Capturing heterogeneity: The role of a study area's extent for estimating mean throughfall

Alexander Zimmermann^{1*}, Sebastian Voss¹, Johanna Clara Metzger^{2,3}, Anke Hildebrandt^{2,3},
Beate Zimmermann⁴

¹ University of Potsdam, Institute of Earth and Environmental Science, Potsdam, Germany

² Friedrich Schiller University Jena, Institute of Geoscience, Jena, Germany

³ Max Planck Institute for Biogeochemistry, Jena, Germany

⁴ Research Institute for Post-Mining Landscapes, Finsterwalde, Germany

* Corresponding author: alexander.zimmermann.ii@uni-potsdam.de

Abstract

The selection of an appropriate spatial extent of a sampling plot is one among several important decisions involved in planning a throughfall sampling scheme. In fact, the choice of the extent may determine whether or not a study can adequately characterize the hydrological fluxes of the studied ecosystem. Previous attempts to optimize throughfall sampling schemes focused on the selection of an appropriate sample size, support, and sampling design, while comparatively little attention has been given to the role of the extent. In this contribution, we investigated the influence of the extent on the representativeness of mean throughfall estimates for three forest ecosystems of varying stand structure. Our study is based on virtual sampling of simulated throughfall fields. We derived these fields from throughfall data sampled in a simply structured forest (young tropical forest) and two heterogeneous forests (old tropical forest, unmanaged mixed European beech forest). We then sampled the simulated throughfall fields with three common extents and various sample sizes for a range of events and for accumulated data. Our findings suggest that the size of the study area should be carefully adapted to the complexity of the system under study and to the required temporal resolution of the throughfall data (i.e. event-based versus accumulated). Generally, event-based sampling in complex structured forests (conditions that favor comparatively long autocorrelations in throughfall) requires the largest extents. For event-based sampling, the choice of an appropriate extent can be as important as using an adequate sample size.

Keywords

throughfall; sampling; extent; hydrological monitoring; tropical forest; European beech forest

1. Introduction

Accurate and precise estimates of mean throughfall are required for a range of hydrological applications such as interception modeling (e.g. Germer et al., 2006), calculation of solute fluxes (e.g. Macinnis-Ng et al., 2012), and estimation of groundwater recharge (e.g. Domínguez et al., 2016). Most interception models, for instance, contain a parameter that describes the water storage capacity of the canopy (Muzylo et al., 2009). The estimation of this parameter, in turn, critically depends on the quality of throughfall estimates (Germer et al., 2006; Lloyd et al., 1988). To optimize throughfall sampling, a number of studies investigated the effect of sample size, support (i.e. the shape of a collector), and sampling design (Cuartas et al., 2007; Holwerda et al., 2006; Kimmins, 1973; Kostelnik et al., 1989; Price and Carlyle-Moses, 2003; Puckett, 1991; Ritter and Regalado, 2010, 2014; Rodrigo and Àvila, 2001; Ziegler et al., 2009; Zimmermann and Zimmermann, 2014; Zimmermann et al., 2010; Zlindra et al., 2011). Somewhat surprisingly, comparatively little attention has been paid to the role of the extent, i.e. the size of the area under study.

A literature survey among studies dealing with interception modeling, empirical interception estimates, and solute flux calculations revealed that the applied extents vary widely (Figure 1). Most of the studies sampled throughfall in 0.04 ha to 0.25 ha plots. Some investigations, however, even relied on sampled areas $< 40 \text{ m}^2$ (Murakami, 2007; Pryet et al., 2012; Shinohara et al., 2013). With a few exceptions, the studies did not provide a rationale for the chosen extent. In some cases, logistical constraints restricted the extent (e.g. Macinnis-Ng et al., 2012). In other cases, relatively small and apparently homogenous patches were considered representative for a certain forest type or age class (e.g. Pérez-Suárez et al., 2014; Uyttendaele and Iroumé, 2002). However, only a few studies (Loustau et al., 1992; Viville et al., 1993) attempted to provide evidence that the selected extent was indeed representative for the investigated forest stand.

< Figure 1 here please >

The lack of studies that investigated the role of the extent for the estimation of mean throughfall motivated us to conduct a virtual sampling experiment. The experiment is based on simulated spatial fields that were generated using throughfall data from three sites with contrasting forest structure. In the experiment we compared the amount of throughfall sampled with three common extents (cf. Figure 1) against the true mean of the simulated throughfall fields. This approach is superior to previous analyses that were based on empirical data only (Loustau et al., 1992; Manfroi et al., 2006; Viville et al., 1993) because the virtual sampling experiment allows to achieve the following two objectives: 1) test several extents systematically and 2) separate the effect of the extent from the uncertainty associated with sampling.

To our knowledge, this study is the first to reveal which extent is required to capture the heterogeneity of throughfall for a variety of forest stands.

