



### Supplement of

# Seasonality of the particle number concentration and size distribution: a global analysis retrieved from the network of Global Atmosphere Watch (GAW) near-surface observatories

Clémence Rose et al.

Correspondence to: Clémence Rose (c.rose@opgc.fr)

The copyright of individual parts of the supplement might differ from the article licence.

## 1. Effect of cut-off diameters different from 10 nm on the calculation of $N_{tot}$ - Comparison of $N_{tot}$ derived from collocated MPSS and CPC measurements

The influence on  $N_{tot}$  of cut-off diameters different from 10 nm (either lower or higher) was evaluated based on measurements performed with MPSS detecting particles over broad enough size ranges and sufficient annual data availability (> 60%, See

- 5 Sect. 4.1). The effect of including sub-10 nm particles in  $N_{tot}$  was investigated first: the ratio of the particle concentration in the range 3-10 nm (N<sub>3-10</sub>) over  $N_{tot}$  was more specifically calculated using data collected at VAR and SMR, and the same was repeated for the size range 5-10 nm (N<sub>5-10</sub>) using data collected at MEL, LEI, LEI-M and DRN. Because the objective of this analysis was not to deeply investigate the seasonal variations of the contribution of sub-10 nm particles to  $N_{tot}$ , the analysis was performed at the annual scale only, and the variability of this contribution, likely related to the variability of the particle
- sources, was evaluated based on the percentiles of the calculated ratios. As illustrated in Fig. S1, the ratio of  $N_{3-10}$  over  $N_{tot}$  calculated for the two Finnish sites is of the order of 2% on average; it is mostly below 7% (75<sup>th</sup> percentile) and rarely exceeds 17% (90<sup>th</sup> perc.). In contrast, the ratio of  $N_{5-10}$  over  $N_{tot}$  derived for the 4 German sites is slightly higher, 15% on average, and can be as high as 45% (90<sup>th</sup> perc.). A detailed analysis of these observations with respect to the station types is beyond the scope of this study; it is however likely that these last results are related to the contrasting environments of the sites, with
- 15 higher concentrations of sub-10 nm particles (with respect to  $N_{tot}$ ) observed at urban sites in all seasons in connection with traffic. Because none of the stations equipped with a CPC detecting particles down 2.5 nm (ARN, ETL and GSN) are located in an urban area, we believe that the measurements performed at these sites will not lead to a strong overestimation of  $N_{tot}$ . Using the same approach, the effect caused by a lack of measurement in the lowest sizes of the reference range (10-500 nm) was then investigated. The ratios of the particle concentration in the range 10-12 nm (N<sub>10-12</sub>) and 10-20 nm (N<sub>10-20</sub>) over  $N_{tot}$
- 20 were calculated based on data collected at a relatively large subset of stations (27) representative of various conditions (ANB, BIR, CHC, DEM, DRN, DRW, DTC, FKL, HPB, SMR, IPR, KOS, KPS, LEI, LEI-E, LEI-M, MEL, NGL, OPE, PAL, PRG, SSL, TRL, VAR, WAL, ZEP and ZSF). The ratio of  $N_{10-12}$  over  $N_{tot}$  is on average of the order of 2%, mostly below 9% (90<sup>th</sup> perc.), suggesting that such a small cut-point difference is not a major issue for  $N_{tot}$  (Fig. S1). In contrast, the lack of measurement in the sub-20 nm range might have a stronger effect on  $N_{tot}$ , as the contribution of these particles to  $N_{tot}$  is on
- 25 average of the order of 13%, and can be up to 38%. This last observation should be kept in mind, specifically when considering the data from MAD and JFJ, where the lower cut-points are ~15 and ~17 nm, respectively (Table 1). As a last sensitivity test, the particle concentration in the range 500-800 nm (*N*<sub>500-800</sub>) was compared to *N*<sub>tot</sub> using measurements performed at 22 sites (ANB, BIR, DRN, DRW, DTC, FKL, GIF, HPB, SMR, IPR, KOS, LEI, LEI-E, LEI-M, MEL, MSY, NGL, SSL, TRL, WAL, WGG and ZEP). As shown in Fig. S1, *N*<sub>500-800</sub> is on average much lower compared to *N*<sub>tot</sub> (ratio of the
- 30 order of 0.1% on average). This clearly indicates that larger particles contribute little to the particle number concentration in all environments, and suggests, in turn, that the higher cut-off of 500 nm used for the derivation of  $N_{tot}$  from MPSS data has no strong effect on the results.



