Rothamsted Repository Download

A - Papers appearing in refereed journals

Zimmermann, B., Zimmermann, A., Lark, R. M. and Elsenbeer, H. 2010. Sampling procedures for throughfall monitoring: a simulation study. *Water Resources Research.* 46, p. W01503.

Click Here for Full Article

Sampling procedures for throughfall monitoring: A simulation study

Beate Zimmermann,[1] Alexander Zimmermann,[2] Richard Murray Lark,[3] and Helmut Elsenbeer[1,2]

Received 26 January 2009; revised 9 June 2009; accepted 3 September 2009; published 6 January 2010.

[1] What is the most appropriate sampling scheme to estimate event-based average throughfall? A satisfactory answer to this seemingly simple question has yet to be found, a failure which we attribute to previous efforts' dependence on empirical studies. Here we try to answer this question by simulating stochastic throughfall fields based on parameters for statistical models of large monitoring data sets. We subsequently sampled these fields with different sampling designs and variable sample supports. We evaluated the performance of a particular sampling scheme with respect to the uncertainty of possible estimated means of throughfall volumes. Even for a relative error limit of 20%, an impractically large number of small, funnel-type collectors would be required to estimate mean throughfall, particularly for small events. While stratification of the target area is not superior to simple random sampling, cluster random sampling involves the risk of being less efficient. A larger sample support, e.g., the use of trough-type collectors, considerably reduces the necessary sample sizes and eliminates the sensitivity of the mean to outliers. Since the gain in time associated with the manual handling of troughs versus funnels depends on the local precipitation regime, the employment of automatically recording clusters of long troughs emerges as the most promising sampling scheme. Even so, a relative error of less than 5% appears out of reach for throughfall under heterogeneous canopies. We therefore suspect a considerable uncertainty of input parameters for interception models derived from measured throughfall, in particular, for those requiring data of small throughfall events.

Citation: Zimmermann, B., A. Zimmermann, R. M. Lark, and H. Elsenbeer (2010), Sampling procedures for throughfall monitoring: A simulation study, *Water Resour. Res.*, *46*, W01503, doi:10.1029/2009WR007776.

1. Introduction

[2] The redistribution of rainfall in forest canopies results in throughfall patterns which show substantial spatial variability [*Levia and Frost*, 2006], though large differences exist among forest ecosystems [*Lloyd and Marques*, 1988; *Holwerda et al.*, 2006]. The search for sampling schemes that provide reliable estimates of mean throughfall, given this variability, has attracted some interest in recent decades [*Helvey and Patric*, 1965; *Kimmins*, 1973; *Lloyd and Marques*, 1988; *Kostelnik et al.*, 1989; *Puckett*, 1991; *Thimonier*, 1998; *Rodrigo and Àvila*, 2001, *Holwerda et al.*, 2006]. In this context, particular attention was given to sample support and, to a lesser extent, sampling design. Following *de Gruijter et al.* [2006], we define the sample support as the shape, size and orientation of the sampling units. A sampling design assigns a probability of selection to a set or sequence of sampling units in the sampling universe. The combination of a sampling design and an estimator for a target parameter, such as mean throughfall of an area, is called a sampling strategy [*de Gruijter et al.*, 2006]. As suggested by the same authors, we refer to the entire plan, which combines all the decisions and information pertinent to the acquisition, recording and processing of data, as the sampling scheme.

[3] Robust sampling schemes have been identified for some problems, e.g., the estimation of interception loss over extended periods from temporally accumulated samples. For instance, the random relocation of small, funnel-type collectors during the sampling period, subject to some assumptions, improves the efficiency of sampling estimates by increasing the spatial coverage [*Lloyd and Marques*, 1988; *Holwerda et al.*, 2006]. However, for the common problem of estimating throughfall for particular events, there is no consensus on the best strategy, but there is recognition that estimates of this variable are subject to considerable sampling error, and that these errors will propagate when the estimates are used in subsequent modeling. For this reason *Hutjes et al.* [1990] called for improved accuracy and precision of the estimates of throughfall.

[4] Various practical questions about how best to sample throughfall for individual events remain unresolved. Regarding the sample support, several empirical studies addressed the potential benefits of trough- versus funnel-type collectors [e.g., *Kostelnik et al.*, 1989], but no general

[1]Smithsonian Tropical Research Institute, Balboa, Panama.
[2]Institute of Geoecology, University of Potsdam, Potsdam, Germany.
[3]Rothamsted Research, Harpenden, UK.

0043-1397/10/2009WR007776$09.00

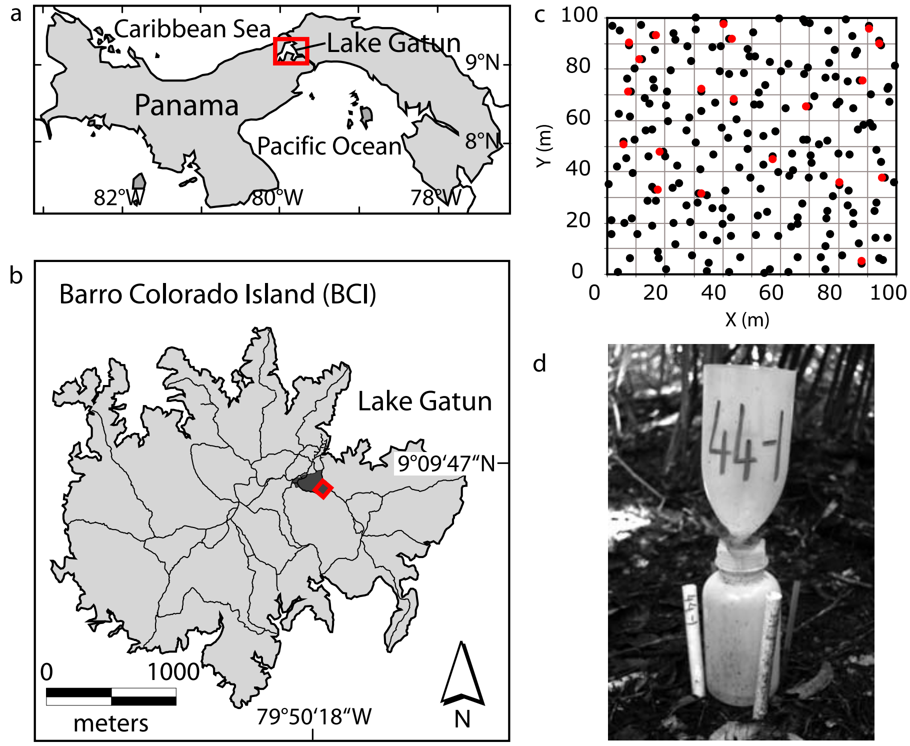

Figure 1. Location of the research area (a) in Panama and (b) on Barro Colorado Island. In Figure 1b the dark gray area indicates the position of Lutz Creek catchment, the lines refer to the trail system, and the red square indicates the location of the sampling area. (c) The close-up view of the 1 ha sampling area shows the location of sampling points (dots). Different colors of dots indicate collector placement according to different sampling strategies, that is, black dots refer to locations that were chosen according to a design-based sampling strategy, whereas red dots mark locations selected by a model-based strategy. (d) The photo shows one of the funnel-type samplers.

recommendations as to the best type of intercepting device have emerged so far [*Thimonier*, 1998]. Although we might, a priori, anticipate advantages for a trough collector that integrates short-range variation in throughfall, which contributes to the spatial variation of funnel-based measurements. Furthermore, most studies used a simple random sampling design [*Bruijnzeel and Wiersum*, 1987; *Lloyd and Marques*, 1988; *Puckett*, 1991; *Gash et al.*, 1995; *Lin et al.*, 1997; *Schellekens et al.*, 1999; *Carlyle-Moses et al.*, 2004; *Keim et al.*, 2005; *Holwerda et al.*, 2006; *A. Zimmermann et al.*, 2007, 2008], and the potential benefits and drawbacks of alternative designs such as stratified random sampling or cluster sampling have not been assessed.

[5] The comparative assessment of different sample supports or sampling designs, or both, is potentially costly and labor intensive. This is because the comparison of particular supports under particular designs requires a set of field observations which allows us to estimate standard errors of sample means under the alternative scheme. Previous studies of which we are aware have therefore been limited in the generality of conclusions they can support. Here we propose an alternative approach. We still use empirical observations, but we use these to derive a spatial statistical model of throughfall which can then be used to simulate observations with different support sizes, and under different sampling designs. This approach was already used in a study on soil monitoring [*Papritz and Webster*, 1995], but we are not aware of its application to the current problem.

[6] In this paper, we first describe the spatial statistical model that we use, and we explain how it is fitted to a set of throughfall data during several separate events in a study plot. We then address the following research questions by using the original data sets (question 1) and the fitted models (questions 2 and 3): (1) How many small, funnel-type collectors are required to estimate event-based mean throughfall at prescribed error limits at our tropical rain forest site? (2) How much can be gained by increasing the sample support, i.e., by using trough-type collectors; which factors influence the potential benefits? (3) Are there gains or losses associated with the use of stratified or a cluster random sampling, respectively, versus simple random sampling? Finally, we interpret our results regarding their implications for interception modeling.

2. Methods

2.1. Site Description

[7] We investigated throughfall patterns in a square 1 ha plot located in the Lutz Creek Catchment (9° 9′N, 79° 51′W) on Barro Colorado Island, Panama (Figure 1). The island was isolated from the mainland during the creation of Lake Gatun, which is part of the Panama Canal. Its climate is characterized by distinct wet and dry seasons. The wet season lasts approximately from May to mid-December. Total annual rainfall averages 2651 ± 441 mm (mean ± 1sd, n = 83, data from 1925 to 2007, Smithsonian Tropical

Research Institute, Environmental Science Program). The vegetation is classified as tropical semideciduous moist forest [*Foster and Brokaw*, 1982]. Ten percent of the canopy tree species are dry season deciduous [*Croat*, 1978]. The forest in the study area is secondary growth of more than 100 years of age with an unevenly distributed understory. Stand height is 25–35 m with few emergents approaching 45 m. For a detailed description of the stand characteristics we refer to *A. Zimmermann et al.* [2009].

2.2. Sampling and Instrumentation

[8] Our throughfall sampling scheme comprises a design-based and a model-based sampling component. Regarding the former component, the sampling units are selected by probability sampling and inference is based on the sampling design; for the latter part, statistical inference is based on the model. The design-based component consists of a stratified simple random sampling with compact geographical stratification [*de Gruijter et al.*, 2006]. We divided our 1 ha plot into 100 square subplots of length 10 m. In each of these subplots we randomly selected two throughfall sampling points (Figure 1). We then chose 20 of these sampling points at random, and selected an additional sampling point 1 m away in a random direction (model-based sampling component). These additional 20 points increased the number of sampling locations separated by short lag distances, which is important for estimating the shape of the variogram model near the origin [*B. Zimmermann et al.*, 2008].

[9] Throughfall samples were collected on an event basis during two campaigns. The first lasted from August to November 2007 (n = 42 events). The second covered the dry and the subsequent transition season in 2008 and lasted from February to June (n = 14 events). To qualify for an event, at least 0.6 mm of rainfall had to be recorded, which had to be preceded and followed by at least 2 hours without rainfall. Our observations indicate that throughfall completely ceased within these 2 hours, and we usually started sampling 2.5 h after the end of rainfall. Since throughfall sampling took at least 3 h, all events of our data set are separated by more than 5 h. Rainfall was continuously recorded with two tipping bucket rain gauges in a clearing 300 m from the throughfall sampling site. This distance between rainfall and throughfall measuring sites might be problematic for the calculation of interception loss due to the small-scale variability of convective rainfall. Nevertheless, this problem is irrelevant in the context of the present study.