2. Methods

2.1 General overview

Our study involves three main methodological steps. The first one comprises the collection of throughfall data. Section 2.2 provides details on the sampling approach, the field sites, and the throughfall data. The second step involves the generation of simulated throughfall fields which are based on the data obtained in step 1. Section 2.3 gives a brief overview on the construction of these fields. The third step comprises the virtual sampling of the simulated fields with varying spatial extents and sample sizes. The calculations involved in the virtual experiments are described in Section 2.4.

2.2 Throughfall sampling and data

In our study, we sampled throughfall at three sites which feature distinct forest structures. Data acquisition at all three sites followed the same sampling scheme. Throughfall was collected on an event basis with 350 small funnel-type collectors (each with a receiving area of 113 cm²). The sampling design comprises a design-based and a model-based sampling component (cf. Zimmermann and Zimmermann, 2014). The design-based component ($n = 200$ sampling points) consists of a stratified simple random sample with compact geographical stratification (de Gruijter et al., 2006). The model-based sampling component comprises additional 150 sampling points located in predetermined distances to selected design-based points. The design-based sampling guarantees an unbiased estimate of the spatial mean. The model-based sampling optimizes the estimation of the variogram near the origin (cf. Voss et al., 2016).

At site 1, throughfall was sampled in a 1 ha plot within a 5-year old secondary forest located in central Panama (Figure 2a). The plot is positioned on a planar, 20° slope. Its stand structure is very dense with 18440 stems ha⁻¹ in the diameter at breast height (DBH) class between 1 cm and 5 cm. Tree height varies between 2 m and 6 m. The plot contains 72 tree species with a DBH >1 cm. Three species dominate the young secondary forest stand in number: *Conostegia xalapensis* (46%), *Vismia baccifera* (19%), and *Vismia macrophylla* (13%). Dataset 1 (site 1) comprises 35 events. From the event data we generated 10 weekly and 1 bulk dataset.

< Figure 2 here please >

At site 2, throughfall was sampled in a 1 ha plot located in a more than 130-year old secondary forest (Foster and Brokaw, 1996; Kenoyer, 1929) on Barro Colorado Island, central Panama (Figure 2b). The plot is positioned on a planar, 10° slope. Its stand structure is very heterogeneous: patches of thick understory alternate with relatively open areas dominated by

single large trees. Stand height is between 25 m and 35 m with some tall trees approaching 45 m. A botanical survey revealed that the plot contains 89 tree species > 5 cm DBH. The most abundant tree species are *Faramaea occidentalis* (12%), *Alseis blackiana* (10%), *Macrocnemum roseum* (9%), and *Heisteria concinna* (7%). Dataset 2 (site 2) comprises 36 events from which we generated 10 weekly and 1 bulk dataset.

At site 3, throughfall was sampled during snow-free conditions in a 1 ha plot of mixed beech forest located in the Hainich National Park in Thuringia, Germany (Figure 2c). The site is part of the AquaDiva Critical Zone Exploratory (Kuesel et al., 2016). The plot, which is positioned in flat terrain, has a heterogeneous stand structure: patches with a high density of young trees alternate with more open areas that are dominated by large and > 100 year old trees. The substantial amount of dead wood in the stand is a characteristic feature of this forest which has not been managed for more than two decades. Stand height averages 19 m, the largest trees approach around 40 m. The plot contains 7 tree species. The most abundant species are *Fagus sylvatica* (66.8 %), *Acer pseudoplatanus* (19.8 %), and *Fraxinus excelsior* (9 %). Dataset 3 (site 3) comprises 30 events. From this data we generated 8 weekly and 1 bulk data set.

2.3 Generation of simulated fields

We generated simulated fields for the event, weekly and bulk datasets. The calculation of the fields can be summarized in 6 steps (Figure 3):

1) Exploratory data analysis (step 1, Figure 3). In this step we tested for trends and underlying asymmetry (cf. Kerry and Oliver, 2007). Event or aggregated datasets with trends were omitted from further analysis. Datasets with an underlying asymmetry (i.e. a skew of the underlying or primary distribution) were transformed before further processing. The

transformation is necessary because even robust variogram estimators can only deal with contaminated normal data (Lark, 2000).

2) Variogram estimation by method-of-moments (step 2, Figure 3). This step of the analysis is necessary for the following separation of outliers from the underlying distribution. First, we calculated the experimental variogram with the classical variogram estimator defined by Matheron (1962):

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

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 $\mathbf{h}$. Second, we fitted three models (exponential, spherical, pure nugget) to the experimental variogram, and chose the one with the smallest sum of squared residuals from the fit. Third, we assessed the fitted model by leave-one-out cross-validation; that is, $z(\mathbf{x}_i)$ is left out and estimated by ordinary kriging to yield a prediction $\hat{Z}(\mathbf{x}_i)$ and the kriging variance $\sigma_{K,i}^2$. With these data we calculated the statistic $\theta(\mathbf{x})$ (Lark, 2000) to check whether the semivariance was overestimated due to an undue influence of outlying values:

$$\theta(\mathbf{x}_i) = \frac{\{\hat{Z}(\mathbf{x}_i) - z(\mathbf{x}_i)\}^2}{\sigma_{K,i}^2}. \quad (2)$$

