Rainfall redistribution in a tropical forest: Spatial and temporal patterns

Alexander Zimmermann,[1] Beate Zimmermann,[2] and Helmut Elsenbeer[1,2]

Received 21 September 2008; revised 19 July 2009; accepted 4 August 2009; published 11 November 2009.

[1] The investigation of throughfall patterns has received considerable interest over the last decades. And yet, the geographical bias of pertinent previous studies and their methodologies and approaches to data analysis cast a doubt on the general validity of claims regarding spatial and temporal patterns of throughfall. We employed 220 collectors in a 1-ha plot of semideciduous tropical rain forest in Panama and sampled throughfall during a period of 14 months. Our analysis of spatial patterns is based on 60 data sets, whereas the temporal analysis comprises 91 events. Both data sets show skewed frequency distributions. When skewness arises from large outliers, the classical, nonrobust variogram estimator overestimates the sill variance and, in some cases, even induces spurious autocorrelation structures. In these situations, robust variogram estimation techniques offer a solution. Throughfall in our plot typically displayed no or only weak spatial autocorrelations. In contrast, temporal correlations were strong, that is, wet and dry locations persisted over consecutive wet seasons. Interestingly, seasonality and hence deciduousness had no influence on spatial and temporal patterns. We argue that if throughfall patterns are to have any explanatory power with respect to patterns of near-surface processes, data analytical artifacts must be ruled out lest spurious correlation be confounded with causality; furthermore, temporal stability over the domain of interest is essential.

Citation: 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.

1. Introduction

[2] Spatial patterns of several hydrological and biogeochemical processes at the forest floor have been linked to throughfall patterns, such as the distribution of soil water content [*Schume et al.*, 2003], seepage water and ion fluxes [*Manderscheid and Matzner*, 1995], decomposition of organic material [*Möttönen et al.*, 1999], root water uptake [*Bouten et al.*, 1992], and root growth [*Ford and Deans*, 1978]. The verdict is, however, not unequivocal. *Raat et al.* [2002] and *Shachnovich et al.* [2008] found no relationship between throughfall patterns and the distribution of the soil water content, which they attributed to the homogenizing effect of the forest floor. Nonetheless, this list hints at the potential impact of throughfall patterns on hydrological and biological processes in forest ecosystems. It appears that the strength of the link between throughfall and processes at or below the forest floor critically depends on the spatial patterns of throughfall and its temporal persistence.

[3] Spatial patterns are, of course, intimately linked to spatial variability, which by all accounts is rather high for throughfall [*Levia and Frost*, 2006], though large differences exist among forest ecosystems; latitudinal and management gradients may explain some of them [*Lloyd and Marques*, 1988; *Hölscher et al.*, 1998; *Möttönen et al.*, 1999; *Keim et al.*, 2005; *Holwerda et al.*, 2006]. Throughfall measurements in managed temperate forests, for example, vary from 0 to 100% of incident precipitation [*Lloyd and Marques*, 1988], whereas a range of 0 to 200% is representative for many tropical forests [*Lloyd and Marques*, 1988; *Holwerda et al.*, 2006]. Some forests, however, exhibit even more extreme distributions with a range between 0 and 1000% [*Cavelier et al.*, 1997]. Such skewed frequency distributions emerge because some leaf morphologies, e.g., leaves with drip tips, and canopy structures, e.g., drip points on inclined branches, favor the concentration of throughfall; consequently, other below-canopy areas must be drier. As to spatial structure expressed in terms of autocorrelation, the reported correlation lengths vary as widely as the few forest ecosystems investigated. Some researchers found no autocorrelation over the studied lag distances [*Loustau et al.*, 1992; *Bellot and Escarre*, 1998], whereas others detected spatial correlations in the range of 3 to 10 m [*Möttönen et al.*, 1999; *Keim et al.*, 2005; *Staelens et al.*, 2006]. Surprisingly, *Loescher et al.* [2002] detected autocorrelation of throughfall measurements over a distance of 43 m in a wet tropical forest, which they attributed to canopy gaps and individual tree crowns. It is tempting to interpret this broad range of recorded autocorrelation patterns as a reflection of differences in forest structure. And yet, the high variability of throughfall at small spatial scales should focus attention on the appropriateness of the design [*Skøien and Blöschl*, 2006] and the data analytical tools [*Zimmermann et al.*, 2008b] on which those correlation lengths depend.

[4] Temporal persistence seems to be required to initiate moisture gradients strong and persistent enough to trigger biotic and abiotic responses. The results of the few studies

[1]Institute of Geoecology, University of Potsdam, Potsdam, Germany.
[2]Smithsonian Tropical Research Institute, Balboa, Ancón, Panama.

0043-1397/09/2008WR007470$09.00

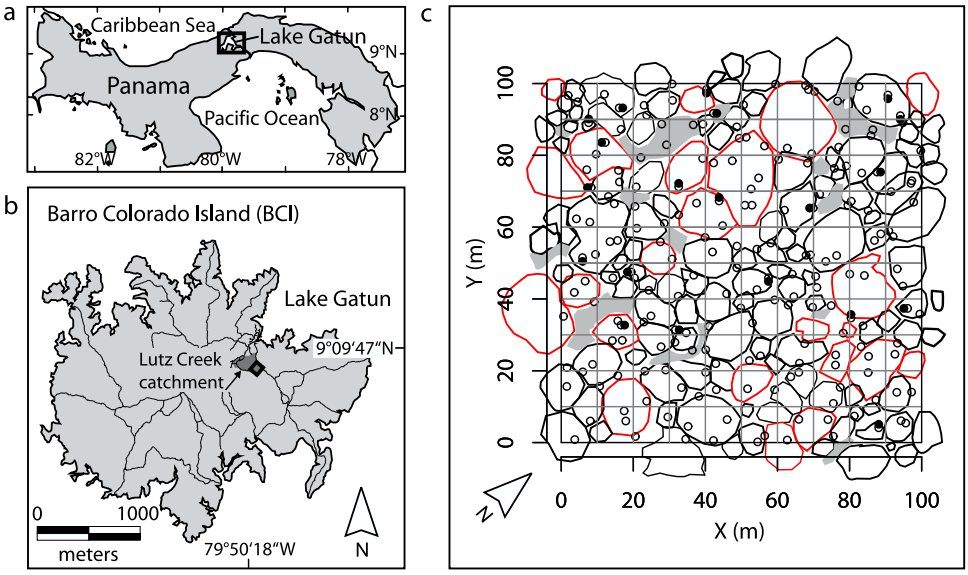

Figure 1. Location of the research area in (a) Panama and (b) Barro Colorado Island. The lines refer to the trail system and the square indicates the location of the sampling area. (c) The close-up view shows the structure of the tree crowns in the sampling area, red outlines refer to trees that are dry season deciduous, and gray areas indicate locations of small-scale disturbances due to tree and branch fall. Areas that do not belong to any of the former categories contain trees with indistinguishable crowns (e.g., due to lianas). The map is based on an aerial photograph from April 2008 and ground survey data (for details on the crown photograph, we refer to the work of *Jansen et al.* [2008]). Locations of throughfall sampling points and the square subplots are also shown; open circles refer to locations that were chosen according to the design-based component, whereas solid circles mark locations selected by the model-based component of the applied sampling scheme.

dealing with the temporal persistence of throughfall patterns vary from no persistent between-storm throughfall patterns [*Lin et al.*, 1997] to stability over several months [*Raat et al.*, 2002; *Staelens et al.*, 2006; *Zimmermann et al.*, 2008a].

[5] This brief overview points to rather diverging results, and the lack of explanations for observed throughfall autocorrelation structures and temporal persistence, despite their relevance for a range of biotic and abiotic processes [*Keim et al.*, 2005], motivated us to seek answers to the following research questions:

[6] 1. Do putative spatial and temporal patterns of throughfall reflect a natural phenomenon or the choice of variogram estimators?

[7] 2. Does the redistribution of rain in a heterogeneous tropical rain forest canopy result in a detectable throughfall autocorrelation structure over the studied lag distances?

[8] 3. Do throughfall measurements show temporal persistence?

[9] To answer these questions is not only important for several aspects in hydrology but may equally influence the work of biogeochemists and ecologists as near surface water in forests controls key chemical reactions (e.g., nitrification-denitrification rates [*Raat et al.*, 2002]) and plant growth (e.g., survival of seedlings [*Engelbrecht and Kursar*, 2003]).

2. Materials and Methods

2.1. Site Description

[10] We investigated throughfall patterns in a 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 main land during the creation of Lake Gatun, which is part of the Panama Canal. The area is characterized by a rough topography with steep slopes of up to 40° [*Dietrich et al.*, 1982]. The 1-ha plot, however, is situated on a more gentle northwest facing hillside with an inclination of 13.1 ± 7.3° (mean ± 1sd, 10 by 5 m grid, n = 210). The climate of Barro Colorado 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]. Deciduousness in the area is a complex phenomenon; some trees can be leafless for several weeks to months every year, while others drop their leaves only in particular dry years [*Foster and Brokaw*, 1982]. Because a few tree species lose their leaves in June–July, deciduousness is not a mere dry season phenomenon, which adds another aspect to the within-canopy dynamics. 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. The stand characteristics in our plot (Table 1) closely resemble data of other tree censuses in the area [*Thorington et al.*, 1982]. Canopy openness is an exception in that our measurements show a higher variability compared to another study from Barro Colorado Island in which canopy gaps were excluded [*Harms et al.*, 2004].

Table 1. Stand Characteristics of the 1-ha Study Area[a]

Number of stems, 1 cm ≤ dbh < 5 cm	1779
Number of stems, 5 cm ≤ dbh < 10 cm	658
Number of stems, 10 cm ≤ dbh < 20 cm	278
Number of stems, 20 cm ≤ dbh < 40 cm	143
Number of stems, 40 cm ≤ dbh < 60 cm	37
Number of stems, dbh ≥ 60 cm[b]	24
Number of tree species, dbh ≥ 5 cm	98
Basal area, dbh ≥ 5 cm, (m^2)	35.3
Canopy tree crown diameter (mean ± 1sd, m, n = 131)	9.7 ± 4.5
Canopy openness (median ± MAD, %, n = 200)	3.0 ± 4.1

[a]Abbreviations: dbh, diameter at breast height; sd, standard deviation; MAD, median absolute deviation.
[b]The maximum dbh is 130 cm.

2.2. Sampling and Instrumentation

[11] Our throughfall sampling scheme comprises a design-based and a model-based sampling component (Figure 1c). 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 10 m side length. In each of these subplots, which represent the strata, we randomly allocated two throughfall sampling points. We then chose 20 of these sampling points at random, and selected an additional sampling location 1 m away in a random direction. This model-based component increased the number of sampling points at short lag distances, which is important for estimating the shape of the variogram model near the origin.

[12] Throughfall samples were collected on an event basis from August 2007 to October 2008 (n = 91 events; Appendix A, Table A1). The sampling covered both peaks of the 2007 and 2008 wet season and the dry and subsequent transition season in 2008. Events were separated by at least two hours without rain and had to accumulate at least 0.6 mm of rainfall with a minimum mean intensity of 1.2 mm h^{-1}. According to this definition 198 events occurred during the study period. Logistical constraints and small scale variability of rainfall did not permit the sampling of all events. Rainfall was continuously recorded with two Hobo® tipping bucket rain gauges (orifice of 182 cm^2, 0.2 mm tip resolution); additionally, we used 5 manual read out collectors (orifice of 113 cm^2). All rainfall was recorded in a clearing 300 m from the throughfall sampling site. The throughfall collectors (n = 220) consisted of a 2-l polyethylene sampling bottle and a funnel. The receiving area of each collector was 113 cm^2; hence, total sampling area in the 1-ha plot summed up to 2.49 m^2. A polyethylene net with 0.5 mm mesh width on the bottom of the funnel prevented measuring errors due to organic material and small animals.

