Dataset name: **iCUPE Dataset (DS) from Deliverable 4.1.1:**

A blueprint for novel proxy variables integrating in-situ and satellite remote sensing data with an exemplary dataset

Author(s) and affiliations: **Pauli Paasonen[1], Aino Ovaska[1], Jon Atherton[1,2], Chao Zhang[1,2], Albert Porcar-Castell[1,2], Tuomo Nieminen[1], Antti Manninen[3], Ville Vakkari[3], Antti Hyvärinen[3], Jakob Boyd Pernov[4], Andreas Massling[4], Alison Beamish[5], Sabine Chabrillat[5], Steffen M. Noe[6,7], Vito Vitale[8], Tuukka Petäjä[1]**

1. Institute for Atmospheric and Earth System Research, University of Helsinki, Finland
2. Department of Forest Sciences, University of Helsinki, Finland
3. Finnish Meteorological Institute, Finland
4. Department of Environmental Science, iClimate, Aarhus University, Denmark
5. Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Section Remote sensing and Geoinformatics
6. Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Estonia
7. Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Estonia
8. Italian National Research Council CNR, Italy

Place and date:

Helsinki, Finland, 11.12.2020

Pauli Paasonen
University of Helsinki, Institute for Atmospheric and Earth System Research (INAR) – physics
P.O. Box 64, 00014 University of Helsinki
E-mail: pauli.paasonen@helsinki.fi
Phone: +358 50 415 5252

Abstract

In iCUPE Task 4.1 we derived proxies for three central parameters related to Polar pollution and biosphere – condensation sink (CS), mixing layer heigh (MLH) and gross primary production (GPP). The aim was to develop proxies based on variables that can be retrieved from satellite data or global reanalysis data in order to generate fields for the studied parameters that cover much larger areas than the available in-situ data. However, we did not develop such fields for the proxy variables during the project, but advanced significantly the work towards that goal. We developed proxies for CS at three Finnish measurement stations and for MLH at one station, which we present as the data sets of this Deliverable. We tested the performance of the CS proxy against observations at the Villum station in Greenland and observed that the proxies did not capture the CS in Greenland in a sufficient manner. Thus, we estimate that publishing a spatially more representative proxy fields based on the currently available in-situ data would not serve the purpose, since their validation against in-situ data would not be possible in as much detail as needed. Since the concept for the proxies for these two variables is ready, we expect that representative Polar proxies with better spatial coverage can be made functional in future by applying emerging data from more numerous Polar measurement sites. For GPP, we did not develop a functional proxy, but investigated carefully the different parameters that can be retrieved from satellites and applied for such a proxy. Additionally, we conducted a detailed in-situ spectral characterization of dominant Arctic vegetation communities to inform upcoming hyperspectral satellite missions and explored low-tech proxies for validation in remote Arctic field sites.

1. Condensation sink

CS describes the rate at which condensable vapour molecules condense on existing aerosol particle surfaces (Kulmala et al., 2001). A CS proxy with high enough spatial and temporal coverage makes it possible to better predict and model the removal of condensable gas phase species from the atmosphere and estimate their deposition. CS is one of the variables that characterise the occurrence of new particle formation events (Kulmala et al., 2017).

We calculate daily average condensation sinks with the method described in Kulmala et al (2012) from particle number size distribution data obtained at three measurement sites in Finland (SMEAR II station in Hyytiälä, 61.847463° N, 24.294795° E; SMEAR I station in Värriö, 67.755944° N, 29.610137° E; Pallas station, 67.96715° N, 24.11233° E) and Villum Research Station situated on the Danish military base, Station Nord, in Northeastern Greenland (81° 36’ N, 16° 40’ W). The Finnish stations represent sub-Arctic areas, whereas Villum is located 900 km from the North Pole and 600 km from Zeppelin Mountain Station on Svalbard, thus Villum represents a “natural high-Arctic station”. The particle number size distributions are measured with Scanning Mobility Particle Sizers or Differential Mobility Particle Sizers (SMPS and DMPS, for observations in Finland: Aalto et al., 2001; observations in Villum, Wiedensohler et al., 2012).