In case of a correct variogram the median of $\theta(\mathbf{x})$ is 0.455 (Lark, 2000). Fourth, in cases where the median of $\theta(\mathbf{x})$ was significantly less or greater than 0.455 we repeated the calculation of the experimental variogram with the less efficient (Lark, 2000) but robust estimator introduced by Dowd (1984):

$$2\hat{\gamma}_D(\mathbf{h}) = 2.198\{\text{median}(|z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})|)\}^2 \quad i = 1, 2, \dots, N(\mathbf{h}), \quad (3)$$

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 $\mathbf{h}$. Finally, we applied the statistic $\theta(\mathbf{x})$ again to check whether outliers still influenced the semivariance estimate.

3) Identification and removal of spatial outliers (step 3, Figure 3). In this step of our analysis we calculated the standardized error of cross-validation $\varepsilon_S(\mathbf{x})$ (Bárdossy and Kundzewicz, 1990; Lark, 2002) with the variogram obtained in step 2. The standardized error of cross-validation is defined as follows:

$$\varepsilon_S(\mathbf{x}_i) = \frac{\hat{Z}(\mathbf{x}_i) - z(\mathbf{x}_i)}{\sigma_{K,i}} \quad (4)$$

with the kriged value $\hat{Z}(\mathbf{x}_i)$, the observed value $z(\mathbf{x}_i)$ and the kriging standard error $\sigma_{K,i}$. According to Bárdossy and Kundzewicz (1990), it is plausible to assume that $\varepsilon_S(\mathbf{x})$ follows a standard normal distribution. Consequently, we can use the z statistic to classify a value as an outlier. In the present study, points in the sampled dataset with a standardized error of cross-validation smaller than -2.576 ($\alpha/2 = 0.005$) were removed as spatial outliers before further calculations. After outlier removal we checked the bivariate distribution of the data by means of **h**-scattergrams (Webster and Oliver, 2007). These plots show all pairs of measurements that are separated by a pre-specified distance and should be produced for a number of distance classes. In case the **h**-scattergrams displayed no outlying values we estimated the variogram parameters using residual maximum likelihood (REML). All remaining datasets were discarded from further analysis.

4) Estimation of variogram parameters by residual maximum likelihood (REML) (step 4, Figure 3). Maximum likelihood based estimation methods are state of the art in variogram estimation (Lark et al., 2006). Compared to method-of-moment based variogram estimation, REML is more efficient and the fitted model is estimated independently of lag bins (Lark et al., 2006). Since REML is sensitive to outliers, we used the outlier-free datasets generated by the steps above for variogram modeling. Subsequently, we checked the reliability of the REML model with the statistic $\theta(\mathbf{x})$. This procedure provided us a variogram model that

described the spatial variation of the primary process (throughfall data without outliers) not the variation of the secondary process (outliers originating from drip points).

5) Analysis of outlier spatial pattern and modeling of outlier distribution (step 5, Figure 3). Following Zimmermann et al. (2010) we checked the assumption that the outliers were randomly distributed. Next, we obtained the distribution of the outliers by pooling spatial outliers from all datasets of a study site in order to obtain a sufficient sample size. All outliers were standardized by first subtracting their mean value and then dividing by their standard deviation for any given dataset. In case the histogram of the pooled standardized outliers did not match a normal distribution we estimated an empirical probability density function by kernel density estimation (cf. Zimmermann et al., 2010).

6) Generation of throughfall fields (step 6, Figure 3). Our check for trends, a correct variogram, a bivariate normal distribution, and the spatially random distribution of outliers reduced the original event data to 29 (site 1), 25 (site 2), and 26 (site 3). The weekly datasets reduced to 8 (site 1), 9 (site 2), and 7 (site 3). The bulk datasets appeared to be less problematic and hence were all kept for analysis. For all these datasets, we obtained simulated random fields by applying the function `RFsimulate`. The function is implemented in the R-package `RandomFields` (Schlather et al., 2016). For the simulations we used the mean value and the variogram model parameters which we obtained with REML. The generated fields have the dimension of a square grid of interval 0.1 m and edge-length 100 m. To obtain realistic simulated throughfall fields we contaminated the simulations by randomly replacing a fraction of grid cells, which equaled the fraction of outliers in the original dataset, with values sampled from the distribution of outliers. Prior to installation, the standardized outlier value was rescaled to an outlying value for the dataset of interest by multiplying it by the outlier standard deviation for that dataset, and by adding the corresponding outlier average value.

< Figure 3 here please >

2.4 Virtual sampling experiments

To investigate the effect of the extent (i.e. the size of a plot) on the estimation of mean throughfall we conducted a virtual sampling experiment that involved plots of edge length 10 m, 30 m, and 50 m. The selected plot sizes represent typical extents of previous interception and solute deposition studies (Figure 1).

