



On the measurement uncertainty of Hirst-type volumetric pollen and spore samplers

Simon Adamov · Natalie Lemonis · Bernard Clot · Benoît Crouzy ·
Regula Gehrig · Marie-José Graber · Christine Sallin · Fiona Tummon 

Received: 7 August 2020 / Accepted: 1 September 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Hirst-type volumetric spore traps are used across the globe and are the current standard instrument for monitoring pollen. While they suffer from various issues related to sampling, measurement, and human error, they are, nevertheless, relatively cost-efficient and robust instruments that have been in use for decades. They are also the only reference against which newer instruments can be directly evaluated and it is thus important to understand and quantify all errors to make fair comparisons. Here, we investigate the variability across three Hirst-type traps run in parallel for three months during the main pollen season in Payerne, Switzerland. A variety of temporal resolutions is studied. Overall, daily average values show median relative differences of 16% for total pollen (inter-quartile range (IQR) = 8–32%) and

between 23 and 67% for the top three taxa considered. The values are identical for total pollen when only concentrations > 10 pollen grains/m³ are considered, but decrease notably for the individual taxa investigated, with median relative differences ranging 14–30%. At the 2-hourly resolution, there is considerably more variability between the three samplers, with median relative differences of 42% for total pollen (IQR = 20–88%) and 62–120% for the top three taxa, respectively. Again, when the lowest concentrations are not included, these differences decrease somewhat to 40% for total pollen (IQR = 19–78%) and 41–56% for the top three taxa, respectively. Observations of low concentrations below 10 pollen grains/m³ therefore have a large impact on the measurement uncertainty: for daily average total pollen concentrations there are differences of over 100% between the three samplers, and for individual taxa differences reach up to 200%, the maximum value possible in terms of pairwise comparisons.

Simon Adamov and Natalie Lemonis have contributed equally to the analyses.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10453-021-09724-5>.

S. Adamov · N. Lemonis · B. Clot ·
B. Crouzy · R. Gehrig · M.-J. Graber ·
C. Sallin · F. Tummon (✉)
Federal Office of Meteorology and Climatology
MeteoSwiss, Payerne, Switzerland
e-mail: fiona.tummon@meteoswiss.ch

N. Lemonis
Now at J'aime ma planète, Geneva, Switzerland

Keywords Pollen · Hirst-type volumetric sampler · Error quantification · Measurement uncertainty · Temporal resolution · Operational monitoring

1 Introduction

Airborne pollen and fungal spores have important impacts on human health (de Weger et al., 2013; Pawankar et al., 2013) and agriculture (Cunha et al., 2016; Oteros et al., 2014; Savary et al., 2019), and serve as important indicators for identification of invasive species (Karrer et al., 2015; Sikoparija et al., 2009) as well as climate change effects (Zhang et al., 2015; Ziello et al., 2012; Ziska et al., 2019). As such, networks to monitor these particles have been established in many countries over the past decades (Buters et al., 2018). At present, the majority of these networks employ manual methods based on Hirst-type volumetric samplers (Hirst, 1952). These samplers draw air in at a rate of 10L/min and airborne particles impact onto a rotating tape, which is usually collected on a weekly basis and manually analysed using light microscopy.

Standards to ensure comparability within and across networks have been established for routine monitoring of airborne pollen and fungal spores (CEN/EN 16,868, 2019; Galán et al., 2014). Nevertheless, methods based on Hirst-type samplers are known to suffer from a number of shortcomings, including collection efficiency of Hirst-type traps (Mandrioli et al., 1998), errors in estimates of the flow rate used to calculate concentrations (Oteros et al., 2017), variability of particle deposition onto the collection tape (Käpylä & Penttinen, 1981; Sikoparija et al., 2011; Tormo Molina et al., 1996), issues related to sub-sampling (Comtois et al., 1999; Cotos-Yáñez et al., 2013; Frenz, 1999; Rojo et al., 2019), and human errors in counting and identifying pollen grains (Carinanos et al., 2000; Galán et al., 2014; Pedersen & Moseholm, 1993). Efforts to standardise counting methods and control quality across laboratories are an important step towards reducing human error (Galán et al., 2014; Oteros et al., 2013; Sikoparija et al., 2017; Smith et al., 2019).

While several studies have compared measurements from separate devices of the same kind, there have been relatively few comparisons of more than two devices and often these have been at considerable distance from one another (e.g. Ekebom et al., 1997; Gottardini & Cristofolini, 1997; Irdi et al., 2002; Meiffren, 1988; Mitakakis & McGee, 2000; Pedersen & Moseholm, 1993; Rizzi et al., 1992). Tormo Molina et al. (2013) compared five traps at the same site and

found significant differences in terms of the timing of peak pollen concentrations when hourly data were considered; however, no such differences were evident in daily averages. In contrast, Rojo et al. (2019) compared three devices in Munich, Germany, and two traps in Zrenjanin, Serbia. Although ratios of mean pollen concentrations between the instruments were mostly close to 1, they found that the standard deviations of daily average ratios varied significantly (on average 34%); a possible explanation for why no significant differences was found in various previous studies that only considered mean pollen concentrations rather than the variability of these values (Rojo et al., 2019).

The drums of Hirst-type traps rotate behind a 2-mm inlet at 2 mm/hour with samples deposited over a length of approximately 4 mm and hourly variability is thus smoothed (Käpylä & Penttinen, 1981). While these traps are most commonly used to provide daily average values, they have also been used for studies at much higher temporal resolutions, up to the hourly scale (e.g. Berggren et al., 1995; Borycka & Kasprzyk, 2018; Clot, 2001; Prtenjak et al., 2012; Sarda Estève et al., 2018). Such timescales are particularly relevant for assessing new automatic pollen monitoring techniques, which provide data at resolutions of hours to minutes (e.g. Crouzy et al., 2016; Oteros et al., 2015; Sauvageat et al., 2020; Sauliene et al., 2019). However, it is important to fully understand all sources of errors related to traditional manual measurements, and particularly to investigate the error at various temporal resolutions, to ensure that any comparisons with newer automatic instruments are robust and fair.

In this study, we compare measurements from three Hirst-type traps placed side by side in Payerne, Switzerland, from the pollen season of 2013. While these observations date back somewhat, counting methods have not changed significantly since and the data and results remain valid. It is the first time that three samplers measuring in parallel at the same location have been investigated for a range of different temporal sampling resolutions, the error associated with these resolutions and for different pollen concentration classes estimated, and the limits of Hirst-type trap measurements explored. New automatic systems and the forecast models, which use the real-time data they produce, are compared to these manual Hirst measurements for validation/evaluation

purposes. Understanding the intrinsic uncertainties of the Hirst traps is therefore crucial.

2 Material and methods

2.1 Experimental site

Three Hirst-type traps (Burkard Manufacturing, UK) were co-located on the roof of the Swiss Federal Office of Meteorology and Climatology MeteoSwiss building in Payerne, Switzerland. The roof is flat and at a height of 7 m above the surrounding ground. The traps were placed in a line, approximately 3 m from each other. The study ran from 26 March to 25 July 2013 and the three Hirst-type traps operated simultaneously for this entire period although only the observations from 1 April through to 30 June are considered here, covering the main season of the most important allergenic pollen taxa in Payerne.

2.2 Pollen measurements

The three Hirst-type traps (numbered 2, 4, and 8) were calibrated and set for seven-day sampling using Melinex tape coated with silicon (Lanzoni, Italy). After collection each week, tapes were cut into seven daily segments and mounted onto microscope slides. Two longitudinal scans were counted on each slide using an Olympus microscope with a $600\times$ magnification (4 and 8 mm from the top edge of the slide) and, at a later stage when the operational standard had changed to lower magnification (in order to increase the scanned surface), two additional lines (5 and 9 mm from the top edge of the slide) were counted at $400\times$ magnification. In total, an area of 8.64% of the slide was analysed. The scanned surface excluded the upper and lower parts of the field of view, where particles can easily be missed by the counter because they appear only briefly in the field of view. The final pollen concentration was obtained by taking the average concentration from each of the four lines (calculated according to CEN, 2019, item 6.3.6). Three individuals were responsible for the counts with the counts for each trap being split between the three (i.e. not one specific person per trap). This was important to reduce the impact of any human error.

The volume of air sampled was checked each week using a standard Burkard flowmeter each time the

drums were exchanged. Later, investigations showed that as pumps age, such flowmeters do not always measure correctly (Oteros et al., 2017). Laboratory studies with a more accurate electronic flowmeter at the Payerne observation site showed that the traps operated with a mean flow rate of 13.5 L/min. An average correction factor of 1.35 was thus applied to the flow rates to adapt the flows to the standard 10 L/min. While this may change the absolute pollen concentrations obtained, in this study we compare the results from the three traps; therefore, this correction factor has little impact on any conclusions obtained.

The following pollen taxa were referred to as “all pollen taxa” when evaluated together: *Alnus*, *Betula*, *Carpinus*, *Castanea*, *Corylus*, Cupressaceae, *Fagus*, *Fraxinus*, *Juglans*, *Picea*, *Pinus*, *Plantago*, *Platanus*, Poaceae, *Populus*, *Quercus*, *Rumex*, *Salix*, *Taxus*, *Ulmus*, and Urticaceae/Cannabaceae. The sum of all counted pollen is reported in the “total pollen” category. We also selected the five most frequently occurring pollen taxa during the three-month study period for an analysis based on single pollen types, namely *Betula* spp., *Fraxinus* spp., Poaceae, *Pinus* spp., and *Quercus* spp. (Table 1). Thereof, *Betula* spp., *Fraxinus* spp., and Poaceae are referred to as the “top three taxa”.

2.3 Statistical analysis

The pollen concentration data were aggregated over different time steps: 2-h, 3-h, 6-h, 12-h, and daily averages to investigate how error varies depending on the averaging window. At the highest temporal resolution considered (2-h), the data had no missing values. Since the pollen season depends on the plant taxa, a simple uniform definition was chosen for the period of occurrence of a particular taxa using only the time steps for which at least one of the three traps measured that specific taxa. All other time steps when no pollen of that taxa was present were disregarded for the analyses of individual taxa. This included any days that may have occurred during the season where all three traps measured no pollen, for example, due to strong precipitation events. Note that this may result in a slight overestimation of the error calculated since agreement between the three traps for any days when all three devices did not measure any pollen is not taken into account.

