Spatial patterns of throughfall isotopic composition at the event and seasonal timescales

Scott T. Allen a,*, Richard F. Keim a, Jeffrey J. McDonnell b,c,d

a School of Renewable Natural Resources, Louisiana State University, Baton Rouge, LA, USA
b Global Institute for Water Security, University of Saskatchewan, Saskatoon, Canada
c School of Geosciences, University of Aberdeen, Aberdeen, Scotland, UK
d Dept of Forest Engineering, Resources and Management, Oregon State University, Corvallis, OR, USA

A R T I C L E   I N F O

Article history:
Received 22 May 2014
Received in revised form 20 October 2014
Accepted 13 December 2014
Available online 20 December 2014

This manuscript was handled by Laurent Charlet, Editor-in-Chief, with the assistance of Fereidoun Rezanezhad, Associate Editor

Keywords:
Stable isotopes
Spatiotemporal variability
Sampling error
Net precipitation
Canopy interception

S U M M A R Y

Spatial variability of throughfall isotopic composition in forests is indicative of complex processes occurring in the canopy and remains insufficiently understood to properly characterize precipitation inputs to the catchment water balance. Here we investigate variability of throughfall isotopic composition with the objectives: (1) to quantify the spatial variability in event-scale samples, (2) to determine if there are persistent controls over the variability and how these affect variability of seasonally accumulated throughfall, and (3) to analyze the distribution of measured throughfall isotopic composition associated with varying sampling regimes. We measured throughfall over two, three-month periods in western Oregon, USA under a Douglas-fir canopy. The mean spatial range of δ18O for each event was 1.6‰ and 1.2‰ through Fall 2009 (11 events) and Spring 2010 (7 events), respectively. However, the spatial pattern of isotopic composition was not temporally stable causing season-total throughfall to be less variable than event throughfall (1.0‰; range of cumulative δ18O for Fall 2009). Isotopic composition was not spatially autocorrelated and not explained by location relative to tree stems. Sampling error analysis for both field measurements and Monte-Carlo simulated datasets representing different sampling schemes revealed the standard deviation of differences from the true mean as high as 0.45‰ (δ18O) and 1.29‰ (d-excess). The magnitude of this isotopic variation suggests that small sample sizes are a source of substantial experimental error.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Stable isotopes of water are useful for tracing the movement of water through catchments (Kendall and McDonnell, 1998). The isotopic composition of precipitation and other end members are used as input signals for estimating streamwater sources (Klaus and McDonnell, 2013), transit time (Tetzlaff et al., 2011), plant water sources (Goldsmith et al., 2012), and multiple other applications. However, precipitation rarely infiltrates or runs off without some preceding fractionating process (Gat and Tzur, 1968). Rainfall from the open sky (gross precipitation; Pg) is intercepted by vegetation canopies resulting in some evaporation (interception loss). The remainder reaches the soil as throughfall (TF), composed of both water that bypasses or is temporarily intercepted by the canopy, or as stemflow that runs down the bark surfaces. Consequently, interception changes numerous characteristics of precipitation (Levia et al., 2011), including its isotopic composition. Throughfall is generally heavier isotopically than Pg, but variable within and among events (Ikawa et al., 2011; Kato et al., 2013; Saxena, 1986). Kubota and Tsuboyama (2003) showed that ignoring the isotopic difference between TF and Pg is a source of error in storm hydrograph separation. Accordingly, TF isotopic composition is a more appropriate input concentration for models that employ stable isotopes of precipitation for water tracing.

Event-mean isotopic differences between TF and Pg have been the focus of previous studies (e.g., Ikawa et al., 2011; Saxena, 1986), but few studies to date have examined the spatial variability of throughfall isotopic composition within and among events. Recent work has shown that, within short measurement periods, the range in isotopic composition at individual TF sampling locations can exceed the difference between open precipitation and TF (Allen et al., 2014; Brodersen et al., 2000; Kato et al., 2013).
This intra-event spatial variability is substantial and may parallel the well-known spatial variability observed for TF amount (Raat et al., 2002).

At the plot scale, TF amount is often autocorrelated spatially (Gerrits et al., 2010; Keim et al., 2005; Loescher et al., 2002) and temporally (Gerrits et al., 2010; Keim et al., 2005; Raat et al., 2002; Staelens et al., 2006). This has been largely attributed to canopy architecture (Gerrits et al., 2010; Staelens et al., 2006; Whelan and Anderson, 1996). Studies of the spatial structure of TF amount have led to improved sampling as well as realization of how TF variability affects subsurface hydrology (Coenders-Gerrits et al., 2013; Hopp and McDonnell, 2011; Raat et al., 2002). Because spatial patterns in TF amount are temporally persistent, repeated sampling in fixed or too few locations can result in large anomalies in estimates of the mean. Thus randomly relocating TF collectors between events (roving collectors) has been advocated as a way to reduce measurement uncertainty (Holwerda et al., 2006; Ritter and Regalado, 2014).