In our experiment, we randomly positioned the virtual plots 10,000 times in simulated 1 ha throughfall fields. At each plot position, we took a random sample of 10,000 points with a support of 0.1 m by 0.1 m and calculated the mean of this sample. Following this procedure we obtained a distribution of sample means for the respective plot size and simulated field. From this distribution of sample means we took the 2.5% and the 97.5% percentile ($P_{2.5}$ and $P_{97.5}$) and calculated the absolute deviation of the percentiles from the true mean $\bar{z}$ of the simulated field (the true mean $\bar{z}$ is the average of all values of the simulated field). In a next step we took the maximum of both deviations and expressed the value as a percentage of the true mean $\bar{z}$ which provided the absolute percentage error (*APE*):

$$APE = \frac{\max(|\bar{z} - P_{2.5}|, |\bar{z} - P_{97.5}|)}{\bar{z}} * 100 \quad (5)$$

Because throughfall spatial patterns depend on forest structure and event size (Zimmermann and Zimmermann, 2014), we conducted our virtual sampling experiment based on data of three differently structured forests and a wide range of event sizes (single events and accumulated data).

The analysis described so far involves a sample size of $n = 10,000$ for each extent position. This large sample size enables investigation of the effect of the extent without the influence of

the uncertainty due to sampling support and sample size (cf. Zimmermann and Zimmermann, 2014). To investigate the joint influence of extent and sample size, we repeated the analysis described above for a variety of sample sizes (5, 10, 25, 50, 100, 200, 300, and 500). For reasons of consistency, we worked with the same support (i.e. 0.1 m by 0.1 m). This analysis allowed us to assess the relative importance of extent versus sample size.

3. Results

3.1 Characteristics of the throughfall data: event size, variability, and spatial patterns

The event sizes differ between the three datasets. The largest events were sampled at site 1, followed by site 2 and 3 (Figure 4a). These differences manifest themselves in the weekly and bulk data. The three datasets also show pronounced differences in the variability of the data (Figure 4b). Events sampled at site 1, 2, and 3 show median coefficients of variation (CV) of 51 %, 55 %, and 27 %, respectively. Weekly and total bulk data show the same pattern, though CVs are much lower than for the event data.

< Figure 4 here please >

Our data further display clear differences of the effective range (Figure 4c). The forest at site 1 with its relatively uniform structure (Figure 2a) generates throughfall spatial patterns with a short range (Figure 4c). The more complex forests (site 2 and 3, Figure 2b and 2c), in contrast, produce comparatively long-ranging throughfall patterns (Figure 4c). The structure of a forest, however, is not the only control of the throughfall autocorrelation distance. An increasing event size or an accumulation of the data always results in a decreasing range of the throughfall variograms (cf. Figure 4a and 4c). Hence, the differences in the ranges between site 2 and 3 (Figure 4c) rather reflect differences in sampled throughfall volumes (Figure 4a) than differences in forest structure.

3.2 The role of the extent for mean throughfall estimates

The choice of the extent has an influence on the accuracy of the mean throughfall estimates, but its importance changes depending on several factors. First, we examine the effect of the extent for differently structured forests. In complex structured forests (sites 2 and 3), sampling small events (≥ 1 mm and < 5 mm) over a small extent (e.g. 10 m by 10 m plot) results in substantial deviation of the mean throughfall estimate from the true mean (8 to 70 %, Figure 5b and 5c). In contrast, in the simply structured forest (site 1), the errors are somewhat smaller, and range between 2 and 20 % (Figure 5a). For the same event class (≥ 1 mm and < 5 mm), sampling over a larger extent (a 50 m by 50 m plot) decreases the error of the mean throughfall estimate considerably: In complex forests (site 2 and 3), the errors range only between 2 and 12 % (Figure 5b and 5c); for the simply structured forest (site 1), the errors take values between 1 and 3 % (Figure 5a). These data provide evidence that in forests with a complex structure (and comparatively long-ranging throughfall patterns, cf. Figure 4c), sampling small plots might cause substantial uncertainty of the mean throughfall estimate.

< Figure 5 here please >

Second, we investigate the effect of the extent at different throughfall amounts. For instance, if we consider the combination of a complex structured forest (site 3) and a small extent (10 m by 10 m plot), our data indicate that the deviation from the true mean throughfall decreases dramatically with an increasing amount of throughfall (Figure 5c): For very small events (< 1 mm) we detected errors of 45 to 110 %, events of the size class ≥ 1 mm and < 5 mm are associated with errors of 16 to 70 %, events of the size class ≥ 5 mm and < 10 mm show errors of 10 to 30 %, and events ≥ 10 mm display errors between 4 and 21 %. Similar trends can be observed for other combinations of forest types and extents, though the errors can be substantially lower. For instance, if we consider the simply structured forest (site 1), an extent

of 30 m by 30 m, and the same four event classes, the range of errors decreases as follows: 2 to 31 %, 2 to 8 %, 1 to 5 %, and 1 to 8 %, respectively. These data illustrate that the sampling of small events (that are often associated with comparatively long-ranging throughfall patterns, cf. Figure 4c) requires sampling comparatively large plots.