During measurement, the particles are dried (in Villum naturally due to transition from cold ambient air to heated cabin), which decreases their size. To determine the CS in ambient humidity we use hygroscopic growth parametrisations described by Laakso et al. (2004) for measurements in Finland and by Zhou et al. (2001) for measurements in Villum. The measurement instruments and measurement heights are collected in table 1. The data from Hyytiälä and Värriö is available at https://avaa.tdata.fi/web/smart/smear.

Table 1. *Measurement instruments for variables used in condensation sink proxy.*

	Hyytiälä	Värriö	Pallas	Villum
Particle number size distribution	SMPS	SMPS	SMPS	SMPS (Wiedensohler et al., 2012)
CO	Horiba APMA 370 at 16.8 m height	Horiba APMA 360 at 9 m height	Picarro G2401 at 7 m height	Teledyne API T300U
Ait temperature	Pt100 thermistor at 16.8 m height	Pt100 thermistor at 2.2 m height	Vaisala HUMICAP pt100 at 4 m height	Vaisala HMP155 at 3 m height

For deriving the proxy, we use simultaneous in-situ measurements of carbon monoxide concentration (CO) and temperature (T) at each of the three sites. CO concentration acts as a tracer for aerosols formed in combustion processes, which form the main anthropogenic source of particles. Since CO is relatively long-lived in the atmosphere (typical residence time of 2 months; Khalil and Rasmussen, 1990), CO is assumed to represent both primary and secondary aerosol originating from combustion. Air temperature is applied for representing biogenic secondary organic aerosol, which is driven by temperature dependent emissions of organic vapours. The temperature dependence of aerosol concentrations in size ranges relevant for CS in various parts of the world, stemming from the secondary growth of particles to those sizes, has been reported e.g. by Paasonen et al. (2013 and 2018).

Instead of applying directly the observed CO concentration, we applied as a tracer for combustion aerosol in the proxy the difference (ΔCO) between measured CO and the 5th percentile during the specific time of year. This stems from the observation that CO has a clear seasonal cycle (Fig. 1), which is not observed in CS. The seasonal cycle probably arises from the lack of radiation leading to low winter time concentrations of OH radical that is the main sink for CO (Khalil and Rasmussen, 1990), which causes CO to accumulate. Since we use CO only as a tracer for aerosols formed in combustion processes, this background CO does not correspond to higher CS values. ΔCO removes this seasonality. For the Villum measurements there was not enough temporally overlapping measurements of CS and CO to determine the seasonal cycle of CO minima. Thus, we did not derive a proxy for Villum, but investigate later on how well the proxies derived for Finnish stations perform in Villum.

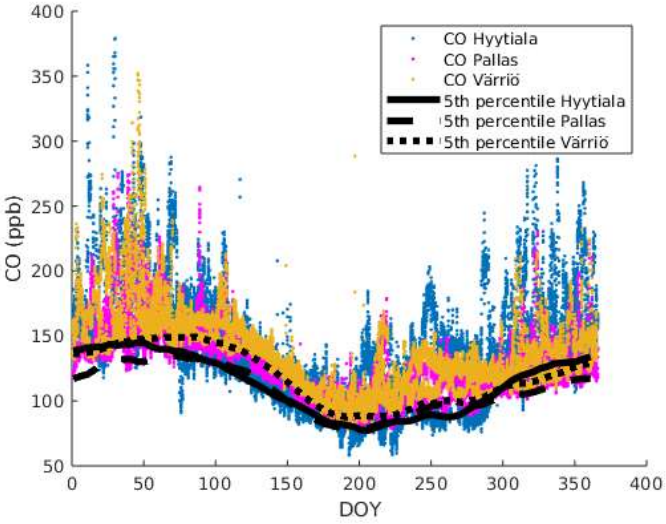

Figure 1. *The annual cycle of carbon monoxide concentration at the measurement stations and their 5-day 5th percentile as moving mean.*