2.3. Measurement of Canopy Openness

[13] In order to facilitate the interpretation of throughfall data we measured canopy openness at all design-based throughfall sampling positions (n = 200). These measurements are based on hemispherical photographs, which we analyzed with Gap Light Analyzer 2.0 [*Frazer et al.*, 1999]. We used the full spectral resolution (RGB images) because the use of all three bands results in more discernable detail in sunlit image areas [*Jonckheere et al.*, 2005]. The hemispherical photographs were acquired using a Nikkon Coolpix 4500 digital camera with a Nikkon FC-E8 0.21x fish-eye lens. The camera was mounted on a tripod in approximately 0.5 m height and leveled horizontally. We performed all photographs under overcast conditions to minimize the anisotropy of the sky radiance. All photographs were taken between the 21st and 24th September 2007. For each of the photographs we determined the canopy openness of 6 different zenith angles: 1.9°, 3.8°, 5.0°, 7.5°, 10° and 20°. We then determined which of the different image sections correlated best with our throughfall measurements. For this analysis we used only events that were sampled in a time frame of less than a month from the date of canopy photography to minimize bias due to changing canopy structures. We found that the angle of 5.0° showed the best results; hence we used canopy openness data of this image section for further calculations.

2.4. Analysis of Throughfall Data

2.4.1. Software

[14] For all calculations, we used both MatLab® and the language and environment of R, version 2.2.6. [*R Development Core Team*, 2004], and here primarily the libraries geoR [*Ribeiro and Diggle*, 2001] and gstat [*Pebesma*, 2004]. For calculating experimental variograms with the estimator proposed by *Genton* [1998], we applied the Fortran code of *Rousseeuw and Croux* [1993].

2.4.2. Data

[15] For the analysis of the spatial correlation of throughfall measurements we used the first 60 events of the study period. We chose this subset of events because it comprises all seasonal changes of the canopy. We assumed that in case of a similar autocorrelation structure for those 60 events the analysis of further events would not provide new insights (see section 3.2.2). The data set for spatial analysis consists of 13200 observations of which 44 (0.33%) are missing values.

[16] The analysis of the temporal correlation of throughfall measurements is based on the whole data set (n = 91 events) which contains 20020 observations including 94 (0.47%) missing values. The inclusion of all 91 events was necessary to capture the temporal autocorrelation structure (see section 3.3.2). The analysis of temporal correlations requires data that is independent of event size; therefore we used standardized throughfall, $\tilde{T}$, which we calculated as:

$$\tilde{T}_{c_i,E_i} = \left(\frac{T_{c_i,E_i} - median(T_{E_i})}{MAD(T_{E_i})} \right) \quad (1)$$

where T_{c_i,E_i}, i = 1, 2, . . ., n denotes throughfall at collector c_i and event E_i, i = 1, 2, . . ., n; *median* (T_{Ei}) is the median throughfall of all collectors during event E_i, and *MAD* (T_{Ei}) is the median absolute deviation from the median based on all collectors during event E_i.

2.4.3. Models to Describe Spatial and Temporal Correlation

[17] Our objective is to obtain variogram models that describe the spatial and temporal dependence among throughfall measurements. The calculation of spatial and temporal variograms differs insofar as our analysis of spatial data is based on omnidirectional variograms, whereas we used a directional variogram to describe the temporal correlation of throughfall measurements (Figure 2). Apart from this difference all steps in the exploratory and geostatistical analysis are very similar, thus we do not describe the analytical procedure separately.

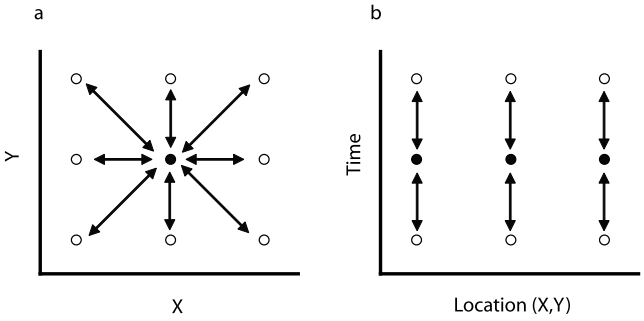

Figure 2. Schematic illustration of the approach for variogram modeling in (a) space and (b) time. We used omnidirectional variograms to investigate spatial patterns, whereas our temporal analysis is based on directional variograms. For reasons of clarity we show only lag pairs (arrows) of selected data (solid points).

[18] The variogram models have to take into account that throughfall data can be asymmetric. In general, we can distinguish two cases of asymmetry [*Lark*, 2000; *Kerry and Oliver*, 2007a, 2007b]. First, asymmetry can arise from a long tail of values in the underlying (primary) process. Second, data with outliers might be regarded as the superposition of two distinct processes in which outliers that belong to another (secondary) process contaminate the underlying distribution. In both cases, the overall data distribution may appear strongly skewed.

[19] Since the effect of a skewed underlying distribution on the variogram differs from that of single outliers, data processing and variogram modeling should account for these differences [*Kerry and Oliver*, 2007a, 2007b]. In case of a skewed underlying distribution, it is recommended to transform the data, whereas the presence of outliers in the data should be tackled either by their removal or the use of robust variogram estimators [e.g., *McBratney and Webster*, 1986; *Lark*, 2000; *Kerry and Oliver*, 2007a, 2007b].

[20] In throughfall data sets we expect to find large outliers in the spatial but also in the temporal domain. Outliers in the spatial domain occur due to sampling locations beneath drip points of leaves or inclined stems. In the temporal domain we expect these points to be responsible for outlying values because drip points are not active during all events; hence, it is likely that these points show large temporal fluctuations with throughfall magnitude. Since outliers in both domains might not approach the center of the distribution even after transformation, we have to decide whether or not to remove them. In practice, the separation of throughfall data according to an underlying and a contaminating process is not straightforward. On the one hand, throughfall occurs locally concentrated as drip points which appear extreme compared to neighboring sampling locations. On the other hand, there is no justifiable threshold to separate drip points from the underlying process. In such cases, where the removal of contaminants, i.e., outliers, is not possible without risking the exclusion of genuine observations, and where the spatial statistics of the underlying process is of interest, robust techniques should be considered to estimate the variogram [*Lark*, 2000]. We therefore use robust variogram estimators in addition to *Matheron*'s [1962] estimator, and we assess the results by means of cross validation as proposed by *Lark* [2000] and *Kerry and Oliver* [2007a, 2007b]. Since departure from normality of the underlying distribution may introduce bias into robust estimators [*Lark*, 2000], we perform a comprehensive exploratory data analysis and transform data if necessary prior to variogram estimation.

2.4.4. Exploratory Data Analysis

[21] Our exploratory data analysis and the following geostatistical analysis consist of several steps. To provide a brief overview of our data analytical methods we summarize important steps and decisions in a flowchart (Figure 3). For exploratory data analysis we first visualized the univariate data distribution by diagnostic plots (histograms, quantile-quantile plots, box plots), and calculated the coefficient of skewness. To determine the need for transformation, we used the octile skew [*Brys et al.*, 2003], denoted "skew_8," which is a measure of the asymmetry of the first (O_1) and seventh octile (O_7) of the data about the median:

$$skew_8 = \frac{(O_7 - median) - (median - O_1)}{(O_7 - O_1)}. \quad (2)$$

[22] The advantage of the octile skew compared to the conventional coefficient of skew lies in its insensitivity to extreme values. We used this robust approach to accommodate any outliers that may inflate the coefficient of skewness even though the underlying distribution has a Gaussian shape. Since in this case the skewness cannot be removed by data transformation [*Kerry and Oliver*, 2007b], we wanted to transform our data only if the underlying distribution departs from normality. A rule of thumb [*Rawlins et al.*, 2005] suggests to transform the data if the octile skew is larger than 0.2 or smaller than −0.2. We applied this rule and transformed the data if necessary to square roots or to common logarithms ($\log_{10}$). Since the spatial data contained zero values and the temporal data comprised negative numbers (equation (1)) we added a constant of 1 and 3 to the data, respectively, to make all values just positive prior to transformation.

[23] In order to explore in detail the influence of extreme values on the experimental variograms, we also displayed the bivariate data distribution by means of **h**-scattergrams [*Webster and Oliver*, 2001]. These are scatterplots of point pairs separated by a fixed distance, which are produced for a number of lag classes. Outliers may appear in some, but not necessarily all, of those classes. Where they occur, nonrobust semivariance estimates are potentially too large. Besides this benefit of **h**-scattergrams in displaying extreme values, they can also be used to assess the plausibility of the effective range of a variogram model [*Zimmermann et al.*, 2008b].

[24] Since a geostatistical analysis requires second-order stationarity, we produced diagnostic plots (plots of the data, which were divided into quintiles; plots of the data versus the coordinates) to explore the data for nonstationarity of the mean that may be caused by local trends according to

$$z(\mathbf{x}) = \mu(\mathbf{x}) + \varepsilon(\mathbf{x}), \quad (3)$$

where $z(\mathbf{x})$ is the observed variable at location $\mathbf{x}$, $\mu(\mathbf{x})$ is the local mean, i.e., it represents a deterministic drift of the variable at location $\mathbf{x}$, and $\varepsilon(\mathbf{x})$ is the random component at location $\mathbf{x}$ that should be normally distributed with zero mean and that satisfies the second-order stationarity.

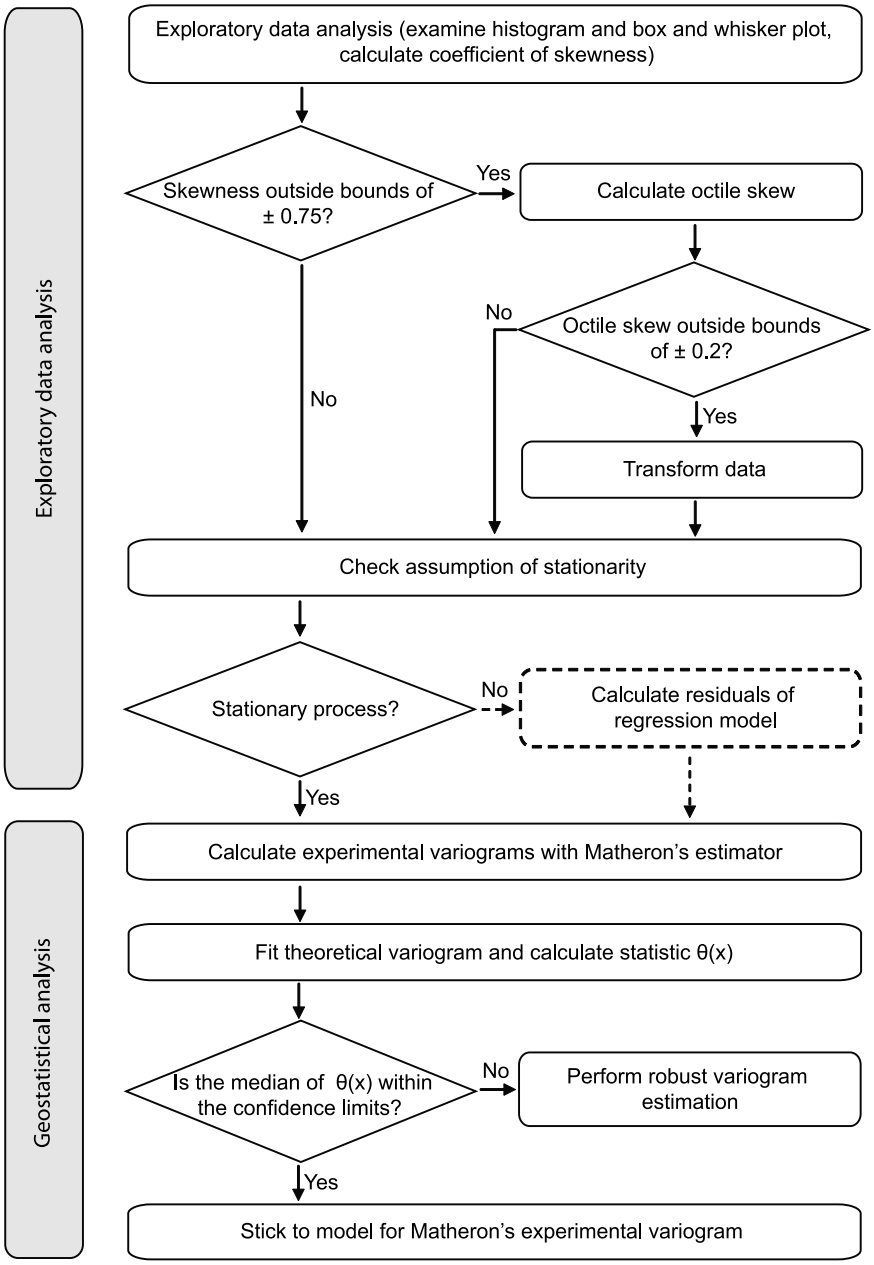

Figure 3. Summary of main steps and decisions involved in analyzing our throughfall data, which incorporates the recommendations for analyzing spatial data as given by *Kerry and Oliver* [2007a] and *Zimmermann et al.* [2008b]. The dashed box highlights an operation not applicable to our data, i.e., the assumption of second-order stationarity was valid for our throughfall data sets. If Matheron's estimator performs satisfactorily in studies of spatial structure, we recommend estimating the variogram parameters by residual maximum likelihood, particularly in the framework of spatial prediction.