Generally, our results demonstrate that increasing the size of the extent can be a very effective measure to ensure representative estimates of mean throughfall. Furthermore, our calculations indicate that in particular throughfall sampling in complex forests benefits from an increase of the extent. A plot size of 50 m by 50 m, for instance, guarantees to limit the deviation from the true mean throughfall to a maximum of 10 % even for complex structured forests and very small events (≥ 1 mm).

3.3 The relative importance of extent and sample size

So far we have examined the role of the extent without considering the influence of the sample size (because we used a sample size of $n = 10,000$ at each plot position). For selected events, we now vary both extent and sample size. These calculations (summarized in Figure 6) indicate that the relative importance of extent and sample size depends on the autocorrelation distance of the throughfall data. In situations where throughfall shows no spatial dependence or a very short range (Figure 6a), the extent has little influence on the deviation of the throughfall estimate from the true mean throughfall, which is solely controlled by sample size. The absolute error of the throughfall estimate is, of course, also influenced by the variability of the data (often expressed as the coefficient of variation).

< Figure 6 here please >

In situations where throughfall shows a comparatively long autocorrelation distance (Figure 6b and 6c), the extent has a substantial influence on the sampling error. For datasets with a

pronounced spatial structure in the throughfall pattern, there is a point at which an increase of the sample sizes has no effect on the accuracy of the mean throughfall estimate anymore. Beyond this point, the error of the throughfall estimate can only be decreased by increasing the extent. Interestingly, the point at which the sample size has no effect anymore can be reached at very small sample sizes (cf. Figure 6c). These findings indicate that the relative importance of the extent compared to the importance of the sample size increases with increasing spatial structure in the throughfall data.

4. Discussion

To date the importance of the extent for the sampling of representative estimates of mean throughfall has not been adequately assessed. This is partly because previous studies on throughfall sampling strategies almost exclusively worked with empirical data (Holwerda et al., 2006; Kimmins, 1973; Kostelnik et al., 1989; Price and Carlyle-Moses, 2003; Puckett, 1991; Ritter and Regalado, 2010, 2014; Rodrigo and Àvila, 2001; Ziegler et al., 2009; Zlindra et al., 2011). Empirical sampling approaches, however, cannot be used to assess the role of the extent because the sampling usually involves only one extent, a single plot position, and a limited sample size. This study is based on a virtual sampling experiment involving three extents, 10,000 randomly chosen plot positions, and a variety of sample sizes. The results of our calculations, which provide new insights into the role of the extent for estimating mean throughfall, have three practical implications.

First, previous recommendations on the size of throughfall sampling areas (Clarke et al., 2010; DVWK, 1986) may require a revision. For instance, in the manual of the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests), Clarke et al. (2010) recommended to use sample areas of 500 m² – 2500 m² for the monitoring of solute deposition and throughfall quantity. While an extent of 2500 m²

seems to be acceptable for most situations (Figure 5 and 6), a sample area of 500 m² is not. Of course, small sampling areas can still provide representative numbers, e.g. in situations where the monitoring is based on accumulated data (Figure 5). However, in regions that receive comparatively little rainfall even long sampling intervals (i.e. biweekly or monthly sampling) may contain too small amounts of precipitation (i.e. < 10 mm) to decrease errors in mean throughfall estimates to an acceptable level (Figure 5).

Second, current protocols on solute deposition and throughfall monitoring focus primarily on the number and the support of samplers to capture the spatial heterogeneity of a given forest stand (Clark et al., 2010; Hansen et al., 2013). Future versions of these protocols may also consider the extent as an important parameter to optimize throughfall sampling. For instance, an event-based sampling in a forest with an uneven stand structure may not only require a relatively large sampling effort (Clark et al., 2010; Hansen et al., 2013; Zimmermann and Zimmermann, 2014) but also a relatively large plot (e.g. a plot ≥ 0.25 ha). Provided that the site of interest receives enough rain, the size of the plot can be further adapted to the desired temporal aggregation of the throughfall data. Generally, studies working with accumulated data require smaller plots than studies relying on event-based data (Figure 5).

Third, previous recommendations to select those patches of a forest with a structure “representative” for a certain region or forest ecosystem (DVWK, 1986) may have little practical relevance. This is because the decision, which properties of a forest patch are typical, also depends on the spatial scale. For instance, at small spatial scales a patch with a closed canopy and a relatively uniform appearance might be representative, whereas at the scale of an entire forest stand there is no reason to exclude a treefall gap. Consequently, sampling within relatively large plots (e.g. 1 ha) may reduce the risk of excluding typical structural elements of the studied forest ecosystem. However, plots should be smaller than variations in local rainfall patterns (cf. Tetzlaff and Uhlenbrook, 2005) to avoid external influence on

throughfall measurements. The upper limit of the plot size therefore depends on the local rainfall regime, which makes general recommendations difficult.

