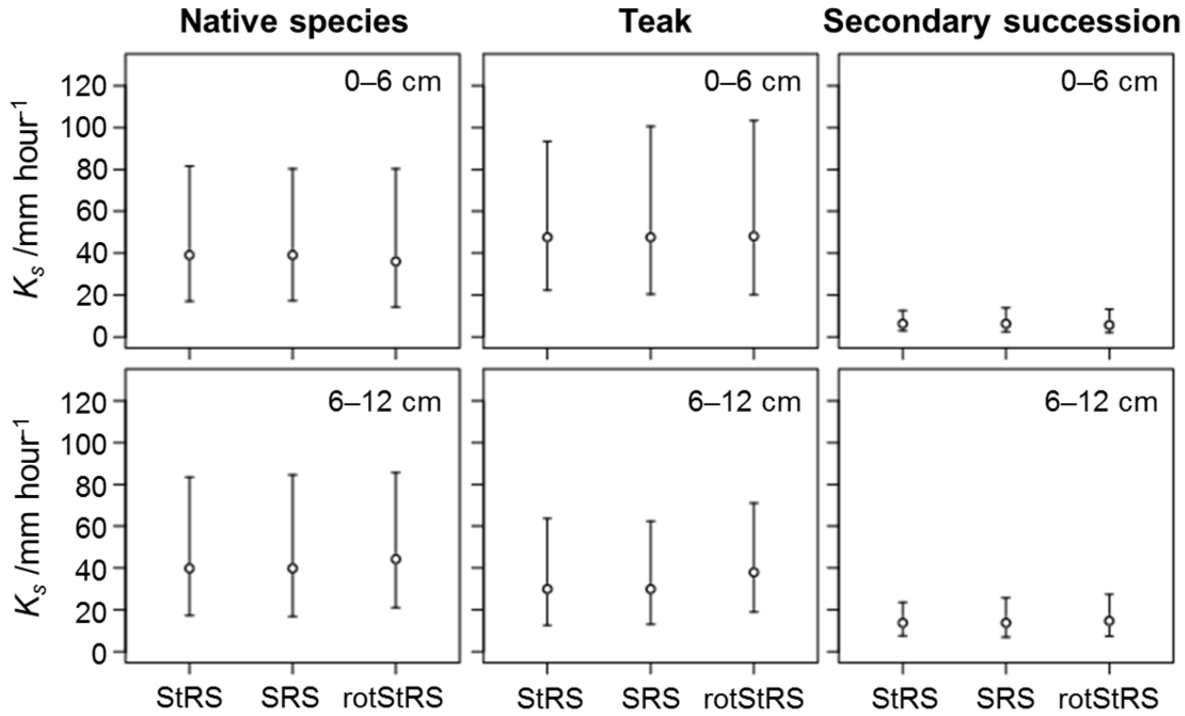
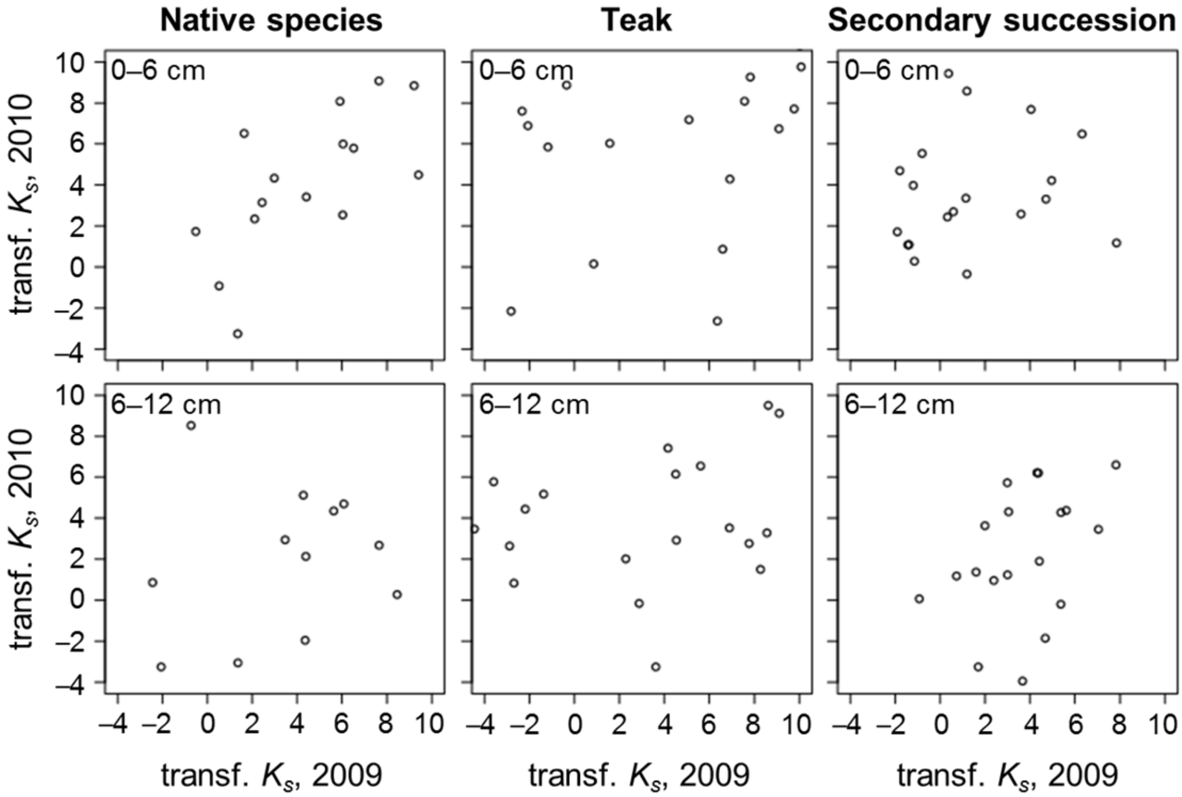

Native species
Teak
Secondary succession
0–6 cm
6–12 cm
K_S /mm hour^{-1}
0
20
40
60
80
100
120
StRS
SRS
rotStRS

Figure 5



Native species
Teak
Secondary succession
0–6 cm
6–12 cm
transf. K_s, 2010
transf. K_s, 2009

TABLES

Table 1 Spatial variance and pooled within-stratum variance for the three different catchments and both depths. Abbreviations are V_{sp} for the spatial variance, V_{pool} for the pooled within-stratum variance

Catchment	Depth /cm	V_{sp} /(mm hour^{-1})^2	V_{pool} /(mm hour^{-1})^2
Native species	0–6	13.9	14.4
Teak	0–6	21.3	18
Secondary succession	0–6	11.4	9
Native species	6–12	13.8	9.9
Teak	6–12	17.7	18.7
Secondary succession	6–12	7.8	6.5