[10] The throughfall collectors (n = 220) consisted of a 2 L polyethylene sampling bottle and a funnel (Figure 1). The receiving area of each collector was 113 cm^2; so the total sampling area in the 1 ha plot was 2.49 m^2. A polyethylene net with a 0.5 mm mesh at the bottom of the funnel minimized measurement errors due to organic material and insects.

2.3. Sample Size Requirements for Funnel-Type Collectors

[11] In order to compare sample size requirements for our sample support size with published values from the literature, we computed the required sample sizes using all original throughfall events (n = 56) with the widely used formula for a confidence interval at level $1 - \alpha = 0.95$ (standard formula from here on) such that the relative error (defined below) is smaller than a prespecified limit r

$$n = \left(\frac{u_{1-\alpha/2} \cdot \hat{S}(z)}{r \cdot \hat{\bar{z}}}\right)^2, \tag{1}$$

where $u_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution, and $\frac{\hat{S}(z)}{\hat{\bar{z}}}$ is the estimated coefficient of variation. The relative error is defined as

$$\left|\frac{\left(\hat{\bar{z}} - \bar{z}\right)}{\bar{z}}\right|, \tag{2}$$

where $\hat{\bar{z}}$ is the estimated sample mean and $\bar{z}$ is the true mean, i.e., the population mean of a deterministic variable in a system, which is not available in practice. The relative error has been used frequently in throughfall studies along with the standard requirement that it should not exceed 5% or 10% [*Kimmins*, 1973; *Rodrigo and Ávila*, 2001; *Holwerda et al.*, 2006]. Due to sampling error no strategy can guarantee this, but we can require that, for example, the estimated mean achieves this standard with a given probability.

2.4. Model of Spatial Data and Its Estimation

2.4.1. Model

[12] Our objective is to analyze the data described in section 2.2 so as to generate a model which is appropriate for generating observations from different sampling designs and support sizes. Such a model must account for the spatial dependence among such measurements (nearby measurements are more likely to be similar than measurements further apart), and for the fact that throughfall data may include extreme (upper tail) values from a few isolated spots where throughfall is concentrated. These outliers occur due to sampling locations beneath drip points of leaves or inclined stems [*A. Zimmermann et al.*, 2009]. For the purpose of this study, we propose that n measurements of throughfall correspond to a random variate $\boldsymbol{Z}$ which has the distribution:

$$\begin{aligned} Z &\sim N\{\boldsymbol{\mu}, \boldsymbol{\Gamma}\} \text{ with probability } 1 - p_c \\ &\sim D \text{ with probability } p_c, \end{aligned} \tag{3}$$

where $N\{\boldsymbol{\mu}, \boldsymbol{\Gamma}\}$ denotes a normal random variate with a stationary mean value in the vector $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Gamma}$ which we assume can be described parametrically by some bounded variogram model $\gamma(\mathbf{h})$, such that the element of $\boldsymbol{\Gamma}$ that corresponds to observations at any two sites $\mathbf{x}_i$ and $\mathbf{x}_j$ (i≠j) is given by $\gamma(\infty) - \gamma(\mathbf{x}_i - \mathbf{x}_j)$. The distribution D denotes some random contaminating process. The probability of a contaminating event at any site is p_c, and we assume that the occurrence of a contamination at any two locations is an independent process (an assumption we test in practice). This separation of the throughfall frequency distribution into a "normal" and a "contaminating" process is arbitrary and requires a statistical criterion (equation (7)). It is, however, unavoidable because not

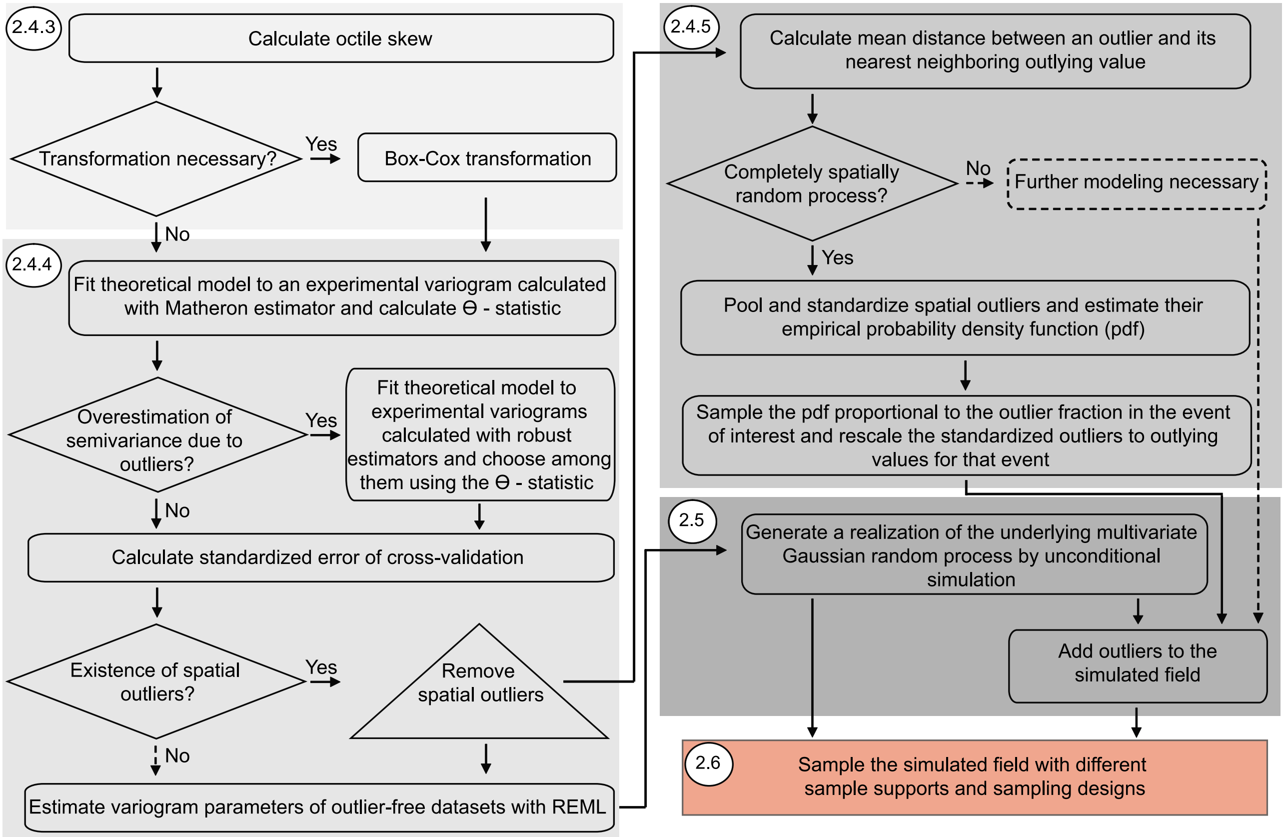

Figure 2. Summary of all steps involved in modeling the spatial throughfall data so as to simulate realistic random fields of throughfall magnitude, which were subsequently sampled with different sample supports and selected sampling designs. The encircled numbers correspond to the sections of this paper in which the particular steps are described in detail. Dashed arrows and dashed boxes highlight decisions or operations, which never applied to or which were not necessary for the modeling of the presented data (e.g., all selected data sets contained spatial outliers).

only the Gaussian assumption has to be satisfied before variogram modeling and geostatistical simulations, but also the final model needs to incorporate outlying values which cannot be forced to the center of the distribution even after transformation [*A. Zimmermann et al.*, 2009]. In principal we should like to model spatiotemporal variation of throughfall for our purposes. However, the spatial statistics of our throughfall measurements for different events were markedly different (Table 2); indicating that joint space-time modeling assuming coregionalization [*Papritz and Webster*, 1995] or some other simple space-time covariance structure [e.g., *Jost et al.*, 2005] would not be feasible. Moreover, drip points are not active during all events [*A. Zimmermann et al.*, 2009]. For this reason we modeled the different events and simulated sampling from them independently.

2.4.2. Fitting the Model

[13] We obtained parameters for our model from throughfall measurements for 14 events which were selected from the overall data set of 56 events. The selected events cover the sampled range of mean throughfall volumes (0.6–79.4 mm). Another selection criterion required that the throughfall measurements of a particular event display, after the removal of outliers, a bivariate normal distribution. This is important because the estimation of the covariance parameters with residual maximum likelihood, the robust variogram estimators and Gaussian simulations assume a multivariate Gaussian distribution, which cannot be verified [*Pardo-Igúzquiza*, 1998]. *Chilès and Delfiner* [1999] therefore suggested, in the framework of Gaussian simulation, that at least the bivariate distributions of the (transformed) data could be checked (see section 2.4.3). In sections 2.4.3 to 2.5 we describe the procedure to obtain simulated throughfall fields which are the cornerstone for our simulation study. The individual steps of this procedure are summarized in Figure 2.

2.4.3. Exploratory Data Analysis

[14] To ensure that our estimates are efficient [*Webster and Oliver*, 2007] and to justify the assumption of an underlying multivariate normal process, we must transform our data, where necessary, so that the histogram appears consistent with a normal distribution. In addition to inspecting the histogram, the skewness of the data is a useful indicator of the plausibility of the assumption of normality [*Webster and Oliver*, 2007]. However, we want to transform our data only if this is necessary to justify the assumption that the underlying spatially correlated process is normal, and the standard coefficient of skewness is susceptible to the effects of outliers. We therefore computed a robust alternative, the octile skew [*Brys et al.*, 2003], skew_8. This is a measure of the

asymmetry of the first (O_1) and seventh octile (O_7) of the data about the median

$$\mathrm{skew}_8 = \frac{(O_7 - \mathrm{median}) - (\mathrm{median} - O_1)}{(O_7 - O_1)}, \quad (4)$$

and hence, is insensitive to extreme values. We use the octile skew as a diagnostic, because a variable with the distribution given by (3) is likely to be skewed if the means of the underlying normal and contaminating processes are different. *Rawlins et al.* [2005] recommended as a rule of thumb that data need to be transformed if the octile skew is larger than 0.2 or smaller than −0.2. We adopted this approach and used Box-Cox transformations [*Box and Cox*, 1964] to find the most appropriate transformation to achieve Gaussian behavior if the octile skew was outside this recommended interval.

[15] In the context of techniques that require a nonverifiable multivariate Gaussian distribution, it was recommended [*Chilès and Delfiner*, 1999; *B. Zimmermann et al.*, 2008] to check at least the bivariate data distribution, both to detect deviations from the Gaussian shape and to reveal spatial outliers. The latter cannot be ruled out simply by the inspection of univariate data distributions because spatial outliers might be unusual only compared to their close neighbors, but not in a histogram of the whole data set [*Lark*, 2002]. As a simple visualization of the bivariate distribution we used scatterplots, **h** scattergrams [*Webster and Oliver*, 2007], of point pairs separated by a fixed distance, which should be produced for a number of distance classes. In a previous paper [*A. Zimmermann et al.*, 2009] we showed that the occurrence of outliers in some, but not all, of those classes results in overestimation of the semivariance by nonrobust estimators.