Table 1 Frequency of occurrence (number of days with pollen present), maximum concentration, average concentration, and pollen integral for the 21 most frequent pollen types observed

in Payerne from April–June 2013. Values for each of the three traps are shown and calculated using the period of occurrence for each taxa (as described in Sect. 2.3)

Taxa	Frequency of occurrence (# Days)			Maximum daily concentration (Pollen/m ³)			Average daily concentration (Pollen/m ³)			Pollen integral (Pollen*Day/m ³)		
	T2	T4	T8	T2	T4	T8	T2	T4	T8	T2	T4	T8
<i>Fraxinus</i>	48	50	51	2578	2748	3420	97	100	130	8682	8881	11,572
Poaceae	68	74	68	679	624	738	66	66	74	5908	5885	6557
<i>Betula</i>	51	67	61	726	568	999	36	31	48	3173	2728	4242
<i>Quercus</i>	59	65	63	676	597	802	38	32	42	3357	2811	3695
<i>Pinus</i>	56	57	59	182	166	185	19	18	21	1673	1582	1864
Cupressaceae	57	69	67	148	105	230	9	9	14	810	824	1244
<i>Fagus</i>	51	53	58	125	101	145	8	8	10	694	688	904
<i>Populus</i>	32	33	36	150	131	236	6	6	10	537	535	849
Urticaceae/Cannabaceae	33	36	48	87	78	100	8	7	9	685	666	819
<i>Taxus</i>	34	46	44	155	165	174	6	6	8	503	516	718
<i>Picea</i>	48	46	45	109	96	147	5	4	5	646	358	489
<i>Rumex</i>	61	60	61	20	23	24	3	4	4	298	344	346
<i>Carpinus</i>	33	35	31	39	30	65	3	2	4	292	172	329
<i>Salix</i>	41	45	43	27	18	31	2	2	3	206	194	296
<i>Alnus</i>	46	39	40	53	44	41	3	3	3	263	285	286
<i>Platanus</i>	31	27	36	47	67	71	2	2	2	158	190	211
<i>Plantago</i>	48	51	51	12	12	9	2	2	2	163	161	176
<i>Juglans</i>	36	36	34	15	20	26	1	2	2	103	158	160
<i>Ulmus</i>	18	18	19	19	11	28	1	1	1	81	59	127
<i>Corylus</i>	38	33	30	19	12	21	2	1	1	145	116	119
<i>Castanea</i>	2	3	4	1	1	2	0	0	0	1	2	4
Total	891	943	949	5867	5617	7494	317	306	393	28,378	27,155	35,007

The top five pollen taxa were selected based on the highest pollen integral (sum over the period of measurement). In addition, five concentration classes were defined: 0–9 pollen grains/m³, 10–19 pollen grains/m³, 20–49 pollen grains/m³, 50–99 pollen grains/m³, 100–299 pollen grains/m³, and ≥ 300 pollen grains/m³, with the aim of better understanding whether different pollen concentrations have an impact on the sampling and measurement error. The mean concentration of all individual pollen taxa across all three traps was calculated to define which group each particular time step belonged to. After an initial residual analysis, the measurements were converted into logarithmic concentrations for the statistical comparison.

Prior to all analyses, the normality of the data was tested with a complete residual analysis, inspecting QQ plots, Tukey–Anscombe plots, and Index plots. Thereby assessing whether the residuals are independent; normally distributed; have a constant variance and a mean of zero. The respective parametric or nonparametric tests have been chosen accordingly. The correlation between traps was calculated pairwise with three different approaches: Pearson, (1895) correlation coefficients for linear correlation between traps, Spearman, (1904), and Kendall, (1938) for rank-based correlations. The latter two are nonparametric tests and do not assume normality of the data. Parametric one-way ANOVA (Fisher, 1919) tests were also conducted. This test, however, imposes

several assumptions on the underlying data, including that they are independent and have normally distributed errors with a constant variance and a mean of zero. The robust Kruskal–Wallis rank test (Kruskal & Wallis, 1952) was therefore also applied since it has less strict assumptions: the samples must be mutually independent, measurements must be continuous on at least an ordinal scale and all traps should have the same shape distributions. Depending on the temporal averaging window, certain assumptions were violated and the appropriate choice of the statistical procedure is important. This is highlighted where relevant in the text.

Both ANOVA and Kruskal–Wallis are considered omnibus tests, testing whether the variance among the three different traps is greater than the unexplained variance (or the variance within a dataset from a particular trap). If the resulting explained variance is low, one can assume that the three traps are similar. Both omnibus tests can be followed-up by multiple pairwise tests between the traps. For the standard one-way ANOVA, Tukey's honest significant difference (HSD) test (Tukey, 1949) was applied to measure the differences of the mean. This original version of Tukey HSD assumes normally distributed data with constant variance. The rank-based Kruskal–Wallis test was followed-up by pairwise comparisons and computation of simultaneous confidence intervals for the estimated effects using the *nparrcomp* package (Konietschke et al., 2015) with the Tukey method. All statistical analyses were carried out using the R Language (R Core Team, 2021).

3 Results and Discussion

3.1 Pollen season and measurement frequency

An overview of the main characteristics of the 20 most frequent pollen taxa during the three-month study period is provided in Table 1. 2013 was a year with very high peak concentrations of *Fraxinus*, while the intensity of most other pollen taxa was in the range of the mean or slightly below, although the pollen season of *Betula* and *Fraxinus* was shorter than normal. *Fagus* was less frequently observed than average.

The consistency of measurement frequency among the three traps was considered for a range of timescales (Table 2). This metric gives a simple indication of how

often pollen were observed by the traps, but does not provide information on how well they agree in terms of the measured pollen concentrations of pollen counted (see the further analyses below).

For total pollen, the agreement between the three traps is excellent for both the 24- and 12-hourly averages, with all three traps measuring pollen 99% and 98% of the time, respectively. At the 6-hourly level, this value degrades to 93% and when 2-hourly averages are considered, it drops further to 85%. This means that if one were to consider 2-hourly averages using a single trap, there is an approximately 15% chance that no pollen will be observed even if it is present in the air.

The percentage of time steps for which all three traps measure an individual pollen taxa is considerably lower than for total pollen. Even for daily averages of the five most frequently occurring pollen taxa, this value ranges from 50% (*Betula*) to 84% (Poaceae). This indicates that there is markedly more error when considering individual pollen taxa, and there are much larger chances that a particular pollen will not be observed in a single trap despite it being present in the atmosphere. This is likely a result of the sub-sampling procedures used or human error in the counting process, or both. Pollen that is easy to recognise during times with lower pollen diversity on the slide, such as Poaceae, seems to be more consistently observed. Likewise, the duration and shape of the period of occurrence for each taxa also likely plays a role in the consistency across traps, with pollens that have shorter periods of occurrence and more days with lower concentrations also showing less consistency across traps (e.g. *Fraxinus*, *Betula*).

When one considers higher temporal resolutions, the percentage of time that a particular pollen taxon is observed across all three traps is even lower, ranging from 41 to 60% for the 2-hourly averages of the five most frequent taxa. This suggests that at such temporal resolutions, there is very high uncertainty when considering a single pollen taxon from a single trap, with a considerable chance (up to as much as 60%) of missing those observations entirely. Some of the potential reasons for these large differences are considered in the following sections.

Table 2 Frequency of measurement (in %) across the three traps for various temporal averaging periods for the five most frequent taxa as well as total pollen. The total pollen values are calculated for the period considered in this study (1 April–30 June 2013), while for the individual pollen taxa values are calculated for the period where at least one trap measured that specific pollen taxa (see Sect. 2.3 for details)

	Total	<i>Fraxinus</i>	Poaceae	<i>Betula</i>	<i>Quercus</i>	<i>Pinus</i>
<i>24-Hour Averages</i>						
All traps measured	99	61	84	50	61	76
Two traps measured	1	14	10	24	23	12
One trap measured	0	24	6	26	17	12
<i>12-h Averages</i>						
All traps measured	98	56	77	48	61	74
Two traps measured	2	19	14	25	20	10
One trap measured	0	25	10	27	19	16
<i>6-Hour Averages</i>						
All traps measured	93	52	71	43	51	64
Two traps measured	6	22	13	22	26	16
One trap measured	1	25	15	34	23	20
<i>2-Hour Averages</i>						
All traps measured	85	51	60	42	41	55
Two traps measured	11	17	18	20	25	18
One trap measured	4	32	22	39	34	27

3.2 Total pollen

3.2.1 Daily averages

There is a good agreement between the three traps in terms of measured daily average total pollen concentrations (Fig. 1a). The medians and IQR spread of the measurements are similar (Fig. 1b), as well as the distribution of daily concentrations (Fig. 1c–e). The standard deviations (text boxes included in Figs. 1c–e) are also comparable for traps 2 and 4; however, the value for trap 8 is higher than the other two traps (37% above the average of the other two). Nevertheless, results from three different pairwise correlation estimates suggest good agreement between the traps, with the 95% confidence interval for the correlation factors ranging from 0.72 to 0.98 (Table 3). This is also confirmed with the results of the ANOVA, which suggest no significant differences between traps for daily average total pollen concentrations (Table 4). Median pairwise differences for daily average total pollen are 16% (IQR = 8–32%) (Table A1, Supplementary Material).

Separating the daily average total pollen data into different concentration groups, it becomes clear that the differences between traps are similar, except for the smallest concentration group (0–9 pollen grains/m³) for which they become considerably larger (Fig. 2). For these low concentrations, the median

relative differences between traps are on average 112%, for total pollen, while for concentrations above or equal to 10 pollen grains/m³ the median differences range between 21 and 34%. The larger differences at low concentrations are in part related to the fact that even differences of 1–2 pollen grains can result in large percentage difference. However, it is also because at these low concentrations the traps do not always measure pollen at the same time (see Table 2 and Figure A1 supplementary material). This is not entirely unexpected, since it is known that at low concentrations counting and sampling errors become more important (Comtois et al., 1999; Gottardini et al., 2009; Oteros et al., 2013; Sikoparija et al., 2011). The pairwise mean relative differences for all concentrations above or equal to 10 pollen grains/m³ of 26% (standard deviation (SD) = 33%) agree with previous studies. Rojo et al., (2019) found a mean standard deviation of 34% for the ratio between traps for values ≥ 10 pollen grains/m³ while Comtois et al., (1999) determined an overall error of about 30% when counting a sub-sample of four longitudinal lines, but the error was highly correlated with the amount of pollen counted, decreasing with an increasing number of pollen counted pollen. Furthermore, Pedersen and Moseholm (1993) found relative errors of 47–73% over a 20-day observation period for four different pollen types, with the proportion of the error resulting