In our proxy we assume two unconnected sources of particles that contribute to the condensation sink, leading to a proxy with a form

$$CS_{proxy} = a_1 T + b_{ave} \Delta CO \quad (1.1)$$

The determination of the constants a_1 and b_{ave} is described below. In Figure 2a we show the relation between CS and ΔCO at different temperature bins. To achieve a better result, we create a routine that finds and removes outliers in each bin to achieve the best possible fits. The routine makes first a simple linear least squares fit to the data points. Next, it calculates for which data points the ratio of the observation and the fitted line falls out of two–sigma standard deviation and leaves those points out as outliers. Then the routine makes a fit to the remaining data points by minimizing the sum of the data point specific ratios between observed and fitted values (using multiplicative inverse for ratios below one) with Matlab function *fminsearch.* By applying this minimization instead of linear least squares, the routine weights similarly the errors of the fitted value equalling to 200 % and 50 % (or 1000 % and 10 %) of the observed CS, regardless of the value of CS.

The slopes that describe CS dependence on ΔCO (Fig. 2a), are quite constant at all temperatures. On the other hand, intercepts of the fits are also constant at lower temperatures but start to increase for higher temperature bins. From this we conclude that the intercepts describe how CS relates to temperature. We define parameter b_{ave} as the average of the slopes weighted with the p-value of the bin. This acts as the CO dependant part of the proxy. We then recalculate new intercepts based on the constant slope (Fig. 2b). This shows that at temperatures below 5°C the ecosystem emissions are low and only at higher temperatures biogenic secondary organic aerosols start to contribute to CS. Therefore, we calculate the average intercept for $T < 5°$ C, similarly as we did for the slopes, and a weighted fit for the higher temperatures. These form the temperature dependant part of the proxy ($a_1(T)$). The final form of the CS proxy can be seen in equation 1.1.

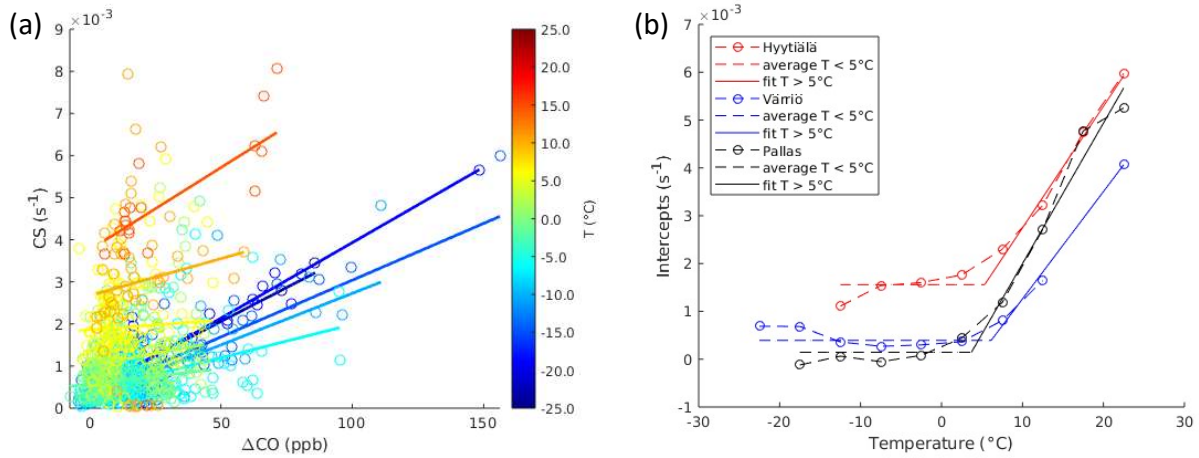

Figure 2. (a) *Observed CS in Värriö as a function of the difference between CO and 5th percentile shown in Fig. 1. Colours and fits represent 5°C temperature bins, where blue-green is temperatures <5°C and yellow-red >5°C.* ***(b)*** *Intercepts from fittings in panel (a) recalculated based on forced average slope* b_{ave}*, as a function of temperature. Average and fit are used to define* $a_1(T)$*, which acts as the temperature dependant part of the proxy.*