Multiple interacting processes affect TF. Consequently the relationship between amount and isotopic composition is complex (Brodersen et al., 2000; Saxena, 1986). One important process controlling isotopic composition appears to be time-varying transmission of precipitation (that varies in isotopic composition through the event) through the canopy (i.e., the ‘selection’ process; Brodersen et al., 2000; DeWalle and Swislocki, 1994; Ikawa et al., 2011; Kato et al., 2013). Evaporative fractionation (Kato et al., 2013) and isotopic exchange (Kendall, 1993; Saxena, 1986) can occur and result in complex relations with spatial variability in TF amount. Understanding this spatiotemporal variability of TF isotopic composition is critical for using TF as an input value or TF amount. Understanding this spatiotemporal variability of TF amount is often autocorrelated spatially (Gerrits et al., 2010; Keim et al., 2005; Loescher et al., 2002) and temporally (Gerrits et al., 2010; Keim et al., 2005; Raat et al., 2002; Staelens et al., 2006). This has been largely attributed to canopy architecture (Gerrits et al., 2010; Staelens et al., 2006; Whelan and Anderson, 1996). Studies of the spatial structure of TF amount have led to improved sampling as well as realization of how TF variability affects subsurface hydrology (Coenders-Gerrits et al., 2013; Hopp and McDonnell, 2011; Raat et al., 2002). Because spatial patterns in TF amount are temporally persistent, repeated sampling in fixed or too few locations can result in large anomalies in estimates of the mean. Thus randomly relocating TF collectors between events (roving collectors) has been advocated as a way to reduce measurement uncertainty (Holwerda et al., 2006; Ritter and Regalado, 2014).

While the spatiotemporal variability of TF amount has been addressed and has led to comprehensive analyses of sampling errors in TF amount (Ritter and Regalado, 2010; Zimmermann et al., 2010), strategies to sample throughfall for isotopic composition are poorly developed. Such study requires consideration of a suite of controlling processes that likely differ from those controlling TF amount.

Here we characterize the spatiotemporal structure of coupled amount and isotopic composition variability of throughfall. Specifically, we measured throughfall amount and isotopic composition over two three-month periods to address three objectives: (1) to quantify the spatial variability of isotopic composition at the event scale, (2) to determine if there are persistent controls over the variability and how these affect variability of seasonally accumulated throughfall, and (3) to analyze the distribution of measured TF isotopic composition associated with varying number of collectors and using fixed versus roving collectors.

2. Methods

2.1. Site description

This study was conducted in Watershed 1 (WS1) of the H.J. Andrews Experimental Forest in the western Cascade Range of Oregon, which has been extensively described in previous studies (e.g., Lutz and Halpern, 2006; Rothacher, 1965). The steeply-sloped, 96 ha basin was clear-cut harvested in the late 1960s and is now covered with a dense canopy dominated by Douglas-fir (Pseudotsuga menziesii). The climate is typical of the Pacific Northwest; mean annual precipitation exceeds 2000 mm with 80% falling between October and April. All TF collection was at plots at 500 m elevation.

Humidity and rainfall intensity were measured at the H.J. Andrews benchmark meteorological station about 500 m from the study plots.

Two separate experiments were conducted. Experiment 1 (Exp1) took place on a relatively flat area of predominantly Douglas-fir forest coinciding with a 75 m long section of a study transect established in 1962 (Halpern and Dyrness, 2010) and the site of multiple ecological and hydrological experiments (e.g., Bond et al., 2002; Halpern and Franklin, 1990). Experiment 2 (Exp2) used a pair of 12 × 5 m plots as a north-aspect-plot (NAP) and a south-aspect-plot (SAP), 100 m apart on steep opposite slopes of WS1. Canopy cover for SAP and NAP was 92% and 95% respectively, estimated using Fusion software (United States Forest Service Remote Sensing Applications Center, Salt Lake City, UT) from a LiDAR flight in August, 2008. Canopy cover for Exp1 was not quantified but was similar to Exp2. Three P 5 collectors were located within 180 m of all sampling plots in a 0.04 ha clearing surrounded by trees 15–20 m tall.

2.2. Sample collection

Experiment 1, described by Allen et al. (2014), was conducted between October and December 2009. Sampling was by 13 TF collectors with 9.5 cm diameter openings placed along the transect under Douglas-fir trees, randomly with respect to boles, crowns, and other collectors. Storms were sampled per event and collected after precipitation and the majority of drip ceased. Logistical constraints caused some sampling periods to consist of multiple consecutive storm events, yielding 11 collection periods (1.1–1.11). Intra-event dry periods never exceeded 2 days.

Experiment 2 was conducted between April and July 2011 and used higher spatial density of collectors than Exp1 to better characterize spatial patterns. Thirty-six collection points were established at random positions in each plot. Eighteen TF collectors were used at each 12 × 5 m plot and were randomly relocated among the 36 fixed locations for each sampling period. Inter-collector distances ranged from 0.2 m to 12 m with a mean distance of 4.3 m. Both P 5 and TF were collected with 2L polyethylene bottles attached to 15.5 cm diameter funnels about once per week, depending on precipitation. Within the sampling periods, there was never more than one calendar day without rain. There was a total of seven Exp2 collection periods (2.1–2.7; Table 1).