Fig. S1 Ratios of the particle number concentration in different size ranges  $(N_x)$  over  $N_{tot}$ .  $N_x$  denotes  $N_{3-10}$ ,  $N_{5-10}$ ,  $N_{10-12}$ ,  $N_{10-20}$  and  $N_{500-800}$ , which correspond to the concentrations in the size ranges 3-10, 5-10, 10-12, 10-20 and 500-800 nm, respectively. The markers represent the median of the ratios, the lower and upper limits of the solid error bars indicate the 1<sup>st</sup> and 3<sup>rd</sup> quartiles, respectively, and the lower and upper limits of the dashed error bars indicate the 10<sup>th</sup> and 90<sup>th</sup> percentiles of the ratios, respectively.



10 Fig. S2 Comparison of N<sub>tot</sub> derived from collocated CPC (N<sub>CPC</sub>) and MPSS (N<sub>MPSS</sub>) measurements.

Table S1 Comparison of  $N_{tot}$  derived from collocated CPC and MPSS measurements. For each site, the coefficient of determination obtained between the values of  $N_{tot}$  derived from each instrument as well as the equation of the linear fit shown in Fig. S2 are reported.

Station	R <sup>2</sup>	Fit equation
MSY	0.76	$N_{MPSS} = 0.50 \times N_{CPC} + 1034$
HPB	0.84	$N_{MPSS} = 0.79 \times N_{CPC} + 453$
PUY	0.87	$N_{MPSS} = 1.30 \times N_{CPC} + 61$
SMR	0.96	$N_{MPSS} = 0.85 \times N_{CPC} + 15$
VAR	0.94	$N_{MPSS}$ = $1.07 \times N_{CPC}$ - $30$
PAL	0.74	$N_{MPSS} = 0.89 \times N_{CPC} + 89$

#### 2. Coverage criteria for the calculation of N<sub>tot</sub> annual and seasonal statistics

Similar to the analysis reported in Sect. 4.1 at the annual scale, the effect of long gaps in the data was investigated at the seasonal scale, for which the exclusion of non-contiguous weeks was also tested (this was not tested at the annual scale due to the high number of possible combinations). As shown in Fig. S3, the most pronounced variability is, again, observed for the

- 5 polar sites, specifically those in the Southern Hemisphere (TRL and NMY) during MAM, consistent with a strong variation of  $N_{tot}$  during this time of the year (see Fig. 7). Besides these two sites, noticeable variability is also seen for ETL in DJF, and is also likely explained by the variability of  $N_{tot}$  at the site during this period. Otherwise, the effect of reduced data availability on the seasonal statistics seems to be limited, at least for datasets with up to 6 missing weeks. The median values derived from the original and from the reduced datasets are mostly within a factor of 1.5 and they tend to be homogeneously distributed
- 10 around 1. Regarding the configuration of the missing period (contiguous vs non-contiguous weeks), there is no clear effect up to 3 weeks missing, and for longer periods, contrasting results are obtained depending on the season. Indeed, when simulating the absence of 4-6 weeks, the exclusion of non-contiguous weeks leads to broader dispersion of the medians in DJF, but to smaller dispersion during MAM (particularly visible for TRL and NMY), and there is no difference in JJA and SON. The effect of excluding individual hourly averages was also evaluated. Figure S4 shows the results obtained when the data
- 15 availability was artificially decreased to 75, 60, 50 and 25%. For each target data availability, individual data points to exclude were randomly selected, and the test was performed 25 times to get more insight into the variability of the results. In a similar way as previously done for the longer gaps, the medians and percentiles derived from the shorter datasets were compared to those of the original dataset. Changes in the statistics are logically more obvious when decreasing the amount of data but remain relatively limited (at both scales, annual and seasonal), with all ratios between 0.8 and 1.3, homogeneously distributed
- 20 around 1. It appears that the impact of sporadic/short gaps on the statistics is less pronounced compared to the exclusion of longer «blocks» of data resulting in similar data availability (e.g. 6 weeks *vs* 50% at the seasonal scale and 24 weeks *vs* 50% at the annual scale). The effect is nonetheless, on average, more pronounced at the seasonal scale, and particularly for MAM, and is again the strongest for the polar sites, which experience the most amplitude in the seasonal cycle.



