Supplement of Atmos. Meas. Tech., 17, 921–941, 2024
https://doi.org/10.5194/amt-17-921-2024-supplement

Atmospheric Measurement Techniques
Open Access
EGU

Supplement of

Long-term aerosol particle depolarization ratio measurements with HALO Photonics Doppler lidar

Viet Le et al.

Correspondence to: Viet Le (viet.le@fmi.fi)

S1 Saturation at cloud base

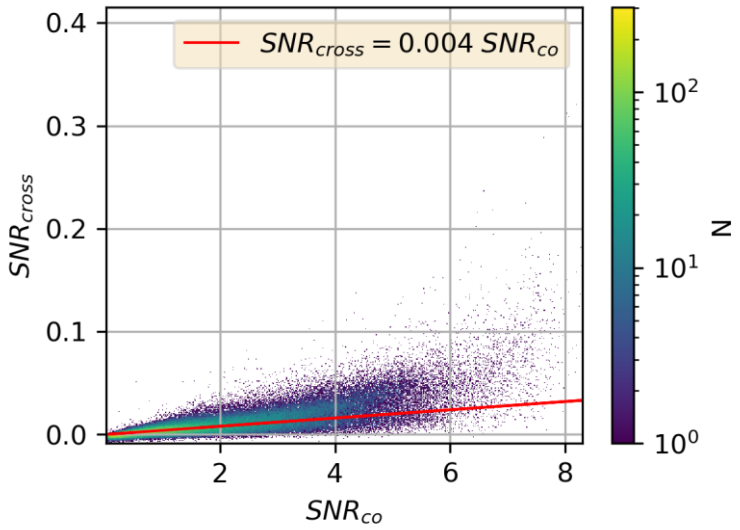


Figure S1: SNR_{co} and SNR_{cross} measured by Utö-32XR at the cloud base.

S2 Post-processing

For a successful fit, aerosol- and hydrometeor- free (background or noise-only) range gates need to be identified in the profile. Firstly, the data is averaged every hour, and the SNR_{co} and SNR_{cross} profiles are then decomposed by stationary wavelet transform (Nason and Silverman, 1995) with the wavelet bior2.6 using Pywavelets (Lee et al., 2019). Next, the variance of the noise in the SNR is removed by applying a hard threshold shrinkage function using universal thresholding (Donoho and Johnstone, 1994) to the approximation and detail coefficients from level 1 to level 4 resulted from the wavelet transform (Nason and Silverman, 1995). The SNR is then reconstructed using inverse stationary wavelet transform (Nason and Silverman, 1995). Finally, the background range gates are identified as those having reconstructed SNR_{co} values less than the standard deviation of the instrument noise floor divided by the squared root of the number of profiles averaged in that hour. The 2nd order polynomial fit can then be performed on those noise-only data points identified in the profile. The fit follows Eq. (S1) and the correction is then applied to the entire SNR profile, similar to (Vakkari et al., 2019).

$$SNR_z = a + h_z \cdot b + c \cdot h_z^2 , \quad (S1)$$

where SNR_z and h_z are the background SNR and height at each range gate, z, and a, b, c are the parameters of the fit for each profile.

Next, the aerosol-only range gates are identified based on the result of the AI algorithm applied to the original non-averaged data. If 80 percent of original the non-averaged measurements at a range gate in an hour were identified as aerosol by the Aerosol Identification algorithm, then that 1-hour averaged range gate is identified as aerosol measurement. Finally, following Vakkari et al., 2021, the bleed-through corrected δ of aerosol can be calculated following Eq. (3) in the manuscript, utilizing the estimated bleed through obtained in Table 3.

Uncertainty in δ of aerosol

The background and the bleed through corrections include uncertainties that will propagate through to the uncertainty in the final retrieved δ of aerosol. A Bayesian approach utilizing Stan (Carpenter et al., 2017) has been used to estimate these uncertainties. The full Bayesian inference is obtained through a sampling method based on Hamiltonian Monte Carlo simulation (Betancourt and Girolami, 2015), which is a form of Markov chain Monte Carlo sampling (van Ravenzwaaij et al., 2018). The resulting posterior distribution is used to estimate the mean and uncertainty of the parameters and generate the posterior predictive distribution of the variables.