For sampling during both experiments, the volume of water accumulated in each collector was measured and a zero headspace subsample was taken with a 20 ml glass vial for isotope analysis.

2.3. Analyses

We calculated interception loss as the difference between mean P 5 and TF for each event, and a volume weighted mean over all events for seasonal interception loss. Although stemflow occurs, we assumed it was not substantial in Douglas-fir forest (Link et al., 2004; Rothacher, 1963).

All isotope data are expressed in terms of δ, calculated as:

$$\delta = \left( \frac{R_{\text{sample}}}{R_{\text{V-SMOW}}} - 1 \right) \times 1000/\%$$

(1)

where V-SMOW is the Vienna Standard Mean Ocean Water (Coplen et al., 2002) and R is 18O/16O or 2H/1H. Water samples from Exp1 were analyzed for δ18O and δ2H on an off-axis integrated cavity output laser spectrometer (Los Gatos Research LWIA, Mountain View, CA) by the Institute for Water and Watersheds Collaboratory (Corvallis, Oregon). Accuracy was 0.18 ± 0.07‰ and -1.02 ± 0.92‰ (mean ± standard error) for δ18O and δ2H respectively, calculated as deviation of a third standard from a two-point calibration line.
The accuracy of sample amount, $\delta^{18}O$, and $d$-excess were calculated and compared to the respective population means; this was iterated 10,000 times to the respective population means; this was iterated 10,000 times to reach asymptotic behavior. For Exp2, measurements associated with individual events versus cumulative sampling (Exp1 and Exp2), roving versus fixed location collectors for cumulative sampling (Exp1 only; Exp2 dataset used roving collectors so we could not simulate fixed sampling), and for single versus multiple (maximum 10) collectors (Exp1 and Exp2).

In each iteration of Monte Carlo simulations, a subset of collectors (N = 1,2, ..., 10) was selected at random and weighted mean sample amount, $\delta^{18}O$, and $d$-excess were calculated and compared to the respective population means; this was iterated 10,000 times, sufficient for reaching asymptotic behavior. For Exp2, measurements were transformed as deviations from plot means (NAP or SAP) rather than whole population mean. For cumulative sampling, the set of N collectors used across all events of each season was either fixed (fixed cumulative TF), or randomly resampled by event (roving cumulative TF). Although similar analyses have been used in TF amount studies with up to hundreds of collectors (e.g., Ritter and Regalado, 2014). Specific comparisons were errors associated with individual events versus cumulative sampling (Exp1 and Exp2), roving versus fixed location collectors for cumulative sampling (Exp1 only; Exp2 dataset used roving collectors so we could not simulate fixed sampling), and for single versus multiple (maximum 10) collectors (Exp1 and Exp2).

We estimated the sampling error distributions for TF amount, $\delta^{18}O$, and $d$-excess using Monte-Carlo simulations to answer the question of how sampling error is dependent on sampling strategy (e.g., Ritter and Regalado, 2014). Specific comparisons were sample amount, $\delta^{18}O$, and $d$-excess were calculated and compared to the respective population means; this was iterated 10,000 times, sufficient for reaching asymptotic behavior. For Exp2, measurements were transformed as deviations from plot means (NAP or SAP) rather than whole population mean. For cumulative sampling, the set of N collectors used across all events of each season was either fixed (fixed cumulative TF), or randomly resampled by event (roving cumulative TF). Although similar analyses have been used in TF amount studies with up to hundreds of collectors (Ritter and Regalado, 2014; Zimmermann et al., 2010), we limited simulations to maximum 10 collectors to simulate the sample sizes more common to isotope studies.

2.3.2. Throughfall sampling simulation

We estimated the sampling error distributions for TF amount, $\delta^{18}O$, and $d$-excess using Monte-Carlo simulations to answer the question of how sampling error is dependent on sampling strategy (e.g., Ritter and Regalado, 2014). Specific comparisons were errors associated with individual events versus cumulative sampling (Exp1 and Exp2), roving versus fixed location collectors for cumulative sampling (Exp1 only; Exp2 dataset used roving collectors so we could not simulate fixed sampling), and for single versus multiple (maximum 10) collectors (Exp1 and Exp2).

In each iteration of Monte Carlo simulations, a subset of collectors (N = 1,2, ..., 10) was selected at random and weighted mean sample amount, $\delta^{18}O$, and $d$-excess were calculated and compared to the respective population means; this was iterated 10,000 times, sufficient for reaching asymptotic behavior. For Exp2, measurements were transformed as deviations from plot means (NAP or SAP) rather than whole population mean. For cumulative sampling, the set of N collectors used across all events of each season was either fixed (fixed cumulative TF), or randomly resampled by event (roving cumulative TF). Although similar analyses have been used in TF amount studies with up to hundreds of collectors (Ritter and Regalado, 2014; Zimmermann et al., 2010), we limited simulations to maximum 10 collectors to simulate the sample sizes more common to isotope studies.