Fig. S3 Variability of  $N_{tot}$  seasonal statistics in reduced datasets. For each investigated gap length and configuration (i.e. 1 to 8 weeks, contiguous or not), all possible combinations of weeks to exclude were tested, and in each case the ratio between the newly derived median of  $N_{tot}$  and that derived from the original dataset was calculated (circles). The upward and downward triangles provide insight into the range of variability. The upfacing triangles represent the ratio between the maximum value of the 75<sup>th</sup> percentile of  $N_{tot}$  obtained from the reduced datasets and the 75<sup>th</sup> percentile calculated from the original time series. The downfacing triangles represent the 25<sup>th</sup> percentile from the original dataset divided by the minimum of the 25<sup>th</sup> percentile.



Fig. S4 Variability of  $N_{tot}$  annual and seasonal statistics in reduced datasets. For each period (year or season), individual hourly averages were randomly excluded to reach four target data availabilities to be investigated (25%, 50%, 60% and 75%). In each case, the test was repeated 25 times, and the ratio between the newly derived median of  $N_{tot}$  and that derived from the original dataset was calculated (circles). The upward and downward triangles provide insight into the range of variability. The upfacing

5 triangles represent the ratio between the maximum value of the 75<sup>th</sup> percentile of  $N_{tot}$  obtained from the reduced datasets and the 75<sup>th</sup> percentile calculated from the original time series. The downfacing triangles represent the 25<sup>th</sup> percentile from the original dataset divided by the minimum of the 25<sup>th</sup> percentile.



Fig. S5 Statistics regarding the length of the interruption periods observed in the different datasets at the annual scale. The sites are sorted by ascending order of their annual data availability, which is reported at the top of each panel.



#### 3. Seasonal variation of the total particle number concentration and size distribution

Fig. S6 Geometric standard deviation of the modes. For each site and season, the thicker bar represents the standard deviation of the Aitken mode ( $\sigma_{m,1}$ ) and the thinner one that of the accumulation mode ( $\sigma_{m,2}$ ). In addition, the values at the top of each panel indicate the site-specific variability of geometric standard deviation, with the italicized text corresponding to  $\sigma_{m,2}$ . The meaning of the abbreviations used for the footprint is the following: RB for rural background and U for urban. Details regarding the calculation of the site-specific variability of the modes characteristics are available in Sect. 5.1.