2.4.5. Geostatistical Analysis

[25] First we calculated experimental variograms using the estimator by *Matheron* [1962]:

$$2\hat{\gamma}_M(\mathbf{h}) = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}\{z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})\}^2, \qquad (4)$$

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 $\mathbf{h}$. To ensure that the selection of lag classes does not unduly influence the variogram, we used regular lag classes. That is, for the spatial analysis we calculated the semivariance at 2-m lags with a lag tolerance of 0.99 m over half the maximum separation distance; we only used a semivariance estimate if its number of contributing point pairs was at least 30. For the temporal analysis we calculated the semivariance at 1-day lags, but in contrast to the spatial analysis, we used semivariance estimates to a distance of up to 70% of the temporal extent because of the continuous large number of point pairs in these lag classes (>1700 point pairs). Four variogram models (exponential, Gaussian, spherical, pure

nugget) were fitted to the experimental variograms by ordinary least squares, and we selected the model which had the minimum average sum of squares from the fit.

[26] Observations that qualify as outliers potentially cause an overestimation of the sill variance. *Lark* [2000] showed how cross validation (or a validation subset) helps to assess the applicability of the standard variogram estimator by using a statistic $\theta(\mathbf{x})$:

$$\theta(\mathbf{x}) = \frac{\{z(\mathbf{x}) - \hat{Z}(\mathbf{x})\}^2}{\sigma^2_{\mathbf{K},\mathbf{x}}}, \quad (5)$$

where $z(\mathbf{x})$ is the observed value at location $\mathbf{x}$, $\hat{Z}(\mathbf{x})$ is the kriged estimate and $\sigma^2_{\mathbf{K},\mathbf{x}}$ the kriging variance. If kriging errors follow a Gaussian distribution, $\theta(\mathbf{x})$ will be distributed as χ^2 with one degree of freedom. Since the median of the standard χ^2 distribution with one degree of freedom is 0.455, the median of $\theta(\mathbf{x})$ is also 0.455 when a variogram appropriate for interpolating intrinsic data is used. A sample median significantly ($\alpha = 0.05$) 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 $\theta(\mathbf{x})$ for our data sets, we proceeded as follows. First, we performed 1000 unconditional simulations to predict the values of throughfall magnitude. Second, we used cross validation to assess each simulation using the statistic $\theta(\mathbf{x})$; that is to say we computed the median of $\theta(\mathbf{x})$. Third, we determined the 2.5% and 97.5% percentiles of the $\theta(\mathbf{x})$-median distribution to approximate 95% confidence limits. For data sets whose 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].The Cressie-Hawkins' estimator is given by:

$$2\hat{\gamma}_{CH}(\mathbf{h}) = \frac{\left\{\frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}|z(\mathbf{x}_i) - z(\mathbf{x}_i+\mathbf{h})|^{\frac{1}{2}}\right\}^4}{0.457 + \frac{0.494}{N(\mathbf{h})} + \frac{0.045}{N^2(\mathbf{h})}}. \quad (6)$$

The Dowd estimator is

$$2\hat{\gamma}_D(\mathbf{h}) = 2.198\{median(|y_i(\mathbf{h})|)\}^2, \quad (7)$$

where $y_i(\mathbf{h}) = z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})$, $i = 1, 2, \ldots, N(\mathbf{h})$; and Genton's estimator is given by

$$2\hat{\gamma}_G(\mathbf{h}) = \left(2.219\{|y_i(\mathbf{h}) - y_j(\mathbf{h})|; i<j\}\binom{H}{2}\right)^2, \quad (8)$$

with $y_i(\mathbf{h})$ defined as for equation (7) and H = integer part $(n/2) + 1$, $n = N(\mathbf{h})$.

[27] Because the robust estimators listed above differ in their susceptibility to skewness in the underlying distributions, and particularly in their resistance to outliers, *Lark* [2000] recommended the same cross-validation procedure as described above to choose among them. He suggested to use the estimator with the $\theta(\mathbf{x})$ median closest to the expectation of 0.455, and to use the efficient Genton's estimator if none of the robust estimators are markedly better by this criterion. We adopted this recommendation; for an in-depth discussion of robust variogram estimation we refer to the work of *Lark* [2000].

3. Results

3.1. Event Characteristics

[28] During the 14-month study period we sampled 91 events, which yielded 1089 mm of throughfall (Appendix A, Table A1). These 91 events vary greatly with respect to throughfall magnitude (0.1–77.4 mm), storm intensity (maximum 30-minute rainfall intensity between 0.6 and 74.8 mm h^{-1}), and antecedent wetness conditions (antecedent rainfall of previous 3 days between 0 and 88.4 mm). Since our data set not only comprises 46% of all rain events during the study period but also covers consecutive wet seasons and one dry season, we assume that our data reflects the variety of meteorological characteristics typical for the research area.

3.2. Spatial Analysis

3.2.1. Exploratory Data Analysis

[29] For the sampled events, a correlation exists between the magnitude of an event and the octile skew (Figure 4a) that decreases asymptotically with increasing event size. As a consequence, the shape of the underlying distribution of throughfall magnitude, which is taken into account when calculating the octile skew (equation (2)), changes from strongly right skewed for small events to Gaussian for large rain storms (Figure 4b). This shift of frequency distributions coincides with a decline of spatial variability with increasing throughfall magnitude, as the median absolute deviation (MAD)/median ratio depends on median throughfall magnitude (Figure 4a, inlet; Appendix A, Table A1).

[30] A closer examination of our data (Figure 4a; section A, Table A1) reveals the processes that produce distinct shapes of throughfall frequency distributions. Small events often exhibit both a high octile skew (>0.2) and a high skew (>1.5); hence, their skewness is produced by an actually skewed underlying distribution, and not only by single outliers. We attribute the underlying skewness to the patchiness of throughfall as a function of vegetation density; that is to say, many outliers in small events originate from spots with high canopy openness (Figure 5). Canopy openness, however, does not explain all outlying values. That is, during some small events we observed drip from big leaves, which we consider responsible for highly outlying values regardless of canopy closure above the sampling point (Figure 5b). Many larger events belong to another type of frequency distribution (Figure 4b, example C) that shows a low octile skew (<0.2) but a high skewness (>1.5). Their high skewness results from a few outliers far from the centre of the data. Our field observations indicate that some of these outliers originate from stemflow drip points and leaf drip tips, which concentrate water to amounts of up to 1000% of incident precipitation. The low underlying skew of large events indicates the diminishing influence of canopy openness on the throughfall frequency distribution. The correlation of canopy openness and throughfall magnitudes corroborates this observation in that the importance of canopy openness as a predictor of throughfall magnitude decreases with increasing event size (Figure 6).

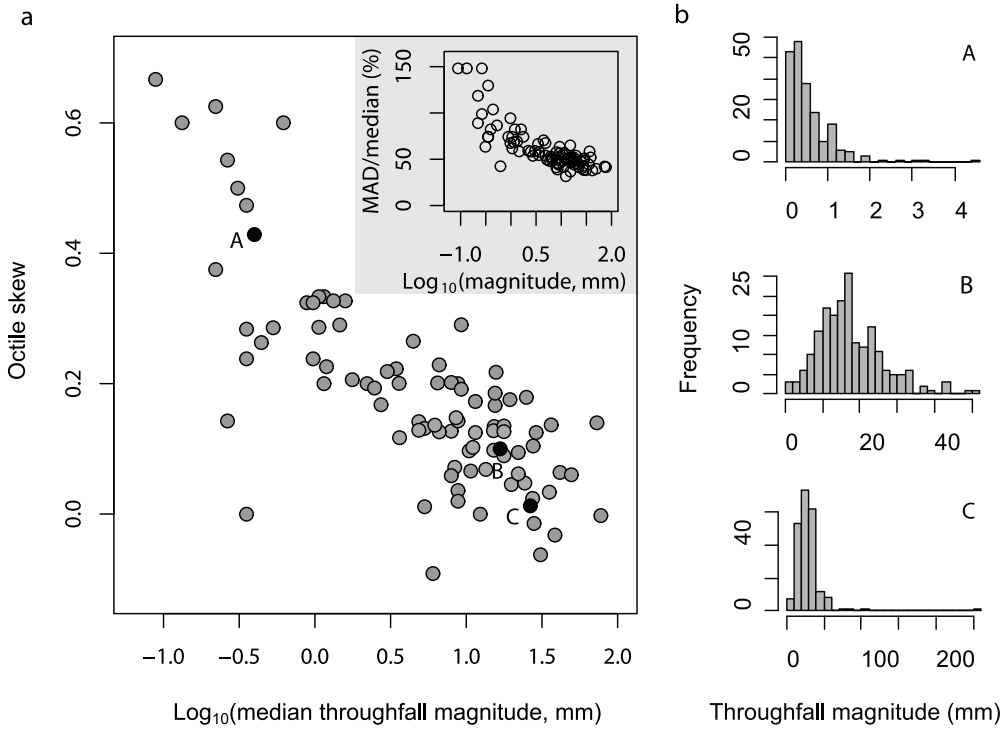

Figure 4. (a) The octile skew as a function of median throughfall magnitude. Octile skew and throughfall magnitude are nonlinearly correlated (Spearman rank correlation coefficient, $\rho = -0.73$, $p < 0.001$, $n = 91$). The black circles represent events along a gradient of throughfall magnitudes, whose frequency distributions are shown at the right-hand side of this figure. (b) The frequency distributions of examples A (event 43), B (event 22), and C (event 16) exhibit distinct octile skews (A, high; B and C, low) and skews (A and C, high; B, low). The inlet shows the MAD (median absolute deviation)/median throughfall ratio as a function of the median throughfall.