2.3.3. Variograms

Variograms were used to assess the spatial dependence of TF amount, $\delta^{18}O$, and $d$-excess. Experimental variograms (Keim et al., 2005) were created with the variance, $\gamma(h)$, defined as:

$$\gamma(h) = \frac{1}{2N(h)} \sum (Z(x+h) - Z(x))^2$$

where $Z(x)$ is the TF amount, $d$-excess, or $\delta^{18}O$ at location x, and $h$ is the lag distance. The variogram was calculated for each season and aspect, and the results were compared to theoretical variograms.

The accuracy of sample amount, $\delta^{18}O$, and $d$-excess were calculated and compared to the respective population means; this was iterated 10,000 times, sufficient for reaching asymptotic behavior. For Exp2, measurements were transformed as deviations from plot means (NAP or SAP) rather than whole population mean. For cumulative sampling, the set of N collectors used across all events of each season was either fixed (fixed cumulative TF), or randomly resampled by event (roving cumulative TF). Although similar analyses have been used in TF amount studies with up to hundreds of collectors (Ritter and Regalado, 2014; Zimmermann et al., 2010), we limited simulations to maximum 10 collectors to simulate the sample sizes more common to isotope studies.
where $z_x$ is the parameter of interest (amount, d-excess, or $\delta^{18}O$) at point $x$, $z_{x+h}$ is the value at another point separated by lag $h$, and $n$ is the number of pairs. The maximum lag distance calculated was half of the longest lag to avoid edge effects. Variance was calculated for 20 bins of equal $n$.

3. Results

3.1. Overview

Precipitation differed between Exp1 and Exp2. Events in Exp1 were longer duration and higher amount than more frequent and smaller events in Exp2. In Exp1, rain occurred 75% of days (mean 11.2 mm/day), and in Exp2, rain occurred 71% of days (mean 6.3 mm/day). Some events were missed because of logistical constraints so not all events are consecutive. Measured $P_g$ was 477 mm and 382 mm for Exp1 and Exp2, respectively (Table 1).

Interception loss was greater during Exp2, with 36% for the SAP versus 14% loss in Exp1. For Exp1, event-mean interception loss was not strongly related to event size ($r^2 = 0.45$, $p = 0.87$) whereas $\Delta d$-excess was positively correlated with event size in Exp2 ($r^2 = 0.80$, $p < 0.01$). For both Exp1 and Exp2, $\Delta d$-excess was not correlated with event-mean percent interception loss ($r^2 = 0.01$, $p = 0.8$; $r^2 = 0.33$, $p = 0.18$). While $\Delta d$-excess was positively correlated with precipitation intensity in Exp1 ($r^2 = 0.67$, $p < 0.01$), it was not for Exp2 ($r^2 = 0.13$, $p = 0.42$). For Exp1, $\Delta d$-excess was $-0.4 \pm 1.1\%$ (mean ± SD) with range among collectors of $5.0 \pm 3.4\%$ (mean ± SD). For Exp2, mean $\Delta d$-excess was $-0.9 \pm 0.9\%$ ranging from $-2.1\%$ (events 2.5 and 2.7) to $0.3\%$ (2.3). The range of $d$-excess among collectors by event was $3.9 \pm 1.2\%$.

Coinciding with the higher Exp2 interception loss, mean Exp2 TF $\Delta^{18}O$ tended to be greater ($0.67 \pm 0.02\%$, mean ± pooled SE), than Exp1 $\Delta^{18}O$ ($0.32 \pm 0.02\%$, mean ± pooled SE). The $\Delta$-excess during Exp2 ($-0.80 \pm 0.06\%$, mean ± pooled SE) was more negative than the $\Delta d$-excess for Exp1 ($-0.36 \pm 0.13\%$, mean ± pooled SE).

Despite the SAP having an average of 10.0% ($P_g$) more interception loss than observed at the NAP, the two plots had generally similar mean values of $\delta^{18}O$ and d-excess (Table 1). T-tests showed $\Delta^{18}O$ was positive for every sampling period $(0.7 \pm 0.6\% \text{e})$ and the range of spatial variation was $1.2 \pm 0.4\% \text{e}$. Event means for Exp2 $\Delta^{18}O$ ranged from $0.2\%$ to $1.8\% \text{e}$ (Table 1). In both Exp1 and Exp2, $\Delta^{18}O$ was not strongly correlated with event size ($r^2 = 0.05$, $p = 0.52$ and $r^2 = 0.30$, $p = 0.21$ for Exp1 and Exp2, respectively), event-mean percent interception loss ($r^2 = 0.01$, $p = 0.78$; $r^2 = 0.18$, $p = 0.34$), or precipitation intensity ($r^2 < 0.01$, $p > 0.99$; $r^2 = 0.10$, $p = 0.50$). Among events in both Exp1 and Exp2, event-mean absolute value of $\Delta^{18}O$ was negatively correlated with event amount ($r^2 = 0.25$, $b = -0.01\% \text{mm}^{-1}$, $p = 0.03$), so there was a slightly greater deviation of TF $\delta^{18}O$ from $P_g$ with small events.