The 2nd order polynomial model of SNR is similar to Eq. (S1), but with the error component added:

$$SNR_r = a + h_r \cdot b + c \cdot h_r^2 + \epsilon\,, \qquad (S2)$$

where $\epsilon \sim Normal(0, \sigma^2)$.

Equation S2 is equivalent to $SNR_r \sim Normal(a + h_r \cdot b + c \cdot h_r^2,\ \sigma)$, or the so-called likelihood $p(SNR \mid a, b, c, \sigma, h_r)$. The standard non-informative prior was used for all parameters:

$$p(a) \sim Normal(0, 10^3)$$
$$p(b) \sim Normal(0, 10^3)$$
$$p(c) \sim Normal(0, 10^3)$$
$$p(\sigma) \sim Gamma^{-1}(0.001, 0.001)$$

The posterior distributions of the parameters are then obtained by sampling through Stan.

$$p(a, b, c, \sigma \mid SNR, h) \propto p(SNR \mid a, b, c, \sigma)\, p(a, b, c, \sigma)$$

Four different sampling chains (van Ravenzwaaij et al., 2018) are utilized with 2000 samples for each chain (excluding the 500 warmup/burn-in samples), and convergence diagnostics have been carried out to ensure that the chains converged. The sampling results represent the posterior distribution of all the parameters.

From these posterior distributions, the mean and standard deviation of the parameters can be calculated. The fitted background SNR distribution at all heights can be generated by sampling from the posterior predictive distribution. It is done through selecting random draws from the posterior, which are then plugged into Eq. (S2) to calculate the posterior predictive samples. The mean and standard deviation of the fitted background SNR at each height can be estimated by calculating the mean and standard deviation of the generated samples at each height.

The samples of the corrected SNR distribution at each height are calculated by dividing the original SNR by the samples from the estimated background SNR distribution at each height. The samples of the final δ of aerosol distribution at each height are then calculated following Eq. (3) assuming the bleed-through follows a normal distribution with mean and standard deviation from Table 3. The mean and uncertainty of the final δ of aerosol is estimated by calculating the mean and standard deviation of the samples at each height.