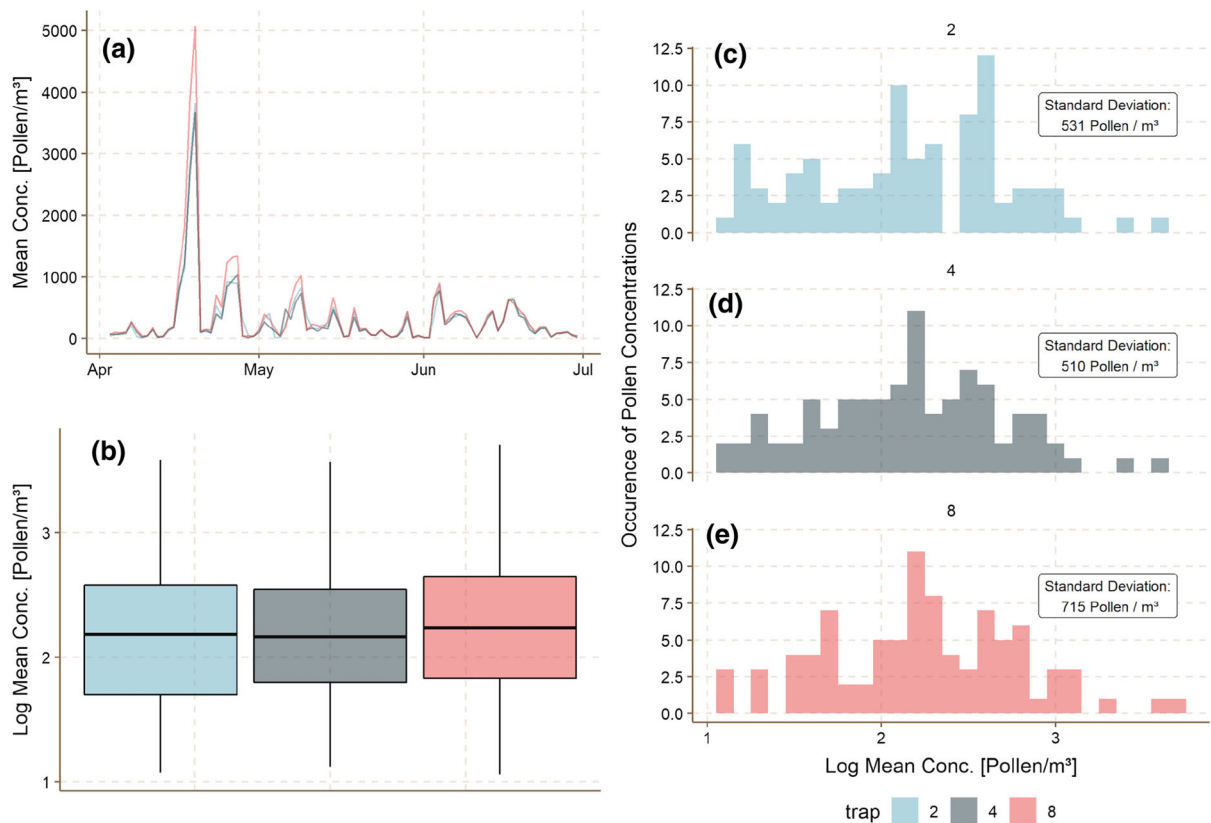


Fig. 1 Daily average total pollen concentrations for all three traps. **a** time series for the 2013 measurement period considered in this study, **b** box-whisker plots, and **c–e** histograms of all values for traps 2, 4, and 8, respectively, with bin widths of 0.1 on the log scale

Table 3 Correlations between the three pairs of traps for total pollen concentrations for both daily and 2-h averages

	Trap 2 vs. Trap 4		Trap 2 vs. Trap 8		Trap 4 vs. Trap 8	
	Total	Poaceae	Total	Poaceae	Total	Poaceae
<i>Daily Averages</i>						
R-Pearson	0.75–0.88	0.9–0.96	0.79–0.92	0.89–0.95	0.96–0.98	0.95–0.98
ρ -Spearman	0.86–0.94	0.91–0.97	0.88–0.95	0.91–0.96	0.95–0.98	0.94–0.98
τ -Kendall	0.72–0.87	0.73–0.9	0.74–0.89	0.72–0.88	0.8–0.93	0.78–0.93
<i>2-Hour Averages</i>						
ρ -Spearman	0.81–0.85	0.79–0.85	0.81–0.85	0.78–0.83	0.90–0.93	0.81–0.86
τ -Kendall	0.64–0.72	0.62–0.71	0.64–0.71	0.60–0.69	0.72–0.79	0.63–0.73

The 95% confidence intervals are given for the three different correlation estimates. *Note that visual assessment of the underlying data suggests that the assumption of normality does not hold for the 2-hourly resolution; hence, the R-Pearson correlation factors are not shown here

from the technicians contributing 2–13% of the total uncertainty.

The high variability in the daily average total pollen at concentrations less than 10 pollen grains/m³ is also

confirmed by the ANOVA analysis. The percentage of explained variance (PES) across the traps is less than 1% for this concentration range, with the remainder related to unexplained variance, or variance “internal”

Table 4 Results from the ANOVA test applied to the six concentration groups for both daily and 2-hourly averages

	Factor	Daily averages				2-hourly averages*			
		DF	SumSq	Pr(> F)	PES	DF	SumSq	Pr(> F)	PES
Total	Trap	2	3	44.90%	0.67%	2	57	0.03%	0.52%
	Residual	264	442			3114	10,991		
Group 0–9	Trap	2	7	0.09%	0.67%	2	173	0.00%	0.81%
	Residual	1950	1035			13,746	21,155		
Group 10–19	Trap	2	7	0.00%	10.30%	2	20	0.00%	1.35%
	Residual	327	61			2025	1463		
Group 20–49	Trap	2	2	1.31%	4.00%	2	36	0.00%	3.16%
	Residual	237	48			2217	1104		
Group 50–99	Trap	2	1	0.08%	11.1%	2	8	0.00%	3.03%
	Residual	108	8			1077	256		
Group 100–299	Trap	2	1	0.41%	8.33%	2	9	0.00%	3.37%
	Residual	96	11			1164	258		
Group ≥ 300	Trap	2	1	21.50%	4.00%	2	11	0.00%	4.00%
	Residual	66	24			654	297		

DF = degrees of freedom, SumSq = sum of the squares of the residuals, $Pr(> F)$ = the probability that the F -test statistic is < 0.05 , PES = percentage explained variance. Pr ($> F$) values highlighted in bold are significant at the 95% confidence level. *Note that visual assessment of the underlying data suggests that the assumption of normality does not hold for the 2-hourly averages for concentrations between 0 and 9, findings were similar by robust methods not shown here

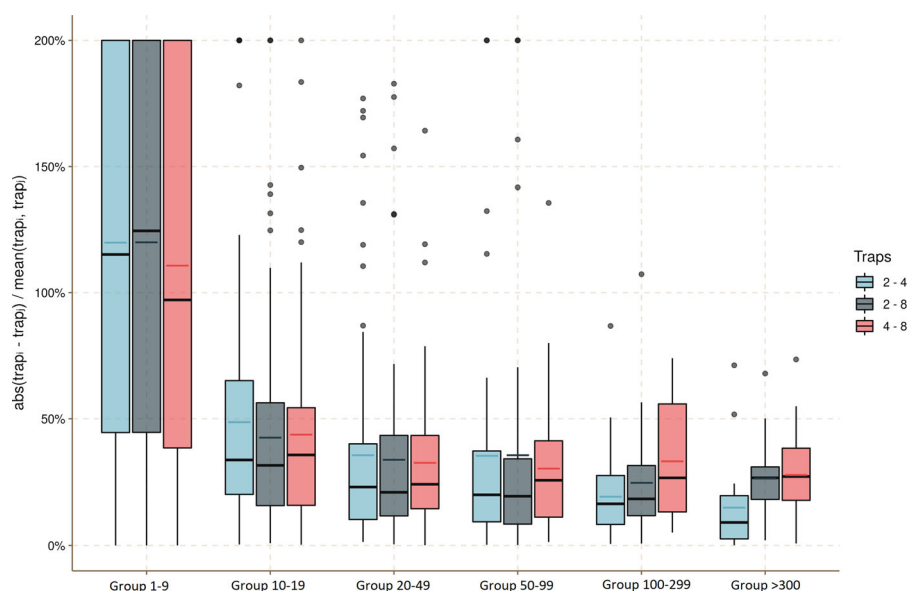


Fig. 2 Box plots of the percentage differences between pairs of traps for daily average concentrations of all pollen taxa (individually) separated into six different concentration groups (from left to right: 0–9, 10–19, 20–49, 50–99, 100–299,

and ≥ 300 pollen grains/m³). The thick black lines in the boxes indicate the median values, the coloured lines the means, while the boxes cover the inter-quartile range. Note: by definition the maximum percentage difference cannot be greater than 200%

to each individual trap. For higher daily average total pollen concentrations, there is less internal variance, that is the data are less noisy, and a larger proportion of the variability is explained by the differences between traps, although this does not exceed 11% for any of the concentration groups. The ANOVA results also show that the *F*-test statistic is less than 5% (the *Pr* ($> F$) value), suggesting that the null hypothesis that the traps are similar should be rejected. This is confirmed by the robust *np*comp test, which indicates that there are significant differences between the three traps for all daily average concentration groups for total pollen, except for the highest group with > 300 pollen grains/ m^3 (see Table A2, supplementary material).

While the results for total pollen daily averages indicate no significant differences between the three traps, when separated into concentration groups, they do become significant. This may simply be a result of selecting smaller portions of data for which there is, by definition, less variability within the group itself, making the differences between traps more obvious (and significant). While a somewhat artificial separation, it is important to understand that the uncertainty across the measurements does seem to vary depending on the concentration (Fig. 2) and this needs to be taken into consideration when comparing either data from two or more Hirst-type traps or other instruments.

3.2.2 Two-hourly averages

There are considerably larger differences between the three traps in terms of two-hourly averages for all pollen (Fig. 3). As for the daily averages, the differences between traps are largest at low concentrations; however, the percentage differences between traps are well over 50% up until 50 pollen grains/ m^3 . Even at higher concentrations, the differences are larger than for the daily averages, with considerably more spread in the IQR. For the very lowest concentrations, the median percentage differences are 200% (by definition the largest value possible when comparing two traps) indicating very high variability between traps. These results are confirmed by the ANOVA test, which shows that the explained variance between traps is lower than the daily averages, with PES values less than 4% for all concentrations. This indicates that there is more residual variance or variability within the data from each trap than between the traps. In addition, the *F*-test statistic shows that the traps are likely

different from one another (Table 4), and the robust Tukey HSD contrast test confirms this, with statistically significant differences between traps for all concentrations (Table A2 supplementary material).

Median relative differences for total pollen are 42% (IQR = 20–88%) and 62–120% for the top three taxa, respectively (Table A1, Supplementary Material). Again, when the lowest concentrations are not included, these differences decrease to 40% for total pollen (IQR = 19–78%) and 41–56% for the top three taxa, respectively. A large proportion of the uncertainty thus stems from the measurements at low concentrations. When there are less than 10 pollen grains/ m^3 , uncertainty for the median relative differences of 2-hourly averages is over 112% (IQR: 43–200%). This is not entirely unexpected, since it is known that at low concentrations counting and sampling errors become more important (Comtois et al., 1999; Gottardini et al., 2009; Oteros et al., 2013; Sikoparija et al., 2011).

3.3 Individual Pollen Taxa

The five most frequently occurring pollen taxa in Payerne during the 2013 season were *Fraxinus*, Poaceae, *Betula*, *Quercus*, and *Pinus* (see Table 1). For the sake of brevity, in this section we focus mainly on Poaceae since this taxon is one of the most frequently occurring pollens and also one causing significant allergy in Switzerland (Burbach et al., 2009). Results for the other most frequently occurring taxa are shown in the supplementary material (Figures A5–A6).