For Exp1 and Exp2, TF $\Delta d$-excess was not consistently positive or negative among events (Table 1) and varied significantly among collectors within events (Fig. 1). For Exp1, $\Delta d$-excess was negatively correlated with event size ($r^2 = 0.38$, $p = 0.04$) whereas $\Delta d$-excess was positively correlated with event size in Exp2 ($r^2 = 0.80$, $p < 0.01$). For both Exp1 and Exp2, $\Delta d$-excess was not correlated with event-mean percent interception loss ($r^2 = 0.01$, $p = 0.8$; $r^2 = 0.33$, $p = 0.18$). While $\Delta d$-excess was positively correlated with precipitation intensity in Exp1 ($r^2 = 0.67$, $p < 0.01$), it was not for Exp2 ($r^2 = 0.13$, $p = 0.42$). For Exp1, $\Delta d$-excess was $-0.4 \pm 1.1\%$ (mean ± SD) with range among collectors of $5.0 \pm 3.4\%$ (mean ± SD). For Exp2, mean $\Delta d$-excess was $-0.9 \pm 0.9\%$ ranging from $-2.1\%$ (events 2.5 and 2.7) to $0.3\%$ (2.3). The range of $d$-excess among collectors by event was $3.9 \pm 1.2\%$.

Coinciding with the higher Exp2 interception loss, mean Exp2 TF $\Delta^{18}O$ tended to be greater ($0.67 \pm 0.02\%$, mean ± pooled SE), than Exp1 $\Delta^{18}O$ ($0.32 \pm 0.02\%$, mean ± pooled SE). The $\Delta$-excess during Exp2 ($-0.80 \pm 0.06\%$, mean ± pooled SE) was more negative than the $\Delta d$-excess for Exp1 ($-0.36 \pm 0.13\%$, mean ± pooled SE).

Despite the SAP having an average of 10.0% ($P_g$) more interception loss than observed at the NAP, the two plots had generally similar mean values of $\delta^{18}O$ and d-excess (Table 1). T-tests showed

![Fig. 1. Distribution plots of throughfall (TF) deviation from gross precipitation ($P_g$) for: amount (TF/$P_g$), $\Delta^{18}O$ ($\Delta = \text{TF} - P_g$), and $\Delta d$-excess. White dots are individual observations, in black are kernel-smoothed histograms, and gray 'X's are means.](image-url)
no significant differences \((p > 0.05)\) between NAP and SAP d-excess for any event.

### 3.2. Spatial heterogeneity and pattern stability

Spatial variability of TF varied by event and between Exp1 and Exp2. For Exp1, the shapes of the distributions and ranges in TF amount, \(\Delta \delta^{18}O\), and \(\Delta d\)-excess varied among events (Fig. 1). During Exp2, both the shapes and ranges of distributions of TF amount were more similar among events (Fig. 1). Although the means varied, the ranges of \(\Delta \delta^{18}O\) and \(\Delta d\)-excess were also fairly consistent among events during Exp2 (Fig. 1). Range of variation in \(\delta^{18}O\) and \(d\)-excess were more similar among events than for individual events. We interpret this as patch stability because there were a few locations distinct from the rest. In contrast, there was a steeper trend across the collectors in Exp2 which indicated a more continuous spectrum of variation (general stability). There was a weaker pattern in the NAP due to a wider range of variation at each collector. For Exp2, caution must be used in interpreting time-stability plots because each location had a different combination of contributing events.

For both Exp1 and Exp2, the spatial patterns in TF amount were temporally stable (Fig. 2). For Exp1, certain collectors tended to have consistently higher or lower amounts than others, however the range of variation for most collectors overlapped substantially (i.e., the trend from lowest to highest amount was mostly flat other than the tails). We interpret this as patch stability because there were a few locations distinct from the rest. In contrast, there was a steeper trend across the collectors in Exp2 which indicated a more continuous spectrum of variation (general stability). There was a weaker pattern in the NAP due to a wider range of variation at each collector. For Exp2, caution must be used in interpreting time-stability plots because each location had a different combination of contributing events.

For Exp1 and Exp2 (at both the NAP and SAP), there was a statistically significant collector effect among rank-transformed values (Kruskal–Wallis test, \(p < 0.0001\)). For Exp2, TF amount at the SAP was more temporally stable than at the NAP as indicated by tighter distribution around the mean (Fig. 2). For \(\delta^{18}O\) and \(d\)-excess, there was no collector effect for Exp1 or Exp2 on either plot (Kruskal–Wallis test, \(p > 0.1\)).

### 3.3. Cumulative TF

The range of cumulative TF among collectors was lower than the range over all individual events, and the reduction in range was greater for \(\delta^{18}O\) and \(d\)-excess (Fig. 3) than amount. The large range in cumulative TF amount was caused by temporal stability, so that cumulative amount range was only 10% lower than the mean range for individual events. In contrast, the range of cumulative \(\Delta \delta^{18}O\) was 36% lower than the mean range of individual events; the range of cumulative \(\Delta d\)-excess was 62% lower than the mean range of individual events. The reduction in range from individual events to cumulative totals in isotopic composition indicates unstable spatial patterns that smoothed variability over time.

### 3.4. Simulated throughfall sampling variability

The TF sampling simulation showed that standard deviations of simulated, cumulative TF amount were less using roving rather than fixed-position samplers (Fig. 4A). Standard deviations for cumulative amount sampling at fixed locations were not substantially less than that of individual events.