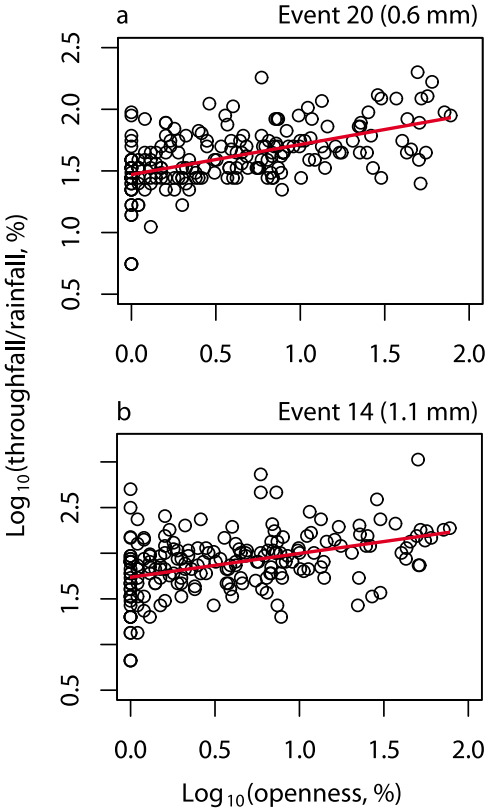

Figure 5. Exemplary relationships between canopy openness (%) and relative throughfall (throughfall/rainfall, %) for two small events, (a) event 20 and (b) event 14. The lines indicate linear regression models (event 20: $r^2 = 0.29$, $p < 0.001$, $n = 200$; event 14: $r^2 = 0.17$, $p < 0.001$, $n = 200$).

[31] Even though drip points clearly distort the univariate distribution, their influence on the experimental variogram depends on the lag classes in which they occur as can be illustrated with two examples (Figure 7). Event 22, which is slightly skewed (Appendix A, Table A1), displays no outliers in its bivariate distribution, whereas event 16 comprises one large outlier. This gets conspicuous at lag distances of more than 4 m (Figure 7); hence, we expect the experimental variogram to be sensitive to this value at the 4–8 m lag. Event 16 was not transformed because the octile skew was below 0.2; however, even after a $\log_{10}$ transformation, outliers were still present in the data as revealed by diagnostic plots. The latter example shows that the octile skew in general successfully indicates transformation needs. This statistical criterion, however, cannot replace a careful examination of the data, and in some cases a transformation might be reasonable even though the octile skew is below 0.2 or above −0.2 (e.g., event 50; Appendix A, Table A1).

3.2.2. Geostatistical Analysis

[32] According to the $\theta(\mathbf{x})$ statistic [*Lark*, 2000], the nonrobust Matheron estimator (equation (4)) overestimated the semivariance in 55% of the analyzed data sets (n = 60 events) as a consequence of outliers in the data (Appendix A, Table A2). In these cases, we applied robust estimators and found that none of them behaved superior, though Dowds' and Gentons' estimator were often closer to the expectation of the median of $\theta(\mathbf{x})$ than the Cressie-Hawkins estimator.

[33] When we computed the semivariance for each pair of points, x_i and x_j as

$$\gamma(x_i, x_j) = \frac{1}{2}\{z(\mathbf{x}_i) - z(\mathbf{x}_j)\}^2, \tag{9}$$

and plotted these values as boxplots against lag distance classes, which produces a graph called the "variogram cloud" [*Webster and Oliver*, 2001] (Figure 8a), it became obvious that a single extreme outlier has a strong influence on the shape of that cloud. In our example (event 16), the first extremely large semivariance values occur around the 8–10 m lag (Figure 8a); not surprisingly, the Matheron experimental variogram starts to overestimate the semivariance at precisely the same distance classes (Figure 8b). The resulting seemingly strong autocorrelation structure vanishes when experimental variograms are based on robust estimators (Figure 8b). A transformation of data sets that display extreme outliers (Figure 4, example C) appears to be only partly successful. Transformations reduce the influence of outliers on the semivariance cloud (Figure 8c), and nonrobust estimates of the experimental variogram do not produce any artificial autocorrelation structure. According to the $\theta(\mathbf{x})$ statistic [*Lark*, 2000], however, Matheron's estimator still overestimates the semivariance, which is reflected in the considerable difference between sill variances calculated with nonrobust and robust variogram estimators (Figure 8d).

[34] The variogram analysis for the subset of 60 events showed a consistent pattern, that is to say, throughfall displayed either weak or no spatial autocorrelations over the studied lag distances (Appendix A, Table A2; Figure 9). Since all 60 events exhibited similar throughfall autocorrelation patterns, we rule out any influence of deciduousness on these patterns.

3.3. Temporal Analysis

3.3.1. Exploratory Data Analysis

[35] Similar to the spatial data, the temporal data set shows large outlying values. This is not surprising in view of our observations that drip points are not active during all events. Their irregular activation results in large temporal fluctua-

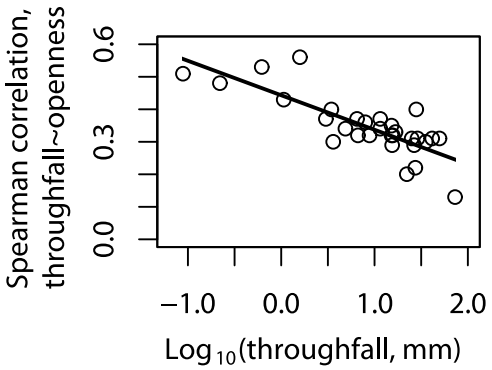

Figure 6. Canopy openness as a predictor of throughfall magnitude (as Spearman rank correlation coefficients) plotted against throughfall magnitude. The line indicates a linear regression model ($r^2 = 0.64$, $p < 0.001$, $n = 30$). Note: only events that were sampled in a time frame of less than a month from the date of canopy photography were used for this analysis.

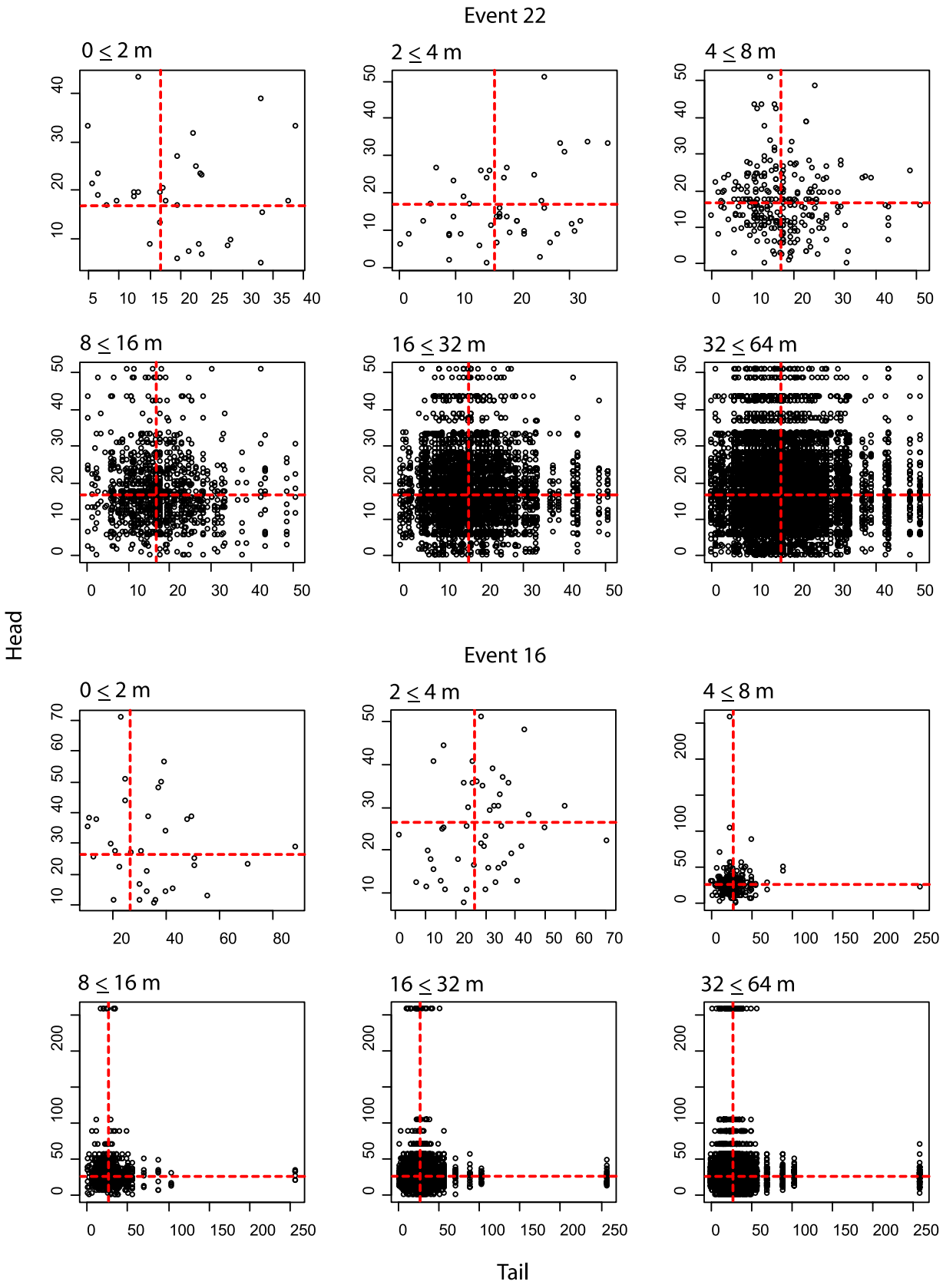

Figure 7. Examples of **h**-scattergrams that reveal contrasting bivariate data distributions (events 22 and 16) for selected lag classes. The plots show all pairs of throughfall measurements at locations **x** separated by a certain lag-distance class (shown at the top of each graph); the value at the start of the distance vector **h**, $z(\mathbf{x})$, is called the tail value, and the value at the end of the distance vector, $z(\mathbf{x} + \mathbf{h})$, is the head value. The vertical and horizontal dashed lines correspond to the population median.

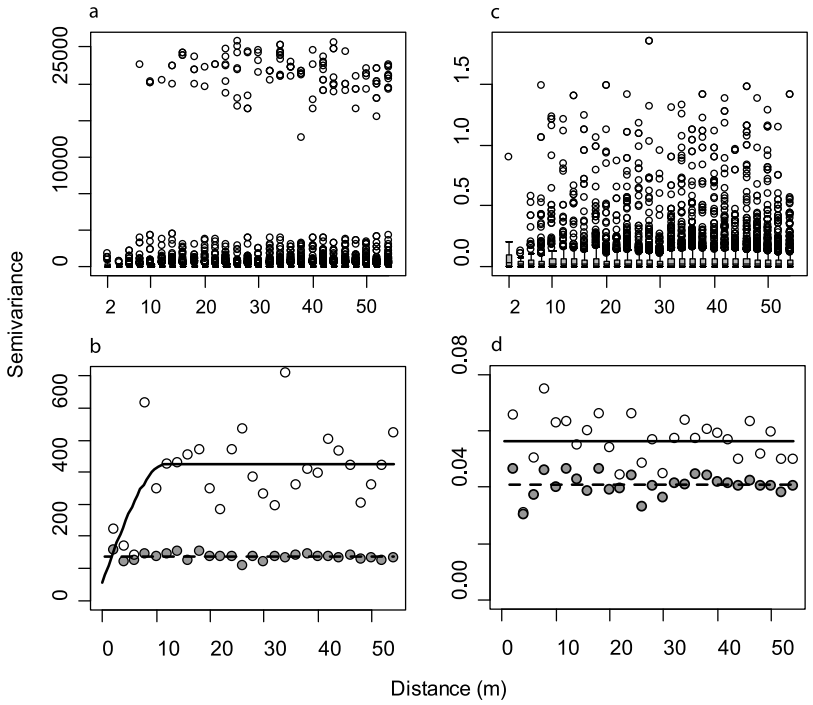

Figure 8. Boxplots of semivariances for all point pairs (equation (9)) plotted against lag distance classes for two examples: (a) raw data of event 16 and (c) log-transformed data of the same event. (b and d) The corresponding experimental and theoretical variograms based on the nonrobust Matheron estimator (equation (4)) (open circles, solid line, respectively) and the robust estimator due to Genton (equation (8)) (solid circles, dashed line, respectively) are shown below.

