Towards harmonization of forest deposition collectors - case study of comparing collector designs

Zlindra D [1], Eler K [2], Clarke N [3], Simončič P [1]

In recent years the harmonization of methods in the frame of the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) operating under the UNECE Convention on Long-range Transboundary Air Pollution (CLRTAP) has been intensified. Among the C-actions of the FutMon project (LIFE07 ENV/D/000218; 2009-2010) the C1-Dep-22(SI) action was established with the goal to harmonize and develop the deposition monitoring procedures and sampling methods. The sampling equipment, spatial design of sampling plots and sampling frequency throughout Europe vary considerably. Therefore a step-by-step approach was made where the harmonized sampling equipment was developed and tested first. The selected collectors were installed at one observation plot of each participating country where measurements of throughfall and bulk deposition were run in parallel with the national collectors for a period of one year. To evaluate the agreement between methods, different statistical analyses were used including Altman-Bland plots, model II regression, and repeated measures ANOVA. Preliminary results from the "Intensive forest ecosystem monitoring plot" plot Brdo in NW Slovenia show a good agreement between national and harmonized bulk (both funnel-type) collectors, while comparison of throughfall measurements indicates systematic bias between harmonized (funnel-type) and national (gutter-type) collectors.

Keywords: Precipitation amount, Precipitation chemistry, Forest monitoring, Harmonization, Collector design

Introduction

The forest monitoring activities within the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) operating under the UNECE Convention on Long-range Transboundary Air Pollution (CLRTAP) are divided in two levels. Level I activities are based on observations of forests in the plots which systematically represent 256 km² each and are less intensive. Level II activities are more intensive, in-depth investigations; many additional important observations are performed on these plots. Locations of Level II plots were not systematically selected by geographical co-ordinates. Since the beginning of forest monitoring within the ICP Forests program the methods and approaches have differed from country to country, especially for Level II activities. Nevertheless, large data series from a long time period have been collected and evaluated. Specifically under the FutMon Life project (LIFE07 ENV/D/000218; 2009-2010), with several actions coordinating quality assurance and quality control of monitoring data (QA/QC) and quality assu-

☐ (1) Slovenian Forestry Institute, 1000 Ljubljana (Slovenia); (2) University of Ljubljana, Biotechnical Faculty, 1000 Ljubljana (Slovenia); (3) Norwegian Forest and Landscape Institute, 1431 Ås (Norway)

@ Daniel Zlindra (Daniel.Zlindra@gozdis.si)

rance in laboratories for all applied methods (http://www.futmon.org/structure-and-organisation), huge steps in the harmonization of forest monitoring in Europe have been made, beginning with large-scale harmonization, laboratory method harmonization and manual harmonization.

Methods to estimate forest deposition have until now rarely been compared. Different collector designs and their spatial representativeness were compared by Reynolds & Leyton (1963), Kostelnik et al. (1989), Reynolds & Neal (1991), Crockford & Richardson (1990) and Bleeker et al. (2003). Ziegler et al. (2009) compared stationary and mobile collectors in tropical rain forest and discovered good agreement at precipitation levels higher than 700 mm. To obtain good agreement they argued for testing and correction for the splash loss of collectors. In the intercomparison of twenty throughfall sampling systems used by different participating countries of the ICP Forests network, Bleeker et al. (2003) found fairly good agreement between systems with large enough collector size and total coverage area but weaker agreement for smaller collectors where spatial variation gave a more pronounced difference from the best estimate (average of all systems). Mean deviation of throughfall quantities was 8% and ranged between -22% and 22.5%. Measurements of concentrations of main nutrients and ionic species mainly had larger deviations than quantities and were especially high for H (mean 21%, max 65%). Concentrations of NO_3^-, K^+ and Kjeldahl-N were generally assessed more accurately. The authors stressed the significantly higher importance of harmonization of field methods (collector type, placement, size, total cover area, geometry of collectors) compared to laboratory methods for accurately assessing the throughfall fluxes. A very comprehensive review of forest deposition monitoring, equipment and sampling design was made by Thimonier (1998) where all the details on deposition monitoring equipment, sampling design and sampling frequency, and spatial representativeness are presented.

In order to compare different national collectors to one reference, the Expert Panel on Deposition of ICP Forests provided a prototype of a collector (so-called "harmonized" collector) fulfilling all the requirements of WMO (2008) for precipitation collectors. The collectors were exposed for one year, except for snow periods when present. The numbers of the collectors per site are prescribed in the ICP Forests Manual Chapter XIV (Clarke et al. 2010), generally depending on type and homogeneity of the forest stand (conifers or broadleaf species, one or two main tree species), age, canopy closure etc. Because of high variability of the

Fig. 1 - The configuration of the collectors in the stand - line A (Photo: D. Zlindra).

designs of the plots and the spatial design of national collectors among the participating countries, other crucial parameters were kept as homogeneous as possible on the country level; *i.e.*, generally only the collectors were different. It was suggested that the spatial design should be as far as possible similar to the spatial design of the national collectors, the sampling frequency of the harmonized collectors should be the same as for national collectors, and also the procedures for cleaning the collectors and bulking the samples for subsequent chemical analyses. This paper describes the sampler comparison for one Level II plot in Slovenia.

Material and Methods

Study Site

The study was conducted on the ICP Forests Level II plot no. 4, which is also a FutMon core plot, located on the Brdo estate in Gorenjska region in north-west Slovenia, latitude 46°17' N, longitude 14°24' E, elevation 471 m a. s. l.