2.4.4. Variogram Estimation

[16] First we calculated experimental variograms using the estimator due to *Matheron* [1962]

$$2\hat{\gamma}_M(\mathbf{h}) = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}\{z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})\}^2, \quad (5)$$

where $z(\mathbf{x}_i)$ is the observed value at location $\mathbf{x}_i$, and $N(\mathbf{h})$ are the pairs of observations that are separated by lag **h.** We fitted four theoretical models (exponential, Gaussian, spherical, pure nugget) to the experimental variogram, and chose the one with the smallest sum of squared residuals from the fit. We assessed the fitted model by cross-validation, in which each observation, $z(\mathbf{x})$, is left out in turn and predicted by ordinary kriging from the rest to yield a prediction, $\hat{Z}(\mathbf{x})$ with a kriging variance $\sigma^2_{\mathbf{K},\mathbf{x}}$. In order to assess whether the data contain outliers that cause an overestimation of the semivariance by this nonrobust estimator, we calculated the statistic $\theta(\mathbf{x})$ [*Lark*, 2000]

$$\theta(\mathbf{x}) = \frac{\{z(\mathbf{x}) - \hat{Z}(\mathbf{x})\}^2}{\sigma^2_{\mathbf{K},\mathbf{x}}}. \quad (6)$$

[17] If kriging errors follow a Gaussian distribution, $\theta(\mathbf{x})$ will be distributed as χ^2 with one degree of freedom. Because the kriging variance, which is the denominator of equation (6), estimates the squared kriging error, the mean of $\theta(\mathbf{x})$ should be 1 if the kriging estimates and their variances have been determined with an appropriate variogram and if the assumption of intrinsic stationarity is plausible. However, the mean of $\theta(\mathbf{x})$ is itself subject to the effects of outlying observations, and so the median is used as a robust diagnostic [*Lark*, 2000]. Since the median of the standard χ^2 distribution with one degree of freedom is 0.455, this is the expected value of the median of $\theta(\mathbf{x})$. A sample median significantly less than 0.455 suggests that kriging overestimates the variance, whereas one which is greater than 0.455 underestimates the variance. In order to compute confidence limits for the median of $\theta(\mathbf{x})$, accounting for the spatial dependence among the kriging errors, we performed 1000 unconditional simulations to predict the values of throughfall magnitude, computed the median of $\theta(\mathbf{x})$, and determined the 2.5% and 97.5% percentiles of the distribution of the median of $\theta(\mathbf{x})$ to approximate 95% confidence limits. If the median of $\theta(\mathbf{x})$ was outside those limits for the variograms based on the Matheron estimator, we calculated robust experimental variograms proposed by *Cressie and Hawkins* [1980], *Dowd* [1984], and *Genton* [1998], and chose among them following *Lark* [2000].

[18] We then used the standardized error $\varepsilon_s(\mathbf{x})$ of cross-validation with the selected variogram [*Bárdossy and Kundzewicz*, 1990] for the identification of outliers at a particular location $\mathbf{x}_o$, which is defined as follows:

$$\varepsilon_s(\mathbf{x}) = \frac{\hat{Z}(\mathbf{x}_o) - z(\mathbf{x}_o)}{\sigma_{\mathbf{K},o}}. \quad (7)$$

[19] If a property at a particular location $\mathbf{x}_o$ is a realization of a quasi point process (i.e., the contaminating process of our model) rather than the background process (i.e., the Gaussian random process of our model) then the variance of the deviation $\{\hat{Z}(\mathbf{x}_o) - z(\mathbf{x}_o)\}$ is likely to be underestimated by the kriging variance $\sigma^2_{\mathbf{K},o}$, which increases the value of $\varepsilon_s(\mathbf{x})$. *Bárdossy and Kundzewicz* [1990] proposed the standardized error to be used to classify a value as an outlier if $\varepsilon_s(\mathbf{x})$ is smaller than −1.96 for a 95% confidence level; in the present study, we set the confidence level to 99%. The standardized error of cross-validation for the identification of spatial outliers was also used by *Meklit et al.* [2008]. Once identified, we removed the outlying values and subsequently estimated the covariance parameters of the uncontaminated data sets by residual maximum likelihood, which is preferred to methods of moments estimation of the variogram as the fitted model is independent of arbitrary decisions we may make about, for example, lag bins [*Lark et al.*, 2006; *B. Zimmermann et al.*, 2008].

2.4.5. Modeling the Outliers

[20] The procedures described so far provide us with a variogram function for the underlying multivariate Gaussian random process that we assume to be realized in those of our throughfall data that do not represent outlying values. We also require a model for these outliers if we are to simulate realistic throughfall data sets. The first issue is whether it is justifiable to assume that the outliers are distributed independently and at random among our sampling points, i.e., whether they represent a completely spatially random (csr) process [*Cressie*, 1993]. We used the mean distance between an outlying value and its nearest neighboring outlying value on the sample grid, $\overline{W}$, as a diagnostic statistic to test the

null hypothesis that the outlying data are a realization of a completely spatially random process. For any given data set, a fraction (n_c) of our observations were identified as outliers. We then calculated $\overline{W}$ for this set of outliers, and used a Monte Carlo analysis to calculate a sample distribution for $\overline{W}$ under the null hypothesis that outliers are csr. In each of 100,000 Monte Carlo runs n_c of the original sample points were allocated to the outlying process independently and at random (which corresponds to our null hypothesis), and we calculated the value of $\overline{W}$ for each resulting pattern. The resulting 100,000 values of $\overline{W}$ provide us with an empirical sample distribution for this statistic under the null hypothesis, conditional on our sampling scheme. The pth and the $(1 - p)$th percentile of this distribution yielded critical thresholds to test the null hypothesis with a P value of $2p$. If our observed value of $\overline{W}$ was less than the lower threshold, then the null hypothesis of a csr distribution was rejected, and spatial clustering of the outliers is indicated. If our observed value of $\overline{W}$ was larger than the upper threshold, then the null hypothesis of csr distribution was also rejected, and we inferred an overdispersion of the process.

[21] The second issue is to identify a statistical distribution of the contaminating process, in our case the distribution of the outlying throughfall data. In order to achieve a sufficient sample size, we pooled the spatial outliers from all events, which were standardized by first subtracting their mean value, then dividing by their standard deviation for any one event. The histogram of these pooled outliers, which represents the contaminating process and from which we wanted to sample possible outlying values, was positively skewed. Rather than making any assumption about the distribution of the spatial outliers, we estimated an empirical probability density function (pdf) by kernel density estimation [*Silverman*, 1978, 1986]. *Silverman* [1986] demonstrated that the choice of the kernel is not as important as the choice of an appropriate smoothing parameter, i.e., the width of the kernel. We selected the Epanechnikov kernel and used the test graph method of *Silverman* [1978] to select the smoothing parameter.

2.5. Simulation of Throughfall Fields

[22] To assess alternative sampling schemes, we generated fields of throughfall values from the spatial statistical model introduced in section 2.4 for different recorded events. Our basic field is a square grid of interval 10 cm, and 1000 rows and columns (i.e., it corresponds to a 1 ha plot). We assume that the basic grid cell of area 100 cm^2 corresponds to the support of our original measurements, a reasonable approximation to funnels with an area of 113 cm^2. Realizations of the random process $\boldsymbol{Z} \sim N\{\boldsymbol{\mu}, \boldsymbol{\Gamma}\}$ were generated by unconditional simulation of a Gaussian random variable using the procedure *GaussRF*, which is implemented in the package *Random Fields* in R [*Schlather*, 2001]. Here the mean vector contained a stationary mean (the sample average throughfall of the data on the event being simulated after removal of the outliers), and the covariance matrix was computed from the residual maximum likelihood (REML) estimate of the variance parameters for the same data.

[23] In order to quantify the effects of outliers on the performance of the sampling strategies to be tested, we sampled both from the uncontaminated throughfall field (described above) and from a simulated field to which we added outliers. When simulating outliers from a particular event we set the number of outliers, n_c, to the number identified for that event. In all cases our null hypothesis that the outliers are csr was retained, so these outliers were substituted at independently and randomly selected nodes of the grid. The value of throughfall for the outlying observation was sampled from the smoothed pdf for outliers using the procedures *density* and *rnorm* in R [*Becker et al.*, 1988]. This standardized outlier value was then rescaled to an outlying value for the event of interest by multiplying it by the outlier standard deviation for that event, and by adding the corresponding outlier average value. For events where the original throughfall data had been transformed the simulated values were then back transformed to the original scale.

[24] The mean of a particular realization of the random process (both for contaminated and uncontaminated fields) was computed from all 10^6 nodes. This is our notional population mean that we wish to estimate by sampling. A single value from the simulated grid could be treated as a value obtained by measurement with a funnel-type collector. The simulated values for adjacent cells on the grid could be aggregated to simulate values collected with a device of corresponding support.

2.6. Sampling From the Simulated Fields

[25] We investigated two aspects of sampling for the estimation of mean throughfall in a plot of fixed size. First, we compared the performance of funnel- and trough-type collectors, that is, we quantified the benefits of an increased sample support. Second, we tested different sampling designs for funnel-type collectors (i.e., the same sample support as our original data).

2.6.1. Procedure to Increase the Sample Support

[26] To obtain simulated values that correspond to trough-type collectors, the simulated throughfall values were first back converted from standard rainfall units (mm) to volumetric, support-dependent data (mL). Next, we summed up neighboring grid cells to achieve the desired increase of support. To simulate troughs, we assumed that their width was 10 cm so that accumulated simulated sample volumes of neighboring grid cells corresponded to a selected trough length. We simulated four lengths of troughs, namely 1 m, 2 m, 4 m, and 10 m. The spatial covariance of our throughfall data showed no directional dependence; hence, the orientation of the troughs is immaterial. In this case all our simulated troughs lay on rows of the grid. Prior to sampling from these aggregated throughfall fields we converted the values to standard rainfall units for comparability with the results of the whole virtual experiment.

2.6.2. Test for Influence of Different Autocorrelation Structures on the Benefit of an Increased Sample Support

[27] Since both the strength of the spatial autocorrelations (i.e., the nugget-to-sill ratio) and the range over which throughfall measurements are correlated (i.e., the effective range) differed among the selected events (for event selection see section 2.4.2), we expect the benefit of troughs versus funnels to vary correspondingly. To test this hypothesis, we proceeded as follows. First, we calculated for every event the range between the 2.5% and 97.5% percentile of the distribution of the 10000 sample means, which we

Table 1. Sample Sizes for Sampling From the Simulated Throughfall Fields

Simple Random Sampling (Funnels and Troughs)	Stratified Simple Random Sampling[a]		Cluster Random Sampling[b]	
	16[c]	25[c]	3[d]	5[d]
15	–	–	15 (5)	–
30	32 (2)	–	30 (10)	30 (6)
50	48 (3)	50 (2)	51 (17)	50 (10)
75	80 (5)	75 (3)	–	75 (15)
100	96 (6)	100 (4)	–	100 (20)
150	144 (9)	150 (6)	–	–
200	192 (12)	200 (8)	–	–

[a]Values in parentheses refer to the number of sample points per stratum.
[b]Values in parentheses refer to the number of sample points per transect.
[c]Number of strata.
[d]Number of transects. The clusters are defined as transects.

obtained from sampling the simulated fields with the different sample supports (funnel-type collectors; trough-type collectors having lengths of 1 m, 2 m, 4 m, and 10 m) and a given sample size. To rule out influences of a particular sample size on the final results, we did these calculations for a small and the largest tested sample size ($n = 30$ and $n = 200$). Next, we computed the ratio of the respective ranges of trough- and funnel-type collectors, for all tested trough lengths. When considering, e.g., event 3, which is one of the selected throughfall events, the use of 30 funnel-type collectors yields an interpercentile range of 0.85, which decreases to 0.72 for the employment of 30 1 m troughs. Hence, in this case the ratio of 1 m troughs to funnels equals 0.85. Clearly, the smaller the ratio the larger is the benefit of the trough- compared to the funnel-type collectors. We finally calculated Spearman rank correlation coefficients to correlate these ratios with the aforementioned characteristics of the spatial correlation structure, namely the effective range and the nugget-to-sill ratio.