In Figure 3 the created proxy is compared to observed CS in Värriö.

The final result is CS dataset (in MS Excel format), which contains the daily time-series for each station in units s^{-1}. For Hyytiälä the time-series covers period 2016-2018, for Värriö 2012-2014 and Pallas 2013-2016.

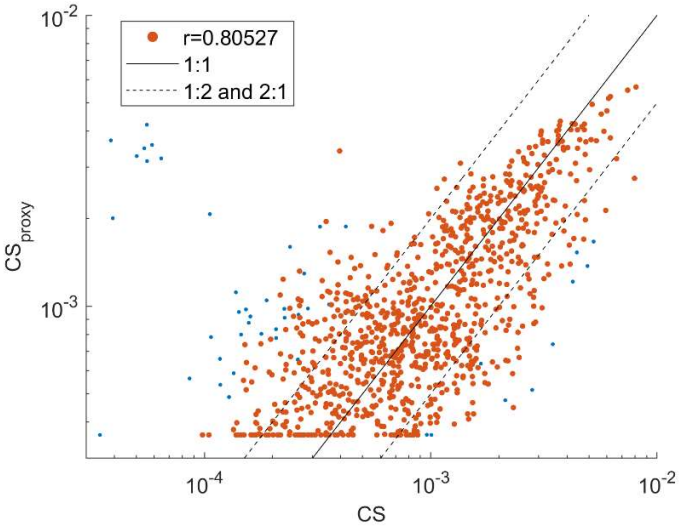

Figure 3. *The created condensation sink (CS) proxy (Eq. 1.1) for Värriö compared to observed CS. The blue dots represent data points that were removed as outliers when determining the relation between CS and ΔCO (Fig. 2a). The presented Person's correlation coefficient r does not include outliers.*

Finally, we investigated how the proxies derived for the Finnish measurement stations perform at the Villum station in Greenland (Fig. 4). All the proxies fail to represent the concentrations at Villum. All the data points are overestimated when applying the Hyytiälä parameters, which is understandable as the

low-temperature low-CO intercept in Hyytiälä (Fig. 2b) is higher than any of the observations at Villum. However, also the parameters for Värriö and Pallas fail to produce any correlation with Villum data.

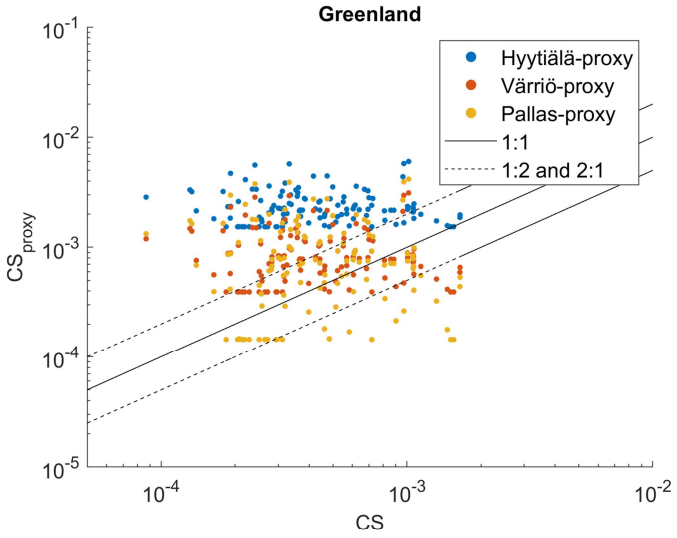

Figure 4. *The values of the condensation sink proxy (Eq. 1.1) with parameters derived for Hyytiälä, Värriö and Pallas compared to CS observed at Villum station in Greenland.*