The effect of the post-processing procedure on δ of aerosol

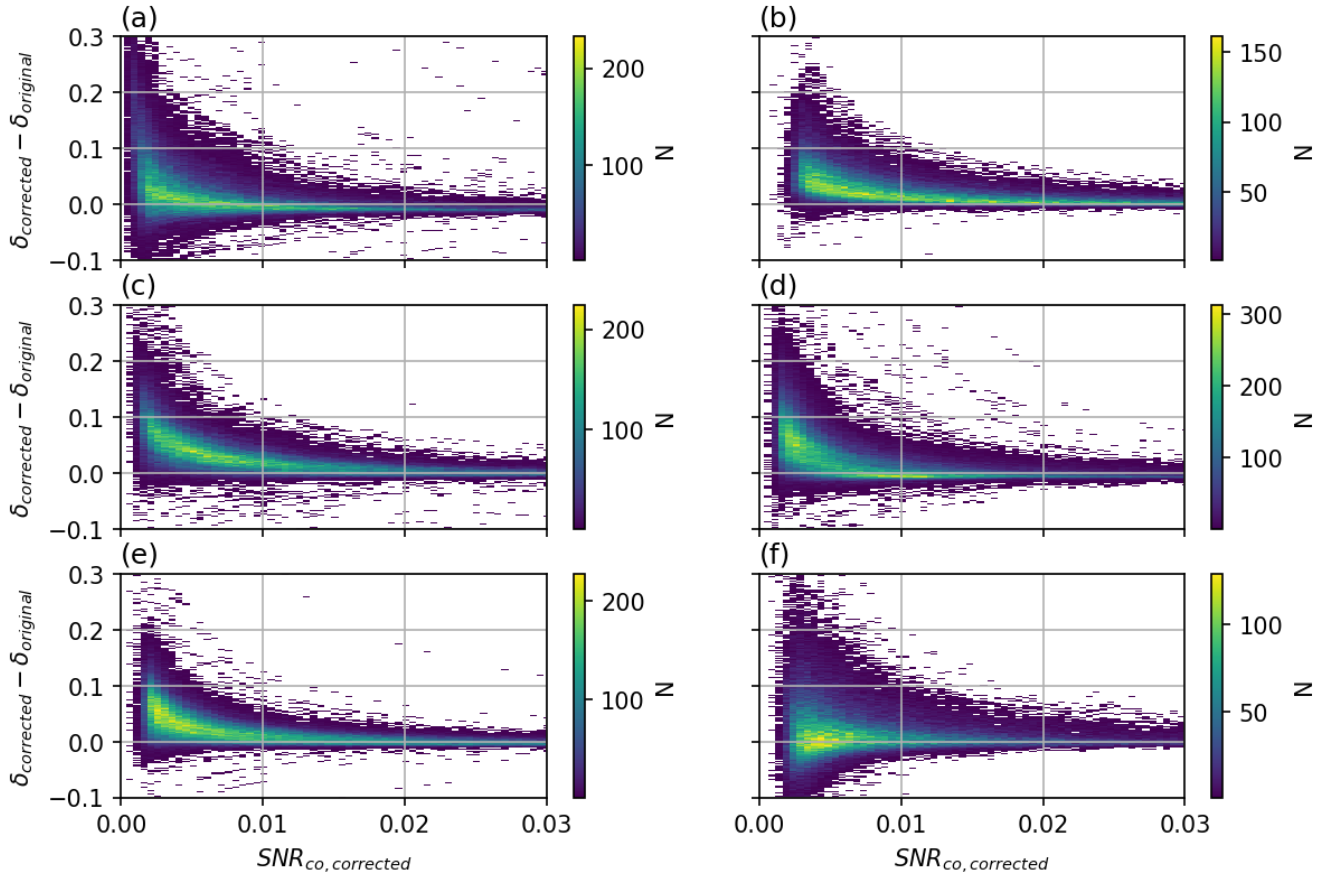


Figure S2: 2D histogram of the changes of δ of aerosol from the post-processing procedure in relation with SNR_{co} in a) Utö-32, b) Utö-32XR, c) Hyytiälä-33, d) Hyytiälä-46, e) Vehmasmäki-53, f) Sodankylä-54

Instrument	% of total aerosol data point with $\lvert\delta_{corrected} - \delta_{original}\rvert > 0.05$	% of total aerosol data point with $\lvert\delta_{corrected} - \delta_{original}\rvert > 0.1$
Utö-32	18.2%	6.6%
Utö-32XR	11.0%	1.9%
Hyytiälä-33	31.8%	9.3%
Hyytiälä-46	24.6%	7.1%
Vehmasmäki-53	15.5%	2.7%
Sodankylä-54	17.9%	5.1%

Table S1: Percentage of aerosol data with the changes of δ resulted from the post-processing procedure. The second column displays the percentage of aerosol data that has δ changed more than 0.05. The third column displays the percentage of aerosol data that has δ changed more than 0.1.

Fig. S2 demonstrates the impact of the background correction to δ of aerosol. For all the instruments except Sodankylä-54, the background correction increases δ of aerosol. The increase is higher for aerosol with weaker SNR_{co}. The effect is negligible for strong aerosol signal with SNR_{co} larger than 0.01. Table S1 shows the percentage of data affected by this background correction. The effect is most prominent for the instruments at Hyytiälä; as 31.8% and 24.6% of the aerosol data have δ changed by 0.05 for Hyytiälä-33 and Hyytiälä-46 respectively. However, significant changes of aerosol δ more than 0.1 is only at 9.3% and 7.1% for these instruments. This result shows the importance of background correction in the retrieval of weak aerosol signals.

Without it, biases from the 2nd order polynomial component in the background would propagate into the δ of aerosol.