The monitoring activities on the selected plot have been performed starting in 2003. The research plot of 10 000 m² (100 x 100 m) area is fenced and encompasses an inner part of 50 x 50 m. The inner part is dedicated to the study of growth, defoliation, forest health and ground vegetation monitoring and research. In the outer part, different collectors are installed *e.g.*, throughfall collectors for precipitation other than snow (gutter-type), throughfall collectors for snow (cornet-type), baskets for litterfall, and soil solution sampler devices (suction cup lysi-

meters - Fig. 1). The main tree species in the plot is Scots Pine (*Pinus sylvestris* L.). The stand belongs to the *Vaccinio myrtilli-Pinetum* forest type, is 80 to 120 years old, and canopy openness is 18%. The prevailing soil unit is *Dystric Cambisols* (FAO 1998) developed from fluvioglacial gravels and sands. The site is slightly exposed to the south-east (210°). The open-field plot (Fig. 2) is located about 150 meters from the main

plot and is dedicated to bulk precipitation measurements, meteorological observations and ozone sampling with passive sampler dosimeters.

Collector designs and sampling procedure

In the outer 25 m borders of the plot, 10 national throughfall collectors of gutter type were placed. They were divided in two lines (named Line A and Line B). Each gutter was a 3 m long tube with 4 cm in diameter. Each tube had three slits in line, 87 cm long and 0.9 cm wide. Each gutter was connected to a 5 litre collecting bottle by silicone tubing. The collecting bottles were placed in the ground in order to minimize temperature influence on the samples. For the purpose of the experiment, 30 additional throughfall collectors of the harmonized design (which were of funnel type) were placed in the near vicinity of the gutters: along every gutter 3 funnels were installed on each side alternately, at a perpendicular distance of 1 meter from the gutter. The funnels were installed at a height of 1.0 m above ground level by means of PVC pipe. The sampling area had a diameter of 159 mm. The funnel construction was made of white polyethylene (PE) and meets the requirements of the relevant Manuals (WMO 2008, Clarke et al. 2010). An especially important detail of the funnel design was the rim, because the rim (vertical on the inner side, very steep slope on the outer side) prevented rain drops splitting and partly (uncontrollably) sprinkling into or out of the funnel. Another very important geometrical attribute of the collector is the

Fig. 2 - The configuration of the collectors in the open field (Photo: D. Zlindra).

height of the vertical part of the funnel in connection with the more gradually sloping part of it, leading into the outflow tube. The height of the vertical part had to be high enough and the gradual part steep enough so that the whole amount of the precipitation could be captured without losses due to splashing. The collecting bottles (5 litres volume) were directly connected to the funnel. Around the funnel a bird ring was mounted, consisting of two rings in two heights formed by fishing line. For avoiding tree litter dropping into the collecting bottles a cone shaped PE mesh with openings of 1 mm was placed in the neck of every funnel. Sampling started in July 2009 and lasted until October 2010, except for the snow period in January and February 2010. Both types of collectors were sampled on the same days, fortnightly. The cleaning procedures were also the same for both types: after the first 14-day period the sampling equipment (funnels, gutters and both types of collecting bottles) was cleaned and washed with ultra pure water in the field, and after the second 14-day period all parts of the sampling equipment, except the gutters, were cleaned and washed in the laboratory.

The amount of precipitation was registered after every 14-day period and for each collector. For the chemical analysis samples were pooled. Samples from harmonized collectors were bulked in such a way that the first 15 of them, placed near line A of national throughfall collectors, contributed to one sample for chemical analyses and the other 15 collectors, placed near line B of national throughfall collectors, contributed to the other. Samples gained by national collectors were also bulked from the A and B lines separately. Before chemical analysis, samples of the first (Period a) and second (Period b) 14-day period were bulked in proportion to their volumes.

In the open-field plot three national bulk collectors were installed. These were funnels, placed in a triangle at a distance of 4 meters from each other. They were installed at a height of 1.45 m above ground level. The funnels had diameters of 230 mm and were made of black polyethylene (PE). The collecting bottles (8 litres volume) were directly connected to the funnel. There were no bird rings around the funnel and no mesh for excluding litter in them.

Also here harmonized bulk collectors (funnel type) were placed in the triangle at a distance of 6.5 meters from each other. The fictive centre of this triangle was the same as for the national collectors, but the end points were 1.5 m outwards. The design of the harmonized collectors was the same as in the stand. Sampling periods and cleaning procedures were the same as for the collectors in the stand.

The designs of the national and harmo-nized bulk deposition collectors were similar except for minor details: the diameter of the national funnel was larger and consequently the collecting bottle was of higher volume to prevent overflow. The vertical edge was in both cases 10 centimetres but the angle to the lower, gradually sloping part of the funnels was 150 and 160 degrees for national and harmonized funnels respectively. Thus the possibility for losses of precipitation was bigger in the case of the national collectors. For both collectors a pipe of suitable diameter was used for placing them at the acceptable height.

The precipitation quantity was registered after every 14-day period for each collector separately. For the chemical analysis samples from the two national and two harmonized collectors were taken separately for every 14-day period. Before chemical analysis, samples of the first (Period a) and second (Period b) 14-day period were bulked in proportion to their volumes.

Data analysis

Brief theory of testing the agreement between methods

When a new measurement method or device is developed it is important to test its agreement with other standardized or already established methods. The agreement is considered adequate when we might replace the old method with the new one. To test for agreement, the same samples are measured using both new and old method. In the case of destructive methods or where there are other reasons which prevent repeatability of measurement on the same samples (space and time variability), a larger number of samples is measured from the same popula-tion. Within the measurement comparison approaches the agreement is evaluated by testing for significant systematic differences. The differences or bias can be fixed (same shift across the range of values) or propor-tional (shift related to the magnitude of a measured variable). What is sought for in method comparison studies is actually not the agreement but the bias (Ludbrook 2002) which results from the statistical foundations itself. At a certain confidence level only the rejection of the zero hypothesis stating the agreement between methods is possible but not its confirmation.