3.3.1 Daily average concentrations

As for total pollen, daily average concentrations for Poaceae are relatively consistent between all three traps. Percentage differences between the pairs of traps are around or below 30% for concentrations greater or equal to 10 pollen grains/ m^3 (Fig. 4). Interestingly, the spread of the percentage differences for concentrations above 20 pollen grains/ m^3 is smaller for Poaceae than for all pollen (Fig. 4 compared to Fig. 2). This may be a result of grasses being more widespread across Switzerland (resulting in a smoother concentration time series) than most other pollen sources and because the blooming season is relatively long, with a variety of species flowering

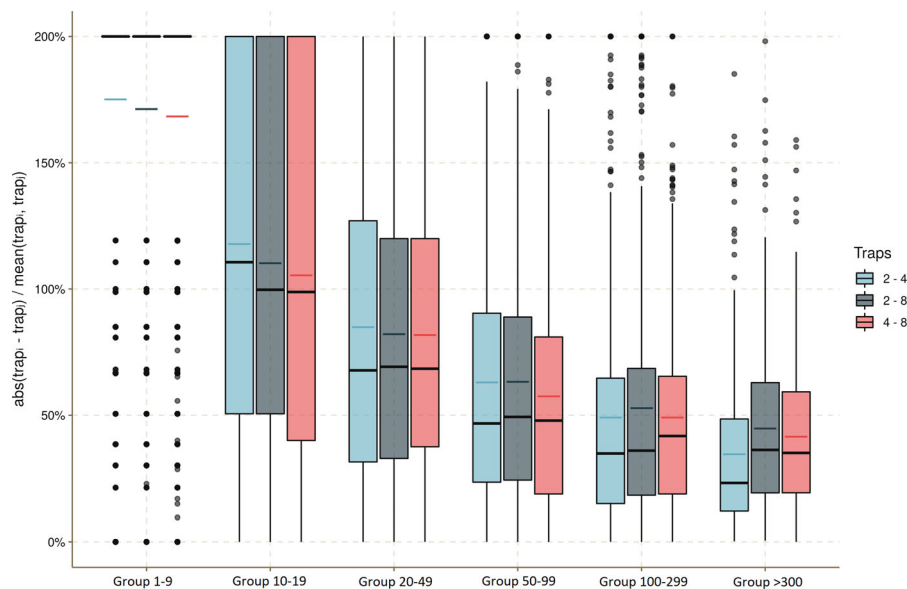


Fig. 3 Box plots of the percentage differences between pairs of traps for 2-hourly average concentrations of all pollen taxa for six different concentration groups (from left to right: 0–9, 10–19, 20–49, 50–99, 100–299, and ≥ 300 pollen grains/ m^3).

The thick black lines in the boxes indicate the median values, the coloured lines the means, while the boxes cover the inter-quartile range. Note: by definition the maximum percentage difference cannot be greater than 200%

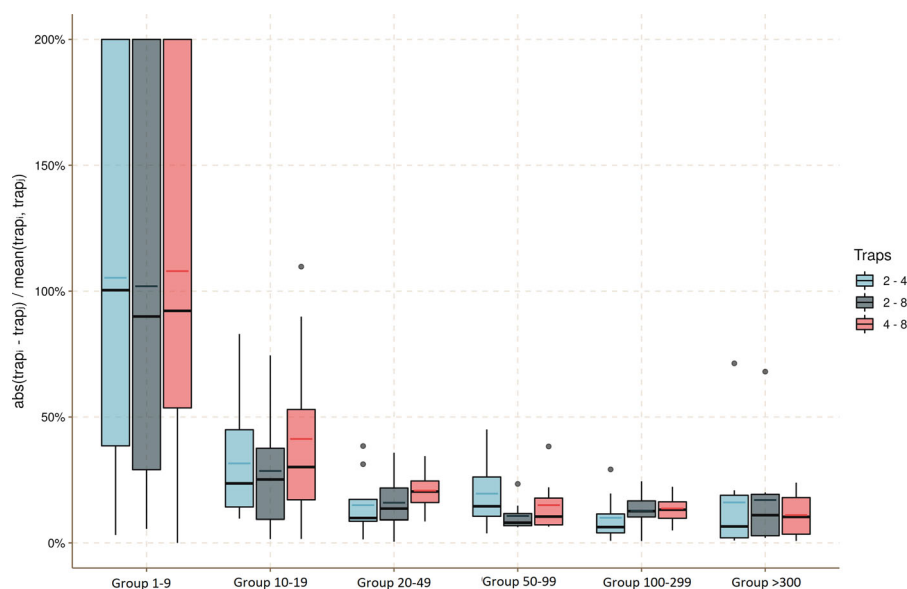


Fig. 4 Box plots of the percentage differences between pairs of traps for daily average concentrations of Poaceae for the six different concentration groups considered (from left to right: 0–9, 10–19, 20–49, 50–99, 100–299, and ≥ 300 pollen grains/ m^3).

The thick black lines in the boxes indicate the median values, the coloured lines the means, while the boxes cover the IQR range. Note: by definition the maximum percentage difference cannot be greater than 200%

one after another, allowing a larger sample size and thus meaningful means/medians. For the lowest concentrations, however, the differences reach close to 100% on average, with a very large spread, indicating

that there is high uncertainty associated with measurements of a single pollen taxon at concentrations below 10 pollen grains/ m^3 .

In terms of daily averages, there is only a very small proportion of the variance explained by the differences between traps (PES values < 1%) for the most frequent pollen taxa in Payerne (Table 4). The ANOVA and robust contrast tests likewise show that for these individual pollen taxa differences between traps are not significant when daily averages are considered (Table 5 and Table A4, supplementary material). This may simply be a result of selecting smaller portions of data for which there is, by definition, less variability within the group itself, making the differences between traps more obvious (and significant).

3.3.2 Two-hourly averages

As for all pollen taxa, there are considerably larger differences for the 2-hourly resolution Poaceae observations compared to the daily averages (Fig. 5). Relative differences are below 50% at concentrations above 50 pollen grains/m³, but considerably larger for concentrations below this level. The differences remain relatively consistent between the pairs of traps, with median values and IQRs being similar; however, the large percentage differences indicate high uncertainty in the measured values. The larger spread in values is in part related to the fact that for individual pollen taxa at the two-hourly resolution there are a

high proportion of data points for which only one of the three traps measured a particular taxon (Table 3). For concentrations below 50 pollen grains/m³, percentage differences between traps are very large, reaching up to 200% for the lowest concentrations (< 10 pollen grains/m³).

Again, as for all pollen taxa, the results from the ANOVA show that less than 1% of the variance for the most frequently occurring taxa results from differences between the traps (PES values in Table 4) and the large majority rather from variance within the data from each trap. For the two most frequent taxa, Poaceae and *Fraxinus*, the *F*-test statistic indicates that there are no significant differences between the traps and this is confirmed by the robust contrast test (Table A3 supplementary material). For *Betula*, *Quercus*, and *Pinus*, the *F*-test statistic indicates that there are potential differences between the traps (Table 5), which are shown to be statistically significant with the robust contrast test for just *Betula* and *Pinus* (Table A3 supplementary material).

3.4 Pollen size

A similar analysis was carried out for pollen size classes, with the aim of understanding whether particle size had any impact on the variability between traps. The 21 most frequently occurring taxa were separated

Table 5 Results from the ANOVA test applied to the five most frequently occurring pollen taxa for both daily and 2-hourly averages

	Factor	Daily averages				2-hourly averages*			
		DF	SumSq	Pr(> F)	PES	DF	SumSq	Pr(> F)	PES
<i>Fraxinus</i>	Trap	2	2	83.90%	0.26%	2	19	21.70%	0.25%
	Residual	144	766			1209	7503		
Poaceae	Trap	2	1	84.90%	0.15%	2	13	20.70%	0.15%
	Residual	207	661			2034	8485		
<i>Betula</i>	Trap	2	4	59.30%	0.66%	2	37	1.58%	0.60%
	Residual	171	602			1377	6108		
<i>Quercus</i>	Trap	2	1	91.10%	0.17%	2	25	4.79%	0.39%
	Residual	195	601			1569	6377		
<i>Pinus</i>	Trap	2	1	71.80%	0.27%	2	24	2.01%	0.53%
	Residual	171	370			1482	4460		

DF = degrees of freedom, SumSq = sum of the squares of the residuals, Pr(> F) = the probability that the *F*-test statistic is < 0.05, PES = percentage explained variance. Pr(> F) values highlighted in bold are significant at the 95% confidence level. *Note that visual assessment of the underlying data suggests that the assumption of normality may not hold for this test at the 2-hourly resolution and results should be treated with some caution. Robust methods produced similar results, but are not shown here

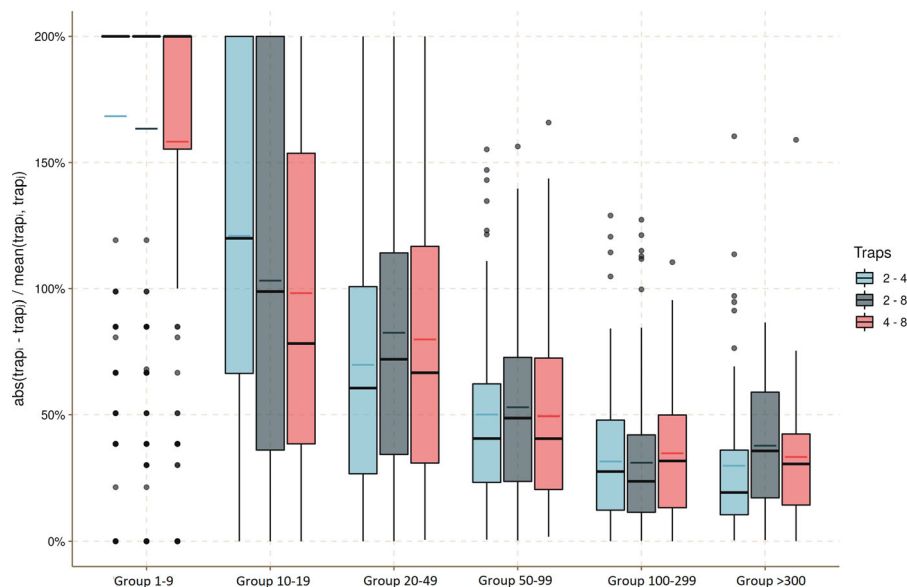


Fig. 5 Box plots of the percentage differences between pairs of traps for 2-hourly average Poaceae for the six different concentration groups considered (from left to right: 0–9, 10–19, 20–49, 50–99, 100–299, and ≥ 300 pollen grains/m³).

into five groups depending on the average diameter as provided by Beug, (2004). The groups were as follows: Group 1: *Urticaceae/Cannabaceae, Castanea*; Group 2: *Alnus, Betula, Corylus, Fraxinus, Platanus, Rumex, Salix*; Group 3: *Plantago, Poaceae, Populus, Taxus, Cupressaceae, Ulmus*; Group 4: *Carpinus, Fagus, Juglans, Quercus*; Group 5: *Pinus, Picea*. Results from the statistical tests show no statistical differences for daily averages; however, when 2-hourly averages are considered there are some statistical differences, notably for the medium-sized particles (see Table A4 supplementary material). However, as mentioned above, there is such high variance within the 2-hourly data that these results should be considered with caution.

4 Conclusions

In this paper, we use data from three Hirst-type traps run in parallel at the same site to provide an estimate of the measurement uncertainty at various temporal resolutions. The campaign set-up, with, for example, several human operators each counting for each of the three traps, means that our results can be expected to

The thick black lines in the boxes indicate the median values, the coloured lines the means, while the boxes cover the IQR range of differences. Note: by definition the maximum percentage difference cannot be greater than 200%

provide a realistic representation of the real-life variability and error within an operational network.