2.6.3. Comparison of Different Sampling Designs for Funnel-Type Collectors

[28] The sampling designs, which we compared to simple random sampling, were chosen to reflect designs used in the literature, or which are of potential practical importance in the future. One is a stratified simple random design with compact geographical stratification; i.e., the plot is divided into equally sized subplots (strata). This ensures that the randomly selected sampling locations cover the target area as fully as possible. It also avoids possible clustering of sample points, which is often the case with simple random sampling, and hence can increase precision. In the case of cluster random sampling, the other sampling design which we compared to simple random sampling, predefined sets of sampling locations instead of individual locations are selected. The resulting spatial clustering of sampling locations has the operational advantage of reducing the travel time between locations in the field, but it reduces precision. Therefore, the rationale for using it is that the operational advantages allow a larger sample size for the same budget, which outweighs the initial loss of precision. In this study, we only tested the performance of cluster versus simple random sampling using identical sample sizes. An overview of the tested samples sizes for all designs, and the respective numbers of strata and clusters are given in Table 1. For further details on the designs we considered we refer to *de Gruijter et al.* [2006].

2.7. Estimation of Recording Time

[29] The throughfall collected in sampling devices is commonly measured manually, because of the costs of equipment to measure it automatically. This means that the benefits of sampling throughfall on a larger support must be weighed against the additional time that it takes for a worker to measure a larger sample. We made measurements of the time to record the expected sample volumes in a funnel, 1, 2, or 4 m trough in an event of particular intensity. We did not test 10 m troughs because they would collect a volume too large for manual handling. We undertook this experiment assuming that mean throughfall corresponded to the lower quartile, the median, and the upper quartile of the average throughfall of each of our 56 events. Fifteen people participated in the experiment, either scientists or their assistants.

3. Results

3.1. Characteristics of the Throughfall Data Sets

[30] In a previous paper [*A. Zimmermann et al.*, 2009] we showed that the frequency distribution of throughfall depends on event size. Whereas large events exhibit large skewness due to stemflow drip points, but a small octile skew, small events have both a large skewness and a large octile skew. This indicates that the underlying (uncontaminated) distribution of throughfall for small events is skewed, and this is not just the effect of outliers. For instance, the smallest of the events which we used for the simulations, event 43, has an octile skew larger than 0.2, which is why we transformed the data before variogram analysis (Table 2).

[31] In addition to this dependence of the shape of the distribution on the mean, the variance and mean are also correlated. The coefficient of variation for the events is largest for event 43, which received the smallest throughfall volume, whereas the smallest CV characterizes event 34 with the maximum mean throughfall (Table 2); this relationship is general [*A. Zimmermann et al.*, 2009]. The percentage of outliers in the original data sets varies slightly between 0.9 and 4.5% (Table 2).

[32] Regarding the spatial structure of throughfall events, we were able to discern three patterns. No spatial autocorrelation was found for events 28 and 34; strong but short-range (<2 m) correlation characterized events 3, 6, 8, 10, and 17; and events 12, 19, 25, 32, 33, 42, and 43 revealed a moderate-to-weak spatial structure but with longer effective ranges. The performance of the variogram models, which were estimated by REML for the data sets without spatial outliers, is indicated by the θ statistic. For all tested events, the median of $\theta(\mathbf{x})$ remains within the confidence limits, and the mean of $\theta(\mathbf{x})$ is very close to its expectation (Table 2).

3.2. Sample Size Requirements for Funnel-Type Collectors

[33] Here we present the results of equation (1), which was used to calculate the sample sizes needed to keep the relative error of the mean below 20%, 10%, or 5%, for the 56 original data sets. In general, due to the larger coefficient of variation of small throughfall events, the required sample

Table 2. Summary Statistics and Variogram Parameters for the Simulated Throughfall Events

	Original Data					Data Without Extreme Values (α = 0.01 for Removement Procedure)											
									Variogram Parameters for REML Models								
Event	Mean (mm)	CV[a] (%)	Skewness	Octile Skew	Outliers (%)	Mean (mm)	CV[a] (%)	Skewness	Model[b]	Nugget	Partial Sill[c]	Sill	Range (m)	Nugget/ Sill	Effective Range[d] (m)	θmed[e]	θmean[f]
43	0.6	103.9	2.9	0.43	2.3	0.5[g]	81.8[g]	1.1[g]	exp[h]	0.1[h]	0.0[h]	0.1[h]	2.5[h]	0.7[h]	7.6[h]	0.357[h]	0.999[h]
3	2.0	92.5	2.6	0.33	2.3	1.8[g]	74.9[g]	1.1[g]	sph[h]	0.0[h]	0.3[h]	0.3[h]	1.9[h]	0.0[h]	1.9[h]	0.464[h]	1.007[h]
17	5.5	71.4	4.6	0.14	4.5	4.9	45.3	0.4	exp	0.0	4.9	4.9	0.4	0.0	1.1	0.448	1.000
32	5.8	68.6	5.5	0.13	0.9	5.5	48.9	0.4	exp	4.8	2.5	7.3	1.8	0.7	5.5	0.516	0.998
28	7.0	53.5	1.4	0.13	2.3	6.6	45.9	0.4	nug	9.2	0.0	9.2	0.0	1.0	0.0	0.530	1.000
10	8.8	60.0	2.6	0.13	3.2	8.2	45.4	0.4	sph	0.0	13.7	13.7	1.6	0.0	1.6	0.446	0.998
6	13.0	70.0	3.7	0.13	4.1	11.8	49.9	0.4	sph	0.0	34.2	34.2	1.7	0.0	1.7	0.509	0.996
42	19.4	69.7	6.7	0.13	2.7	18.1	42.7	0.3	exp	32.5	27.2	59.7	3.9	0.5	11.6	0.481	0.992
33	19.8	74.9	5.5	0.09	2.7	18.2	49.1	0.3	sph	53.4	25.9	79.3	7.5	0.7	7.5	0.436	0.993
12	23.5	53.2	2.9	0.09	2.7	22.2	39.9	0.2	exp	51.7	27.3	79.0	4.3	0.7	13.0	0.513	0.993
25	30.2	68.2	5.6	0.02	2.7	27.9	43.5	0.2	exp	125.7	22.5	148.1	6.1	0.8	18.2	0.399	0.999
8	36.4	51.1	1.1	0.03	2.3	34.9	45.9	0.4	sph	0.0	256.2	256.2	1.1	0.0	1.1	0.409	0.998
19	43.4	50.7	3.2	0.06	3.2	41.1	38.1	−0.1	exp	228.5	17.4	245.9	3.8	0.9	11.5	0.403	1.000
34	79.4	43.8	0.4	0.00	1.4	77.9	41.8	0.1	nug	1058.1	0.0	1058.1	0.0	1.0	0.0	0.425	1.000

[a]Coefficient of variation.
[b]Theoretical variogram model; sph, spherical; exp, exponential; nug, pure nugget.
[c]Sill variance minus the nugget variance.
[d]For exponential model range $\times 3$.
[e]Median of the θ statistics for REML models; E{θmed} = 0.455.
[f]Mean of the θ statistics for REML models; E{θmean} = 1.000.
[g]Untransformed data without outliers.
[h]Data transformed by square roots before variogram analysis.

sizes (equation (1)) increase exponentially with decreasing event size (Figure 3). That is, smaller events generally need more samples for the same error limit. The number of funnel-type collectors, which is needed to keep the relative error below 20%, varies between 18 and 326; about 100 funnels would suffice for events with a median throughfall greater than 1 mm (Figure 3). The sample size for small events (<1 mm) required for relative errors below 10% and 5% can be as high as 1300 and 5200, respectively. If only throughfall events with a median larger than 1 mm are considered, the number of required collectors decreases to about 430 for the 10% and to about 1700 for the 5% error limit (Figure 3).

3.3. Sampling From the Simulated Fields

3.3.1. Sample Size Requirements for Trough-Type Collectors

[34] Employment of 1 m troughs and a 20% error limit involve a necessary sample size of up to 50 for the tested throughfall events (Figure 4). Most samples are required for the two small events, event 3 and event 43, and for event 6, which is of medium size, but has a relatively large contamination with outliers and a relatively strong spatial structure (Table 2). If 2 m troughs were used, less than 15 of them would suffice to keep the relative error below 20% except for the small events (events 3 and 43, see Table 2), which would require up the 30 troughs (Figure 4). Increasing the trough length to 4 or 10 m would reduce the number of collectors further to fewer than 15 in all cases (Figure 4). Sample size requirements drastically increase with a reduction of the error limit to 10%. Up to 150, 100, 75, and 30 troughs of lengths of 1, 2, 4, and 10 m, respectively, would have to be installed (Figure 4); again, the small events require the most samples. If only larger events (>5 mm) are of interest, then less than 50 troughs of 1 and 2 m length, and less than 30 of 4 m length would suffice for an estimated mean within 10% accuracy. A sample size below 15 suffices for 10 m troughs, except for small events or events with a pronounced spatial structure (reasons for the latter are given below); in such cases, up to 30 troughs of 10 m length would have to be employed.

[35] It appears to be impossible to keep the relative error below 5% even with the employment of troughs as for all lengths we considered some events would have to be measured with more than 200 samplers (Figure 4). If only events having more than 5 mm mean throughfall and if no, weak or very short spatial autocorrelations are considered, the use of 30 to 50 10 m troughs would suffice for the 5% error limit.

[36] When we correlated the ratio of the interpercentile ranges of trough- and funnel-type collectors with the effective range, we found no correlations for all selected events and both sample sizes (n = 30 and n = 200). We detected, however, a strong relationship between these ratios and the nugget-to-sill ratio. Here, correlation coefficients were −0.9, −0.8, −0.6, and −0.4 for the ratios 1 m troughs to funnels, 2 m troughs to funnels, 4 m troughs to funnels, and 10 m troughs to funnels, respectively. Hence, the relative benefit of trough-type collectors depends on the spatial correlation structure, in our case on the nugget-to-sill ratio; i.e., strong spatial autocorrelations decrease the relative benefits of troughs. Not surprisingly, an increase in trough lengths coincides with a decreasing influence of the spatial autocorrelations. A simultaneous occurrence of long-range and strong spatial autocorrelations likely reduces the benefits of longer troughs, too. As a consequence, prior information on the spatial structure of throughfall would support the decision making regarding appropriate collector types. Troughs will always reduce the sample size for a given error limit, yet this reduction will be smaller in the presence of spatial autocorrelations.

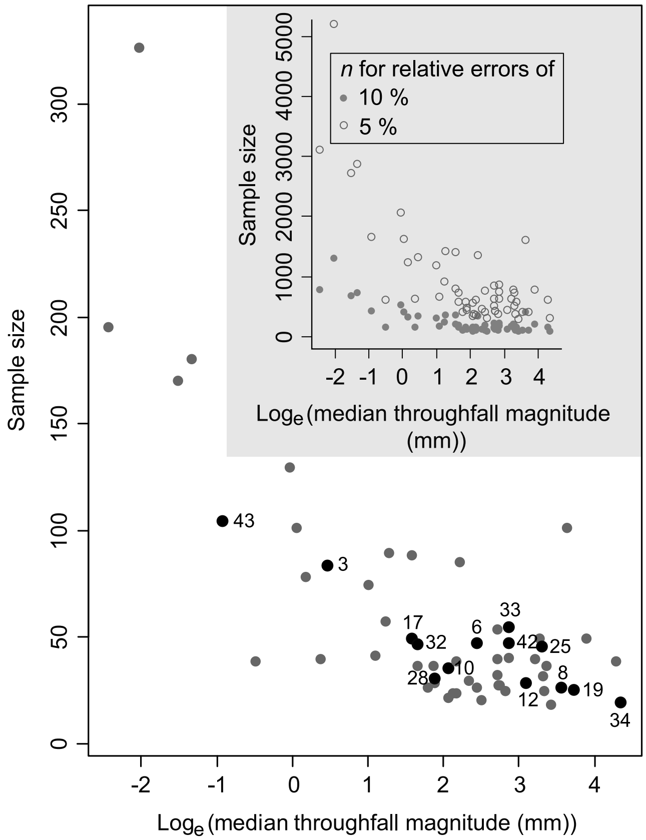

Figure 3. Required sample sizes for funnel-type collectors, which were calculated with equation (1), to keep the relative error of the sample mean below 20% as a function of median throughfall magnitude. Note the log scale for throughfall magnitude. The gray dots represent all 56 sampled throughfall events except the 14 events used for the simulations (black dots with event numbers). The inset shows the required sample sizes for relative error limits of 10% and 5%.