S3 Aerosol Identification algorithm

The detailed steps of the algorithm are explained below, with an example shown in Fig. 7.

1. The first step of the algorithm involves preliminary detection of potential hydrometeors and aerosols from background signals based on β' and SNR_{co}. Result from this step is shown in Fig. 7a.
    - All data points with SNR_{co} larger than one standard deviation of the background SNR_{co} and with β' values less than $10^{-5.5}$ Mm^{-1} sr^{-1} are marked as aerosol. A 2D median kernel was then convolved with these aerosol data points to remove noisy signals due to instrumentation and attenuation.
    - All data points with SNR_{co} larger than three standard deviations of the background SNR_{co} and with $\beta' > 10^{-5.5}$ Mm^{-1} sr^{-1} were marked as hydrometeor. A 2D median and maximum kernel were then convolved with these hydrometeor data points to remove noisy signals due to instrumentation and attenuation.
2. The falling hydrometeor detection step involves separating aerosol in downdrafts due to boundary layer mixing from precipitation using both β' and w. Regions containing both up- and down- drafts are considered to be characteristic of boundary layer mixing, while a region of continuous downdrafts indicates precipitation. The result from this step is shown in Fig. 7d.
    - The updraft proxies are identified by selecting data points that have $w > 1$ m s^{-1}. 2D median and maximum kernels were then convolved to remove noisy signal and expand the updraft proxies.
    - Next, precipitation proxies are selected, having $\beta' > 10^{-7}$ Mm^{-1} sr^{-1} and $w < -1$ m s^{-1}. A 2D median kernel was then convolved with these data points to remove noisy signals. The precipitation proxies that are in the updraft proxies are then removed.
    - All-precipitation regions are identified having $\beta' > 10^{-7}$ Mm^{-1} sr^{-1} and $w < -0.5$ m s^{-1}. A 2D median kernel was then convolved with these data points to remove noisy signals.
    - The updraft region is identified having $w > 0.2$ m s^{-1}. A 2D maximum kernel was then convolved with these data points to increase the size of the updraft region.
    - Finally, precipitation data points are identified by including the precipitation proxies that overlap with any all-precipitation regions but not with updraft regions. A 2D maximum kernel was used for this iterative process.
    - The precipitation data points are then overwritten the aerosol data points from step 1 as hydrometeor.
3. An attenuation correction step sets all observations above clouds and precipitation with their corresponding class since the signal has been heavily attenuated. The result from this step is shown in Fig. 7f.
4. In the final step, a fine-tuned aerosol process is utilized to improve the aerosol class determination accuracy. The final result is shown in Fig. 7h.
    - First, aerosol clusters are identified in the time and height domain using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al., 1996).

- A cluster is flagged as hydrometeor if the mean w of the cluster < -0.5 m s^{-1}.
- The rest of the clusters which are connected to the ground are classified as aerosol.
- The rest of the clusters that have w > -0.2 m s^{-1} are classified as aerosol.
- All clusters that do not satisfy of the previous criteria are classified as undefined.

S4 The effect of relative humidity

Period	Location	Slope	p-value	R squared
Whole data	Utö	-0.073	< 10^{-3}	0.012
	Hyytiälä	-0.14	< 10^{-3}	0.105
	Vehmasmäki	-0.194	< 10^{-3}	0.219
	Sodankylä	-0.165	< 10^{-3}	0.142
May-June	Utö	-0.147	< 10^{-3}	0.045
	Hyytiälä	-0.208	< 10^{-3}	0.166
	Vehmasmäki	-0.169	< 10^{-3}	0.145
	Sodankylä	-0.127	< 10^{-3}	0.057

Table S2: Linear regression analysis summary for δ = Slope × RH + off-set (not shown in the table), where δ is the δ of aerosol and RH is the surface relative humidity (at 2 m a.g.l). The p-values for the slopes are < 10^{-3}, indicates that all the slopes' values are statistically significant (i.e., different than 0). The R squared describes the proportion of variance in δ that can be explained by RH.