For daily average total pollen concentrations, the traps show similar results and are not statistically different from one another. The median pairwise relative differences are 16% for total pollen and between 23 and 67% for individual taxa. The values are identical for total pollen when only concentrations ≥ 10 pollen grains/m³ are considered, but decrease notably for the individual taxa investigated, with median relative differences ranging between 14 and 30%. The pairwise mean relative differences for all concentrations above or equal to 10 pollen grains/m³ of 26% are in agreement with previous studies (e.g. Buters et al., 2012; Comtois et al., 1999; Pedersen & Moseholm, 1993; Rojo et al., 2019). At the 2-hourly resolution, there is considerably more variability between the three samplers, with median relative differences of 42% for total pollen and 62–120% for individual taxa, respectively. At concentrations lower than 10 pollen grains/m³, uncertainty for the median relative differences of 2-hourly averages is over 112%, reaching up to 200% for individual taxa (the maximum value possible for pairwise comparisons). Excluding these observations would obviously reduce uncertainty; however, this is problematic for certain highly

allergenic pollen taxa, such as *Ambrosia*, that are thought to provoke allergic symptoms at concentrations below this limit. Finally, pollen size seems to have little to no influence on the measurement variability and all three traps showed similar behaviour for the size range investigated. It remains to be investigated if the uncertainty found in this study focused on pollen is similar for observations of fungal spores.

The uncertainty in the Hirst-type measurements, particularly at sub-daily scales, to some extent complicates the comparison with automatic pollen monitors, which represent such high temporal resolutions well (e.g. Chappuis et al., 2019). Nevertheless, taking into account the uncertainty associated with each measurement type, manual and automatic, allows more robust evaluations and helps focus on the most relevant temporal resolutions for comparison. Such measurement uncertainties are also of particular importance in trend analyses and for comparing measurements at different sites. Counting a larger surface area of the slides (greater than the 10% counted in this study and considered as standard) may help reduce the error associated with the measurements at low concentrations. However, it remains to be investigated just what surface area needs to be counted to reduce errors to sufficiently low levels, particularly at high temporal resolutions (e.g. 2-hourly).

Acknowledgements This paper is a contribution to the EUMETNET AutoPollen Programme, which is developing a prototype automatic pollen monitoring network in Europe covering all aspects of the information chain from measurements through to communicating information to the public.

References

- Berggren, B., Nilsson, S., & Boëthius, G. (1995). Diurnal variation of airborne birch pollen at some sites in Sweden. *Grana*, 34, 251–259.
- Beug, H.-J. (2004). *Leitfaden der Pollenbestimmung für Mitteleuropa und angrenzende Gebiete*. Verlag Friedrich Pfeil.
- Borycka, K., & Kasprzyk, I. (2018). Hourly pattern of allergenic alder and birch pollen concentrations in the air: Spatial differentiation and the effect of meteorological conditions. *Atmospheric Environment*, 182, 179–192.
- Burbach, G. J., Heinzerling, L. M., Edenharter, G., Bachert, C., Bindeslev-Jensen, C., Bonini, S., Bousquet, J., Bousquet-Rouanet, L., Bousquet, P. J., Bresciani, M., Bruno, A., Canonica, G. W., Darsow, U., Demoly, P., Durham, S., Fokkens, W. J., Giavi, S., Gjomarkaj, M., Gramiccioni, C., ... Zuberbier, T. (2009). GA2LEN skin test study II: Clinical relevance of inhalant allergen sensitizations in Europe. *Allergy*, 64, 1507–1515.
- Buters, J. T. M., Thibaudon, M., Smith, M., Kennedy, R., Rantio-Lehtimäki, A., Albertini, R., Reese, G., Weber, B., Galán, C., Brandao, R., Antunes, C. M., Jäger, S., Berger, U., Celenk, S., Grewling, Ł., Jackowiak, B., Sauliene, I., Weichenmeier, I., Pusch, G., ... Cecchi, L. (2012). Release of Bet v 1 from birch pollen from 5 European countries. Results from the HIALINE study. *Atmospheric Environment*, 55, 496–505.
- Buters, J. T. M., Antunes, C., Galveias, A., Bergmann, K. C., Thibaudon, M., Galán, C., Schmidt-Weber, C., & Oteros, J. (2018). Pollen and spore monitoring in the world. *Clinical Translational Allergy*. <https://doi.org/10.1186/s13601-018-0197-8>
- Cariñanos, P., Emberlin, J., Galán, C., & Dominguez-Vilches, E. (2000). Comparison of two pollen counting methods of slides from a hirst type volumetric trap. *Aerobiologia*, 16, 339.
- CEN/EN 16868:2019 (2019). Ambient air - Sampling and analysis of airborne pollen grains and fungal spores for networks related to allergy networks – Volumetric Hirst method. European Standard, European Committee for Standardisation, Brussels, Belgium, 38p.
- Clot, B. (2001). Airborne birch pollen in Neuchâtel (Switzerland): Onset, peak and daily patterns. *Aerobiologia*, 17, 25–29.
- Comtois, P., Alcazar, P., & Néron, D. (1999). Pollen counts statistics and its relevance to precision. *Aerobiologia*, 15, 19–28.
- Cotos-Yáñez, T. R., Rodríguez-Rajo, F. J., Pérez-González, A., Aira, M. J., & Jato, V. (2013). Quality control in aerobiology: Comparison different slide reading methods. *Aerobiologia*, 29, 1–11.
- Crouzy, B., Stella, M., Konzelmann, T., Calpini, B., & Clot, B. (2016). All-optical automatic pollen identification: Towards an operational system. *Atmospheric Environment*, 140, 202–212.
- Cunha, M., Ribeiro, H., & Abreu, I. (2016). Pollen-based predictive modelling of wine production: Application to an arid region. *European Journal of Agronomy*, 73, 42–54.
- de Weger, L. A., Bergmann, K. C., Rantio-Lehtimäki, A., Dahl, A., Buters, J., Déchamp, C., Belmonte, J., Thibaudon, M., Cecchi, L., Besancenot, J. P., Galán, C., & Waisel, Y. (2013). Impact of pollen. In M. Sofiev & K. C. Bergmann (Eds.), *Allergenic pollen: A review of the production, release, distribution and health impacts* (pp. 161–215). Springer Science+Business Media.
- Ekeboom, A., Nilsson, S., Saar, M., & Van Hage-Hamsten, M. (1997). A comparative study of airborne pollen concentrations of three allergenic types in Tartu (Estonia), Roma/Gotland and Stockholm (Sweden) 1990–1996. *Grana*, 36, 366–372.
- Fisher, R. A. (1919). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399–433.
- Frenz, D. A. (1999). Comparing pollen and spore counts collected with the Rotorod Sampler and Burkard spore trap. *Annals of Allergy, Asthma and Immunology*, 83, 341–349.

- Galán, C., Smith, M., Thibaudon, M., Frenguelli, G., Oteros, J., Gehrig, R., Berger, U., Clot, B., Brandao, R., & EAS QC Working Group. (2014). Pollen monitoring: minimum requirements and reproducibility of analysis. *Aerobiologia*, 30, 385–395.
- Gottardini, E., & Cristofolini, F. (1997). Spring airborne pollen data in two sites in Trentino (Northern Italy): A comparison with meteorological data. *Aerobiologia*. <https://doi.org/10.1007/BF02694508>
- Gottardini, E., Cristofori, A., Vannini, A., & Ferretti, M. (2009). Sampling bias and sampling errors in pollen counting in aerobiological monitoring in Italy. *Journal of Environmental Monitoring*, 11, 751–755.
- Hirst, J. M. (1952). An automatic volumetric spore trap. *Annals of Applied Biology*, 39, 257–265.
- Irdi, G. A., Jones, J. R., & White, C. M. (2002). Pollen and fungal spore sampling and analysis. Statistical evaluations. *Grana*, 41(1), 44–47.
- Käpylä, M., & Penttinen, A. (1981). An evaluation of the microscopical counting methods of the tape in Hirst-Burkard pollen and spore trap. *Grana*, 20, 131–141.
- Karrer, G., Skjøth, C. A., Šikoparija, B., Smith, M., Berger, U., & Essl, F. (2015). Ragweed (*Ambrosia*) pollen source inventory for Austria. *Science of the Total Environment*, 523, 120–128.
- Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30, 81–93.
- Konietschke, F., Placzek, M., Schaarschmidt, F., & Hothorn, L. A. (2015). nparcomp: An R software package for non-parametric multiple comparisons and simultaneous confidence intervals. *Journal of Statistical Software*, 64, 1–17.
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47, 583–621.
- Mandrioli, P., Comtois, P., & Levizzani, V. (1998). *Methods in Aerobiology* (p. 262p). Pitagora Editrice.
- Meiffren, I. (1988). Airborne pollen of Toulouse, southern France. Comparison with Bordeaux and Montpellier. *Grana*, 27, 183–201.
- Mitakakis, T. Z., & McGee, P. A. (2000). Reliability of measures of spores of *Alternaria* and pollen concentrations in air over two towns in rural Australia. *Grana*, 39, 141–145.
- Oteros, J., Galán, C., Alcazar, P., & Dominguez-Vilches, E. (2013). Quality control in bio-monitoring networks, Spanish Aerobiology Network. *Science of the Total Environment*, 443, 559–565.
- Oteros, J., Orlandi, F., García-Mozo, H., Aguilera, F., Ben, D. A., Bonofiglio, T., Abichou, M., Ruiz-Valenzuela, L., Trigo, M. M., Díaz de la Guardia, C., Domínguez-Vilches, E., Msallem, M., Marco Fornaciari, M., & Galán, C. (2014). Better prediction of Mediterranean olive production using pollen-based models. *Agronomy and Sustainable Development*, 34, 685–694.
- Oteros, J., Buters, J., Laven, G., Röseler, S., Wachter, R., Schmidt-Weber, C., & Hofmann, F. (2017). Errors in determining the flow rate of Hirst-type pollen traps. *Aerobiologia*, 33, 201–210.
- Pawankar, R., Canonica, G. W., Holgate, S. T., Lockey, R. F., & Blaiss, M. (2013). World Allergy Organisation (WAO) White Book on Allergy: Update 2013. Milwaukee World Allergy Organisation.
- Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58, 240–242.
- Pedersen, B. V., & Moseholm, L. (1993). Precision of the daily pollen count. Identifying sources of variation using variance component models. *Aerobiologia*, 9, 15–26.
- Prtenjak, M. T., Srnc, L., Peternel, R., Madzarevic, V., Hrga, I., & Stjepanovic. . (2012). Atmospheric conditions during high ragweed pollen concentrations in Zagreb, Croatia. *International Journal of Biometeorology*, 56, 1145–1158.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rizzi, L., Pizzulin, M., & Larese, F. (1992). Comparison between the allergenic airborne pollen in Trieste and at Lozzo di Cadore (Italy) in 1989. *Aerobiologia*, 8, 385–391.
- Rojó, J., Oteros, J., Perez-Badia, R., Cervigon, P., Ferencova, Z., Gutierrez-Bustillo, M., Bergmann, K.-C., Oliver, G., Thibaudon, M., Albertini, R., Rodriguez-De la Cruz, D., Sanchez-Reyes, E., Sanchez-Sanchez, J., Pessi Jukka Reiniharju, A.-M., Saarto, A., Calderon, M.C., Guerrero, C., Berra, D., Bonini, M., Chiodini, E., Fernandez-Gonzalez, D., Garcia, J., Trigo, M.M., Myszkowska, D., Fernandez-Rodriguez, S., Tormo-Molina, R., Damialis, A., Haering, F., Traidl-Hoffmann, C., Severova, E., Caeiro, E., Ribeiro, H., Magyar, D., Makra, L., Udvardy, O., Alcazar, P., Galán, C., Borycka, K., Kasprzyk, I., Newbigin, E., Adams-Groom, B., Apangu, G.P., Frisk, C.A., Skjoth, C., Radišić, P., Šikoparija, B., Celenk, S., Schmidt-Weber, C., & Buters, J. (2019). Near-ground effect of height on pollen exposure. *Environment Research*, 174, 160–169.
- Sarda Estève, R., Baisnée, D., Guinot, B., Petit, J.-E., Sodeau, J., O'Connor, D., Besancenot, J.-P., Thibaudon, M., & Gros, V. (2018). Temporal variability and geographical origins of airborne pollen grains concentrations from 2015 to 2018 at Saclay, France. *Remote Sensing*, 10, 1932.
- Sauliene, I., Sukiene, L., Daunys, G., Valiulis, G., Vaitkevicius, L., Matavulj, P., Brdar, S., Panic, M., Sikoparija, B., Clot, B., Crouzy, B., & Sofiev, M. (2019). Automatic pollen recognition with the rapid-E particle counter: The first-level procedure, experience and next steps. *Atmospheric Measurement Techniques*, 12, 3435–3452.
- Sauvageat, E., Zeder, Y., Auderset, K., Calpini, B., Clot, B., Crouzy, B., Konzelmann, T., Lieberherr, G., Tummon, F., & Vasilatou, K. (2020). Real-time pollen monitoring using digital holography. *Atmospheric Measurement Techniques*, 13, 1–12.
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology and Evolution*, 3, 430–439.
- Sikoparija, B., Smith, M., Skjøth, C. A., Radišić, P., Milkovska, S., & Brandt, J. (2009). The Pannonian plain as a source of *Ambrosia* pollen in the Balkans. *International Journal of Biometeorology*, 53, 263–272.
- Sikoparija, B., Pejak-Sikoparija, T., Radišić, P., Smith, M., & Galán Soldevilla, C. (2011). The effect of changes to the method of estimating the pollen count from aerobiological samples. *Journal of Environmental Monitoring*, 13, 384–390.