Different statistical techniques are used to test for bias significance: regression me-thods, Altman-Bland plot, ANOVA approa-ches, kappa statistics, structural equation modelling etc. The choice is dependent on the type of variables being studied (continu-ous *vs.* categorical - Ludbrook 2002, Bland & Altman 2007), the presence of a "gold-standard" method for comparison (Astrua et al. 2007), the number of replicate measure-ments per sample and the study design.

Calculating the Pearson correlation coeffi-cient (*r*) is often the first choice of investi-gators when comparing methods. Neverthe-less, this measure of strength of a linear rela-tionship between two variables is not an ap-propriate measure for the agreement (Lud-brook 2002, Bland & Altman 1986). We might compute a high *r* but if the data do not lie along the identity (1:1) line the agreement is poor. Correlation also depends on the range of values of the measured quantity.

Secondly, regression analysis is often used. However, ordinary least squares regression (OLS), included in all statistical software, is mostly not the proper choice when compa-ring methods because it assumes fixed values of the independent variable. In terms of method comparison this would mean the exi-stence of a gold-standard method with no measurement errors. This assumption is only rarely met (if ever). In praxis it can be used if measurement errors of the gold-standard method are considerably lower than the errors of the new method and can be con-sidered negligible.

When errors not controlled by the resear-cher are present in both methods then OLS regression underestimates the slope of the re-gression line (Legendre & Legendre 1998). In these cases model II regression ap-proaches (Sokal & Rohlf 1995, Legendre & Legendre 1998) such as Deming regression (also known as major axis regression or least product regression), which allow values of both methods to be attended by random error, should be used. Deming regression, contrary to OLS, minimizes the perpendicu-lar distances between points and regression line (Linnet 1993). Similarly to OLS, it assu-mes constant standard deviations across the range of values. When deviations change considerably, weighted variants of the me-thods are preferable (Linnet 1993). Whatever the regression approach chosen, the fixed and proportional bias of the new method are determined using intercept and slope esti-mates, respectively, together with the confi-dence intervals for each parameter. Signifi-cant fixed bias is ascertained when intercept confidence intervals do not contain 0 and proportional bias is significant when slope confidence intervals do not contain 1.

A simple, yet efficient approach was pro-posed by Altman & Bland (1983) and Bland & Altman (1986). These authors stated that it was the differences between methods that are most useful in searching for bias. By plotting the differences of the methods against their means and computing mean dif-ference and 95% confidence intervals for the difference (*a.k.a.*, limits of agreement) one can easily determine how good the agree-ment between the methods is and if the con-fidence interval is within the reasonable li-mits for the accuracy required from specific measurements. Limits of agreement are cal-

culated as (eqn. 1):

$$D_{avg} \pm z_{2\alpha} SD$$

where D_{avg} is the mean difference between methods, $z_{2\alpha}$ is the standardized normal deviation of two sided P = 0.05 (1.96) and *SD* is the standard deviation of the differences.

In the case of smaller sample size (n), a more conservative approach using the Student *t* distribution should be used (Ludbrook 2002), which is calculated as (eqn. 2):

$$D_{avg} \pm tn_{-1.2\alpha} SD \sqrt{1+(1/n)}$$

where $tn_{-1.2\alpha}$ is the value of the *t* distribution corresponding to two sided P = 0.05 for *n*-1 degrees of freedom and $\sqrt{(1 + 1/n)}$ is the correction for small sample size.

Besides determination of fixed bias (mean difference tested using a paired *t*-test), the Altman-Bland method can also estimate proportional bias if differences are regressed (using OLS) against the means. If no proportional bias is expected, the slope of that regression line will not significantly differ from 0, which is determined by looking at the slope confidence intervals.

The above mentioned approaches hold when the measurements of each method are independent, which means they are performed only once on each sample. Sometimes it is valuable or even only feasible to obtain replicate measurements by each method on each sample. The case of repeated measurements on the same samples is frequent in forest research where a temporal trend (time series) of a quantity is investigated. Different variants of random effects models are used where the variability of differences between methods is divided into within-subject variability and between-subject variability, with the latter estimating the agreement between the methods (Bland & Altman 2007, Myles & Cui 2007).

Evaluation of the collector data

The agreement between harmonized (funnel-type) and national (gutter-type) throughfall deposition collectors was evaluated using different approaches. In this way we also tried to determine the performances of different statistical techniques. For precipitation quantities (volumes) the approach was different than for the chemistry (concentrations and calculated depositions) because volume data were known for individual collectors, in contrast to chemistry where pooling of volume weighted samples of half of the collectors took place. Because volume data were known for individual collectors in each of the sampling periods, the data quality was considerably higher and tests stronger than for the chemistry data where the volume weighted samples of deposits were pooled from half of the collectors giving us two data points for each collector type per sampling period.

Regarding quantities, the following agreements were evaluated: open-field (bulk) precipitation, throughfall precipitation and canopy interception. For the chemical data, agreement between collector types for bulk and throughfall was assessed in terms of concentration, deposition and canopy interaction (throughfall / open-field).

Generally three techniques were used: Deming regression, the Altman-Bland method and repeated ANOVA. The first two methods were used to test for fixed and proportional bias; samples of the different sampling periods were assumed to be independent. For the calculation of limits of agreement according to Bland & Altman (1986), eqn. 2 was used.

Repeated measures ANOVA was used to correct for sample dependence occurring due to repetitive sampling during the observation time. A partly nested ANOVA design was used for throughfall quantities and ratios with sampling period as the within-subjects factor, collector type as the between-subjects factor and collector pair (one gutter with the mean of the corresponding 3 harmonized funnels) as the random factor nested within collector type. The interaction term between collector type and sampling period was also included as the within subject factor. ANOVA for chemistry data was performed by preliminarily computing the mean of two data points of each method per sampling term, since it is difficult to estimate the variability on the basis of two data points, which would result in very large confidence intervals. Mean concentration or deposition or canopy interaction per sampling term was subjected to simple repeated measures (RM) ANOVA with collector type and sampling periods as the factors. With only one replicate per each combination of sampling period and sampling type, calculating the interaction is not feasible and this represents the error term in ANOVA.