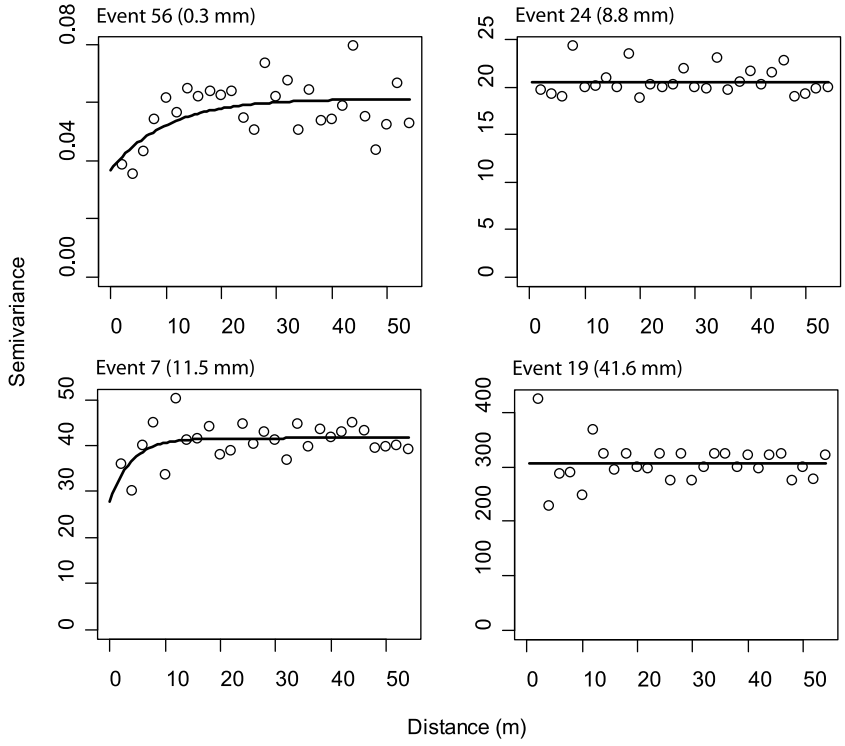

Figure 9. Exemplary variograms across a range of throughfall magnitudes. (left) Events with a weak throughfall autocorrelation structure; (right) variograms with a pure nugget structure. The latter case was detected in 82% of all events.

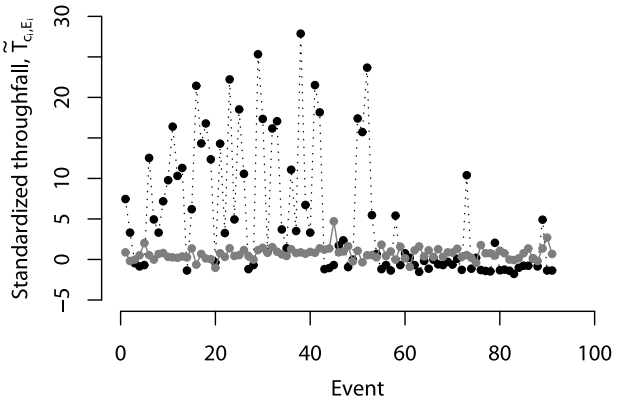

Figure 10. Fluctuations of the standardized throughfall, $\tilde{T}_{c_i,E_j}$, through time at two selected sampling locations. The black dots and the dotted line illustrate the temporal behavior of a drip point, whereas the gray dots and the solid line refer to a sampling location where no drip was detected. Note the disappearance of the drip point after event 60.

tions of the standardized throughfall (Figure 10). Since these outlying values result in both a large skewness (5.88) and a large octile skew (0.2), we $\log_{10}$-transformed the data for further analysis. These data still show some outlying values (skew: 0.32; octile skew: 0.01). They are, however, not particularly numerous; for instance, if we considered all values beyond 1.5 and 3 times the interquartile range from the upper quartile as outliers, the rate of contamination would be 0.0146 and 0.0022 (Figure 11).

[36] To investigate whether our data show local trends we first produced several diagnostic plots (e.g., plots of the data versus the time of sampling) which displayed no trend in the data. Furthermore, we checked if correlations of throughfall measurements between events show similar trends over time regardless of the season. For this analysis we compared events sampled during the transition from dry to wet season (the time of leaf flush) with events that occurred during the peak of the rainy season 2007. This comparison is based on

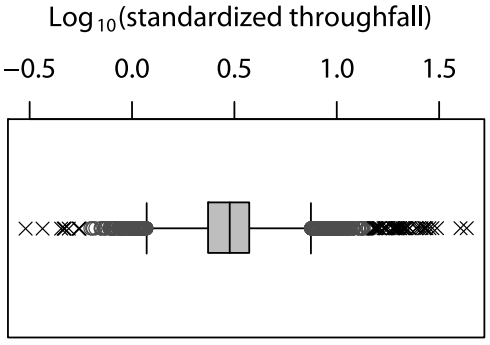

Figure 11. Box-and-whisker plot of the throughfall data used for estimating the temporal variogram ($n = 19926$). The box contains the middle 50% of the data, the horizontal line in this box refers to the median. The whiskers illustrate the distance of 1.5 times the interquartile range from the quartiles. Data points beyond this distance are shown as gray circles, and crosses indicate values with a distance of three times the interquartile range from the quartiles.

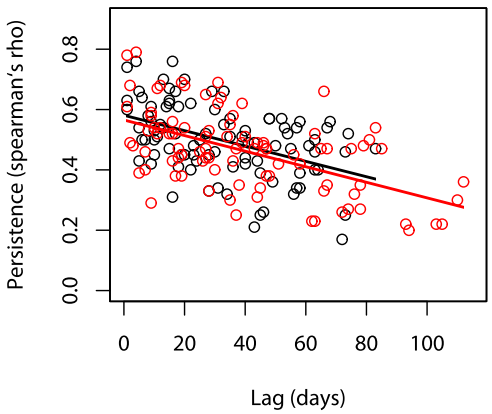

Figure 12. The persistence of throughfall measurements between event pairs ($n = 14$ events, $n = 91$ pairs) as a function of the temporal lag between events for the peak of the rainy season 2007 (black circles) and for the transition from dry to wet season 2008 (red circles). The persistence is expressed with Spearman rank correlation coefficients ρ. The lines indicate linear regression models, the black line refers to the 2007 data ($r^2 = 0.2$, $p < 0.001$, $n = 91$), whereas the red line relates to the 2008 data ($r^2 = 0.31$, $p < 0.001$, $n = 91$).

14 events in both periods of similar throughfall magnitude (Wilcoxon rank sum test, p = 0.98, n = 14). Surprisingly, we did not detect any differences between measurements of the two sampling intervals (Figure 12). This result indicates that deciduousness does not influence temporal correlations of throughfall measurements in our sampling area.

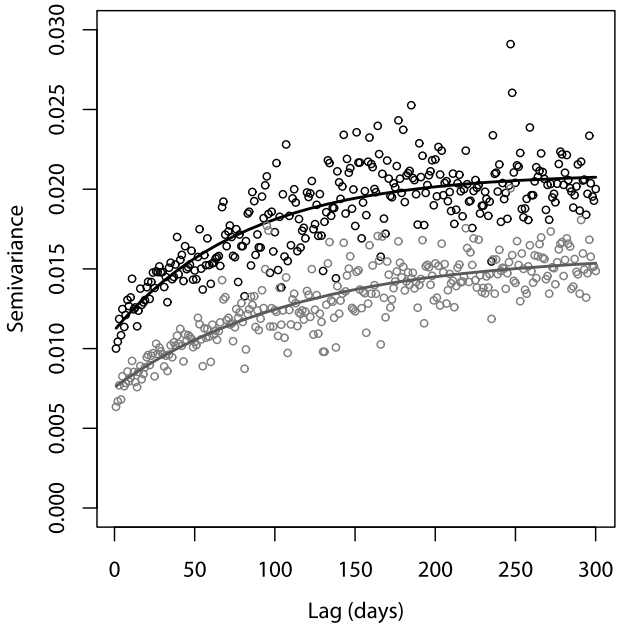

Figure 13. Experimental and theoretical variograms for the temporal data set ($n = 91$ events, $n = 19926$ observations) based on the nonrobust Matheron estimator (equation (4)) (black circles, black line, respectively), and the robust estimator due to Dowd (equation (7)) (gray circles, gray line, respectively). Details of the theoretical variogram model for the latter estimator are given in Table 2.

Table 2. Variogram Model Parameters for Temporal Data[a]

Model	Exponential
Nugget	0.008
Sill	0.016
Range (days)	115
Effective range[b] (days)	344

[a]Model fitted to the experimental variogram, which was estimated with Dowd's estimator.

[b]The effective range R_{eff} for an exponential model is calculated as R_{eff} = range*3.

3.3.2. Geostatistical Analysis

[37] Though the transformed data exhibit only a few outliers that skew the distribution rather modestly, the $\theta(\mathbf{x})$ statistic [*Lark*, 2000] indicated that the nonrobust Matheron estimator overestimated the semivariance (median of $\theta(\mathbf{x})$ = 0.303). Hence we applied robust estimators and found that *Dowd*'s [1984] estimator provided satisfactory results (median of $\theta(\mathbf{x})$ = 0.451). Interestingly, our results show that even a very low fraction of outlying values (<1.5% of the data) influences the sill variance of *Matheron*'s [1962] estimator; the shape of the variogram, however, displays almost no difference between nonrobustly and robustly estimated temporal variograms (Figure 13).

[38] Our data show that throughfall exhibits long-term persistence, that is, measurements at individual sampling points correlate over consecutive wet seasons (Figure 13, Table 2). The variogram indicates that these correlations vanish after one year. This long-term persistence preserves the spatial variability of throughfall through time which is reflected in large differences between dry and wet sampling points. The driest locations in our study area received only 16% of rainfall, whereas extreme wet spots accumulated up to 350% of rainfall during the 14-month sampling period.

4. Discussion

4.1. A Call for Robust Variogram Estimation Techniques

[39] The majority of throughfall studies report asymmetric frequency distributions [*Ford and Deans*, 1978; *Lloyd and Marques*, 1988; *Kostelnik et al.*, 1989; *Cavelier et al.*, 1997; *Bellot and Escarre*, 1998; *Hölscher et al.*, 1998; *Germer et al.*, 2006; *Holwerda et al.*, 2006]. Only a few studies in temperate forests detected Gaussian behavior of throughfall quantity [*Loustau et al.*, 1992; *Möttönen et al.*, 1999; *Keim et al.*, 2005]. Our data indicate that throughfall frequency distributions can be positively skewed for different reasons. Small events show a high skew (Figure 4b, Example A) because throughfall mainly consists of the free throughfall component (Figure 5a), which originates from rain falling in gaps without striking the canopy [*Gash*, 1979; *Loustau et al.*, 1992]. As soon as rainfall magnitude increases, free throughfall alone does not account for all outlying values (Figure 6) because water on leaves begins to coalescence and drip occurs (Figure 5b). Large positive skewness and a large spatial variability of small events have been detected by many studies regardless of forest type [*Bellot and Escarre*, 1998; *Staelens et al.*, 2006]. In contrast, many large events show a high skew (Figure 4b, Example C) due to a few outliers caused by drip points. *Lloyd and Marques* [1988] detected a similar structure in the throughfall frequency distribution of a Brazilian tropical forest and noted the influence of drip points. If the latter are absent, however, large events tend to show a low skewness and a relatively low spatial variability [*Loustau et al.*, 1992; *Bellot and Escarre*, 1998; *Staelens et al.*, 2006], which is also reflected in our data where large events have low octile skews (Figure 4a) and a low MAD/median throughfall ratio (Figure 4a, inlet).