- Sikoparija, B., Galán, C., Smith, M., & EAS QC Working Group. (2017). Pollen-monitoring: Between analyst proficiency testing. *Aerobiologia*, 33, 191–199.
- Smith, M., Oteros, J., Schmidt-Weber, C., & Buters, J. T. M. (2019). An abbreviated method for the quality control of pollen counters. *Grana*, 58, 185–190.
- Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72–101.
- Tormo Molina, R., Munoz Rodriguez, A., & Silva Palacios, I. (1996). Sampling in aerobiology. Differences between traverses along the length of the slide in Hirst spore traps. *Aerobiologia*, 12, 161–166.
- Tormo Molina, R., Maya Manzano, J. M., Fernandez Rodriguez, S., Gonzalo Garijo, A., & Silva Palacios, I. (2013). Influence of environmental factors on measurements with Hirst spore traps. *Grana*, 52, 59–70.
- Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5, 99–114.
- Zhang, Y., Bielory, L., Mi, Z., Robock, A., & Georgopoulos, P. (2015). Allergenic pollen season variations in the past two decades under changing climate in the United States. *Global Change Biology*, 21, 1582–1589.
- Ziello, C., Sparks, T. H., Estrella, N., Belmonte, J., Bergmann, K. C., Bucher, E., Brighetti, M. A., Damialis, A., Detandt, M., Galán, C., Gehrig, R., Grewling, L., Gutierrez Bustillo, A. M., Hallsdottir, M., Kckhans-Bieda, M.-C., De Linares, C., Myszkowska, D., Paldy, A., Sanchez, A., ... Menzel, A. (2012). Changes to airborne pollen counts across Europe. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0034076>
- Ziska, L. H., Makra, L., Harry, S. K., Bruffaerts, N., Hendrickx, M., Coates, F., Saarto, A., Thibaudon, M., Oliver, G., Damialis, A., Charalampopoulos, A., Vokou, D., Heide-marsson, S., Gudjohnsen, E., Bonini, M., Oh, J., Sullivan, K., Ford, L., Brooks, G. D., ... Crimmins, A. R. (2019). Temperature-related changes in airborne allergenic pollen abundance and seasonality across the northern hemisphere: A retrospective data analysis. *The Lancet Planetary Health*, 3(3), e124–e131.

Supplementary Material

24-hour Averages						
Species	Threshold	Mean	SD	q25	Median	q75
<i>Betula</i>	None	98%	78%	30%	67%	200%
<i>Fraxinus</i>	None	93%	78%	23%	66%	200%
Poaceae	None	56%	66%	10%	23%	79%
Total	None	26%	33%	8%	16%	32%
All Species	None	97%	77%	26%	67%	200%
<i>Betula</i>	>= 10 Grains	40%	43%	13%	30%	47%
<i>Fraxinus</i>	>= 10 Grains	32%	34%	14%	25%	34%
Poaceae	>= 10 Grains	20%	19%	7%	14%	23%
Total	>= 10 Grains	26%	33%	8%	16%	32%
All Species	>= 10 Grains	36%	37%	13%	25%	45%
2-hour Averages						
Species	Threshold	Mean	SD	q25	Median	q75
<i>Betula</i>	None	120%	79%	40%	120%	200%
<i>Fraxinus</i>	None	103%	80%	30%	72%	200%
Poaceae	None	91%	76%	25%	62%	200%
Total	None	64%	60%	20%	42%	88%
All Species	None	128%	78%	47%	200%	200%
<i>Betula</i>	>= 10 Grains	75%	61%	26%	56%	111%
<i>Fraxinus</i>	>= 10 Grains	61%	56%	20%	43%	82%
Poaceae	>= 10 Grains	60%	56%	19%	41%	82%
Total	>= 10 Grains	58%	53%	19%	40%	78%
All Species	>= 10 Grains	81%	65%	28%	64%	120%

Table A1: Basic statistics for the top three pollen taxa, total pollen, and the sum of the main species counted (see main text for the list), where SD is the standard deviation and q25 and q75 the 25th and 75th quartile values, respectively.

Species	Traps	Daily Averages				Two-Hour Averages			
		Mean-Difference	Lower	Upper	p-Value	Mean-Difference	Lower	Upper	p-Value
Total	4-2	0.498	0.398	0.599	1.000	0.506	0.476	0.536	0.879
	8-2	0.542	0.440	0.641	0.598	0.536	0.506	0.566	0.008

Species	Traps	Daily Averages				Two-Hour Averages			
		Mean-Difference	Lower	Upper	p-Value	Mean-Difference	Lower	Upper	p-Value
	8-4	0.542	0.440	0.640	0.604	0.530	0.501	0.560	0.043
Group 1_9	4-2	0.071	-0.024	0.165	0.187	0.006	-0.054	0.067	0.968
	8-2	0.152	0.057	0.246	0.001	0.241	0.180	0.302	0
	8-4	0.081	-0.014	0.176	0.112	0.235	0.174	0.296	0
Group 10_19	4-2	-0.082	-0.218	0.055	0.342	0.020	-0.088	0.129	0.9
	8-2	0.256	0.119	0.393	0	0.220	0.111	0.328	0
	8-4	0.337	0.200	0.474	0	0.199	0.091	0.308	0
Group 20_49	4-2	-0.022	-0.189	0.146	0.95	0.034	-0.052	0.120	0.623
	8-2	0.171	0.003	0.339	0.044	0.285	0.199	0.371	0
	8-4	0.193	0.025	0.360	0.02	0.251	0.165	0.337	0
Group 50_99	4-2	-0.051	-0.204	0.101	0.703	-0.056	-0.142	0.029	0.27
	8-2	0.186	0.034	0.339	0.012	0.153	0.068	0.238	0
	8-4	0.238	0.085	0.390	0.001	0.209	0.124	0.294	0
Group 100_299	4-2	-0.167	-0.370	0.035	0.126	-0.089	-0.168	-0.010	0.023
	8-2	0.122	-0.081	0.324	0.329	0.129	0.049	0.208	0
	8-4	0.289	0.086	0.491	0.003	0.217	0.138	0.297	0
Group >300	4-2	-0.012	-0.438	0.414	0.997	-0.040	-0.191	0.111	0.807
	8-2	0.267	-0.159	0.693	0.297	0.249	0.098	0.400	0
	8-4	0.279	-0.147	0.705	0.266	0.289	0.138	0.440	0

Table A2: Results from the robust contrast test between pairs of Hirst-type traps for 2-hourly and daily averages of total pollen for the six different concentration groups considered. The mean difference (log-transformed) as well as the lower and upper bounds of the 95% confidence range of these values are shown. Statistically significant p-values are shown in red.

Species	Traps	Daily Averages				Two-Hour Averages			
		Estimator	Lower	Upper	p-Value	Estimator	Lower	Upper	p-Value
<i>Fraxinus</i>	4-2	0.496	0.361	0.631	0.998	0.484	0.437	0.531	0.697
	8-2	0.539	0.401	0.670	0.794	0.518	0.471	0.565	0.634
	8-4	0.547	0.410	0.678	0.702	0.535	0.488	0.582	0.19

Species	Traps	Daily Averages				Two-Hour Averages			
		Estimator	Lower	Upper	p-Value	Estimator	Lower	Upper	p-Value
Poaceae	4-2	0.479	0.367	0.594	0.908	0.481	0.445	0.518	0.457
	8-2	0.514	0.400	0.627	0.956	0.508	0.471	0.544	0.872
	8-4	0.534	0.419	0.645	0.776	0.526	0.489	0.562	0.223
Betula	4-2	0.545	0.417	0.667	0.694	0.508	0.464	0.551	0.911
	8-2	0.580	0.452	0.698	0.311	0.549	0.505	0.592	0.025
	8-4	0.549	0.422	0.670	0.638	0.542	0.498	0.585	0.065
Quercus	4-2	0.496	0.379	0.613	0.997	0.497	0.456	0.538	0.983
	8-2	0.530	0.411	0.645	0.831	0.533	0.491	0.574	0.15
	8-4	0.525	0.407	0.640	0.878	0.536	0.495	0.577	0.102
Pinus	4-2	0.495	0.371	0.620	0.997	0.483	0.441	0.526	0.619
	8-2	0.541	0.414	0.662	0.736	0.529	0.486	0.571	0.253
	8-4	0.540	0.413	0.661	0.745	0.545	0.502	0.587	0.035

Table A3: Results from the robust contrast test between pairs of Hirst-type traps for 2-hourly and daily average concentrations (log-transformed) for the five most frequently-occurring pollen. The estimator as well as the lower and upper bounds of the 95% confidence range of these values are shown. Statistically significant differences are shown in red.