	Year			DJF			МАМ			IIA			SON		
Station	Med.	10 <sup>th</sup>	90 <sup>th</sup>	Med.	10 <sup>th</sup>	90 <sup>th</sup>	Med.	10 <sup>th</sup>	90 <sup>th</sup>	Med.	10 <sup>th</sup>	90 <sup>th</sup>	Med.	10 <sup>th</sup>	90 <sup>th</sup>
WMO I, A	frica														
RUN	-	-	-	-	-	-	-	-	-	605	196	3924	656	240	2416
WGG	3798	1369	14024	3513	1286	12806	3916	1401	12627	3755	1506	14292	4004	1286	16185
WMO II, A	Asia														
AMY	-	-	-	-	-	-	-	-	-	3175	1630	6280	4247	2257	7659
GSN	-	-	-	-	-	-	2761	1439	5049	-	-	-	-	-	-
LLN	1106	388	2779	810	307	2034	1466	554	2661	1080	380	3214	1121	395	3238
WLG	-	-	-	2332	953	5881	2021	735	5712	1047	414	2646	-	-	-
WMO III,	South Am	erica													
CHC	2644	578	15307	1246	410	5588	2252	526	13827	5483	1303	27475	2557	748	12051
WMO IV, North America, Central America and the Caribbean															
ALT	153	53	426	101	53	235	201	77	354	256	66	686	88	37	379
APP	2555	1146	5309	2684	1014	5613	2815	1146	6522	2339	1304	3984	2593	1070	5064
BND	2125	668	5920	1454	395	4720	3313	1217	7340	1797	681	4914	2222	660	5848
BRW	128	40	599	148	48	354	132	54	417	140	36	1038	107	26	623
CPR	1235	686	3010	-	-	-	-	-	-	1153	629	2688	1372	721	3125
EGB	2594	648	7295	2541	683	7126	1329	281	5071	2701	1120	6464	3893	913	9022
ETL	1116	189	3328	731	135	2742	724	192	3592	1508	707	3723	1159	162	3533
SGP	-	-	-	3496	938	6789	3422	1324	7874	-	-	-	-	-	-
SPL	2159	803	6709	1657	624	5005	2161	672	7562	2773	1410	8373	2062	905	6262
THD	1038	404	2519	1100	394	2817	1006	395	2552	932	399	2252	1162	436	2520
WMO V, S	South-Wes	st Pacific	2												
CGO	559	136	2797	733	292	1803	722	166	3213	282	93	2847	583	170	2797
MLO	409	280	696	414	280	749	376	275	642	438	304	694	408	254	708
SMO	284	175	420	304	200	604	215	142	344	275	177	388	330	221	431
WMO VI, Europe															
ANB	4562	1783	14145	4987	1622	17827	4621	2151	12785	-	-	-	4021	1646	10583
ARN	-	-	-	-	-	-	4996	1668	11950	-	-	-	6332	3575	14475
BIR	1009	232	2878	553	149	1555	998	276	2648	1889	867	4158	963	212	2551
BEO	-	-	-	-	-	-	670	215	1864	-	-	-	-	-	-
CMN	961	261	2212	314	102	931	1089	434	2509	1216	695	2352	-	-	-
DEM	6198	2711	15075	5666	1898	15269	8887	4138	18423	5208	3071	12089	5754	2500	13515

Table S2 Overview of  $N_{tot}$  measurements. For each site, the annual statistics (i.e. median, 10th and 90th percentiles) of  $N_{tot}$  are only reported when corresponding data availability is above 60%, and seasonal statistics are reported when corresponding data availability is above 50%.