In contrast, for isotopic composition, standard deviations were greater (\(d\)-excess) or similar (\(\delta^{18}O\)) with roving compared to fixed position samplers (Fig. 4B and C). For single collectors, the standard deviation around the mean \(\delta^{18}O\) exceeded 0.45‰ for individual events (Exp1) and was as high as 0.29‰ for cumulative \(\delta^{18}O\). Use of either roving or fixed-position collectors reduced the \(\delta^{18}O\) standard deviation as compared to individual events. The reduction in standard deviation with each added collector was greater for individual events than for cumulative. Standard deviations of amount, \(\delta^{18}O\), and \(d\)-excess were greater for Exp1 than Exp2.

### 3.5. Spatial controls over patterns of heterogeneity (Experiment 2 only)

For Exp2, the distance to the nearest tree for each collector was inconsistently and weakly related to TF amount, \(\delta^{18}O\), and \(d\)-excess

![Fig. 2. Time-stability plots. Data plotted were normalized amount, \(\delta^{18}O\), and \(d\)-excess of throughfall for Experiment 1, Experiment 2 south-aspect-plot, and Experiment 2 north-aspect-plot. Each point refers to a single measurement event at a single location, with the values normalized \(x = \frac{1}{c} \). Collectors were ranked by the mean of the dependent variable.](image-url)
for individual events and cumulative means (Table 2). Amount was positively correlated with distance for cumulative mean TF amount on the SAP, although only 23% of variability was explained. Some individual events had significant relationships between distance from stem and amount; however these were not consistent in time or between plots (Table 2). The two events with the strongest relationship between location and amount (SAP only) were the highest rainfall sampling periods; isotopic composition did not show the same relationship for these events.

Amount was weakly spatially autocorrelated, in that variance of TF amount increased with lag to a correlation length of approximately 3 m; this trend was stronger at the SAP. In contrast, there was no detectable spatial autocorrelation for $\delta^{18}$O or d-excess.

4. Discussion

4.1. Spatial variability of event and cumulative TF

The spatial variability of TF $\delta^{18}$O and d-excess was high and not related to TF amount. Others have similarly found high intra-event spatial variation of TF isotopic composition and low correlation with TF amount. Kato et al. (2013) observed ranges for $\delta^{18}$O that exceeded 1‰ for individual events. Brodersen et al. (2000) reported spatial ranges of up to 3‰ for $\delta^{18}$O for samples collected weekly (even with averaging isotopic composition over up to three collectors). This high spatial variation has been hypothesized to be a function of heterogeneity in storage capacities and flow-paths within the canopy (Allen et al., 2014; Brodersen et al., 2000; Kato et al., 2013). However, the lack of correlation between amount and isotopic composition suggests multiple processes controlling variability.

We found event-mean TF was generally heavier in $\delta^{18}$O compared to rainfall, consistent with others (Brodersen et al., 2000; DeWalle and Swistock, 1994; Ikawa et al., 2011; Kato et al., 2013; Liu et al., 2008; Saxena, 1986). Nevertheless, the difference between TF and $P_g$ $^{18}$O was not related to event size or proportional

---

Table 2

<table>
<thead>
<tr>
<th>Collection Period</th>
<th>South Aspect Plot</th>
<th>North Aspect Plot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amount $\delta^{18}$O</td>
<td>d-excess</td>
</tr>
<tr>
<td>2.1</td>
<td>0.49***</td>
<td>0.06</td>
</tr>
<tr>
<td>2.2</td>
<td>0.66***</td>
<td>-0.28</td>
</tr>
<tr>
<td>2.3</td>
<td>0.35</td>
<td>0.15</td>
</tr>
<tr>
<td>2.4</td>
<td>0.16</td>
<td>0.09</td>
</tr>
<tr>
<td>2.5</td>
<td>0.15</td>
<td>-0.12</td>
</tr>
<tr>
<td>2.6</td>
<td>0.38</td>
<td>0.18</td>
</tr>
<tr>
<td>2.7</td>
<td>0.45*</td>
<td>-0.04</td>
</tr>
<tr>
<td>Cumulative Mean</td>
<td><strong>0.48</strong>*</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

* p < 0.10.
** p < 0.05.
*** p < 0.01.
to interception loss. Spatially varying evaporative fractionation seems an unlikely cause for large spatial variations in TF isotopic composition because any isotopic exchange with intra-canopy vapor would mitigate spatial variations in evaporative fractionation (Kendall, 1993). However, fine details of exchange and evaporative processes cannot be resolved with the present data. In Exp1, $^{18}\text{O}$ in TF was not consistently heavier than $^{18}\text{O}$ in $P_g$, suggesting a complex interaction of mixing, exchange, and evaporation. However, in Exp2, which had higher interception loss and lower humidity, $d$-excess was related to event size and interception loss, indicating that evaporation could contribute to resultant isotopic composition. During both Exp1 and Exp2, $d$-excess was not always positive, which cannot be achieved by evaporation alone (Dansgaard, 1964) so mixing and exchange processes also must have occurred.