Species	Traps	Two-Hour Averages				Daily Averages			
		Estimator	Lower	Upper	p-Value	Estimator	Lower	Upper	p-Value
Total	4-2	0.506	0.476	0.536	0.879	0.498	0.397	0.599	1.000
	8-2	0.536	0.506	0.566	0.009	0.542	0.440	0.641	0.598
	8-4	0.530	0.501	0.560	0.042	0.542	0.440	0.640	0.604
Group1	4-2	0.513	0.459	0.567	0.841	0.532	0.378	0.679	0.886
	8-2	0.543	0.489	0.597	0.152	0.563	0.408	0.707	0.607
	8-4	0.531	0.476	0.585	0.387	0.531	0.378	0.678	0.889
Group2	4-2	0.526	0.493	0.559	0.147	0.527	0.424	0.627	0.819
	8-2	0.548	0.515	0.580	0.001	0.541	0.438	0.640	0.626
	8-4	0.522	0.489	0.554	0.263	0.516	0.414	0.616	0.935
Group3	4-2	0.507	0.476	0.538	0.846	0.509	0.408	0.609	0.979
	8-2	0.540	0.509	0.571	0.006	0.549	0.447	0.647	0.500
	8-4	0.534	0.503	0.565	0.026	0.541	0.439	0.640	0.609
Group4	4-2	0.484	0.448	0.520	0.556	0.491	0.380	0.603	0.985
	8-2	0.529	0.493	0.565	0.145	0.522	0.410	0.632	0.893
	8-4	0.546	0.510	0.582	0.006	0.533	0.420	0.642	0.780
Group5	4-2	0.481	0.440	0.521	0.5	0.487	0.367	0.608	0.969
	8-2	0.520	0.479	0.560	0.493	0.521	0.399	0.641	0.918
	8-4	0.539	0.499	0.580	0.059	0.531	0.408	0.650	0.830

Table A4: Results from the robust contrast test between pairs of Hirst-type traps for 2-hourly and daily average concentrations (log-transformed) for the five different pollen size groups (Group 1: *Urticaceae/Cannabaceae*, *Castanea*; Group2: *Alnus*, *Betula*, *Corylus*, *Fraxinus*, *Platanus*, *Rumex*, *Salix*; Group 3: *Plantago*, *Poaceae*, *Populus*, *Taxus*, *Cupressaceae*, *Ulmus*; Group 4: *Carpinus*, *Fagus*, *Juglans*, *Quercus*; Group 5: *Pinus*, *Picea*). The estimator as well as the lower and upper bounds of the 95% confidence range of these values are shown. Statistically significant differences are shown in red.

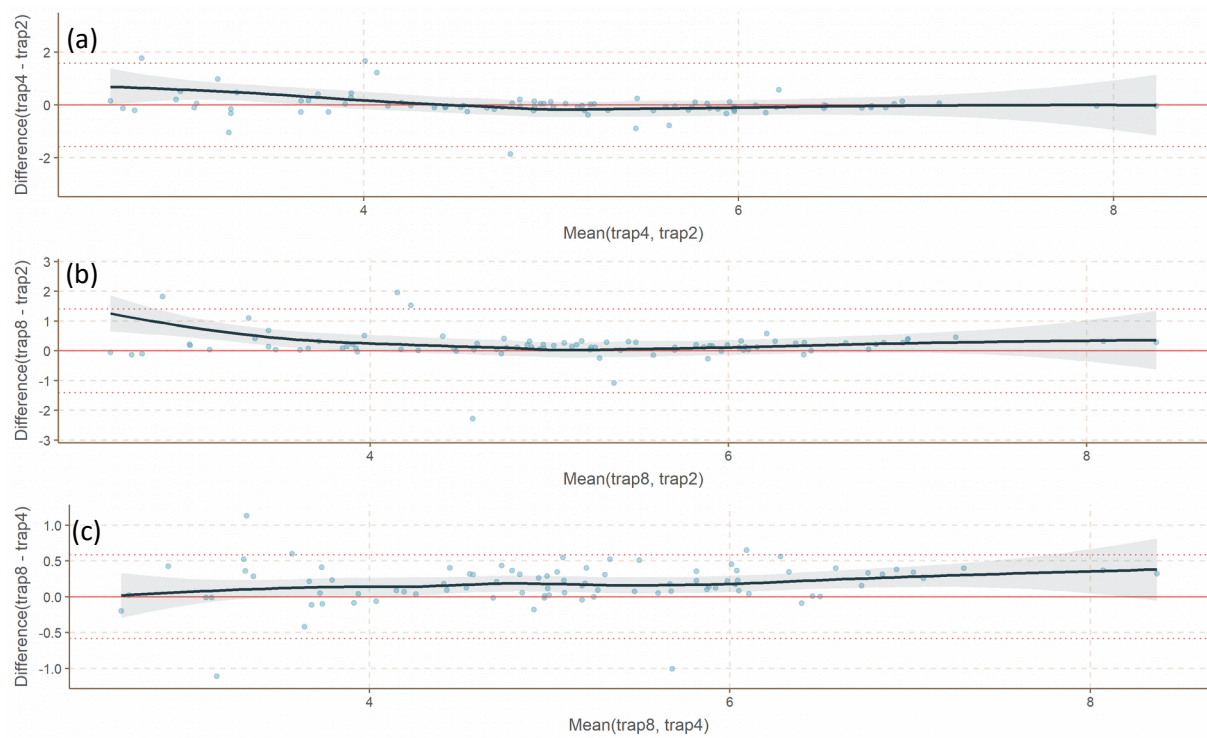


Figure A1: Altman-Bland plots showing the daily average differences in total pollen concentrations (log transformed) between pairs of traps compared to the average of those two traps: (a) trap 2 vs. trap 4, (b) trap 2 vs. trap 8, and (c) trap 4 vs. trap 8. The red dotted lines show ± 1.96 standard deviations from the mean. The grey shaded area and black line show the 95% confidence interval of a LOESS smoother applied to the data.

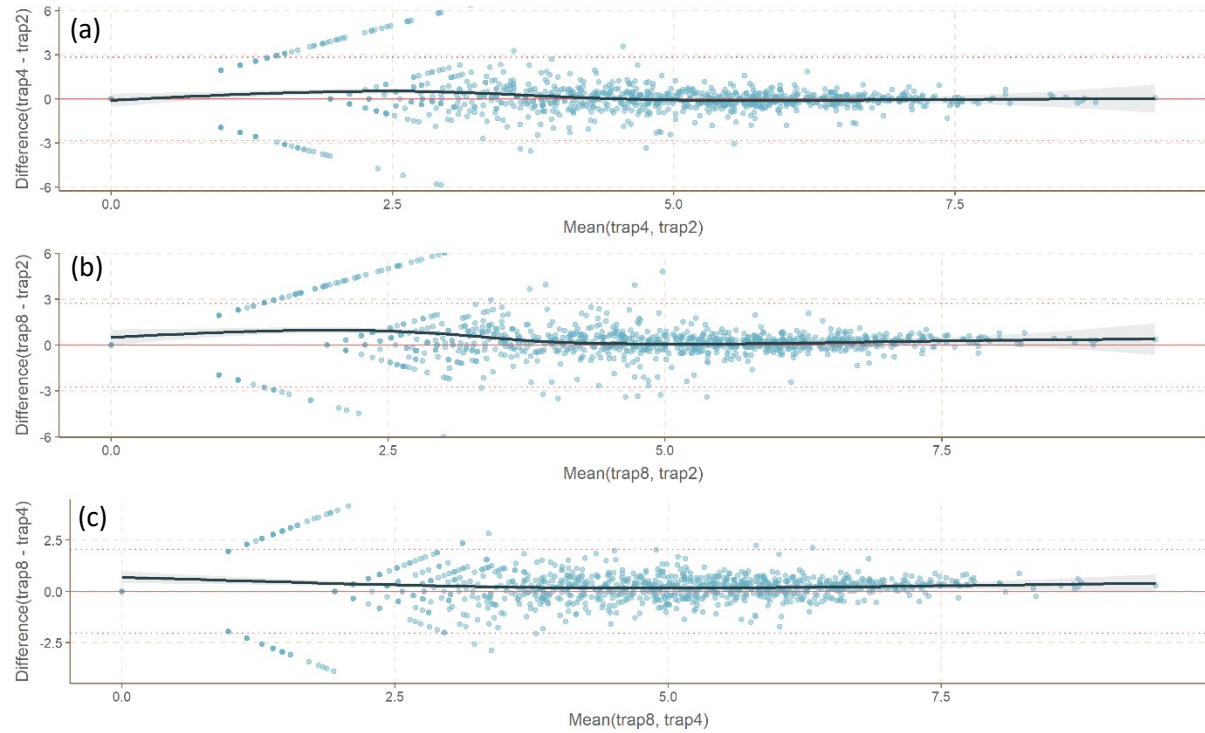


Figure A2: Altman-Bland plots showing the 2-hourly average differences in total pollen counts (log transformed) between pairs of traps compared to the average of those two traps: (a) trap 2 vs. trap 4, (b) trap 2 vs. trap 8, and (c) trap 4 vs. trap 8. The red dotted lines show ± 1.96 standard deviations from the mean. The grey shaded area and black line show a LOESS smoother applied to the data.

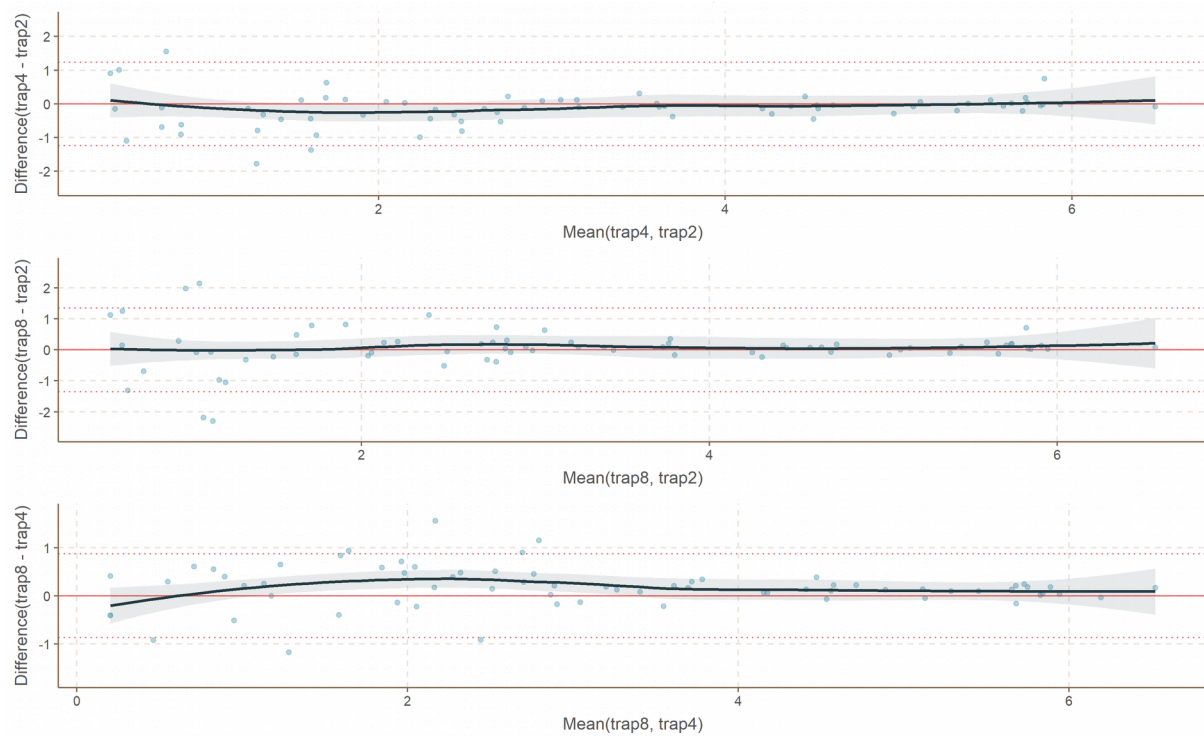


Figure A3: Altman-Bland plots showing the daily average differences in Poaceae counts (log transformed) between pairs of traps compared to the average of those two traps: (a) trap 2 vs. trap 4, (b) trap 2 vs. trap 8, and (c) trap 4 vs. trap 8. The red dotted lines show ± 1.96 standard deviations from the mean. The grey shaded area and black line show a LOESS smoother applied to the data.