DRN	7962	4037	15213	7558	3471	15836	7414	4008	13749	8964	5069	15856	8272	4025	15455
DRW	4579	1993	10148	4414	1596	9953	4411	2335	10394	-	-	-	4097	1979	9051
DTC	3692	1513	10893	2921	1127	7927	3813	1715	11181	4381	2059	17022	-	-	-
FKL	2731	1427	4969	-	-	-	2716	1662	5237	3353	2095	5116	2320	1400	4741
GIF	3331	1218	8253	2658	795	7512	3530	1599	8221	3722	1422	9462	3071	1069	7565
HAC	1004	240	2909	405	150	1517	1193	399	3610	1862	899	3389	637	233	2047
HPB	2737	1192	4836	2055	927	3737	2872	1233	5011	3066	1458	5370	-	-	-
IPR	6810	3057	14382	10071	3577	18774	5849	3171	11066	5757	2868	9581	7528	2845	13641
JFJ	193	56	590	106	32	338	191	72	602	378	162	786	158	62	392
KOS	2690	1111	5159	2162	617	4849	2807	1323	5948	3371	2067	6131	2192	1117	3916
KPS	4798	2747	9422	4426	2502	8090	4983	2661	11650	5003	3241	9505	4768	2779	8269
LEI	5088	2486	10182	4889	2120	9525	5114	2426	10303	5594	3080	12012	4697	2501	9346
LEI-E	8573	3859	18002	7233	2903	16737	9467	4862	17970	10375	5511	21266	6875	3289	14810
LEI-M	10130	4634	21699	9512	3759	22274	10556	5087	21435	8983	4643	19959	11056	5146	22994
MAD	-	-	-	10107	2127	24045	7586	2865	16431	-	-	-	8534	3350	19913
MEL	4434	2154	8361	3769	1727	6936	4602	2154	9538	5278	2871	11767	4154	2219	6643
MSA	-	-	-	-	-	-	4494	1840	13510	4661	1631	13340	2853	861	10283
MSY	3007	1158	8261	1847	805	4712	3252	1407	8232	4247	1806	11519	3049	1384	7465
NGL	2579	1025	5507	1601	637	3145	2246	914	5247	3718	2055	7019	2892	1453	5266
OPE	2412	995	4775	2018	760	4826	3192	1660	5482	2824	1344	5330	1926	891	3367
PAL	356	68	1839	146	51	526	627	128	2296	1050	208	2300	222	47	1183
PDM	-	-	-	330	62	2131	-	-	-	-	-	-	788	223	3403
PRG	6077	2528	13129	5719	2022	12304	5132	2391	11641	7810	3602	15444	6352	2906	13820
PUY	1968	457	5080	785	300	2490	2177	538	5119	2948	978	6566	-	-	-
SMR	1259	430	3074	812	330	2006	1526	620	3865	2011	892	3735	928	336	2136
SNB	1027	291	2562	636	189	1529	1223	417	3096	1737	720	2993	811	270	2093
SSL	1873	491	4448	832	318	2009	2878	935	5671	2794	1281	5182	1599	518	3436
UGR	7477	3234	17694	10910	4512	27024	7601	3672	15668	5719	2652	11533	7588	3180	17262
VAR	391	77	2027	178	48	554	624	142	2164	1355	273	2874	240	53	995
VAV	-	-	-	-	-	-	-	-	-	2148	1281	4936	-	-	-
WAL	3350	1524	6309	2519	1103	5380	3701	1508	7501	4162	2193	7407	3110	1760	5301
ZEP	153	25	573	-	-	-	227	105	468	276	79	981	62	19	342
ZSF	827	228	2550	520	162	1496	1298	302	3425	-	-	-	650	241	1661
WMO VII, Antarctica															
NMY	252	49	783	496	275	1397	186	51	701	62	36	155	293	116	530
SPO	217	33	461	316	225	649	240	134	590	38	25	71	244	83	473
TRL	306	45	761	537	319	1385	155	48	528	55	33	167	375	153	603

#### 4. Focus on CCN-sized particles



Fig. S7 Seasonal statistics of  $N_{50}$ , the particle number concentration in the range 50-500 nm, used as an additional proxy for potential CCN population. The bars represent the median of  $N_{50}$ , and the lower and upper end of the error bars represent the 1<sup>st</sup> and 3<sup>rd</sup> quartiles of the data, respectively. Stations are sorted based on the classification reported in Table 1. The meaning of the abbreviations used for the footprint is the following: RB for rural background and U for urban.