So far, we provided evidence that it is beneficial to adapt the extent of a throughfall sampling plot to both forest structure and the required temporal aggregation of the throughfall data. However, we are aware that other factors influence the spatial variation in throughfall too (cf. Levia and Frost, 2006). For instance, phenological dynamics strongly influence the spatial variation of throughfall (e.g. Helvey and Patric, 1965; Staelens et al., 2006). Another potentially important factor is the type of precipitation (rain, sleet, snow) (Levia and Frost, 2006). To evaluate different throughfall sampling schemes as a function of phenological dynamics and type of precipitation warrants further research.

5. Conclusion

In this study we tested three typical extents (plots of 10 m, 30 m and 50 m edge length) against the true mean of 1 ha simulated throughfall fields. Furthermore, we assessed the relative importance of the extent and sample size. Our key findings are threefold.

(1) The extent (i.e. the size of a plot) is an important parameter to optimize throughfall sampling. The influence of the extent depends on the throughfall autocorrelation distance. The autocorrelation in throughfall data, in turn, depends on the heterogeneity of the forest and the sampled throughfall amount.

(2) Given a sufficient sample size, event-based throughfall sampling in complex structured forests (conditions that favor comparatively long autocorrelations in throughfall) requires plots of ≥ 0.25 ha to ensure mean throughfall estimates to be within ± 10 % of the true mean. In contrast, event-based sampling in simply structured forests requires smaller plots of around 0.1 ha to reach the same error level.

(3) If throughfall shows a pronounced spatial structure, the choice of an adequate extent can be as important as selecting a suitable sample size. In fact, in such a situation, only a joint increase of sample size and extent improves mean throughfall estimates.

6. Acknowledgement

This research was supported by the German Research Foundation (ZI 1300/1-1). Data collection at site 1 was conducted within the framework of the Agua Salud Project, a research initiative sponsored by the HSBC Climate Partnership. We are grateful to Jefferson S. Hall (Smithsonian Tropical Research Institute) for his administrative support at site 1. We also greatly acknowledge the logistical support by Oris Acevedo (Smithsonian Tropical Research Institute) at site 2. Data collection at site 3 was done within the framework of the Collaborative Research Centre “AquaDiva” which is funded by the German Research Foundation (CRC 1076). We thank the Hainich National Park, the Hainich Critical Zone Exploratory site manager Robert Lehmann, and the scientific coordinator Christine Steinhäuser for their support at site 3. We also wish to thank our students for help with collecting field data. Finally, we thank two anonymous reviewers, Editor Tim R. McVicar and the Associate Editor Anthony O'Grady for their constructive comments on an earlier draft of this manuscript.

7. References

- Bárdossy, A., Kundzewicz, Z. W., 1990. Geostatistical methods for detection of outliers in groundwater quality spatial fields, *J. Hydrol.* 115, 343–359
- Clarke, N., Zlindra, D., Ulrich, E., Mosello, R., Derome, J., Derome, K., König, N., Lövblad, G., Draaijers, G.P.J., Hansen, K., Thimonier, A., Waldner, P., 2010. Sampling and analysis of

deposition. 66 pp. Part 14. In: Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. UNECE, ICP Forests, Hamburg. ISBN: 978-3-926301-03-1.

Cuartas, L.A., Tomasella, J., Nobre, A.D., Hodnett, M.G., Waterloo, M.J., Múnera, J.C., 2007. Interception water-partitioning dynamics for a pristine rainforest in Central Amazonia: Marked differences between normal and dry years. *Agric. For. Meteorol.* 145, 69–83.

de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Kotters, M., 2006. Sampling for Natural Resource Monitoring, 332 pp., Springer, Berlin.

Domínguez, C.G., Pryet, A., García Vera, M., Gonzalez, A., Chaumont, C., Tournebize, J., Villacis, M., d'Ozouville, N., Violette, S., 2016. Comparison of deep percolation rates below contrasting land covers with a joint canopy and soil model. *J. Hydrol.* 532, 65–79.

Deutscher Verband für Wasserwirtschaft und Kulturbau e.V. (DVWK), 1986. Ermittlung des Interzeptionsverlustes in Waldbeständen bei Regen. DVWK-Merkblatt 211/1986, ISBN 3-935067-59-3.

Germer, S., Elsenbeer, H., Moraes, J.M., 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.

Hansen, K., Thimonier, A., Clarke, N., Staelens, J., Žlindra, D., Waldner, P., Marchetto, A., 2013. Atmospheric deposition to forest ecosystems. In: *Forest Monitoring – Methods for terrestrial investigations in Europe with an overview of North America and Asia. Developments in Environmental Science* 12, 337–374.