3.3.2. Comparison of Sampling Designs

[37] Simple random sampling versus stratified simple random sampling accounts for only marginal differences in the precision of the estimated means (Figure 5). In contrast, cluster random sampling performs worse when autocorrelation is present; the effective range was particularly important in this respect. Therefore, this type of sampling design was most disadvantageous when applied to events 12, 25, 32, 33, 42, and 43 (Table 2). Event 42, which exhibits both a relatively long correlation length and a comparatively strong spatial autocorrelation, shows the largest difference in precision between cluster random sampling and simple or stratified simple random sampling (Figure 5a). For a given sample size, the number of sample points per transect decreases as the number of transects increases, e.g., from three to five (Table 1); therefore, the number of clusters also matters (Figure 5a). In the absence of or for very short-range spatial autocorrelation, sampling designs do not differ (Figure 5b). In summary, in the presence of spatial autocorrelation neighboring observations contribute less information, which makes a cluster random sampling less efficient.

3.3.3. Influence of Outliers on Sample Size

[38] When using funnel-type collectors, the contamination with outliers drastically increases the sample size (Figure 6). For example, whereas 15–30 samples would suffice to keep the relative error below 20% for most but small events in the absence of outliers, their existence in real-world data sets inflates the number of required collectors (Figure 6). For the same reason it is impossible to achieve an error limit of 10% for funnel-type collectors as

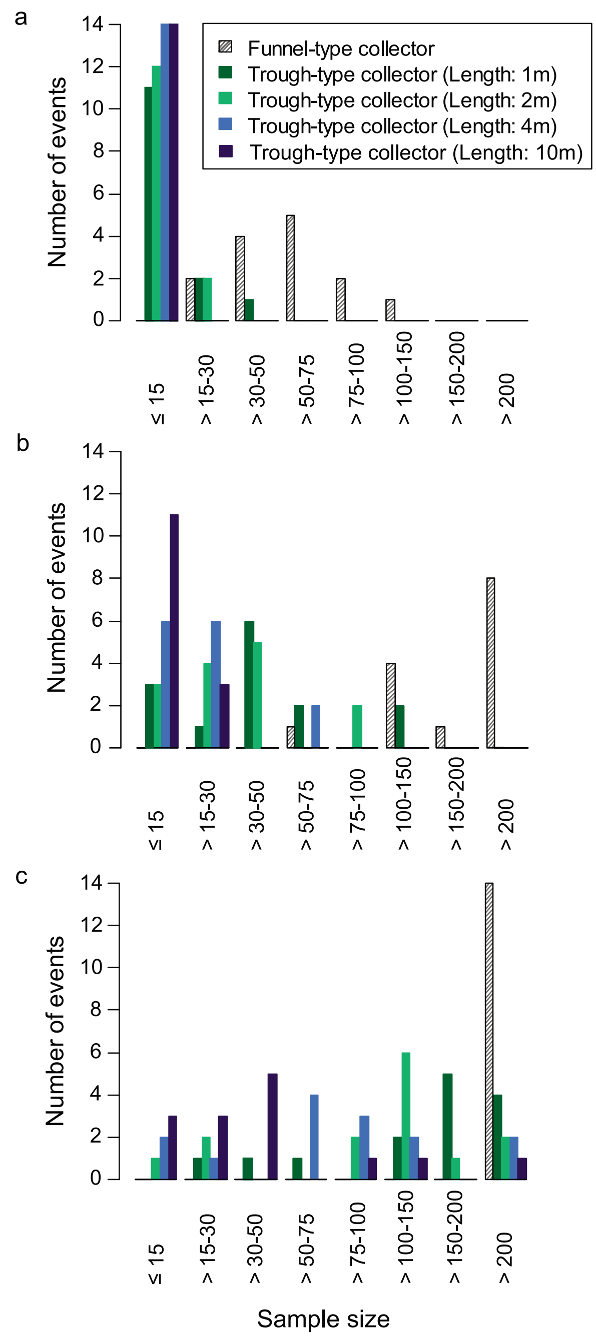

Figure 4. Number of throughfall events out of the 14 events used for the simulations that require ≤15, >15–30, >30–50, >50–75, >75–100, >100–150, >150–200, and >200 funnel- and trough-type collectors. The limits of the relative error of the mean are (a) 20%, (b) 10%, and (c) 5%.

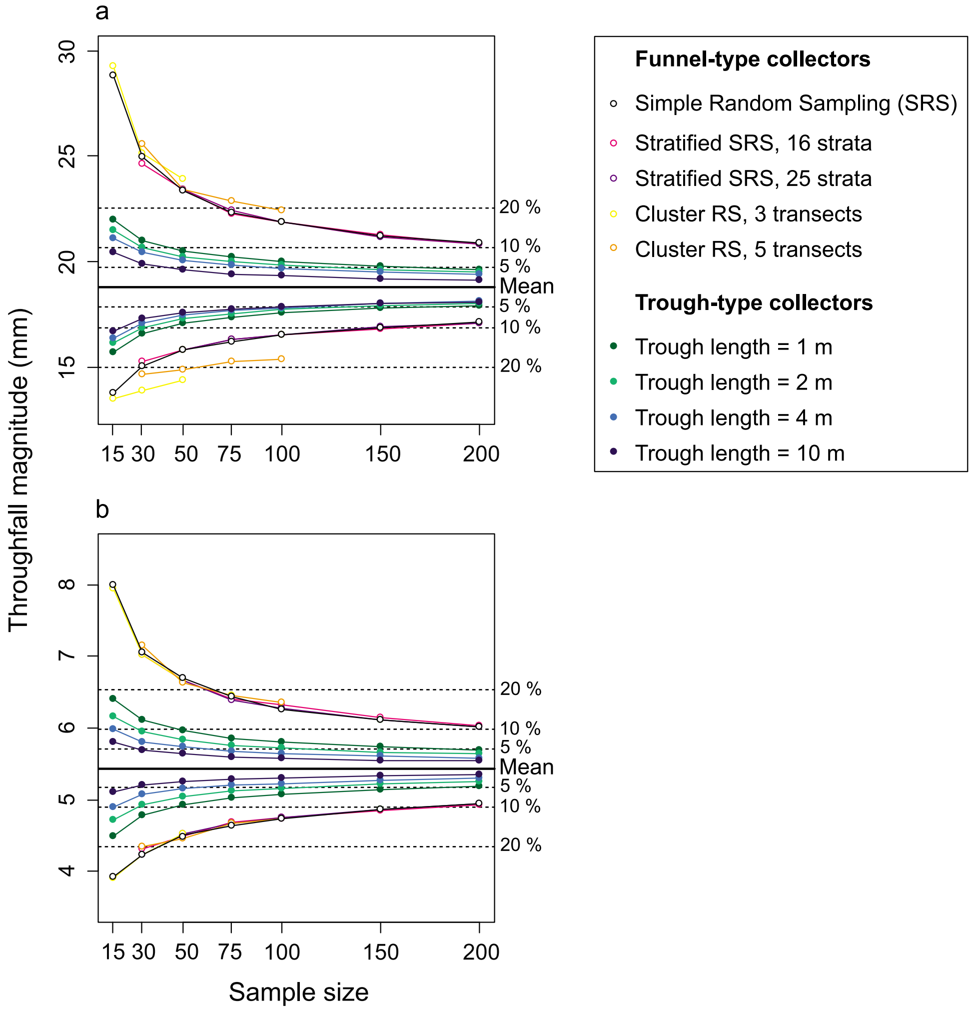

Figure 5. The 2.5% and 97.5% percentiles of the distribution of 10,000 sample means for all tested sample sizes and sampling strategies. The dashed lines indicate the mean ±5%, the mean ±10%, and the mean ±20% relative error of the mean. The bold line represents the true mean, which is the mean of the simulated throughfall field. Events (a) 42 and (b) 17 are shown.

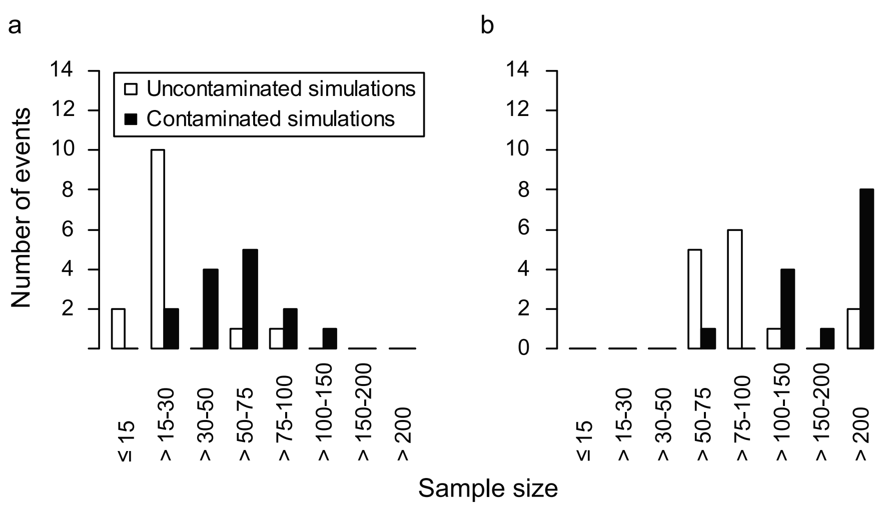

Figure 6. Number of throughfall events out of the 14 events used for the simulations that require ≤15, >15–30, >30–50, >50–75, >75–100, >100–150, >150–200, and >200 funnel-type collectors, both for the simulation without and with outliers, to achieve a prespecified error limit. We chose limits of the relative error of the mean of (a) 20% and (b) 10%.

Table 3. Required Time to Record Throughfall Volumes Intercepted by Funnel- and Trough-Type Collectors for the Spectrum of Throughfall Magnitudes in the Study Area and Error Limits of 20% and 10%[a]

	Funnel-Type Collectors		Trough-Type Collectors (Length of 1 m)		Trough-Type Collectors (Length of 2 m)		Trough-Type Collectors (Length of 4 m)	
	RE = 20%	RE = 10%	RE = 20%	RE = 10%	RE = 20%	RE = 10%	RE = 20%	RE = 10%
Maximum sample size[b]	139	505	36	139	25	94	16	56
First quartile recording time[c] (min)	40	144	14	55	13	47	15	51
Median recording time[c] (min)	38	139	27	105	26	99	29	102
Third quartile recording time[c] (min)	29	107	41	159	45	169	54	188

[a]Here RE is relative error.
[b]Required sample size to estimate mean throughfall for the whole frequency distribution, i.e., including small events.
[c]Recording time for one collecting device times the maximum sample size. Throughfall event size corresponds to the first quartile, the median, and the third quartile of the frequency distribution of the sampled throughfall events.

most throughfall events would have to be sampled with more than 200 samplers. In contrast, troughs are much less affected by extreme values (Figure 5), and the presence of outliers hardly increases the required sample sizes for all error limits. In the majority of cases, the contamination with outliers does not even produce a switch from one sample size class to the next higher level (not shown). This insensitivity of troughs derives from the a priori integration of unusually large or small data points because of their larger collecting area. Therefore, the extremely high throughfall amounts under drip points do not result in such a severe overestimation of mean throughfall as it happens when funnel-type collectors are employed (Figure 5).