25



Fig. S8 Scatter plots of  $N_{50}$  as a function of  $N_{tot}$  (hourly averages) for the different station types: a. polar sites, b. urban sites, c. 20 other lowland sites and d. mountain sites. The color of each pixel indicates the number of data points (hourly averages) falling into its area (all pixels have equal area on a log-log scale). The linear fit performed on the logarithm of the data, separately for each period (year and seasons), is also presented. The statistics of the ratio between  $N_{50}$  and  $N_{tot}$  calculated for each of these periods are in addition shown for each station type in the insert of the corresponding panel; the markers represent the median of the ratios, and the lower and upper limits of the error bars indicate the 1<sup>st</sup> and 3<sup>rd</sup> quartile, respectively.



Fig. S9 Scatter plots of  $N_{100}$  (a., b., c.) and  $N_{50}$  (d., e., f.) as a function of  $N_{tot}$  (hourly averages). The different footprints represented within each station type are highlighted.

Table S3 Connection between  $N_{50}$ , the particle number concentration in the range 50-500 nm, used as an additional proxy for the CCN population, and  $N_{tot}$ . For each station type and season, the equation of the linear fit performed on the logarithm of the data is reported in the second column, and the corresponding coefficient of determination in the third column. Note that based on the corresponding p-values, all correlations were found to be significant at the 95% confidence level (p < 0.05).

Station type / season	Fit equation	R <sup>2</sup>
Polar sites		
Year	$\log_{10}(N_{50}) = 0.93 \times \log_{10}(N_{tot}) - 0.23$	0.76
Winter	$\log_{10}(N_{50}) = 0.95 \times \log_{10}(N_{tot}) - 0.24$	0.81
Spring	$\log_{10}(N_{50}) = 0.70 \times \log_{10}(N_{tot}) + 0.26$	0.48
Summer	$\log_{10}(N_{50}) = 0.99 \times \log_{10}(N_{tot}) - 0.26$	0.66
Fall	$\log_{10}(N_{50}) = 0.97 \times \log_{10}(N_{tot}) - 0.32$	0.72
Urban sites		
Year	$\log_{10}(N_{50}) = 0.87 \times \log_{10}(N_{tot}) + 0.09$	0.67
Winter	$\log_{10}(N_{50}) = 0.99 \times \log_{10}(N_{tot}) - 0.33$	0.77
Spring	$\log_{10}(N_{50}) = 0.79 \times \log_{10}(N_{tot}) + 0.39$	0.61
Summer	$\log_{10}(N_{50}) = 0.73 \times \log_{10}(N_{tot}) + 0.65$	0.59
Fall	$\log_{10}(N_{50}) = 0.92 \times \log_{10}(N_{tot}) - 0.11$	0.69
Other lowland sites		
Year	$\log_{10}(N_{50}) = 0.95 \times \log_{10}(N_{tot}) - 0.15$	0.71
Winter	$\log_{10}(N_{50}) = 1.03 \times \log_{10}(N_{tot}) - 0.40$	0.79
Spring	$\log_{10}(N_{50}) = 0.90 \times \log_{10}(N_{tot}) - 0.03$	0.61
Summer	$\log_{10}(N_{50}) = 0.77 \times \log_{10}(N_{tot}) + 0.52$	0.55
Fall	$\log_{10}(N_{50}) = 1.00 \times \log_{10}(N_{tot}) - 0.33$	0.72
Mountain sites		
Year	$\log_{10}(N_{50}) = 0.96 \times \log_{10}(N_{tot}) - 0.25$	0.86
Winter	$\log_{10}(N_{50}) = 0.94 \times \log_{10}(N_{tot}) - 0.29$	0.88
Spring	$\log_{10}(N_{50}) = 0.95 \times \log_{10}(N_{tot}) - 0.21$	0.85
Summer	$\log_{10}(N_{50}) = 0.94 \times \log_{10}(N_{tot}) - 0.08$	0.85
Fall	$\log_{10}(N_{50}) = 0.94 \times \log_{10}(N_{tot}) - 0.26$	0.86