Complexity in isotopic signal suggests a complex suite of processes within the canopy. Because numerous canopy processes affect isotopic composition, it is unreasonable to expect simple relationships; this is supported by the results of multiple previous studies (Brodersen et al., 2000; DeWalle and Swistock, 1994; Ikawa et al., 2011; Kato et al., 2013; Liu et al., 2008). However, even without such expectations, this study yielded a surprising lack of inter-relationships between the rainfall amount and intensity and the spatial variation in $d$-excess and $^{18}\text{O}$. The effects of evaporation, isotopic exchange, and residual water (Allen et al., 2014) should all have greater effects in smaller storms with lower intensity.

It has been observed elsewhere that storm-total TF is isotopically heavier than $P_g$ and that this is a function of ‘selection’ (DeWalle and Swistock, 1994; Gat and Tzur, 1968). That is, the variable transmission of precipitation with temporally varying isotopic composition. The justification is that storms often become isotopically lighter throughout the event due to the rainout effect (Kendall and Caldwell, 1998); this results in the last rain increment, the lightest, remaining on the canopy after the storm and not contributing to the net TF isotopic composition (Brodersen et al., 2000; DeWalle and Swistock, 1994; Gat and Tzur, 1968). However, this simple model cannot explain all variability because precipitation does not always become isotopically lighter throughout an event (Coplen et al., 2008).

Selection would affect spatial variability because TF is the output from the spatiotemporally varying mixture of canopy components involved in the interception process, such as the splash from foliage and branches (Dunkerley, 2009), drip points from bark (Herwitz, 1987), or foliar storage overflow. All of these components have different residence times (Levia et al., 2011) and thus different isotopic compositions, despite relatively small storage capacities, because the isotopic composition of precipitation can change rapidly (Coplen et al., 2008; Pangle et al., 2013). As these pools contribute to varying flowpaths within and from the canopy, the output would have a spatially and temporally varying isotopic composition.

### 4.2. Temporal stability and heterogeneity of accumulated seasonal throughfall

For both experiments, the spatial pattern of TF amount was mostly stable (Fig. 2), consistent with other studies (Gerrits et al., 2010; Keim et al., 2005; Raat et al., 2002; Staelsens et al., 2006; Whelan and Anderson, 1996). We observed both patch stability (Exp1) and general stability (Exp2 at the SAP) for TF amount; Exp2 NAP had elements of both for TF amount. Patch stability without general stability would suggest there are stable locations that have different controls over TF generation than at unstable locations (Allen et al., 2014); this could be a tool in distinguishing locations with TF generated from different within canopy processes, such as drip points (stable) versus splash droplets (Dunkerley, 2009). General stability suggests that there is a more continuous range in variation which we speculate may be related to variation in local canopy storage. However, the differences in trends between Exp1 and Exp2 time-stability plots could also have resulted from the greater number of collectors and fewer number of events in Exp2.

Relative isotopic composition was generally not temporally stable. One potential cause of the lack of temporal stability in isotopic composition is that the canopy retains pre-event moisture (Allen et al., 2014). Another possibility would be varying times in activation of flow paths associated with a complex selection effect. Although time-stability plots are indicative of such a process, further insights may require sampling with greater temporal resolution or directly sampling pools within the canopy to better understand TF generation and the mixing processes within the canopy.

For cumulative TF, the decrease in variation from event sampling for TF isotopic composition exceeded that of amount because of these differences in stability. The inter-event variation without persistent locations of heavy or light isotopes is an important finding regarding sampling the soil–water end member because it suggests reduced spatial variation in the soil.

### 4.3. Spatial controls over TF variability

The stability of TF amount suggests that there are deterministic controls over spatial variability, likely related to canopy architecture (Staelsens et al., 2006). Our observed correlation length scale of around 3 m is consistent with studies that found canopy measurements related to amount variation (Gerrits et al., 2010; Staelsens et al., 2006). Keim et al. (2005) observed similar correlation lengths but did not observe correlation between amount and position with respect to stem. Isotopic composition was not auto-correlated at the scale we measured or related to collector distance from stem. Even though TF amount was significantly related to distance from stem on average, this was driven by a few specific events. Distance from stem may have been only weakly correlated with TF parameters because distance from stem is not necessarily the most appropriate index of canopy structure. Others have found better relationships between TF measurements and alternate indicators such as canopy cover (Staelsens et al., 2006) or LAI (Fleischbein et al., 2005).

The first two events of Exp2, which were the largest events, had the strongest relationship between TF and distance from stem. The relationship between distance from stem and relative amount was stronger for larger storms, which conflicts with previous findings of weaker spatial patterns with large storms (Bouten et al., 1992; Keim et al., 2005; Marin et al., 2000). Although weekly precipitation was large in Exp2, individual events contributing to each sampling period generally were small and possibly insufficient for developing persistent drip points (Herwitz, 1987).