As a descriptive measure of disagreement between collectors the mean relative deviation (*d*) in % was computed using the following equation (eqn. 3):

$$d = \sum_{i=1}^{j} \left(\frac{p_{harmonized} - p_{national}}{p_{national}} \right)^{\frac{1}{j}}$$

where $p_{harmonized}$ and $p_{national}$ are the volumes or chemical parameters analysed (*i*) and *j* is the number of sampling periods taken into consideration.

Results and Discussion

Quantities of precipitation

Open field (bulk) precipitation

On the basis of Deming regression and the Altman-Bland method, statistically significant fixed and proportional bias of bulk collectors was shown (Fig. 3); the intercept of the Deming regression was -3.342 (confidence interval, C.I.: -4.852, -1.874) and the slope was 1.082 (C.I.: 1.054, 1.111). The significant bias is the result of two influential data points with large precipitation (>150 mm per period) where the volumes measured by harmonized collectors were larger than the volumes found in the national collectors. If these points are not regarded the bias is not significant. No clear explanation for the deviation at larger precipitation volumes can be given. Due to the national collector being shallower than the harmonized collector, larger splashing of rain out of gauges might have occurred at more intense rainfall (Thimonier 1998). Moreover, the larger collecting area and larger collector height of the national collectors can change the wind flow around the collector, resulting in lower catch (Rodda & Smith 1986, Sevruk et al. 1994). However, despite statistical significance the mean deviation of the bulk collectors is so small (< 1%) that we regard it as irrelevant. If we take mean volumes measured by both collectors as the
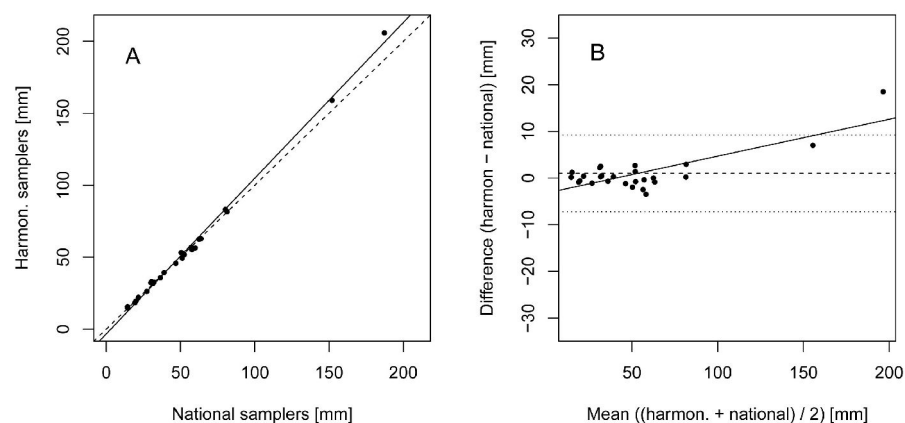

Fig. 3 - Scatter plot with Deming regression (A) and Altman-Bland plot (B) showing the agreement in volumes between national and harmonized bulk collectors.

best estimate, the accuracy of both collectors is well within the 10% accuracy suggested by Bleeker et al. (2003).

Throughfall precipitation

Due to the different design of the through-fall collectors in our study (gutter *vs.* funnel), higher disagreement was expected in the measurements in the forest stand. A scatter plot with Deming regression and an Altman-Bland plot are shown in Fig. 4, both indicating supposedly fixed and proportional negative bias of the harmonized collector compared to the national gutter collector. The mean difference between collectors was 5.3 mm which is significantly different than 0 (p>0.001). In relative terms, this amounts to 9.4% smaller volumes measured by the harmonized collectors compared to the gutters. The slope of the Deming regression line was 0.814 (C.I.: 0.792, 0.836) and the slope of the regression line of differences on means from the Altman-Bland plot was -0.203 (C.I.: -0.230, -0.176), both indicating significant proportional bias. When calculating RM ANOVA, the harmonized collectors again measured significantly lower precipitation quantities than the national gutter collectors (p=0.0012).

A few other investigations showed larger volumes measured by gutters (Kostelnik et al. 1989, Crockford & Richardson 1990) but studies showing no difference (Reynolds & Leyton 1963) or inverse results (Reynolds & Neal 1991) also exist. There might be several reasons for this relatively important disagreement between collectors in our case. As pointed out by Bleeker et al. (2003) the wrongly calculated (or measured) area of collectors very frequently leads to overestimation or underestimation of precipitation. Falsely determined area can also easily explain proportional bias in either of the collectors. The area should be defined with 0.5% precision according to the World Meteorological Organisation (after Thimonier 1998) or 2% precision according to the ICP Forests Manual (Clarke et al. 2010). The design of our gutters is similar to the design of the collector number 11 in the study of Bleeker et al. (2003). For this collector the largest negative deviation (-6.4 %) between the collector area reported by the participating country and the area measured was observed, which shows the difficulty of properly measuring the collecting area of this type of collector. Bleeker et al. (2003) emphasized that when the collector rim is wide enough the inclusion of the area of half the rim width might contribute to better estimates, but for this type of gutter it is difficult to even determine what is the collector rim and what is not. The part of the gutter between the slits can also contribute to the capture area of the collector. Additional research, best with some simulation measure-
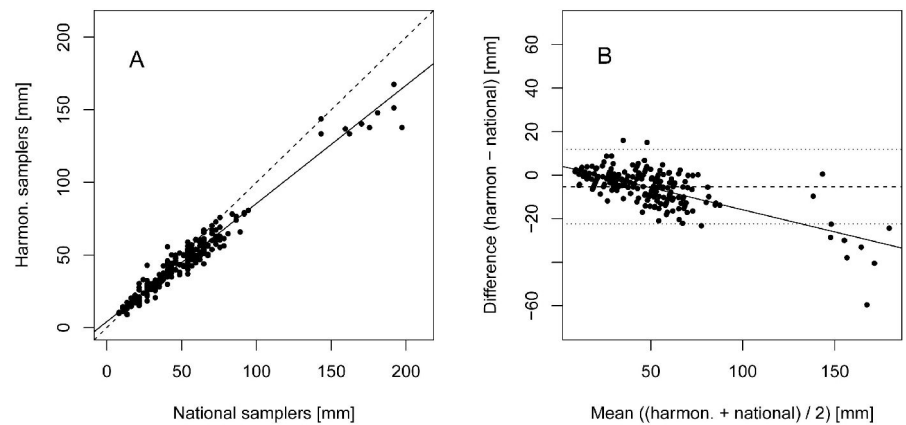
ments, would be needed to correct the collecting area estimates.