3.3.4. Estimation of Recording Time

[39] In order to give an estimate of actual recording times in the field, the time required to empty a certain collector has to be combined with the maximum required sample size for that collector. Based on the 14 simulation events we determined the exact sample size needed for a particular collector and error limit. The results are summarized in Table 3 and show the following trends. At the achievable error limits of 10% and 20%, small events involve an overall recording time for troughs – regardless of their length – which amounts to about one third of the time which would be required to service funnel-type collectors (Table 3). This gain in time for troughs decreases to about three quarters for medium-sized events, and turns into a time loss for larger events, in particular for the 4 m trough (Table 3).

4. Discussion

4.1. Sample Size Requirements for Funnel-Type Collectors

4.1.1. Sample Size as Affected by Forest Type

[40] The standard formula (equation (1)) to calculate the sample size, which is needed to keep the relative error of the mean below a certain percentage, has been applied in a number of throughfall studies. Large differences exist in the number of required collectors depending on forest type [*Kimmins*, 1973]. For instance, throughfall under mixed hardwood canopies can be estimated within 10% mean relative error with as few as 11 fixed collectors at the 95% confidence level [*Puckett*, 1991]. Similarly, 9–11 fixed gauges suffice to keep the relative error below 10% and 22–23 fixed funnels are needed for a 5% error limit in a holm oak forest [*Rodrigo and Àvila*, 2001]. In contrast, *Holwerda et al.* [2006], working in a lower montane rain forest, calculated that 100 fixed collectors are necessary to keep relative error below 10% when the overall coefficient of variation, i.e., for cumulative throughfall after a number of sampling occasions, is considered.

4.1.2. Sample Size as Affected by Event Size

[41] It is a well-established fact that the coefficient of variation increases with decreasing throughfall depth [*Kimmins*, 1973; *Loustau et al.*, 1992b; *Viville et al.*, 1993; *Bellot et al.*, 1999; *Rodrigo and Àvila*, 2001; *Holwerda et al.*, 2006; *Staelens et al.*, 2006]. Accordingly, several investigators have emphasized that the rainfall pattern must be considered when selecting the number of collectors [*Rodrigo and Àvila*, 2001]; e.g., a much larger number of gauges is necessary to sample precipitation events <2 mm than >4 mm [*Price and Carlyle-Moses*, 2003]. Likewise, *Helvey and Patric* [1965] stated that small storms are hardest to sample accurately; others pointed out that the results for the smallest events should be viewed with caution [*Link et al.*, 2004]. *Kimmins* [1973] concluded that the planning of most throughfall studies suffers from an inadequate knowledge of the spatial and temporal variance of throughfall and advocated pilot studies. However, the calculated sample sizes resulting from of such a prestudy will fluctuate considerably with the length of the monitoring period or the event sizes for which they are computed [*Thimonier*, 1998]. Therefore, it should cover a period long enough to represent the regional precipitation regime or the part of it which is of interest for the study.

4.1.3. Sample Size Requirements for an Event-Based Throughfall Monitoring in the Studied Rain Forest

[42] Considering the present study as a pilot study for an old secondary forest in the humid tropics, the use of funnel-type collectors for a continuous, event-based throughfall monitoring implies either an unrealistically high sample size or a considerable error in the estimation of the throughfall mean. Even with the employment of 200 collectors, this error exceeds 10% for the majority of events (Figure 4). Hence, with this sample size, the estimated throughfall mean potentially fluctuates up to 20% around its true value, and the error might even be higher for very small events. Since the presence of outliers distorts the distribution toward its right tail, this fluctuation manifests itself in a more severe over- than underestimation of the mean throughfall magnitude (Figure 5).

4.2. Sampling From the Simulated Fields

4.2.1. Sample Size Requirements for Trough-Type Collectors

[43] The long-lasting debate in the throughfall literature on the use of funnel- versus trough-type or other large collecting devices has not yet come to a conclusion. On the one hand, some authors stated that the sampling strategy and the number of gauges, rather than type, is more important in gaining the true measure of throughfall input in forested ecosystems [*Helvey and Patric*, 1965; *Lloyd and Marques*, 1988; *Neal*, 1990; *Reynolds and Neal*, 1991; *Thimonier*, 1998; *Price and Carlyle-Moses*, 2003]. On the other hand, theoretical considerations and limited empirical evidence led others to assume that gauges with larger orifices might collect a more representative sample because they intercept a larger canopy area [*Kostelnik et al.*, 1989; *Loustau et al.*, 1992b; *Holwerda et al.*, 2006; *McJannet et al.*, 2007]. For instance, *Crockford and Richardson* [1990] showed in an empirical study that the variance of throughfall caught in small funnel-type gauges was much greater than it was for troughs, and that a much higher sample size would have to be used for the same level of precision; they therefore considered troughs to provide the best estimates of mean throughfall. In the present study we showed that a larger support will always reduce the sample size and that the degree of this reduction will depend on the spatial autocorrelation structure. In general, when considering the usefulness of an increased sample support, the contribution of small scale to the overall variability is of particular importance. For instance, if a property varies modestly over a short range but wildly over the extent of the study area, a larger support is likely to be of little use. If, in contrast, most variation occurs at a small scale, an increase in support offers considerable benefits. Due to the dominance of either very short scale or of weak to nonexisting spatial autocorrelations of our throughfall data, the large reduction in sample size with increasing support in the present study is not surprising (Figure 4). This obvious benefit, however, has to be interpreted with caution since the employment of troughs is accompanied by some difficulties. Whereas the practical problem of a splash out from troughs [e.g., *Neal*, 1990] seems to be surmountable [e.g., *McJannet et al.*, 2007], troughs are potentially prone to larger wetting and detritus errors in comparison to funnels, in particular if the installation angle of a trough-type collector is too low for allowing fast drainage, or if there are too few field visits for the removal of detritus. Moreover, the extra logistical effort associated with the installation and maintenance of troughs has to be taken into account. This is particularly true for nonautomatic sampling. Our results show that the employment of trough-type collectors results in large recording times for large events (Table 3) which could, however, be reduced if an additional larger measuring cylinder than the 1 L cylinder used in our experiment was carried to the field. If the overall recording time is considered, which is determined by the recording time times the number of collectors required to go below a certain error limit, troughs are more efficient for all but large events (Table 3). Moreover, our comparison does not take into account the walking time between collectors, which may even further increase the efficiency of recording throughfall collected in troughs compared to funnel-type collectors. Differences in the overall recording time among trough lengths appear to be small. If wetting errors are indeed a function of trough length as hypothesized by *Thimonier* [1998] shorter troughs are to be preferred.

[44] To conclude, we showed that the decision concerning the support is not straightforward. Troughs reduce the sample size and integrate outliers, but their relative benefits depend on the spatial correlation structure and the precipitation regime. They are most useful if short-scale variation dominates. When they have to be recorded manually, the gain in time depends on event size. Due to the large number of shorter troughs, which is needed to keep the relative error of the mean at least below 10%, automatically recording systems would likely consist of clusters of troughs connected to a tipping bucket, e.g., similar to the arrangement used by *McJannet et al.* [2007]. But even with such a layout, a relative error of about 5% or less appears to be beyond reach for highly variable ecosystems such as the tropical forest we looked at.

4.2.2. Comparison of Sampling Designs

[45] In most throughfall investigations, collectors were placed following a simple random sampling procedure [e.g., *Bruijnzeel and Wiersum*, 1987; *Lloyd and Marques*, 1988; *Puckett*, 1991; *Gash et al.*, 1995; *Lin et al.*, 1997; *Schellekens et al.*, 1999; *Carlyle-Moses et al.*, 2004; *Keim et al.*, 2005; *Holwerda et al.*, 2006; *A. Zimmermann et al.*, 2007, 2008]. Some kind of systematic sampling was used by *Lawrence and Fernandez* [1993] and *Whelan et al.* [1998], whereas *Rodrigo and Àvila* [2001] and *Fleischbein et al.* [2005] implemented sampling procedures that may be considered cluster random sampling; the latter study provides no information about the cluster selection procedure. The trough system used by *McJannet et al.* [2007] was replicated within their study areas, but the reader does not get to know the rationale for the choice of the locations. In general, since the sampling scheme does have implications for the interpretation and comparability of the results, we strongly suggest that authors provide sufficient information about their sampling designs, the extent of their study area, and the support of the collecting devices.

[46] Our sampling design comparison reveals that stratification, which aims at dividing the investigation area into more homogeneous subunits, has no effect on the accuracy of the estimated mean in our 1 ha plot. This can be attributed to the fact that the range of the spatial autocorrelation is quite small compared to the extent, and throughfall volumes at nearby locations can differ as much as those further apart. Furthermore, even with a sample size of 200 the chance of a repeated random choice of unfavorable close points is very small due to the enormous number of possible sampling locations, which amount to 10^6. In contrast, we showed that for cluster random sampling the placement of collectors along transects at distances as close as 1 m can be disadvantageous in the presence of spatial autocorrelations. The use of longer transects than the tested 20 m likely eliminates this negative effect, but the rationale of cluster random sampling, reduce walking time between sampling locations, would not be valid any more. In summary, simple random sampling appears to be a good choice in the absence of a pronounced spatial structure. If existence of the latter involves extensive clusters of similar throughfall amounts, a simple random site selection possibly overrepresents certain spots. This scenario

appears rather unlikely since the majority of studies on spatial variation of throughfall found merely short autocorrelations in the range of 0 to 10 m [e.g., *Loustau et al.*, 1992b; *Bellot and Escarre*, 1998; *Möttönen et al.*, 1999; *Keim et al.*, 2005; *Staelens et al.*, 2006; *A. Zimmermann et al.*, 2009].