[40] Our results and the above list of studies provide evidence that a positive skew of throughfall frequency distributions is the rule rather than the exception, particularly in tropical forests [*Lloyd and Marques*, 1988]. Interestingly, our data show that sampling locations which represent outliers in the spatial domain are also responsible for extreme values in the temporal domain because outlying values at these points do not occur in every event (Figure 10). Extreme values in both domains influence the nonrobust variogram estimator (spatial data: Figure 8; temporal data: Figure 13) because they cause an overestimation of the sill variance (Figures 8 and 13) and in some cases they even induce spurious autocorrelation structures (Figures 7 and 8). Theoretical studies of the influence of outliers on the variogram showed that the size of the data set is less important than the rate of contamination [*Kerry and Oliver*, 2007b]. Our results confirm these theoretical findings and provide evidence that even very low rates of contamination (<1.5% of the data) potentially distort nonrobust variogram estimates (Figures 8 and 13). This problem cannot be solved by common transformations, such as square root or $\log_{10}$ transformation (Figures 8 and 13), because some outliers in our throughfall data sets, e.g., those originating from drip points, are too far from the centre of the data. Therefore in many cases the use of robust variogram estimators to model spatial throughfall data appears inevitable if one wants to avoid the arbitrary decision of removing outliers prior to geostatistical analysis. It must be reemphasized, however, that an underlying Gaussian distribution is of particular importance for robust variogram estimation so as to avoid bias [*Lark*, 2000]. Our study demonstrates that different processes induce distinct throughfall frequency distributions and that data treatment has to account for these differences; that is, many small throughfall events have to be transformed to square roots or logarithms before geostatistical analysis, whereas large events often require robust variogram estimation.

[41] In the presence of outliers, we mainly chose the robust estimators proposed by *Dowd* [1984] (equation (7)) and *Genton* [1998] (equation (8)), a choice that is justified by the behavior of their influence functions [*Hampel et al.*, 1986]. The performance of these functions can be described by several measures; one such is the gross error sensitivity [*Genton*, 1998; *Lark*, 2000]. This measure indicates that only Dowd's and Genton's estimator are B-robust, that is, the estimators have a bounded asymptotic bias. Matheron's (equation (4)) and the Cressie-Hawkins' estimator (equation (6)), in contrast, are not B-robust because the effect of an outlying value $x = z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})$ is proportional to x^2 and $x^{0.5}$ over all x, respectively [*Lark*, 2000]. Another important measure of robustness is the breakdown point [*Hampel*, 1971]. Dowd's and Genton's estimator show a large robustness as their breakdown point is 0.5 (the maximum of any estimator), which means that half of the pair differences, y_i (**h**), may be replaced by arbitrary large or small values before the estimators $\hat{\gamma}_D$ (**h**) and $\hat{\gamma}_G$ (**h**) collapse [*Lark*, 2000]. In comparison, the Cressie-Hawkins estimator

Table 3. Comparison of Sampling Designs and Throughfall Autocorrelation Ranges From Studies Conducted in a Variety of Forest Ecosystems

Reference	Location	Forest Type	A_{dom}[a] (m^2)	N[b]	Support[c] (cm^2)	Spacing[d]	Effective Range[e] (m)
Bellot and Escarre [1998]	NE Spain	Holm-oak forest	950	50	?	4.4	pure nugget
Keim et al. [2005]	W USA	60-year-old conifer forest	225	94	9.2	1.5	5[f]
Keim et al. [2005]	W USA	old conifer forest	900	94	9.2	3.1	5[f]
Keim et al. [2005]	W USA	60-year-old deciduous forest	304	94	9.2	–	10[f]
Loescher et al. [2002]	Costa Rica	tropical rain forest	25600	56	95	21.4	43[g]
Loustau et al. [1992]	S France	18-year-old pine stand	2500	52	707	6.9	pure nugget
Möttönen et al. [1999]	E Finland	150- to 200-year-old Scots pine forest	10000	181	50	7.4	9
Shachnovich et al. [2008]	Israel	40-year-old pine forest	70	20	55	1.9	?
Staelens et al. [2006]	Belgium	85-year-old beech tree	225	50	158	2.1	3–4
Staelens et al. [2006]	Belgium	85-year-old beech tree	225	48	460	2.2	no stable sill
This study	Panama	tropical rain forest	10000	220	113	6.7	pure nugget[h]

[a]Size of sampling area.
[b]Number of sampling locations.
[c]Receiving area of collector.
[d]The spacing L_s is calculated according to *Skøien and Blöschl* [2006], that is, $L_s = (A_{dom}/N)^{0.5}$.
[e]The effective range R_{eff} is calculated as exponential model: R_{eff} = range*3, Gaussian model: R_{eff} = range*$3^{0.5}$, spherical model: R_{eff} = range.
[f]*Keim et al.* [2005] estimated the range visually from the experimental variogram, no variogram model was fitted.
[g]We report ranges as given by the authors; however, *Loescher et al.* [2002] might have reported the distance parameter of the covariance function instead of the effective range.
[h]We detected a pure nugget structure in 82% of the analyzed events, the remaining events showed weak structures (i.e., high nugget/sill ratios) with highly variable effective ranges.

is not robust by this commonly used criterion in robust statistics as its breakdown point is zero [*Genton*, 1998; *Lark*, 2000].

[42] To summarize, we wish to highlight two critical steps in a geostatistical investigation of throughfall data. First, a careful exploratory data analysis is necessary to determine transformation needs, and to detect outliers in the data set in order to understand their influence on the estimated covariance functions. In particular, the analysis of the bivariate distribution by means of **h**-scattergrams [*Webster and Oliver*, 2001] is a useful tool as it reveals in which lag classes outliers appear and hence, in which lag class the standard estimator (equation (4)) potentially overestimates the variogram (Figures 7 and 8). Second, the $\theta(\mathbf{x})$ statistic [*Lark*, 2000] is a valuable criterion to assess the applicability of the standard variogram estimator (equation (4)); furthermore, it can be used to eventually choose among robust estimators [*Lark*, 2000]. The choice of the variogram estimator depends on the characteristics of the data, because the difference between the Cressie-Hawkins' estimator [*Cressie and Hawkins*, 1980] and the other robust estimators, $\hat{\gamma}_D(\mathbf{h})$ and $\hat{\gamma}_G(\mathbf{h})$, is not always of practical importance as outliers in real data sets are not unbounded [*Lark*, 2000]. Throughfall data sets that contain extreme outlying values, however, may call for highly robust variogram estimators such as Dowd's [*Dowd*, 1984] and Genton's [*Genton*, 1998] estimator. Surprisingly, robust variogram estimation techniques, which have been widely used in hydrology and soil science [e.g., *Shouse et al.*, 1990; *Mohanty et al.*, 1991; *Woodbury and Sudicky*, 1991; *Lark*, 2002; *Sobieraj et al.*, 2004], have yet to be applied in studies that deal with throughfall patterns. Unfortunately, some throughfall studies [e.g., *Shachnovich et al.*, 2008] do not provide sufficient information on the uni- and bivariate distributions, which casts doubt on the choice of the nonrobust estimator (equation (4)).

4.2. An Attempt to Separate Spatial Patterns From Spatial Illusions: Insights From a Global Comparison

[43] A comparison of a variety of forest ecosystems reveals large differences of throughfall autocorrelation patterns (Table 3). Large variations of the sampling scale triplet (Table 3), which comprises spacing, support, and extent [sensu *Blöschl and Sivapalan*, 1995; *Blöschl*, 1999], do not exactly facilitate the interpretation of these patterns. Ideally, sampling and modeling scales are commensurate with the scales of variation of the variable of interest [*Blöschl*, 1999; *Skøien and Blöschl*, 2006]. Since most throughfall studies [*Loustau et al.*, 1992; *Bellot and Escarre*, 1998; *Keim et al.*, 2005; *Staelens et al.*, 2006] detected a large spatial variation over small scales, an adequate sampling scheme would involve a small spacing, small support, and a sufficiently large extent to cover typical forest structures. When sampling scales do not match the scale of the process in question, they will affect the estimated autocorrelation structure; for instance, large spacings led to overestimation of the correlation length [*Russo and Jury*, 1987; *Skøien and Blöschl*, 2006]. Therefore the speculation of *Loescher et al.* [2002] that "large tree canopies and gaps are the source of much of the spatial variance in throughfall volume" does not necessarily explain the observed range. It is more likely that the large throughfall correlation length of more than 40 m [*Loescher et al.*, 2002] reflects the large spacing of the sampling points (Table 3).

[44] *Webster and Oliver* [2001] demonstrated that experimental variograms based on fewer than 50 data are often erratic. They recommended using no fewer than 100, and ideally 150 observations, for a reliable variogram if the variation is isotropic. We support this recommendation and suggest analyzing variograms of several events; if consistent autocorrelation patterns emerge [*Keim et al.*, 2005], an erratic variogram structure can be ruled out. This implies that conclusions drawn from a single variogram based on fewer than 100 observations [e.g., *Loescher et al.*, 2002; *Staelens et al.*, 2006] have to be interpreted with a grain of salt.

[45] *Keim et al.* [2005] noted that the throughfall correlation length corresponds roughly to one crown diameter. Our data (Figure 9; Appendix A, Table A2) do not support this suggestion. We suppose that tree canopy structures may indeed influence throughfall autocorrelation patterns, but in

areas with a highly variable understory those patterns are likely to disappear because in thick understory throughfall is redistributed before it eventually reaches the forest floor, whereas in locations with a sparse understory it is not. This supposition, however, cannot be tested yet because only few studies provide sufficient data [*Möttönen et al.*, 1999; *Keim et al.*, 2005]; that is, other study sites [*Loustau et al.*, 1992; *Bellot and Escarre*, 1998] cannot be compared because their low number of throughfall sampling locations alone (Table 3) could be responsible for the detected pure nugget autocorrelation structure.

4.3. Temporal Persistence of Throughfall: Long-Term Data Sets and Long-Term Persistence

[46] The majority of studies that investigated the persistence of throughfall detected a temporal stability over several months [*Raat et al.*, 2002; *Staelens et al.*, 2006; *Zimmermann et al.*, 2007, 2008a]; our results are no exception (Figure 13; Table 2). They also imply that wet and dry areas persist over consecutive wet seasons, which may have some ecological relevance because moisture gradients at the forest floor influence, for instance, root water uptake [*Bouten et al.*, 1992], nitrification-denitrification rates [*Carnol and Ineson*, 1999; *Raat et al.*, 2002] and arthropod distributions [*Kaspari and Weiser*, 2000].

[47] A comparison of our experimental setup with a range of studies that investigated the persistence of throughfall measurements in temperate [*Raat et al.*, 2002; *Keim et al.*, 2005; *Staelens et al.*, 2006] and tropical forests [*Zimmermann et al.*, 2007, 2008a] reveals large differences with respect to design criteria (e.g., number of sampling points and occasions) and data analytical methods. To our knowledge this study is the first attempt to describe the temporal persistence of throughfall with variograms. Since reliable variogram analyses require relatively large data sets, benefits have to justify the expense necessary to acquire such data sets. So far, various versions of time stability plots have been used to analyze temporal persistence of throughfall [*Raat et al.*, 2002; *Keim et al.*, 2005; *Staelens et al.*, 2006; *Zimmermann et al.*, 2007, 2008a]. These graphs plot standardized throughfall for all sampling points and occasions in order to illustrate whether spots receiving large or small throughfall volumes persist over time [*Keim et al.*, 2005]. Their main drawback is that unlike temporal variograms they do not reveal the length of the time period over which throughfall measurements are correlated (Figure 13; Table 2). Another disadvantage of time stability plots is that they do not differentiate between temporal lags. Therefore these plots cannot be used for the analysis of long-term data sets because as soon as temporal correlations vanish they will detect a lower persistence. In other words, the quality of temporal variograms usually improves with increasing duration of throughfall monitoring due to an increase in contributing point pairs per lag, while time stability plots deteriorate when the observation period exceeds the range of temporal correlations.