Helvey, J.D., Patric, J.H., 1965. Canopy and litter interception of rainfall by hardwoods of Eastern United States. *Water Resour. Res.* 1, 193–206.

- Holwerda, F., Scatena, F.N., Bruijnzeel, L.A., 2006. Throughfall in a Puerto Rican lower montane rain forest: A comparison of sampling strategies. *J. Hydrol.* 327, 592–602.
- Kerry, R., Oliver, M.A., 2007a. Determining the effect of asymmetric data on the variogram. I. Underlying asymmetry. *Comput. Geosci.* 33, 1212–1232.
- Kimmins, J.P., 1973. Some statistical aspects of sampling throughfall precipitation in nutrient cycling studies in British Columbian coastal forests. *Ecology* 54, 1008–1019.
- Kostelnik, K.M., Lynch, J.A., Grimm, J.W., Corbett, E.S., 1989. Sample Size requirements for estimation of throughfall chemistry beneath a mixed hardwood forest. *J. Environ. Qual.* 18, 274–280.
- Kuesel, K., Totsche, K.U., Trumbore, S.E., Lehmann R., Herrmann, M., 2016. How deep can surface signals be traced in the critical zone? Merging biodiversity with biogeochemistry research in a central German Muschelkalk landscape. *Frontiers in Earth Science* 4, 32.
- 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.
- Lark, R.M., Cullis, B.R., Welham, S.J., 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.
- Levia, D.F., Frost, E.E., 2006. Variability of throughfall volume and solute inputs in wooded ecosystems. *Prog. Phys. Geogr.* 30, 605–632.
- Lloyd, C.R., Gash, J.H.C., Shuttleworth, W.J., 1988. The measurement and modelling of rainfall interception by Amazonian rain forest. *Agric. For. Meteorol.* 43, 277–294.

- Loustau, D., Berbigier, P., Granier, A., El Hadj Moussa, F., 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.
- Macinnis-Ng, C.M.O., Flores, E.E., Müller, H., Schwendenmann, L., 2012. Rainfall partitioning into throughfall and stemflow and associated nutrient fluxes: land use impacts in a lower montane tropical region of Panama. *Biogeochemistry*, 111, 661–676.
- Manfroi, O.J., Kuraji, K., Suzuki, M., Tanaka, N., Kume, T., Nakagawa, M., Kumagai, T., Nakashizuka, T., 2006. Comparison of conventionally observed interception evaporation in a 100-m² subplot with that estimated in a 4-ha area of the same Bornean lowland tropical forest. *J. Hydrol.* 329, 329–349.
- Matheron, G., 1962. *Traité de géostatistique appliquée*, tome i: Mémoires du bureau de recherches géologiques et minières. Pairs Ed. Tech. 14.
- Murakami, S., 2007. Application of three canopy interception models to a young stand of Japanese cypress and interpretation in terms of interception mechanism. *J. Hydrol.* 342, 305–319.
- Muzylo, A., Llorens, P., Valente, F., Keizer, J.J., Domingo, F., Gash, J.H.C., 2009. A review of rainfall interception modelling. *J. Hydrol.* 370, 191–206.
- Pérez-Suárez, M., Arredondo-Moreno, J.T., Huber-Sannwald, E., Serna-Pérez, A., 2014. Forest structure, species traits and rain characteristics influences on horizontal and vertical rainfall partitioning in a semiarid pine-oak forest from Central Mexico. *Ecohydrol.* 7, 532–543.
- Price, A.G., Carlyle-Moses, D.E., 2003. Measurement and modelling of growing-season canopy water fluxes in a mature mixed deciduous forest stand, southern Ontario, Canada. *Agric. For. Meteorol.* 119, 69–85.

- Pryet, A., Domínguez, C., Tomai, P.F., Chaumont, C., d'Ozouville, N., Villacís, M., Violette, S., 2012. Quantification of cloud water interception along the windward slope of Santa Cruz Island, Galapagos (Ecuador). *Agric. For. Meteorol.* 161, 94–106.
- 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.
- Ritter, A., Regalado, C.M., 2010. Investigating the random relocation of gauges below the canopy by means of numerical experiments. *Agric. For. Meteorol.* 150, 1102–1114.
- Rodrigo, A., Àvila, A., 2001. Influence of sampling size in the estimation of mean throughfall in two Mediterranean holm oak forests. *J. Hydrol.* 243, 216–227.
- Schlather, M., Malinowski, A., Oesting, M., Boecker, D., Stokorb, K., Engelke, S., Martini, J., Ballani, F., Moreva, O., Menck, P.J., Gross, S., Ober, U., Berreth, C., Burmeister, K., Manitz, J., Morena, O., Ribeiro, P., Singleton, R., Pfaff, B., R Core Team, 2016. *RandomFields: Simulation and Analysis of Random Fields*. R package version 3.1.8, <http://CRAN.R-project.org/package=RandomFields>.
- Shinohara, Y., Komatsu, H., Kuramoto, K., Otsuki, K., 2013. Characteristics of canopy interception loss in Moso bamboo forests of Japan. *Hydrol. Process.* 27, 2041–2047.
- Staelens, J., De Schrijver, A., Verheyen, K., Verhoest, N.E.C., 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.
- Tetzlaff, D., Uhlenbrook, S., 2005. Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data. *Hydrol. Earth Syst. Sci.* 9, 29–41.