4.3. Implications for Interception Monitoring

[47] Particularly in data-sparse regions, rainfall interception model parameters are usually derived by calibration against measured throughfall volumes. For instance, the free throughfall coefficient, p, has frequently been determined as the slope of the regression line of throughfall on gross rainfall for small storms [e.g., *Bruijnzeel and Wiersum*, 1987; *Gash et al.*, 1995; *Schellekens et al.*, 1999; *Link et al.*, 2004; *Fleischbein et al.*, 2005]. Likewise, break point regression of throughfall on rainfall has been used to identify P_g', the gross rainfall required to saturate the canopy [e.g., *Link et al.*, 2004; *Herbst et al.*, 2008]. Last, the canopy capacity S can be derived from the negative intercept of the regression line of throughfall on gross rainfall [e.g., *Leyton et al.*, 1967; *Bruijnzeel and Wiersum*, 1987; *Hutjes et al.*, 1990; *Lankreijer et al.*, 1993; *Gash et al.*, 1995; *Schellekens et al.*, 1999; *Marin et al.*, 2000; *Fleischbein et al.*, 2005; *Germer et al.*, 2006]. Needless to say, the accuracy of these derived canopy parameters depends on the accuracy of the rainfall and throughfall measurements from which they were derived. This was commented on in several studies, which we will discuss from the throughfall perspective. For instance, *Lloyd et al.* [1988] concluded that relatively large errors in the measurement of the interception loss have to be accepted, which translate into similarly large errors in the derivation of the forest structure parameters. *Bruijnzeel and Wiersum* [1987] pointed out that even if the application of a larger number of throughfall collectors resulted in lower estimates of interception loss, possibly through the inclusion of a representative number of drip points, many studies, including their own one, may have overestimated interception losses. In a study of rainfall interception in an Amazon rain forest, *Sellers et al.* [1989] even suggested that measurement errors of throughfall are such that interception can be estimated reliably only on a periodic, not weekly or event basis. *Vrugt et al.* [2003] studied the identifiability of interception model parameters from measurements of throughfall and canopy storage, and concluded that throughfall measurements are of limited use for the derivation of these parameters because of their considerable uncertainty. Our results suggest that their findings may be partly due to an insufficient accuracy of their throughfall measurements. Uncertainty in throughfall measurements prompted *Hutjes et al.* [1990] to call for an increase in the accuracy of the observations of net rainfall as a prerequisite for any further improvement of interception models. Considering the results of the present study, it appears rather difficult to meet these legitimate requirements. We expect the uncertainty of canopy parameter estimates whose determination relies on small throughfall events, e.g., P_g' and p, to be particularly high. In forest ecosystems characterized by variations in throughfall magnitudes as large as in ours, the maximum achievable accuracy of estimated mean throughfall for small events is limited to about 10%. The high number of funnel-type collectors to achieve this accuracy, and the rather infrequent use of large supports (an exception is the trough system used by *McJannet et al.* [2007]), suggest that the relative errors of the estimated throughfall means may frequently exceed even 10%. Unfortunately, the roving collector approach to reduce the random error, as suggested by *Bruijnzeel and Wiersum* [1987], is of limited use in forest where the throughfall distribution is skewed to the right, e.g., due to drip points or to the overall skewness of small events [*A. Zimmermann et al.*, 2009]. In these cases, the likely overestimation of the mean cannot be reduced by changing the collector locations. To make things even more complicated, the necessary accuracy of throughfall measurements for the derivation of interception model parameters remains largely unknown, and calls for relative error limits of the throughfall mean of 5% or 10% [e.g., *Kimmins*, 1973; *Rodrigo and Àvila*, 2001] have yet to be evaluated against model requirements. This would be an interesting task for future studies. The sensitivity of model outcomes to the model parameters has already been demonstrated [e.g., *Loustau et al.*, 1992a; *Schellekens et al.*, 1999; *Liu*, 2001]. In case the achievable accuracy of throughfall measurements is too low to identify some of these parameters in certain forest ecosystems, alternatives to their determination should be considered. Similar thoughts led *Liu* [2001] to call for more studies that link interception model parameters, such as storage capacity and free throughfall coefficient, to remotely sensed data or surface area indexes of leaves, branches, and stems. For instance, *Nieschulze et al.* [2009] demonstrated the suitability of high-resolution optical space-borne imagery for predicting regional interception. Another successful example of using alternative data is given by *Vrugt et al.* [2003]. They showed that canopy water storage dynamics derived from the attenuation of a microwave signal contained sufficient information for the identification of interception model parameters with a high degree of confidence. Particularly in data-sparse regions, however, where high-end techniques are not routinely available, interception modeling will likely continue to rely on monitoring throughfall dynamics.

5. Conclusions

[48] We summarize our results by answering the research questions posed in the introduction:

[49] 1. Due to an increasing coefficient of variation with decreasing throughfall event size, the required sample sizes increase exponentially for small events. For a 20% error limit, a sample size of up to 300 funnel-type collectors is required, which reduces to about 100 for events larger than a few millimeters. A more accurate estimation of mean throughfall, approaching error limits of 10% or 5%, involves the deployment of up to 1300 and 5200 funnels, respectively. Considering larger events only, about 430 and 1700 collectors are needed to approach error limits of 10% and 5%, respectively. Therefore, funnel-type collectors coincide with either an unrealistic high sample size or a considerable error in the estimation of mean throughfall in the studied tropical forest.

[50] 2. Using 10 cm wide trough-type collectors, the 20% error limit entails the need for up to 50 troughs of 1 m length. The number of required collectors reduces to no more than 30 for a trough length of 2 m and furthermore to

less than 15 for 4 and 10 m long troughs. Up to 150, 100, 75, and 30 troughs of 1, 2, 4, and 10 m length are needed to keep the relative error below 10%. If only throughfall amounts of larger events without strong autocorrelation have to be determined, about 15–30 long and 30–50 shorter troughs would suffice. For the 5% error limit, even the employment of long troughs results in overall large sample sizes, which can exceed 200 for the smallest events. In general, the relative benefits of trough-type collectors depend on the spatial correlation structure. That is to say, the reduction of sample size due to the larger support of troughs is smaller in the presence of spatial autocorrelations. In addition, the gain in time for a manual handling of troughs in comparison to funnels depends on event size; they seem to be least beneficial for large events. Finally, a clear advantage of troughs is the integration of unusually large or small values in their larger collecting area, thereby reducing the risk of an overestimation of the mean due to drip points, which occur frequently in the studied ecosystem.

[51] 3. Regarding the sampling design, simple random sampling is a good choice for monitoring throughfall because of its prevalently short-scale variation. The use of cluster random sampling tends to be less efficient if the distances between monitoring sites are within the range of the spatial autocorrelations.

[52] 4. As to interception modeling, we postulate a high uncertainty of canopy parameters whose identification relies on small throughfall events, most notably the free throughfall coefficient and the amount of gross rainfall required to saturate the canopy. The accuracy of throughfall measurements, which is necessary to determine these parameters with a sufficient degree of confidence, has yet to be identified. If this accuracy turns out to be out of reach, alternatives for their estimation will have to be considered.

[53] In general, it would be desirable if we were able to link forest structures to spatial fields of throughfall; unfortunately, the (published) database worldwide does not suffice for such an attempt. Therefore we call for proper studies on the spatial structure of throughfall that incorporate a sufficient number of sampling points and that span a variety of forest ecosystems in order to reassess our findings for forests with a possibly different spatial structure.

[54] **Acknowledgments.** This research was sponsored as part of The HSBC Climate Partnership, a partnership between HSBC and The Climate Group, Earthwatch Institute, Smithsonian Tropical Research Institute, and WWF. A.Z. and H.E. acknowledge support from the German Research Foundation (El 255/6-1). B.Z. and H.E. acknowledge partial support by the German Academic Exchange Service (DAAD). R.M.L.'s input is part of the program for Mathematical and Computational Biology at Rothamsted Research, sponsored by the U.K. Biotechnology and Biological Sciences Research Council. We thank Luise Neumann-Cosel, Bärbel Ehrig, Silja Hund, Anna Schürkmann, and Frank Bäse for participating in the field work.

References

Bárdossy, A., and Z. W. Kundzewicz (1990), Geostatistical methods for detection of outliers in groundwater quality spatial fields, *J. Hydrol.*, *115*, 343–359, doi:10.1016/0022-1694(90)90213-H.

Becker, R. A., J. M. Chambers, and A. R. Wilks (1988), *The New S Language*, Wadsworth, Pacific Grove, Calif.

Bellot, J., and A. Escarre (1998), Stemflow and throughfall determination in a resprouted Mediterranean holm-oak forest, *Ann. Sci. For.*, *55*, 847–865, doi:10.1051/forest:19980708.

Bellot, J., A. Ávila, and A. Rodrigo (1999), Throughfall and stemflow, in *Ecology of Mediterranean Evergreen Oak Forests*, edited by F. Roda et al., pp. 209–222, Springer, Berlin.

Box, G. E. P., and D. R. Cox (1964), An analysis of transformations, *J. R. Stat. Soc. Ser. B*, *26*, 211–246.

Bruijnzeel, L. A., and K. F. Wiersum (1987), Rainfall interception by a young *Acacia Auriculiformis* (A. Cunn) plantation forest in west Java, Indonesia: Application of Gash's analytical model, *Hydrol. Processes*, *1*, 309–319, doi:10.1002/hyp.3360010402.

Brys, G., M. Hubert, and A. Struyf (2003), A comparison of some new measures of skewness, in *Developments in Robust Statistics*, edited by R. Dutter et al., pp. 98–113, Physica, Heidelberg, Germany.

Carlyle-Moses, D. E., J. S. Flores Laureano, and A. G. Price (2004), Throughfall and throughfall spatial variability in Madrean oak forest communities of northeastern Mexico, *J. Hydrol.*, *297*, 124–135, doi:10.1016/j.jhydrol.2004.04.007.

Chilès, J.-P., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, John Wiley, New York.

Cressie, N., and D. Hawkins (1980), Robust estimation of the variogram, *Math. Geol.*, *12*, 115–125, doi:10.1007/BF01035243.

Cressie, N. A. C. (1993), *Statistics for Spatial Data*, rev. ed., Wiley, New York.

Croat, T. B. (1978), *Flora of Barro Colorado Island*, 943 pp., Stanford Univ. Press, Stanford, Calif.

Crockford, R. H., and D. P. Richardson (1990), Partitioning of rainfall in a eucalypt forest and pine plantation in southeastern Australia: Part I. Throughfall measurement in a eucalypt forest: Effect of method and species composition, *Hydrol. Process.*, *4*, 131–144, doi:10.1002/hyp.3360040204.

de Gruijter, J. J., D. J. Brus, M. F. P. Bierkens, and M. Knotters (2006), *Sampling for Natural Resource Monitoring*, 332 pp., Springer, Berlin.

Dowd, P. A. (1984), The variogram and kriging: Robust and resistant estimators, in *Geostatistics for Natural Resources Characterization*, edited by G. Verly et al., pp. 91–106, D. Reidel, Dordrecht, Netherlands.

Fleischbein, K., W. Wilcke, R. Goller, J. Boy, C. Valarezo, W. Zech, and K. Knoblich (2005), Rainfall interception in a lower montane forest in Ecuador: Effects of canopy properties, *Hydrol. Processes*, *19*, 1355–1371, doi:10.1002/hyp.5562.

Foster, R. B., and N. V. L. Brokaw (1982), Structure and history of the vegetation of Barro Colorado Island, in *The Ecology of a Tropical Forest: Seasonal Rhythms and Long Term Changes*, edited by E. G. Leigh et al., pp. 67–81, Smithson. Inst., Washington, D. C.

Gash, J. H. C., C. R. Lloyd, and G. Lachaud (1995), Estimating sparse forest rainfall interception with an analytical model, *J. Hydrol.*, *170*, 79–86, doi:10.1016/0022-1694(95)02697-N.

Genton, M. G. (1998), Highly robust variogram estimation, *Math. Geol.*, *30*, 213–221, doi:10.1023/A:1021728614555.

Germer, S., H. Elsenbeer, and J. M. Moraes (2006), Throughfall and temporal trends of rainfall redistribution in an open tropical rainforest, southwestern Amazonia (Rondônia, Brazil), *Hydrol. Earth Syst. Sci.*, *10*, 383–393.

Helvey, J. D., and J. H. Patric (1965), Canopy and litter interception of rainfall by hardwoods of eastern United States, *Water Resour. Res.*, *1*, 193–206, doi:10.1029/WR001i002p00193.

Herbst, M., P. T. W. Rosier, D. D. McNeil, R. J. Harding, and D. J. Gowing (2008), Seasonal variability of interception evaporation from the canopy of a mixed deciduous forest, *Agric. For. Meteorol.*, *148*, 1655–1667, doi:10.1016/j.agrformet.2008.05.011.

Holwerda, F., F. N. Scatena, and L. A. Bruijnzeel (2006), Throughfall in a Puerto Rican lower montane rain forest: A comparison of sampling strategies, *J. Hydrol.*, *327*, 592–602, doi:10.1016/j.jhydrol.2005.12.014.

Hutjes, R. W. A., A. Wierda, and A. W. L. Veen (1990), Rainfall interception in the Tai Forest, Ivory Coast: Application of two simulation models to a humid tropical system, *J. Hydrol.*, *114*, 259–275, doi:10.1016/0022-1694(90)90060-B.