The presence of any obstacles for air flow around the collectors (*e.g.*, a bird ring) can also influence precipitation estimates (Sevruk et al. 1994). In our case a bird ring was mounted on the harmonized collectors only, and that could theoretically explain the lower volumes in this collector. However, we do not regard the bird ring as important due to its slender design. The raindrop splash out of the collector is another reason for poor estimates (Bleeker et al. 2003, Thimonier 1998, Ziegler et al. 2009) which should be ruled out in our case due to the collector designs.

Canopy interception

The ratio between throughfall and bulk precipitation represents the inverse of canopy interception. To calculate the interception mean precipitation per period of all collectors of each type was divided by the mean bulk precipitation per period measured by both collector types (Fig. 5). Mean interception for the observation period was -0.81% and 9.05% for gutters and harmonized funnels, respectively. The difference in canopy interception between collectors was signifi-

cant (p<0.001) with national collectors estimating a 9.9 percentage points lower interception of the forest canopy. The value obtained from gutters differs from previously reported interceptions for *Pinus sylvestris* forests (Llorens & Domingo 2007, Nieminen et al. 1999), whereas estimates of interception using harmonized collectors are closer to literature values despite being in the lower range of these reported values. This indicates that gutter precipitation quantities are most probably overestimated, and the values measured with the harmonized equipment are more realistic.

Chemistry data: agreement in concentrations of ions

The data on ion concentrations were generally more variable than quantity data. For bulk deposition no significant disagreement was found when comparing national and harmonized collectors, which was expected from the similar collector designs. For bulk deposition concentrations absolute mean deviation between different types of bulk collector ranged between 0.16% (Ca^{2+}) and 29.1 % (H). Deviation below 5% was found for Na, Ca^{2+}, Mg^{2+}, Cl^-, NO_3^- and SO_4^{2-}, below 15% for K^+, Mn^{3+}, NH_4^+, NO_2^-, total N and

Fig. 4 - Scatter plot with Deming regression (A) and Altman-Bland plot (B) showing the agreement in volumes between national and harmonized throughfall collectors.
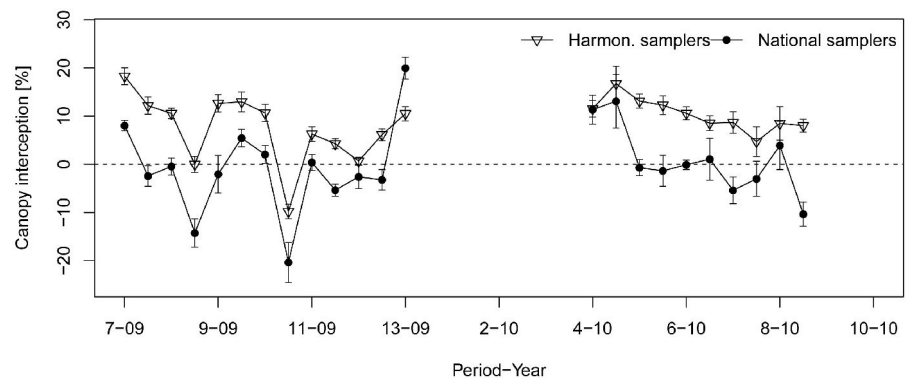

Fig. 5 - Canopy interception derived from national and harmonized throughfall collectors. Means ± standard errors are presented.

Tab. 1 - Summary table of volume-weighted mean concentrations and range of concentrations for harmonized and national throughfall deposition collectors. Relative and absolute deviation of harmonized collectors with respect to national collectors (negative values indicate lower measured concentrations in harmonized collectors). Significance of fixed and proportional bias between collectors according to the Altman-Bland method. (*) for alkalinity the unit is $\mu eq\ L^{-1}$ and for conductivity $\mu S\ cm^{-1}$.