- Uyttendaele, G.Y.P., Iroumé, A., 2002. The solute budget of a forest catchment and solute fluxes within a *Pinus radiata* and a secondary native forest site, southern Chile. *Hydrol. Process.* 16, 2521–2536.
- Viville, D., Biron, P., Granier, A., Dambrine, E., Probst, A., 1993. Interception in a mountainous declining spruce stand in the Strengbach catchment (Vosges, France). *J. Hydrol.* 144, 273–282.
- Voss, S., Zimmermann, B., Zimmermann, A., 2016. Detecting spatial structures in throughfall data: The effect of extent, sample size, sampling design, and variogram estimation method. *J. Hydrol.* 540, 527–537.
- Webster, R., Oliver, M.A., 2007. *Geostatistics for Environmental Scientists (Statistics in Practice)*, 2nd ed. Wiley, Chichester.
- Ziegler, A.D., Giambelluca, T.W., Nullet, M.A., Sutherland, R.A., Tantasarin, C., Vogler, J.B., Negishi, J.W., 2009. Throughfall in an evergreen-dominated forest stand in northern Thailand: Comparison of mobile and stationary methods. *Agric. For. Meteorol.* 149, 373–384.
- Zimmermann, A., Zimmermann, B., 2014. Requirements for throughfall monitoring: The roles of temporal scale and canopy complexity. *Agric. For. Meteorol.* 189, 125–139.
- Zimmermann, B., Zimmermann, A., Lark, R.M., Elsenbeer, H., 2010. Sampling procedures for throughfall monitoring: A simulation study. *Water Resour. Res.* 46, W01503.
- Zlindra, D., Eler, K., Clarke, N., Simončič, P., 2011. Towards harmonization of forest deposition collectors - case study of comparing collector designs. *iForest Biogeosciences and Forestry* 4, 218–225.

Figure 1. The distribution of extents (i.e. plot sizes) of interception and solute flux studies which were published during the period 1970 – 2015 (69 studies, for details see Supplementary material S1). The grey lines refer to extents tested in this study.

Figure 2. Impression of the throughfall sampling sites from which data was used in this study.

Figure 3. Summary of steps involved in modeling spatial throughfall data so as to simulate random throughfall fields. Dashed arrows and dashed boxes highlight decisions or operations, which never applied to or which were not necessary for the modeling of the present datasets.

Figure 4. Characteristics of event-based and accumulated throughfall data of site 1, 2 and 3. Figures show the distribution of a) throughfall amount, b) coefficient of variation of throughfall amount, and c) the effective range.

Figure 5. Deviation of mean throughfall estimates from the true mean for three extents and a variety of throughfall amounts expressed as the absolute percentage error (APE). Figures show results for a) site 1, b) site 2, and c) site 3. Note the data shown here are based on a sample size of $n = 10000$ per plot position. Hence, the deviation from the true mean is solely related to the position of the plots not to sampling based uncertainty.

Figure 6. Deviation of mean throughfall estimates from the true mean for three extents and a variety of sample sizes expressed as the absolute percentage error (APE). Figures show results for three selected throughfall events of different size (mean: mean throughfall amount), variability (cv: coefficient of variation), and autocorrelation structure (range: effective range).

Figure 1. The distribution of extents (i.e. plot sizes) of interception and solute flux studies which were published during the period 1970 – 2015 (69 studies, for details see Supplementary material S1). The grey lines refer to extents tested in this study.

Figure 2. Impression of the throughfall sampling sites from which data was used in this study.

Figure 3. Summary of steps involved in modeling spatial throughfall data so as to simulate random throughfall fields. Dashed arrows and dashed boxes highlight decisions or operations, which never applied to or which were not necessary for the modeling of the present datasets.

Figure 4. Characteristics of event-based and accumulated throughfall data of site 1, 2 and 3.

Figures show the distribution of a) throughfall amount, b) coefficient of variation of throughfall amount, and c) the effective range.

Figure 5. Deviation of mean throughfall estimates from the true mean for three extents and a variety of throughfall amounts expressed as the absolute percentage error (APE). Figures show results for a) site 1, b) site 2, and c) site 3. Note the data shown here are based on a sample size of $n = 10000$ per plot position. Hence, the deviation from the true mean is solely related to the position of the plots not to sampling based uncertainty.

Figure 6. Deviation of mean throughfall estimates from the true mean for three extents and a variety of sample sizes expressed as the absolute percentage error (APE). Figures show results for three selected throughfall events of different size (mean: mean throughfall amount), variability (cv: coefficient of variation), and autocorrelation structure (range: effective range).