Jost, G., G. B. M. Heuvelink, and A. Papritz (2005), Analysing the space-time distribution of soil water storage of a forest ecosystem using spatio-temporal kriging, *Geoderma*, *128*, 258–273, doi:10.1016/j.geoderma.2005.04.008.

Keim, R. F., A. E. Skaugset, and M. Weiler (2005), Temporal persistence of spatial patterns in throughfall, *J. Hydrol.*, *314*, 263–274, doi:10.1016/j.jhydrol.2005.03.021.

Kimmins, J. P. (1973), Some statistical aspects of sampling throughfall precipitation in nutrient cycling studies in British Columbian coastal forests, *Ecology*, *54*, 1008–1019, doi:10.2307/1935567.

Kostelnik, K. M., J. A. Lynch, J. W. Grimm, and E. S. Corbett (1989), Sample size requirements for estimation of throughfall chemistry beneath a mixed hardwood forest, *J. Environ. Qual.*, *18*, 274–280.

Lankreijer, H. J. M., M. J. Hendriks, and W. Klaassen (1993), A comparison of models simulating rainfall interception of forests, *Agric. For. Meteorol.*, *64*, 187–199, doi:10.1016/0168-1923(93)90028-G.

Lark, R. M. (2000), A comparison of some robust estimators of the variogram for use in soil survey, *Eur. J. Soil Sci.*, *51*, 137–157, doi:10.1046/j.1365-2389.2000.00280.x.

Lark, R. M. (2002), Modelling complex soil properties as contaminated regionalized variables, *Geoderma*, *106*, 173–190, doi:10.1016/S0016-7061(01)00123-9.

Lark, R. M., B. R. Cullis, and S. J. Welham (2006), On spatial prediction of soil properties in the presence of a spatial trend: The empirical best linear unbiased predictor (E-BLUP) with REML, *Eur. J. Soil Sci.*, *57*, 787–799, doi:10.1111/j.1365-2389.2005.00768.x.

Lawrence, G. B., and I. J. Fernandez (1993), A reassessment of areal variability of throughfall deposition measurements, *Ecol. Appl.*, *3*, 473–480, doi:10.2307/1941916.

Levia, D. F., and E. E. Frost (2006), Variability of throughfall volume and solute inputs in wooded ecosystems, *Prog. Phys. Geogr.*, *30*, 605–632, doi:10.1177/0309133306071145.

Leyton, L., E. R. C. Reynolds, and F. B. Thompson (1967), Rainfall interception in forest and moorland, in *International Symposium on Forest Hydrology*, edited by W. E. Sopper and H. W. Lull, pp. 163–178, Pergamon, Oxford, U. K.

Lin, T. C., S. P. Hamburg, H. B. King, and Y. J. Hsia (1997), Spatial variability of throughfall in a subtropical rain forest in Taiwan, *J. Environ. Qual.*, *26*, 172–180.

Link, T. E., M. Unsworth, and D. Marks (2004), The dynamics of rainfall interception by a seasonal temperate rainforest, *Agric. For. Meteorol.*, *124*, 171–191, doi:10.1016/j.agrformet.2004.01.010.

Liu, S. (2001), Evaluation of the Liu model for predicting rainfall interception in forests world-wide, *Hydrol. Processes*, *15*, 2341–2360, doi:10.1002/hyp.264.

Lloyd, C. R., and A. d. O. Marques (1988), Spatial variability of throughfall and stemflow measurements in Amazonian rainforest, *Agric. For. Meteorol.*, *42*, 63–73, doi:10.1016/0168-1923(88)90067-6.

Lloyd, C. R., J. H. C. Gash, W. J. Shuttleworth, and A. d. O. Marques (1988), The measurement and modeling of rainfall interception by Amazonian rainforest, *Agric. For. Meteorol.*, *43*, 277–294, doi:10.1016/0168-1923(88)90055-X.

Loustau, D., P. Berbigier, and A. Granier (1992a), Interception loss, throughfall and stemflow in a maritime pine stand. Part II. An application of Gash's analytical model of interception, *J. Hydrol.*, *138*, 469–485, doi:10.1016/0022-1694(92)90131-E.

Loustau, D., P. Berbigier, A. Granier, and F. E. H. Moussa (1992b), Interception loss, throughfall and stemflow in a maritime pine stand. Part I. Variability of throughfall and stemflow beneath the pine canopy, *J. Hydrol.*, *138*, 449–467, doi:10.1016/0022-1694(92)90130-N.

Marin, C. T., W. Bouten, and J. Sevink (2000), Gross rainfall and its partitioning into throughfall, stemflow and evaporation of intercepted water in four forest ecosystems in western Amazonia, *J. Hydrol.*, *237*, 40–57, doi:10.1016/S0022-1694(00)00301-2.

Matheron, G. (1962), *Traité de Géostatistique Appliqué*, vol. 1, Bur. de Rech. Géol. et Min., Paris.

McJannet, D., J. Wallace, and P. Reddell (2007), Precipitation interception in Australian tropical rainforests: Part I. Measurement of stemflow, throughfall and cloud interception, *Hydrol. Processes*, *21*, 1692–1702, doi:10.1002/hyp.6347.

Meklit, T., M. Van Meirvenne, S. Verstraete, J. Bonroy, and F. Tack (2008), Combining marginal and spatial outliers identification to optimize the mapping of the regional geochemical baseline concentration of soil heavy metals, *Geoderma*, *148*, 413–420, doi:10.1016/j.geoderma.2008.11.014.

Möttönen, M., E. Järvinen, T. J. Hokkanen, T. Kuuluvainen, and R. Ohtonen (1999), Spatial distribution of soil ergosterol in the organic layer of a mature Scots pine (*Pinus sylvestris* L.) forest, *Soil Biol. Biochem.*, *31*, 503–516, doi:10.1016/S0038-0717(98)00122-9.

Neal, C. (1990), Comments on "Sample size requirements for estimation of throughfall chemistry beneath a mixed hardwood forest," *J. Environ. Qual.*, *19*, 633–634.

Nieschulze, J., S. Erasmi, J. Dietz, and D. Hoelscher (2009), Satellite-based prediction of rainfall interception by tropical forest stands of a human-dominated landscape in central Sulawesi, Indonesia, *J. Hydrol.*, *364*, 227–235, doi:10.1016/j.jhydrol.2008.10.024.

Papritz, A., and R. Webster (1995), Estimating temporal change in soil monitoring: PartII. Sampling from simulated fields, *Eur. J. Soil Sci.*, *46*, 13–27, doi:10.1111/j.1365-2389.1995.tb01809.x.

Pardo-Igúzquiza, E. (1998), Maximum likelihood estimation of spatial covariance parameters, *Math. Geol.*, *30*, 95–108, doi:10.1023/A:1021765405952.

Price, A. G., and D. E. Carlyle-Moses (2003), Measurement and modeling of growing- season canopy water fluxes in a mature mixed deciduous forest stand, southern Ontario, Canada, *Agric. For. Meteorol.*, *119*, 69–85, doi:10.1016/S0168-1923(03)00117-5.

Puckett, L. J. (1991), Spatial variability and collector requirements for sampling throughfall volume and chemistry under a mixed hardwood canopy, *Can. J. For. Res.*, *21*, 1581–1588, doi:10.1139/x91-220.

Rawlins, B. G., R. M. Lark, K. E. O'Donnell, A. M. Tye, and T. R. Lister (2005), The assessment of point and diffuse metal pollution of soils from an urban geochemical survey of Sheffield, England, *Soil Use Manage.*, *21*, 353–362, doi:10.1079/SUM2005335.

Reynolds, B., and C. Neal (1991), Trough versus funnel collectors for measuring throughfall volumes, *J. Environ. Qual.*, *20*, 518–521.

Rodrigo, A., and A. Àvila (2001), Influence of sampling size in the estimation of mean throughfall in two Mediterranean holm oak forests, *J. Hydrol.*, *243*, 216–227, doi:10.1016/S0022-1694(00)00412-1.

Schellekens, J., F. N. Scatena, L. A. Bruijnzeel, and A. J. Wickel (1999), Modelling rainfall interception by a lowland tropical rain forest in northeastern Puerto Rico, *J. Hydrol.*, *225*, 168–184, doi:10.1016/S0022-1694(99)00157-2.

Schlather, M. (2001), Simulation and analysis of random fields, *R News*, *1*, 18–20.

Sellers, P. J., W. J. Shuttleworth, J. L. Dorman, A. Dalcher, and J. M. Roberts (1989), Calibrating the simple biosphere model for Amazonian tropical forest using field and remote sensing data. Part 1: Average calibration with field data, *J. Appl. Meteorol.*, *28*, 727–759, doi:10.1175/1520-0450(1989)028 < 0727:CTSBMF > 2.0.CO;2.

Silverman, B. W. (1978), Choosing the window width when estimating a density, *Biometrika*, *65*, 1–11, doi:10.1093/biomet/65.1.1.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Staelens, J., A. D. Schrijver, K. Verheyen, and N. E. C. Verhoest (2006), Spatial variability and temporal stability of throughfall water under a dominant beech (*Fagus sylvatica* L.) tree in relationship to canopy cover, *J. Hydrol.*, *330*, 651–662, doi:10.1016/j.jhydrol.2006.04.032.

Thimonier, A. (1998), Measurement of atmospheric deposition under forest canopies: Some recommendations for equipment and sampling design, *Environ. Monit. Assess.*, *52*, 353–387, doi:10.1023/A:1005853429853.

Viville, D., P. Biron, A. Granier, E. Dambrine, and A. Probst (1993), Interception in a mountains declining spruce stand in Strengbach catchment (Vosges, France), *J. Hydrol.*, *144*, 273–282, doi:10.1016/0022-1694(93)90175-9.

Vrugt, J. A., S. C. Dekker, and W. Bouten (2003), Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage, *Water Resour. Res.*, *39*(9), 1251, doi:10.1029/2003WR002013.

Webster, R., and M. A. Oliver (2007), *Geostatistics for Environmental Scientists*, 2nd ed., John Wiley, Chichester, U. K.

Whelan, M. J., L. J. Sanger, M. Baker, and J. M. Anderson (1998), Spatial patterns of throughfall and mineral ion deposition in a lowland Norway spruce (*Picea abies*) plantation at the plot scale, *Atmos. Environ.*, *32*, 3493–3501, doi:10.1016/S1352-2310(98)00054-5.

Zimmermann, A., W. Wilcke, and H. Elsenbeer (2007), Spatial and temporal patterns of throughfall quantity and quality in a tropical montane forest in Ecuador, *J. Hydrol.*, *343*, 80–96, doi:10.1016/j.jhydrol.2007.06.012.

Zimmermann, A., S. Germer, C. Neill, A. V. Krusche, and H. Elsenbeer (2008), Spatio-temporal patterns of throughfall and solute deposition in an open tropical rain forest, *J. Hydrol.*, *360*, 87–102, doi:10.1016/j.jhydrol.2008.07.028.

Zimmermann, A., B. Zimmermann, and H. Elsenbeer (2009), Rainfall redistribution in a tropical forest: Spatial and temporal patterns, *Water Resour. Res.*, *45*, W11413, doi:10.1029/2008WR007470.

Zimmermann, B., E. Zehe, N. K. Hartmann, and H. Elsenbeer (2008), Analyzing spatial data: An assessment of assumptions, new methods, and uncertainty using soil hydraulic data, *Water Resour. Res.*, *44*, W10408, doi:10.1029/2007WR006604.

H. Elsenbeer and A. Zimmermann, Institute of Geoecology, University of Potsdam, Karl Liebknecht Str. 24-25, D-14476 Potsdam, Germany.

R. M. Lark, Rothamsted Research, West Common, Harpenden AL5 2JQ, UK.

B. Zimmermann, Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancón, Panama. (zimmermannb@si.edu)