S5 Hyytiälä and Utö on April 2018

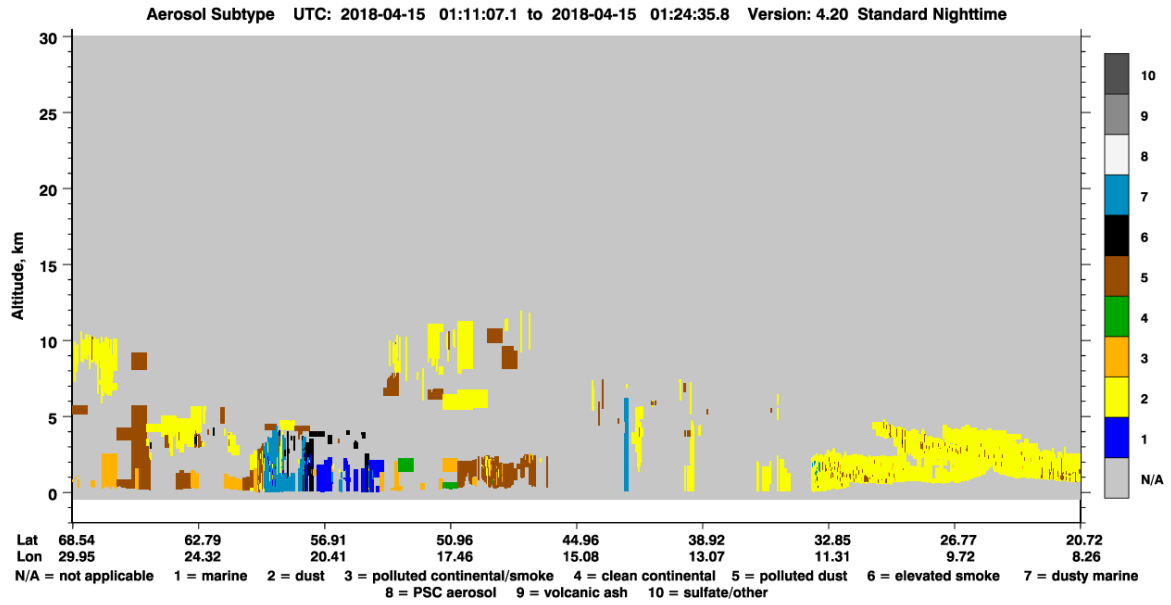


Figure S3: Aerosol subtype V4.2 (Kim et al., 2018; Liu et al., 2019) derived from CALIOP data onboard CALIPSO (Winker et al., 2009) on 2018-04-15. A layer of dust can be observed near Hyytiälä lidar (61.84ºN,24.29ºE).