Parameter	Concentration harmonized collector (mg L⁻¹)*			Concentration national collector (mg L⁻¹)			Mean absolute deviation (mg L⁻¹)	Mean relative deviation (%)	Sig. of fixed bias (p value)	Sig. of proportional bias (p value)
	Mean	**Min**	**Max**	**Mean**	**Min**	**Max**				
H	2.98×10^{-6}	4.28×10^{-7}	5.77×10^{-6}	4.24×10^{-6}	1.08×10^{-6}	1.32×10^{-5}	-1.67×10^{-6}	-38.7	0.003	0.001
Na	0.139	0.009	0.580	0.228	0.043	0.851	-0.066	-25.3	n.s.	n.s.
K	0.627	0.220	5.487	0.900	0.235	6.306	-0.125	-6.4	n.s.	0.001
Ca^2	0.504	0.250	1.534	0.723	0.215	1.783	-0.146	-7.9	n.s.	n.s.
Mg^2	0.091	0.035	0.328	0.123	0.050	0.349	-0.009	-5.8	n.s.	n.s.
Mn^3	0.033	0.025	0.084	0.043	0.027	0.084	-0.001	-1.5	n.s.	n.s.
$N\text{-}NH_4$	0.233	0.011	1.099	0.345	0.040	1.206	-0.063	-22.7	0.015	n.s.
NO_3^-	1.351	0.027	4.567	1.675	0.014	4.620	-0.138	0.2	0.024	n.s.
NO_2^-	0.012	0.006	0.050	0.013	0.005	0.026	0.003	29.6	n.s.	n.s.
SO_4^{2-}	0.797	0.335	1.918	0.935	0.355	1.945	-0.025	-1.4	n.s.	n.s.
Cl^-	0.348	0.127	1.019	0.953	0.374	2.800	-0.597	-48.0	0.002	0.001
DOC	6.849	2.885	24.356	9.063	2.630	29.906	-1.317	-8.0	0.015	0.000
Total N	0.856	0.300	2.355	1.058	0.300	2.511	-0.091	-6.4	0.011	n.s.
Alkalinity	17.094	1.000	107.000	13.910	0.350	93.335	6.376	95.5	0.008	n.s.
Conductivity	15.637	7.130	26.620	18.972	7.410	35.430	-3.335	-15.4	0.001	0.001

DOC, and above 15% for H, alkalinity and electroconductivity. Relatively high deviation of H and alkalinity between different collector types was also reported by Bleeker et al. (2003), indicating the difficulty of measuring these two parameters accurately.

An even larger scatter was observed when analyzing throughfall chemistry data, which caused larger problems when evaluating the agreement. Additionally, data sets for individual ions and other compounds were small, causing further problems. Tab. 1 shows the summary agreement analysis for ions and other chemical parameters measured.

No disagreement was found for Na, Ca^{2+}, Mg^{2+}, Mn^{3+}, NO_2^- and SO_4^{2-}. Of these, mean deviations were less than 10 % for Ca^{2+}, Mg^{2+}, Mn^{3+} and SO_4^{2-}. For Na and NO_2^- larger but non significant deviations were detected. There were minor deviations found for K^+, NO_3^- and total nitrogen which were significant with either fixed or proportional bias. Relatively small but persistent deviations were observed for DOC and NH_4 with harmonized collectors measuring lower concentrations. Larger, significant, persistent and more relevant disagreement was found for Cl^- and H^+ concentrations. No final explanation is found for high disagreement for chloride ions, especially if we take into account the relative passivity of this ion in canopy interactions (Parker 1983). According to information from the manufacturer of the plastic pipes gutters are made of such material that no Cl^- ions should be added to the solution. However, in gutters there are no meshes to prevent litter and other particles from entering the solution and chloride could be derived from these; this still has to

be investigated. Significant fixed bias and large positive deviation (95.5%) were found for alkalinity; this deviation however was caused by some very low measurements which magnified the relative deviation.

Seasonal variation in concentrations of ions and other parameters of throughfall precipitation measured by harmonized and national collectors is shown in Fig. 6.

Deviations of volume-weighted mean concentrations from the harmonized collectors relative to the national collectors were negative for all ions except for NO_3^-, where the mean deviation was almost zero, and NO_2^-. There are many possible reasons for lower concentrations in the harmonized collectors and higher concentrations in the national collectors. Gutters are generally more difficult to clean properly after each sampling, which is crucial to remove algal growth and other dirt capable of causing biochemical transformations and/or additions of some ions and compounds (*e.g.*, DOC, Cl^-). Further, due to the larger collecting area per collector and lower collector height gutters may be more efficient in collecting dry deposition (Bleeker et al. 2003). This could also help to explain the Cl results as Cl is often deposited as dry deposition.

Spatial representativeness

In this study the area covered by each collector type was considerably different. The total sampling areas of ten gutters and thirty harmonized collectors were 1850 and 5879 cm^2, respectively. Both coverages are beyond the minimal requirements according to the ICP Forests Manual (Clarke et al. 2010), but only the harmonized collector system is within the limits suggested by Bleeker et al.

(2003), who for spruce forest suggested using at least 26 collectors to reach the accuracy of 10% of the population mean.

Due to their oblong design gutters cover a larger spatial gradient in the canopy above them, which would cause lower variability in concentrations and volumes measured (Crockford & Richardson 1990). However, the 3.18-times larger total area of the harmonized collectors may have compensated for the weakness of spot measurements of harmonized collectors, so that the variability in quantities of the collectors was not significantly different. Mean coefficients of variation in precipitation quantity of national and harmonized collectors were 13.6% and 10.0%, respectively.

Conclusions

Within this study a harmonized collector proposed by the ICP Forests Expert Panel on Deposition was compared with the Slovenian collectors. In terms of quantities and concentrations of ions and compounds only minor disagreement was found for bulk deposition, which was expected due to the similar funnel-type designs of both collectors.