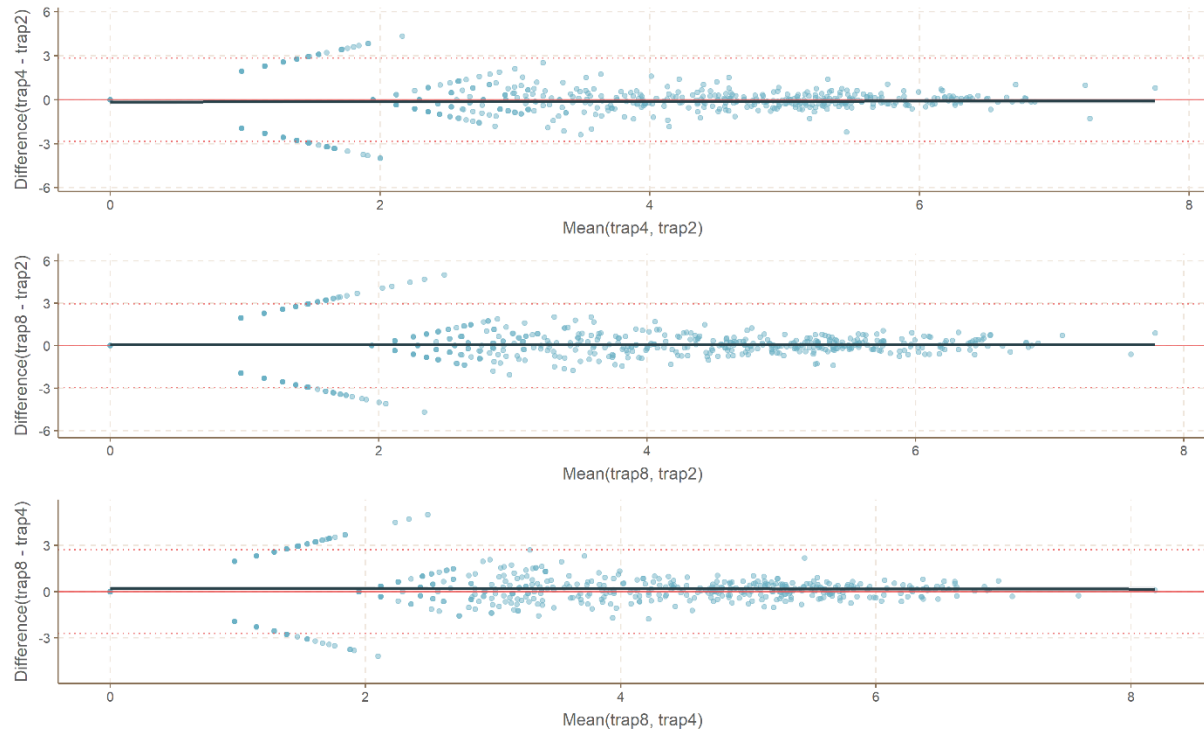


Figure A4: Altman-Bland plots showing the 2-hourly average differences in Poaceae counts (log transformed) between pairs of traps compared to the average of those two traps: (a) trap 2 vs. trap 4, (b) trap 2 vs. trap 8, and (c) trap 4 vs. trap 8. The red dotted lines show ± 1.96 standard deviations from the mean. The grey shaded area and black line show a LOESS smoother applied to the data

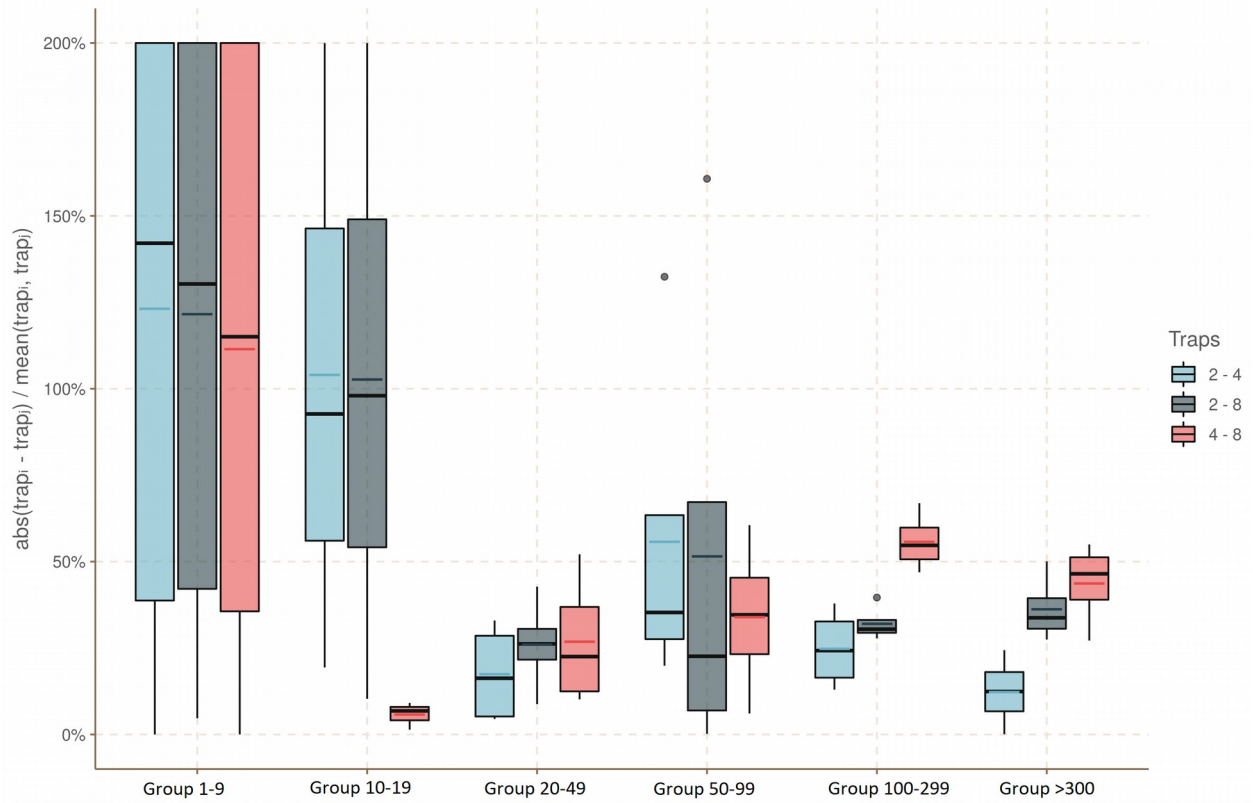


Figure A5: Box plots of the percentage differences between pairs of traps for daily average *Betula* for the six different concentration groups considered (from left to right: 0-9, 10-19, 20-49, 50-99, 100-299, and >300 pollen/m³). The thick black lines in the boxes indicate the median values, the coloured lines the means, while the boxes cover the inter-quartile range.

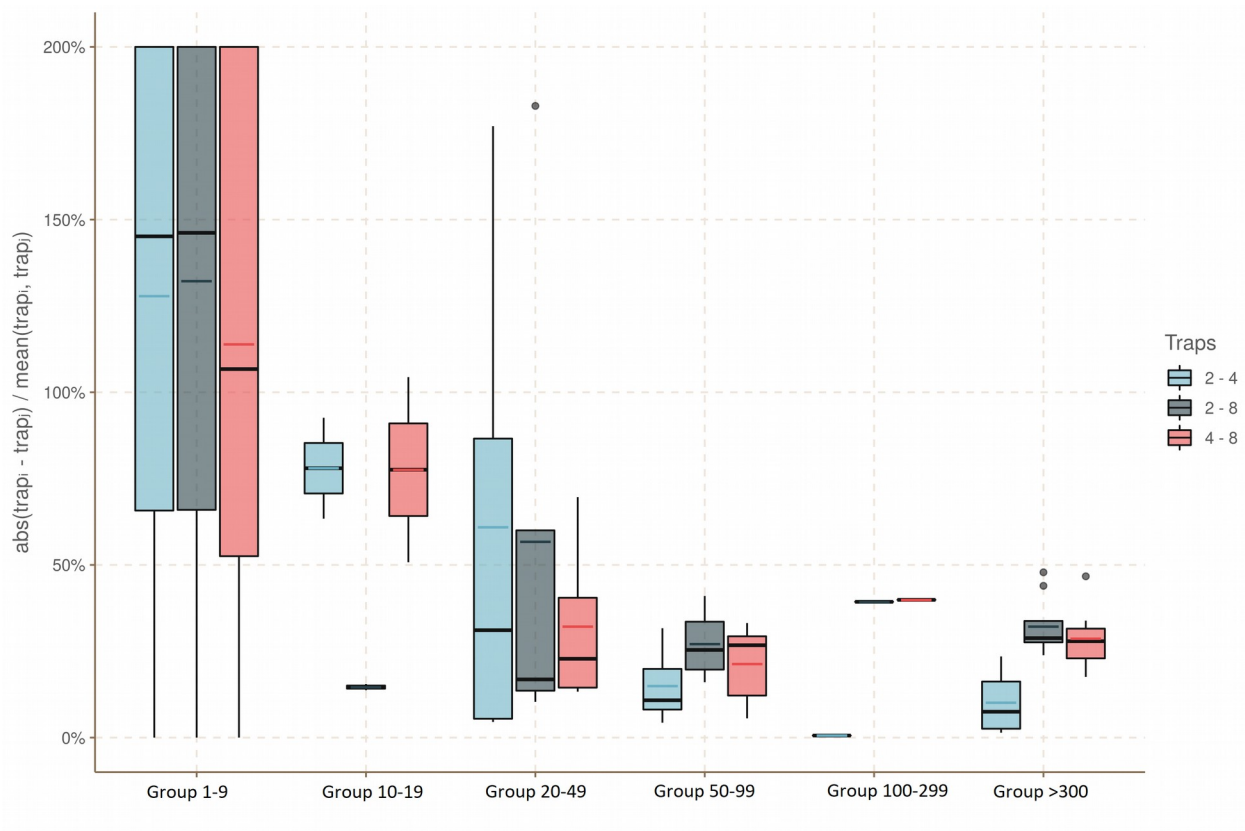


Figure A6: Box plots of the percentage differences between pairs of traps for daily average *Fraxinus* for the six different concentration groups considered (from left to right: 0-9, 10-19, 20-49, 50-99, 100-299, and >300 pollen/m³). The thick black lines in the boxes indicate the median values, the coloured lines the means, while the boxes cover the inter-quartile range.