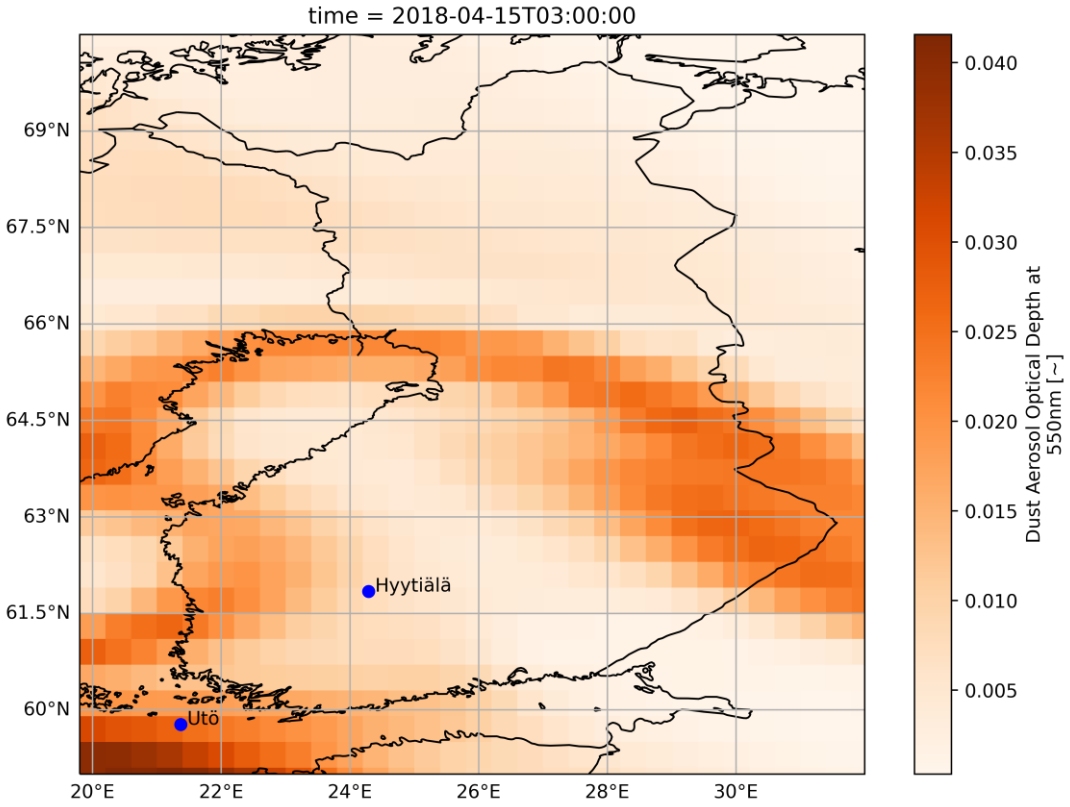


Figure S4: Dust Aerosol Optical Depth at 550 nm wavelength from CAMS model forecast (Benedetti et al., 2009; Morcrette et al., 2009)