For throughfall quantities significant disagreement was found. National collectors measured 9.4% higher volumes than harmonized collectors, which we mainly attributed to the difficulty of properly measuring the real collecting area of the national collector. Proof that the national collectors overestimated the volumes rather than that the harmonized collectors underestimated them was the comparison of interception. No interception (-0.8%) of *Pinus sylvestris* forest was perceived using national collectors, which is very unlikely considering the
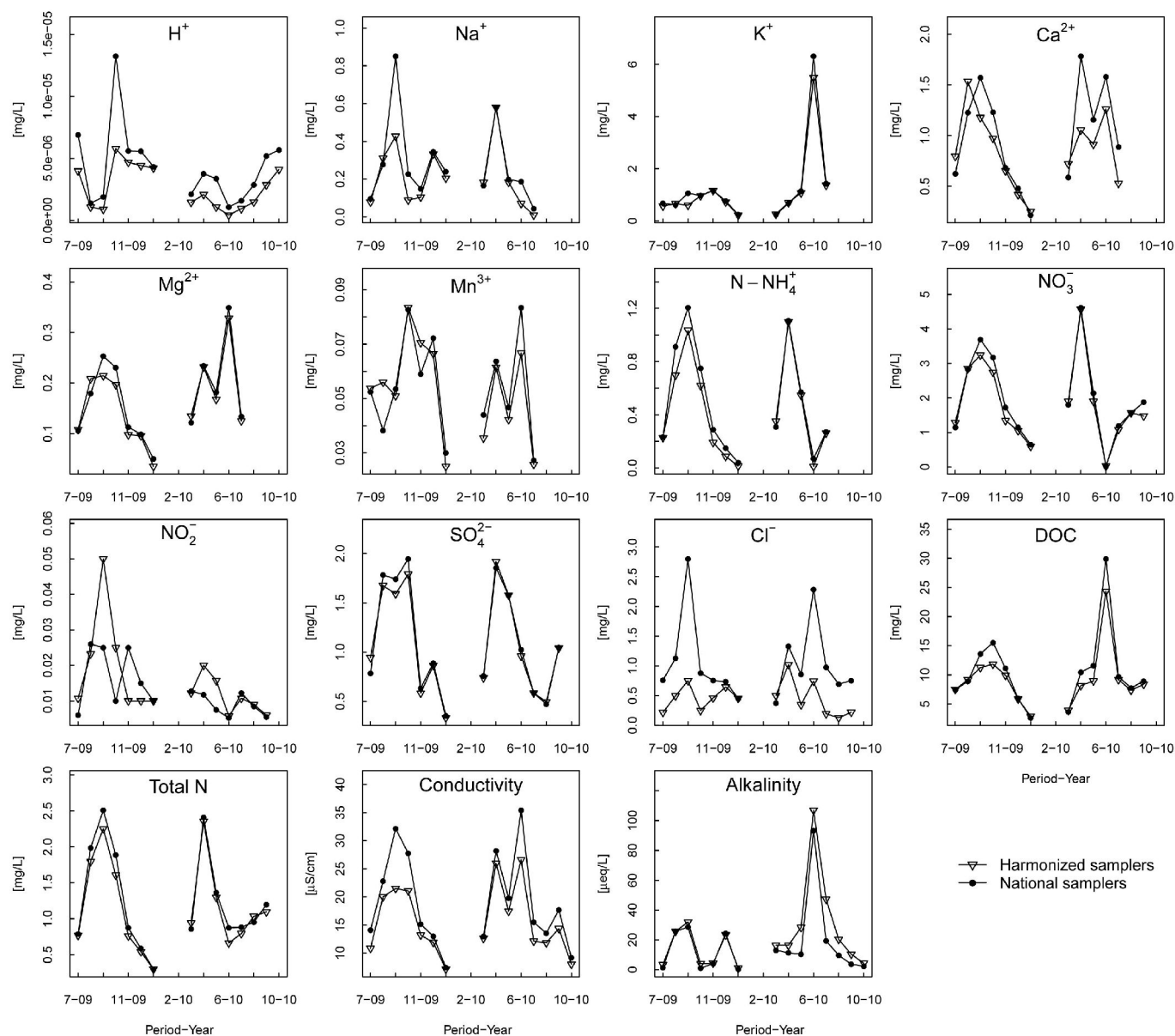
Fig. 6 - Concentrations of ions and other parameters of throughfall precipitation measured by harmonized and national collectors during the investigation period from July 2009 till October 2010.

literature reports. The interception estimated using harmonized collectors was closer to the literature values (9.0%).

Regarding the concentrations of ions and other compounds, gutters again in almost all cases overestimated these concentrations although bias was not significant or not relevant for a large part of the ions and compounds measured (Na, Ca^{2+}, Mg^{2+}, Mn^{3+}, NO_2^-, SO_4^{2-}, K^+, NO_3^-, total N). High and persistent deviations (up to 48%) between collectors throughout the investigation period were found for Cl^-, H and DOC. No final explanation is found yet for the disagreement in some of the chemical parameters, although poor washing of gutters after each

sampling is suspected to cause the observed deviation.

Several statistical approaches were tested within this research. Each of them with certain advantages and disadvantages and it is preferable to test the agreement using several approaches. The Altman-Bland method has proved simple and efficient in visualization of the agreement between methods of measurement but also in inference on the significance of fixed and proportional bias. Deming regression is probably the most general, easy to calculate and should be the first choice in method comparison studies.

The effort has been made across the European intensive forest monitoring net-

work to harmonize the collectors for assessing the quantities and chemistry of bulk and throughfall forest depositions. A large diversity of designs, total collector areas and collector placements exists throughout Europe. Previous studies have shown large differences between currently used national collectors in terms of precipitation quantities and concentrations of ions and other compounds. The funnel-type 16 cm diameter collector has been chosen as the harmonized collector. To check its performance and the agreement with the national collectors, field investigations across Europe coupled with proper statistical data analysis are needed. For the Slovenian case, the harmonized col-

lector system proved to be the better choice for several reasons: more realistic and precise throughfall volume estimates, and easier collector cleaning. For historic data, obtained by national collectors, calibration of results is possible using the results of this comparison. Similar comparisons however will be needed in other forest types with larger spatial variability. The harmonized collector system has the disadvantage of large sample size (3-times as many collectors) which makes the sampling much more labour intensive (up to 4-times more time was needed per one sampling activity) and consequently more expensive. The solution would be fewer collectors (*e.g.*, 10-15), but in that case additional surveys should be performed in this respect if such reduction of sample size is feasible and acceptable in terms of spatial representativeness.