The differences between the NAP and SAP (Exp2) were apparent in the TF amounts. Interception loss was higher, temporal stability of patterns was stronger, the variogram had a more distinct trend, and the relationship with distance from stem was more significant in the SAP than the NAP. Although not quantified, we hypothesize this may have been due to aspect effects on crown forms; trees on the SAP generally had a more conical crown form with gaps whereas the NAP crowns were overlapped resulting in a canopy with more even thickness.

The known controls over TF isotopic variation—selective storage (Brodersen et al., 2000; Ikawa et al., 2011), isotopic fractionation (Ikawa et al., 2011; Kato et al., 2013; Kendall, 1993; Liu et al., 2008), or pre-event moisture (Allen et al., 2014; Gat and Tzur, 1968)—could make TF either isotopically lighter or heavier than $P_g$ depending on the ambient vapor and temporal variation of $P_g$. TF amount can similarly have differential responses to variations in canopy storage; for example, Veneklaas and Van Elk (1990) found
certain locations were drip points in some events but yielded lower amounts due to high canopy storage in other events. For these reasons it is not surprising that patterns of collector-mean TF isotopic compositions were weak and not related to distances from stem.

In contrast with our results, we expected $\delta^{18}O$ and $d$-excess of individual periods would correlate with distance from stem, with a positive correlation or negative correlation depending on storm characteristics. Brodersen et al. (2000) found different TF amounts and isotopic compositions between inner and peripheral canopy space. Differences in canopy thickness and storage would also affect the selection processes (Saxena, 1986) or amount of pre-event canopy moisture (Allen et al., 2014). Similarly, Kato et al. (2013) found that amount was related to distance from stem but isotopic composition was not. We hypothesize that the lack of well-defined spatial controls over isotopic composition arose from sampling intervals that were composed of multiple events, obscuring expected spatial relationships for individual events.

4.4. Implications for TF sampling

Spatial variability of TF isotopic composition must be considered when sampling throughfall as an end member for isotope applications. We have demonstrated that single collector samples can potentially result in errors exceeding 3‰ $\delta^{18}O$. While the sampling simulations suggest errors are more likely <0.9‰ (±2SD) with a single collector, depending on application, this could be a large concern. Kubota and Tsuboyama (2003) showed the TF − $P_c$ differences (generally smaller than our observed variability within TF), could result in errors in separating source contributions to streamflow. Because the isotopic difference between mean TF and $P_c$ can vary substantially in magnitude and direction (Table 1), a correction factor for estimating TF isotopic composition from $P_c$ is not ideal. Thus we recommend use of multiple TF collectors and reporting variation among samples.

For sampling TF amount, roving collectors reduced variability of sample means because the amount pattern was stable. In contrast, the lack of temporal stability in isotopic composition negates the need for roving samplers. Therefore, common strategies for sampling TF amount involving roving or randomly relocating collectors (e.g., Holwerda et al., 2006) may be unnecessary for reducing isotopic uncertainty. The lack of temporal stability in isotopic composition means that fewer collectors are required in cumulative sampling than if there was stability.

The reduced and more consistent spatial variability in Exp2 compared to Exp1 (Fig. 1) indicates that more collectors, longer collection times, and larger collectors smoothed variability. Accordingly, the sampling simulation showed that with Exp2, fewer collectors were required for standard deviations equivalent with Exp1 (Fig. 4). Although not specifically tested in this study, the lack of spatial autocorrelation in isotopic composition makes using larger collectors a reasonable solution, although it would be less appropriate for amount sampling. Longer time periods are not always possible because of study requirements and fractionation of the collected sample becomes more likely unless specifically prevented (Mook and De Vries, 2001).

Designing sampling strategies for TF isotopic composition remains difficult because few studies have been conducted to understand the sources of variability. This study focused on TF isotopic composition in a small area (scale of meters to tens of meters) with a relatively homogeneous canopy. Additional sources of variability at larger scales include variations in forest type (Brodersen et al., 2000; DeWalle and Swistock, 1994) and the effects of topography on rainfall isotopic variability prior to interception (Ingraham, 1998; McGuire et al., 2005).

5. Conclusions

This study showed large spatial variations of isotopic composition for throughfall of both individual events and season totals that do not correspond with throughfall amount. Among collectors located meters to tens of meters apart, the range of $\delta^{18}O$ spatial variability generally exceeded 1‰, was as high as 3.7‰ (Exp1) and 1.9‰ (Exp2) for individual events, and was 1.0‰ for seasonal total (Exp1). Overall, no general controls over the spatial pattern could be inferred. Stable spatial patterns of TF depth were observed but stable patterns of isotopic composition were not. The accumulation of non-persistent spatial patterns of isotopic composition resulted in reduced spatial variability of cumulative seasonal TF. Accordingly, roving samplers reduced variability in cumulative amount sampling but not in isotope sampling. Studies characterizing throughfall as an end member should report the variability among multiple collectors rather than means alone, regardless of spatial extent. Ultimately, the ideal number of collectors or size of collectors is dependent on the level of confidence required and the magnitude of variability observed.

Acknowledgements

Data, facilities, and financial support were provided by HJ Andrews Experimental Forest, funded by the National Science Foundation’s Long-Term Ecological Research Program (DEB 08-23380), US Forest Service Pacific Northwest Research Station, and Oregon State University.

References