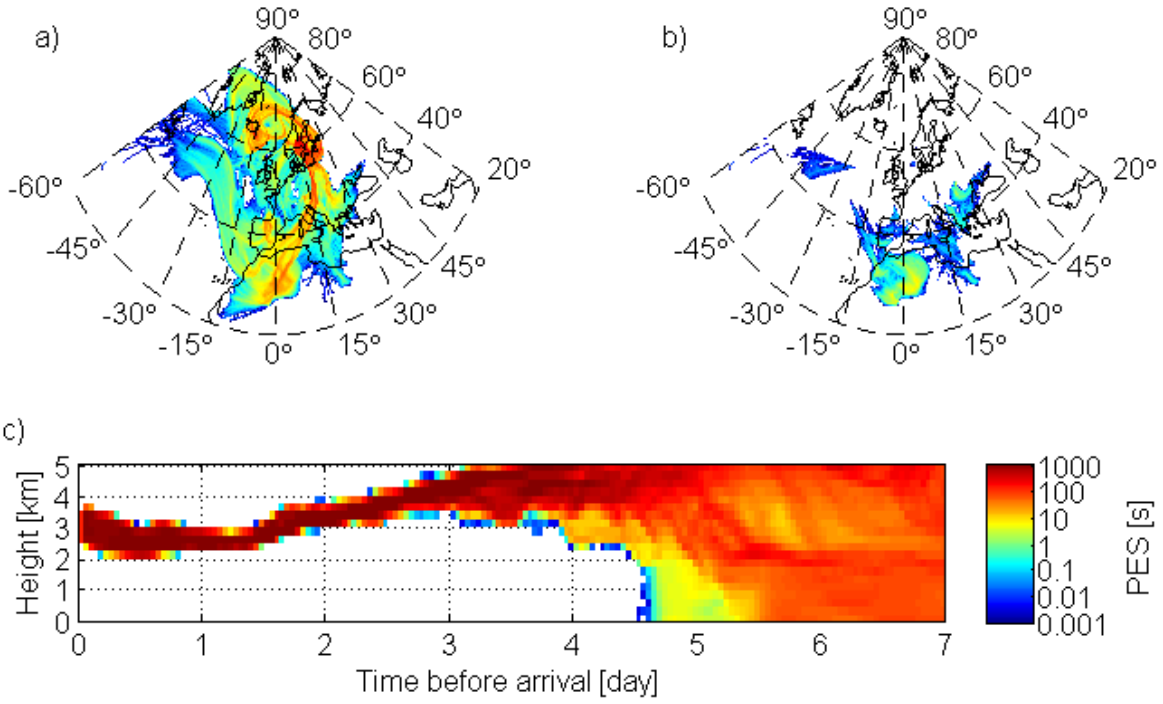
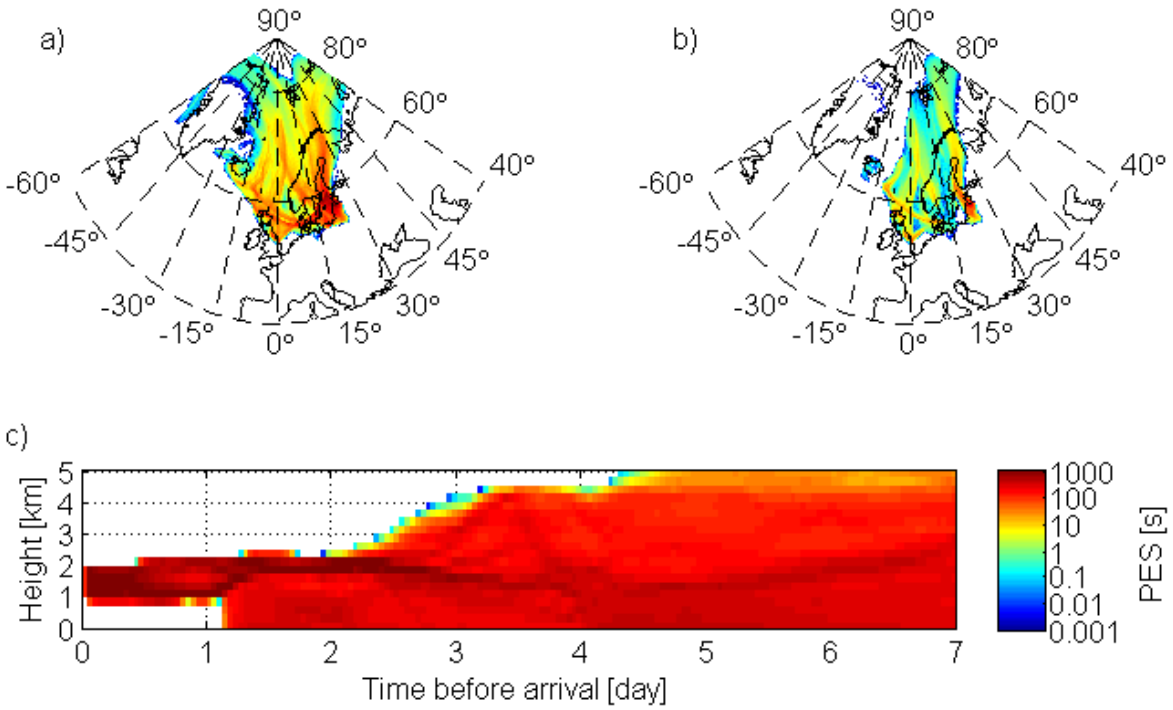


Figure S5: Air mass origin for the elevated layer at 2.5-3.5 km a.g.l. observed on 2018-04-15 at 03 UTC at Hyytiälä. a) PES summed up for all heights for 7 days before arrival at Hyytiälä. b) Sum of PES in the lowest 500 m a.g.l. for 7 days before arrival at Hyytiälä. c) Vertical distribution of PES in the lowest 5 km a.g.l. for 7 days before arrival at Hyytiälä.

S6 Utö on May 2018



135 **Figure S6: Air mass origin for the elevated layer at 1-2 km a.g.l. observed on 2017-05-13 at 18 UTC at Utö. a) PES summed up for all heights for 7 days before arrival at Utö. b) Sum of PES in the lowest 500 m a.g.l. for 7 days before arrival at Utö. c) Vertical distribution of PES in the lowest 5 km a.g.l. for 7 days before arrival at Utö.**