Acknowledgements

References

Altman DG, Bland JM (1983). Measurement in medicine: the analysis of method comparison studies. The Statistician 32: 307-317. - doi: 10.2307/2987937

Astrua M, Ichim D, Pennecci F, Pisani M (2007). Statistical techniques for assessing the agreement between two instruments. Metrologia 44: 385-392. - doi: 10.1088/0026-1394/44/5/015

Bland JM, Altman DG (1986). Statistical methods for assessing agreement between two methods in clinical measurement. Lancet 1 (8476): 307-310. - doi: 10.1016/S0140-6736(86)90837-8

Bland JM, Altman DG (2007). Agreement between methods of measurement with multiple observations per individual. Journal of Biopharmaceutical Statistics 17: 571-582. - doi: 10.1080/10543400701329422

Bleeker A, Draaijers G, van der Veen D, Erisman JW, Möls H, Fonteijn P, Geusebroek M (2003). Field intercomparison of throughfall measurements performed within the framework of the Pan European intensive monitoring program of EU/ICP Forest. Environmental Pollution 125: 123-138. - doi: 10.1016/S0269-7491(03)00142-8

Clarke N, Zlindra D, Ulrich E, Mosello R, Derome J, Derome K, König N, Lövblad G, Draaijers GPJ, Hansen K, Thimonier A, Waldner P (2010). Sampling and analysis of deposition. Part XIV. 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, Germany, pp. 66. (ISBN: 978-3-926301-03-1) [online] URL: http://www.icp-forests.org/Manual.htm

Crockford RH, Richardson DP (1990). Partitioning of rainfall in a eucalypt forest and pine plantation in southeastern Australia. I. Throughfall measurement in aeucalypt forest: effect of method and species composition. Hydrological Processes 4: 131-144. - doi: 10.1002/(ISSN)1099-1085

FAO (1998). World reference base for soil resources. FAO, Rome, Italy, pp. 88.

Kostelnik KM, Lynch JA, Grimm JW, Corbett ES (1989). Sample size requirements for estimation of throughfall chemistry beneath a mixed hardwood forest. Journal of Environmental Quality 18: 274-280. - doi: 10.2134/jeq1989.00472425001800030005x

Legendre P, Legendre L (1998). Numerical ecology (2nd edn). Elsevier, Amsterdam,The Netherlands, pp. 870.

Linnet K (1993). Evaluation of regression procedures for methods comparison studies. Clinical Chemistry 39: 424-432. [online] URL: http://www.ncbi.nlm.nih.gov/pubmed/8448852

Llorens P, Domingo F (2007). Rainfall partitioning by vegetation under Mediterranean conditions. A review of studies in Europe. Journal of Hydrology 335: 37-54. - doi: 10.1016/j.jhydrol.2006.10.032

Ludbrook J (2002). Statistical techniques for comparing measurers and methods of measurement: a critical review. Clinical and Experimental Pharmacology and Physiology 29: 527-536. - doi: 10.1046/j.1440-1681.2002.03686.x

Myles PS, Cui J (2007). Using Altman-Bland method to measure agreement with repeated measures. British Journal of Anaesthesia 99: 309-311. - doi: 10.1093/bja/aem214

Nieminen TM, Derome J, Helmisaari HS (1999). Interactions between precipitation and Scots pine canopies along a heavy-metal pollution gradient. Environmental Pollution 106: 128-137. - doi: 10.1016/S0269-7491(99)00050-0

Parker GG (1983). Throughfall and stemflow in the forest nutrient cycle. Advances in Ecological Research 13: 57-133.

Reynolds ERC, Leyton L (1963). Measurement and significance of throughfall in forest stands. In: "Water relations of plants" (Rutter AJ, Whitehead FH eds). Blackwell Scientific Publications, Oxford, UK, pp. 127-141.

Reynolds B, Neal C (1991). Trough versus funnel collectors for measuring throughfall volumes. Journal of Environmental Quality 20: 518-521. - doi: 10.2134/jeq1991.00472425002000030003x

Rodda JC, Smith SW (1986). The significance of the systematic error in rainfall measurement for assessing wet deposition. Atmospheric Environment 20: 1059-1064. - doi: 10.1016/0004-6981(86)90293-3

Sevruk B, Hertig JA, Tettamanti R (1994). The effect of orifice rim thickness on the wind speed above precipitation gauges. Atmospheric Environment 28: 1939-1944. - doi: 10.1016/1352-2310(94)90334-4

Sokal RR, Rohlf FJ (1995). Biometry: the principles and practices of statistics in biological research (3rd edn). W.H. Freeman, New York, USA, pp. 888.

Thimonier A (1998). Measurement of atmospheric deposition under forest canopies: some recommendations for equipment and sampling design. Environmental Monitoring and Assessment 52: 353-387. - doi: 10.1023/A:1005853429853

WMO (2008). Guide to meteorological instruments and methods of observation (7th edn). World Metereological Organization Publication, no. 8, Geneva, Switzerland. [online] URL: http://www.wmo.int/pages/prog/www/IMOP/publications/CIMO-Guide/CIMO%20Guide%207th%20Edition,%202008/CIMO_Guide-7th_Edition-2008.pdf

Ziegler AD, Giambelluca TW, Nullet MA, Sutherland RA, Tantasarin C, Vogler JB, Negishi JN (2009). Throughfall in an evergreen-dominated forest stand in northern Thailand: comparison of mobile and stationary methods. Agricultural and Forest Meteorology 149: 373-384. - doi: 10.1016/j.agrformet.2008.09.002