[48] Regardless of methodological differences, a comparison of studies which monitored throughfall over several months [*Raat et al.*, 2002; *Staelens et al.*, 2006; *Zimmermann et al.*, 2008a] reveals that deciduous forests in the temperate zone exhibit a relatively low temporal persistence of throughfall measurements. This can be explained by seasonal variations in canopy foliation [*Staelens et al.*, 2006]. In contrast, temperate coniferous forests display a relatively high persistence of throughfall measurements [*Raat et al.*, 2002]. Deciduousness, however, is not restricted to high latitudes, and some tropical forests show a lower persistence of throughfall during leaf flush [*Zimmermann et al.*, 2008a]. Nonetheless, our results (Figures 12 and 13) indicate that phenological dynamics are not necessarily influential. The proportion of deciduous trees and the occurrence of some palm species (e.g., *Orbignya phalerata* Mart.), which are known to strongly affect the redistribution of rainfall due to their seasonal leaf growth and conducive morphology [*Germer et al.*, 2006], may help to explain the diversity of temporal persistence in tropical forests.

5. Conclusions

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

[50] 1. Throughfall data sets show frequently a large positive skew. Outliers in both the spatial and the temporal domain influence the standard, nonrobust variogram estimator, which primarily results in an overestimation of the sill variance and, to some extent, in spurious autocorrelation structures. Common transformations are important for reducing the underlying skewness but cannot guarantee satisfactory results in cases where outliers are far from the centre of the data. In addition to a thorough exploratory data analysis, we propose to apply the θ statistic [*Lark*, 2000] to assess the need for robust variogram estimation when modeling spatial throughfall data. In the presence of outliers, Dowd's [*Dowd*, 1984] and Genton's [*Genton*, 1998] estimators were chosen most frequently by this criterion, which can be explained by their B-robustness and high breakdown points.

[51] 2. The redistribution of rainfall in our study area, a 1-ha plot of tropical forest, results in rather weak or nondetectable throughfall autocorrelation patterns over the studied lag distances. We speculate that this weak or pure nugget spatial structure of throughfall is typical in areas with highly variable understory vegetation because at some locations throughfall is redistributed several times before it eventually reaches the forest floor, whereas in other areas it is not.

[52] 3. Throughfall measurements show a high temporal persistence, and wet and dry areas, respectively, outlast consecutive wet seasons. Seasonality, and hence deciduousness, seems to have no influence on temporal autocorrelations in our research area. Wet spots accumulated up to 350% of incident rainfall, whereas the driest points received only 16% of rainfall during the 14 month study period. The long-term persistence of this spatial heterogeneity may influence biogeochemical processes at the forest floor. This study introduces variogram analysis as a useful tool to assess the temporal persistence of throughfall measurements.

[53] We hope that our work will stimulate research on throughfall spatial and temporal patterns in other places. To improve the understanding of rainfall redistribution in forest ecosystems and its influence on near-surface processes, we have to improve the comparability among study sites. Therefore we propose to standardize the sampling scale triplet, which ideally involves plots of 1 ha and a sufficient number of sampling locations, particularly at small lag distances.

Appendix A

[54] Tables A1 and A2.

Table A1. Event Characteristics and Summary Statistics for Throughfall Data (n = 220)

Event	Date (dd-mm-yy)	Median (mm)	MAD[a] (mm)	MAD[a]/Median Ratio (%)	Skewness	Octile Skew[b]
1	23-08-07	28.1	10.8	38.5	1.19	−0.01
2	25-08-07	15.7	8.2	52.2	1.05	0.22[sqr)]
3	31-08-07	1.6	1.3	82.4	2.58	0.33[sqr)]
4	02-09-07	0.2	0.3	118.6	3.80	0.38[sqr)]
5	03-09-07	0.1	0.1	148.3	2.65	0.67[sqr)]
6	05-09-07	11.5	6.6	57.0	3.71	0.13
7	08-09-07	11.5	5.9	51.3	0.98	0.17
8	09-09-07	35.4	15.7	44.5	1.05	0.03
9	11-09-07	15.5	7.2	46.6	1.70	0.17
10	12-09-07	8.0	3.9	49.4	2.58	0.13
11	13-09-07	15.2	7.5	49.7	4.86	0.10
12	14-09-07	22.1	9.2	41.5	2.86	0.09
13	15-09-07	29.0	14.7	50.9	3.30	0.12
14	17-09-07	1.1	0.7	61.8	4.86	0.29[log)]
15	19-09-07	3.4	2.3	66.5	2.14	0.22[log)]
16	24-09-07	26.5	10.8	40.8	6.95	0.01
17	25-09-07	4.9	2.6	53.9	4.55	0.14
18	27-09-07	15.3	5.6	36.5	4.60	0.13
19	30-09-07	41.6	15.7	37.9	3.23	0.06
20	02-10-07	0.6	0.3	42.4	1.96	0.60[log)]
21	03-10-07	25.0	11.8	47.2	4.28	0.18
22	07-10-07	16.8	7.9	46.8	0.94	0.10
23	12-10-07	49.5	19.7	39.7	6.66	0.06
24	17-10-07	8.8	3.9	44.5	1.00	0.14
25	18-10-07	27.4	11.8	43.0	5.56	0.02
26	19-10-07	6.5	3.3	51.3	2.79	0.20[sqr)]
27	21-10-07	3.0	1.8	58.9	1.44	0.22[sqr)]
28	24-10-07	6.6	3.3	49.4	1.39	0.13
29	25-10-07	3.6	2.0	54.6	8.17	0.20
30	26-10-07	72.9	30.8	42.2	5.60	0.14
31	29-10-07	6.6	3.3	49.4	1.08	0.23[sqr)]
32	30-10-07	5.3	2.6	49.4	5.51	0.13
33	31-10-07	17.7	9.2	51.9	5.50	0.09
34	03-11-07	77.4	32.1	41.5	0.36	0.00
35	04-11-07	1.5	0.9	58.4	1.73	0.29[log)]
36	05-11-07	8.8	4.6	51.9	3.01	0.20[sqr)]
37	07-11-07	6.0	2.9	48.0	1.58	−0.09
38	09-11-07	38.5	20.0	52.0	9.85	−0.03
39	10-11-07	15.1	9.8	65.0	1.60	0.13
40	12-11-07	10.4	5.0	47.7	1.48	0.10
41	13-11-07	2.7	1.5	53.8	7.01	0.17
42	14-11-07	17.7	8.5	48.2	6.73	0.13
43	23-02-08	0.4	0.3	82.4	2.92	0.43[sqr)]
44	21-03-08	1.0	0.9	94.3	3.96	0.24[sqr)]
45	25-03-08	0.1	0.2	148.3	5.31	0.60[sqr)]
46	10-04-08	8.4	3.3	39.0	1.31	0.07
47	29-04-08	5.3	2.6	49.4	1.36	0.01
48	30-04-08	27.6	13.4	48.6	2.26	0.10
49	08-05-08	12.4	3.9	31.8	1.67	0.00
50	11-05-08	4.9	3.3	67.4	5.73	0.13[log)]
51	26-05-08	17.7	7.9	44.5	4.68	0.13
52	27-05-08	9.3	5.2	56.5	7.19	0.29[log)]
53	05-06-08	30.9	11.8	38.1	0.95	−0.06
54	07-06-08	8.0	3.3	41.2	1.24	0.06
55	12-06-08	1.2	1.0	82.4	3.16	0.23[log)]
56	14-06-08	0.3	0.4	148.3	2.69	0.14
57	18-06-08	0.4	0.3	74.1	2.17	0.00
58	25-06-08	8.8	3.9	44.5	1.26	0.04
59	30-07-08	1.1	0.8	68.4	8.70	0.33[log)]
60	31-07-08	19.5	10.5	53.9	1.08	0.18
61	02-08-08	0.3	0.2	63.5	2.58	0.50
62	03-08-08	0.2	0.2	89.0	2.44	0.63
63	06-08-08	3.6	2.1	57.9	1.38	0.12
64	07-08-08	6.2	3.3	53.0	1.51	0.14
65	09-08-08	19.9	9.8	49.4	1.44	0.05
66	14-08-08	0.3	0.3	98.8	5.91	0.54
67	16-08-08	4.5	3.1	70.5	2.85	0.27
68	17-08-08	0.4	0.3	74.1	7.86	0.47
69	21-08-08	8.0	4.6	57.7	2.18	0.20
70	22-08-08	15.5	7.9	50.8	0.80	0.19
71	23-08-08	36.5	21.3	58.4	0.85	0.14
72	24-08-08	2.2	1.3	59.3	6.62	0.20
73	25-08-08	9.3	5.9	63.5	3.78	0.19

Table A1. (continued)

Event	Date (dd-mm-yy)	Median (mm)	MAD[a] (mm)	MAD[a]/Median Ratio (%)	Skewness	Octile Skew[b]
74	28-08-08	0.5	0.5	86.5	3.77	0.29
75	02-09-08	2.5	1.4	58.2	5.17	0.19
76	03-09-08	1.3	0.9	69.2	1.18	0.33
77	06-09-08	1.1	0.8	68.4	1.09	0.20
78	12-09-08	8.6	3.9	45.9	1.30	0.15
79	13-09-08	8.8	3.9	44.5	2.43	0.02
80	18-09-08	24.3	11.8	48.5	0.94	0.05
81	20-09-08	1.1	0.8	74.1	1.88	0.33
82	24-09-08	13.4	6.8	50.7	2.43	0.07
83	25-09-08	22.1	9.8	44.5	2.81	0.06
84	26-09-08	0.9	0.7	74.1	1.93	0.32
85	28-09-08	0.4	0.5	103.8	6.03	0.26
86	30-09-08	0.4	0.5	129.7	2.50	0.24
87	02-10-08	10.7	5.9	55.1	3.28	0.07
88	06-10-08	0.4	0.3	74.1	5.25	0.28
89	09-10-08	11.1	4.6	41.5	2.10	0.10
90	19-10-08	1.0	0.7	67.4	2.06	0.32
91	20-10-08	1.8	1.3	74.1	2.06	0.21

[a]MAD, median absolute deviation.
[b]If the octile skew was >0.2 we transformed the data (unless otherwise noted); transformations are indicated for all events used for spatial analysis (events 1–60): $^{sqr)}$ denotes square root transformation, whereas $^{log)}$ indicates $\log_{10}$ transformation.

Table A2. Results of Variogram Analysis

Event	Estimator[a]	Model[b]	Nugget/Sill Ratio (%)	Effective Range[c] (m)	Median of $\theta(x)$
1	G	Nug	100.00	–	0.369
2	M	Nug	100.00	–	0.425
3	D	Gau	71.91	90.64	0.410
4	M	Nug	100.00	–	0.360
5	M	Exp	75.03	9.81	0.563
6	G	Nug	100.00	–	0.422
7	M	Exp	73.23	9.81	0.457
8	M	Nug	100.00	–	0.347
9	D	Nug	100.00	–	0.438
10	G	Nug	100.00	–	0.435
11	G	Nug	100.00	–	0.494
12	D	Nug	100.00	–	0.450
13	M	Exp	77.28	49.06	0.361
14	M	Nug	100.00	–	0.324
15	M	Nug	100.00	–	0.400
16	G	Nug	100.00	–	0.411
17	D	Nug	100.00	–	0.438
18	G	Nug	100.00	–	0.413
19	G	Nug	100.00	–	0.377
20	M	Nug	100.00	–	0.425
21	G	Nug	100.00	–	0.466
22	M	Nug	100.00	–	0.432
23	D	Nug	100.00	–	0.383
24	M	Nug	100.00	–	0.356
25	D	Nug	100.00	–	0.331
26	D	Nug	100.00	–	0.417
27	M	Nug	100.00	–	0.369
28	M	Nug	100.00	–	0.393
29	CH	Exp	80.92	39.25	0.431
30	D	Nug	100.00	–	0.438
31	D	Nug	100.00	–	0.373
32	CH	Nug	100.00	–	0.515
33	D	Sph	80.99	13.08	0.468
34	M	Nug	100.00	–	0.420
35	M	Nug	100.00	–	0.445
36	M	Nug	100.00	–	0.352
37	M	Nug	100.00	–	0.362
38	CH	Nug	100.00	–	0.451
39	M	Gau	79.80	33.99	0.345
40	G	Nug	100.00	–	0.436
41	CH	Nug	100.00	–	0.474
42	D	Exp	59.76	9.81	0.434
43	D	Nug	100.00	–	0.343
44	M	Nug	100.00	–	0.357

Table A2. (continued)

Event	Estimator[a]	Model[b]	Nugget/Sill Ratio (%)	Effective Range[c] (m)	Median of $\theta(x)$
45	M	Nug	100.00	–	0.369
46	D	Nug	100.00	–	0.400
47	M	Nug	100.00	–	0.335
48	M	Nug	100.00	–	0.382
49	D	Gau	80.23	11.33	0.404
50	M	Nug	100.00	–	0.451
51	G	Nug	100.00	–	0.445
52	D	Nug	100.00	–	0.385
53	M	Nug	100.00	–	0.354
54	M	Nug	100.00	–	0.401
55	M	Nug	100.00	–	0.502
56	CH	Exp	57.47	19.62	0.586
57	D	Nug	100.00	–	0.380
58	D	Sph	76.76	9.81	0.403
59	G	Nug	100.00	–	0.427
60	M	Nug	100.00	–	0.415

[a]Experimental variogram estimators proposed by Matheron (M), Cressie-Hawkins (CH), Dowd (D), and Genton (G).
[b]Selected theoretical variogram model: Exp, exponential; Gau, Gaussian; Sph, spherical; Nug, nugget.
[c]The effective range R_{eff} is calculated as exponential model: R_{eff} = range*3, Gaussian model: R_{eff} = range*$3^{0.5}$, spherical model: R_{eff} = range.

[55] **Acknowledgments.** This research was funded by the German Research Foundation (El 255/6-1). Beate Zimmermann acknowledges support from The HSBC Climate Partnership. We thank Luise Neumann-Cosel, Bärbel Ehrig, Silja Hund, Janine Matthiessen, Anna Schürkmann and Frank Bäse for participating in the field work, and Ben Marchant, Rothamsted Research, for help with some programming tasks. We are indebted to Stephanie Bohlman, Princeton University, and Patrick Jansen, University of Groningen, for providing the aerial photograph of our plot. Moreover, we would like to thank Andreas Papritz, ETH Zürich, for stimulating discussions and advice. Richard Keim, Louisiana State University, and two anonymous reviewers provided very useful comments and convinced us to take a fresh look at our data. Finally, we thank Murray Lark, Rothamsted Research, for his continuous efforts to keep us out of statistical trouble.

References

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

Blöschl, G. (1999), Scaling issues in snow hydrology, *Hydrol. Process.*, *13*, 2149–2175.

Blöschl, G., and M. Sivapalan (1995), Scale issues in hydrological modelling—A review, *Hydrol. Process.*, *9*, 251–290.

Bouten, W., T. J. Heimovaara, and A. Tiktak (1992), Spatial patterns of throughfall and soil water dynamics in a Douglas fir stand, *Water Resour. Res.*, *28*, 3227–3233.

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-Verlag, Heidelberg, Germany.

Carnol, M., and P. Ineson (1999), Environmental factors controlling NO_3^- leaching, N_2O emissions and numbers of NH_4^+ oxidisers in a coniferous forest soil, *Soil Biol. Biochem.*, *31*, 979–990.

Cavelier, J., M. Jaramillo, D. Solis, and D. de León (1997), Water balance and nutrient inputs in bulk precipitation in tropical montane cloud forest in Panama, *J. Hydrol.*, *193*, 83–96.

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

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

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

Dietrich, W. E., D. M. Windsor, and T. Dunne (1982), Geology, climate, and hydrology 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. 21–46, Smithsonian Institution, Washington, D. C.

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, Germany.

Engelbrecht, B. M. J., and T. A. Kursar (2003), Comparative drought-resistance of seedlings of 28 species of co-occurring tropical woody plants, *Oecologia*, *136*, 383–393.

Ford, E. D., and J. D. Deans (1978), The effects of canopy structure on stemflow, throughfall and interception loss in a young Sitka spruce plantation, *J. Appl. Ecol.*, *15*, 905–917.

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, Smithsonian Institution, Washington, D. C.

Frazer, G. W., C. D. Canham, and K. P. Lertzman (1999), Gap Light Analyzer (GLA), Version 2.0: Imaging software to extract canopy structure and gap light transmission indices from true-color fisheye photographs, Simon Frazer University, BC, Canada, and the Institute of Ecosystem Studies, Millbrook, New York.

Gash, J. H. C. (1979), An analytical model of rainfall interception by forests, *Q. J. R. Meteorol. Soc.*, *105*, 43–55.

Genton, M. G. (1998), Highly robust variogram estimation, *Math. Geol.*, *30*, 213–221.

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

Hampel, F. R. (1971), General qualitative definition of robustness, *Ann. Math. Stat.*, *42*, 1887–1896.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseuw, and W. A. Stahel (1986), *Robust Statistics: The Approach Based on Influence Functions*, 502 pp., Wiley, New York.

Harms, K. E., J. S. Powers, and R. A. Montgomery (2004), Variation in small sapling density, understory cover, and resource availability in four neotropical forests, *Biotropica*, *36*, 40–51.

Hölscher, D., T. D. A. Sá, R. F. Möller, M. Denich, and H. Fölster (1998), Rainfall partitioning and related hydrochemical fluxes in a diverse and in a mono specific (*Phenakospermum guyannense*) secondary vegetation stand in eastern Amazonia, *Oecologia*, *114*, 251–257.

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.

Jansen, P. A., S. A. Bohlman, C. X. Garzon-Lopez, H. Olff, H. C. Muller-Landau, and J. S. Wright (2008), Large-scale spatial variation in palm fruit abundance across a tropical moist forest estimated from high-resolution areal photographs, *Ecography*, *31*, 33–42, doi: 10.1111/j.2007.0906-7590.05151.x.

Jonckheere, I., K. Nackaerts, B. Muys, and P. Coppin (2005), Assessment of automatic gap fraction estimation of forests from digital hemispherical photography, *Agric. For. Meteorol.*, *132*, 96–114, doi:10.1016/j.agrformet.2005.06.003.

Kaspari, M., and M. D. Weiser (2000), Ant activity along moisture gradients in a neotropical forest, *Biotropica*, *32*, 703–711.

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.

Kerry, R., and M. A. Oliver (2007a), Determining the effect of asymmetric data on the variogram. I. Underlying asymmetry, *Comput. Geosci.*, *33*, 1212–1232, doi:10.1016/j.cageo.2007.05.008.

Kerry, R., and M. A. Oliver (2007b), Determining the effect of asymmetric data on the variogram. II. Outliers, *Comput. Geosci.*, *33*, 1233–1260, doi:10.1016/j.cageo.2007.05.009.

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.

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.

Lark, R. M. (2002), Modelling complex soil properties as contaminated regionalized variables, *Geoderma*, *106*, 173–190.

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.

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.

Lloyd, C. R., and A. O. Marques (1988), Spatial variability of throughfall and stemflow measurements in Amazonian rainforest, *Agric. For. Meteorol.*, *42*, 63–73.

Loescher, H. W., J. S. Powers, and S. F. Oberbauer (2002), Spatial variation of throughfall volume in an old-growth tropical wet forest, Costa Rica, *J. Trop. Ecol.*, *18*, 397–407, doi:10.1017/S0266467402002274.

Loustau, D., P. Berbigier, A. Granier, and F. E. H. Moussa (1992), Interception loss, throughfall and stemflow in a maritime pine stand: I. Variability of throughfall and stemflow beneath the pine canopy, *J. Hydrol.*, *138*, 449–467.

Manderscheid, B., and E. Matzner (1995), Spatial and temporal variation of soil solution chemistry and ion fluxes through the soil in a mature Norway spruce (*Picea abies* (L.) Karst.) stand, *Biogeochemistry*, *30*, 99–114.

Matheron, G. (1962), *Traité de Géostatistique Appliqué*, Tome 1, Memoires du Bureau de Recherches Géologiques et Minières, Paris.

McBratney, A. B., and R. Webster (1986), Choosing functions for semivariograms of soil properties and fitting them to sampling estimates, *J. Soil Sci.*, *37*, 617–639.

Mohanty, B. P., R. S. Kanwar, and R. Horton (1991), A robust-resistant approach to interpret spatial-behavior of saturated hydraulic conductivity of a glacial till soil under no-tillage system, *Water Resour. Res.*, *27*, 2979–2992.

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.

Pebesma, E. J. (2004), Multivariable geostatistics in S: The gstat package, *Comput. Geosci.*, *30*, 683–691.

Raat, K. J., G. P. J. Draaijers, M. G. Schaap, A. Tietema, and J. M. Verstraten (2002), Spatial variability of throughfall water and chemistry and forest floor water content in a Douglas fir forest stand, *Hydrol. Earth Syst. Sci.*, *6*, 363–374.

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.

R Development Core Team (2004), R: A language and environment for statistical computing, *R Foundation for Statistical Computing*, Vienna, Austria.

Ribeiro, P. J., and P. J. Diggle (2001), GeoR: A package for geostatistical analysis, *R-NEWS Vol. 1*, No 2. (Available at http://cran.r-project.org/doc/Rnews)

Rousseeuw, P. J., and C. Croux (1993), Alternatives to the median absolute deviation, *J. Am. Stat. Assoc.*, *88*, 1273–1283.

Russo, D., and W. A. Jury (1987), A theoretical study of the estimation of the correlation scale in spatially variable fields: 1. Stationary fields, *Water Resour. Res.*, *23*, 1257–1268.

Schume, H., G. Jost, and K. Katzensteiner (2003), Spatio-temporal analysis of the soil water content in a mixed Norway spruce (*Picea abies* (L.) Karst.)-European beech (*Fagus sylvatica* L.) stand, *Geoderma*, *112*, 273–287.

Shachnovich, Y., P. R. Berliner, and P. Bar (2008), Rainfall interception and spatial distribution of throughfall in a pine forest planted in an arid zone, *J. Hydrol.*, *349*, 168–177, doi:10.1016/j.jhydrol.2007.10.051.

Shouse, P. J., T. J. Gerik, W. B. Russell, and D. K. Cassel (1990), Spatial distribution of soil particle size and aggregate stability in a clay soil, *Soil Sci.*, *149*, 351–360.

Skøien, J. O., and G. Blöschl (2006), Scale effects in estimating the variogram and implications for soil hydrology, *Vadose Zone J.*, *5*, 153–167.

Sobieraj, J. A., H. Elsenbeer, and G. Cameron (2004), Scale dependency in spatial patterns of saturated hydraulic conductivity, *Catena*, *55*, 49–77, doi:10.1016/S0341-8162(03)00090-0.

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.

Thorington, R. W., B. Tannenbaum, A. Tarak, and R. Rudran (1982), Distribution of Trees on Barro Colorado Island: A five-hectare sample, in *The Ecology of a Tropical Forest: Seasonal Rhythms and Long Term Changes*, edited by E. G. Leigh et al., pp. 83–94, Smithsonian Institution, Washington, D. C.

Webster, R., and M. A. Oliver (2001), *Geostatistics for Environmental Scientists*, Wiley, Chichester, U. K.

Woodbury, A. D., and E. A. Sudicky (1991), The geostatistical characteristics of the Borden aquifer, *Water Resour. Res.*, *27*, 533–546.

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 (2008a), Spatio-temporal patterns of throughfall and solute deposition in an open tropical rain forest, *J. Hydrol.*, *360*, 87–102, doi:10.16/j.jhydrol.2008.07.028.

Zimmermann, B., E. Zehe, N. K. Hartmann, and H. Elsenbeer (2008b), 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. (zimmermann.alex@yahoo.de)

B. Zimmermann, Smithsonian Tropical Research Institute, MRC 0580-12, Unit 9100 Box 0948, DPO AA 34002-9998, USA.