We derived simple proxies, more specifically functions of CO concentration and temperature, for condensation sink with the data recorded at three measurement stations in Finland. We compared the proxy CS with measured CS obtained at Villum station in northern Greenland. The proxies captured CS at the Finnish stations in adequate manner, but did not reproduce the CS at Villum even qualitatively. Due to the short periods of overlapping CO and CS data from Villum applied in this study, we cannot make conclusions on the reasons for the proxy not capturing the CS variability in high-Polar environment. Our study shows that condensation sink can be parameterised by applying only CO and temperature data. However, the parameters applied in the proxy are location dependent. Thus, a proxy with large spatial coverage would require substantially more long-term (in minimum one year) data sets with simultaneous CS, CO and temperature observations across the domain.

2. Mixing layer height (MLH)

Mixing layer is the lowest layer of atmosphere where emissions are well mixed roughly within an hour (Seibert et al., 2000). This mixing is due to turbulence driven by wind shear or by buoyant air masses warmed in connection to Earth surface that is heated by solar radiation. Some studies refer to mixing layer as (planetary or atmospheric) boundary layer, but these terms may have also different meanings. Low MLH decreases the volume in which the emissions are distributed causing their concentrations to increase. In Polar areas, MLH may remain very shallow and stable for prolonged periods, due to lack of solar radiation in winter time.

The MLH proxy is created using in-situ measurements of global radiation (GRad) and wind speed (WS) from Hyytiälä. For validation, we use MLH from a HALO photonics scanning Doppler lidar measurement

campaign in Hyytiälä 2016-2018 (Vakkari, 2018). The lidar data has been processed as described in Vakkari et al., 2019 and MLH estimated according to Vakkari et al., 2015. Zero values indicate MLH lower than 60 m and MLH is not defined during rain. The global radiation has been measured with Middleton SK08 pyranometer at 18 m (until 8/2017) and EQ08 at 35 m (9/2017 forward). The measurement instrument for wind speed is Vector A101M/L cup anemometer at the height of 16.8 m. We use hourly averages of all variables for deriving the proxies. Radiation and wind speed data are available at https://avaa.tdata.fi/web/smart/smear.

MLH correlates well with radiation (Fig. 5a). During times when radiation is low, e.g. night and winter, also wind speed plays a role (Fig. 5b). The limit GRad<100 (W/m^2) is chosen as it gives the best correlation for MLH as function of wind speed. Above this limit proxy follows directly the linear fit f(GRad) from figure 4a, where the intercept gives $a_{Grad>100}$ and slope $b_{Grad>100}$. To get the low radiation part of the proxy we plot MLH as function of wind speed at different radiation bins (Fig. 5c). As the slopes are constant, we calculate one average slope that will describe the wind dependent part of the proxy (b_{wind}). Next, we recalculate new intercepts for this average slope and fit to the intercepts (Fig. 5d). This is the radiation part (intercept $a_{Grad<100}$ and slope $b_{Grad<100}$) of the low radiation proxy. The final form of the proxy for MLH in meters can be seen in equation 2.1.

$$MLH_{proxy} = \begin{cases} a_{GRad\geq 100} + b_{GRad\geq 100} \cdot GRad, & GRad \geq 100\ W/m^2 \\ a_{GRad\leq 100} + b_{GRad<100} \cdot GRad + b_{wind} \cdot WS, & GRad < 100\ W/m^2 \end{cases} \quad (2.1)$$

The produced proxy is compared to the observed MLH (Fig. 6). The created hourly time-series for Hyytiälä years 2016-2018 is reported in MLH dataset in MS Excel format.

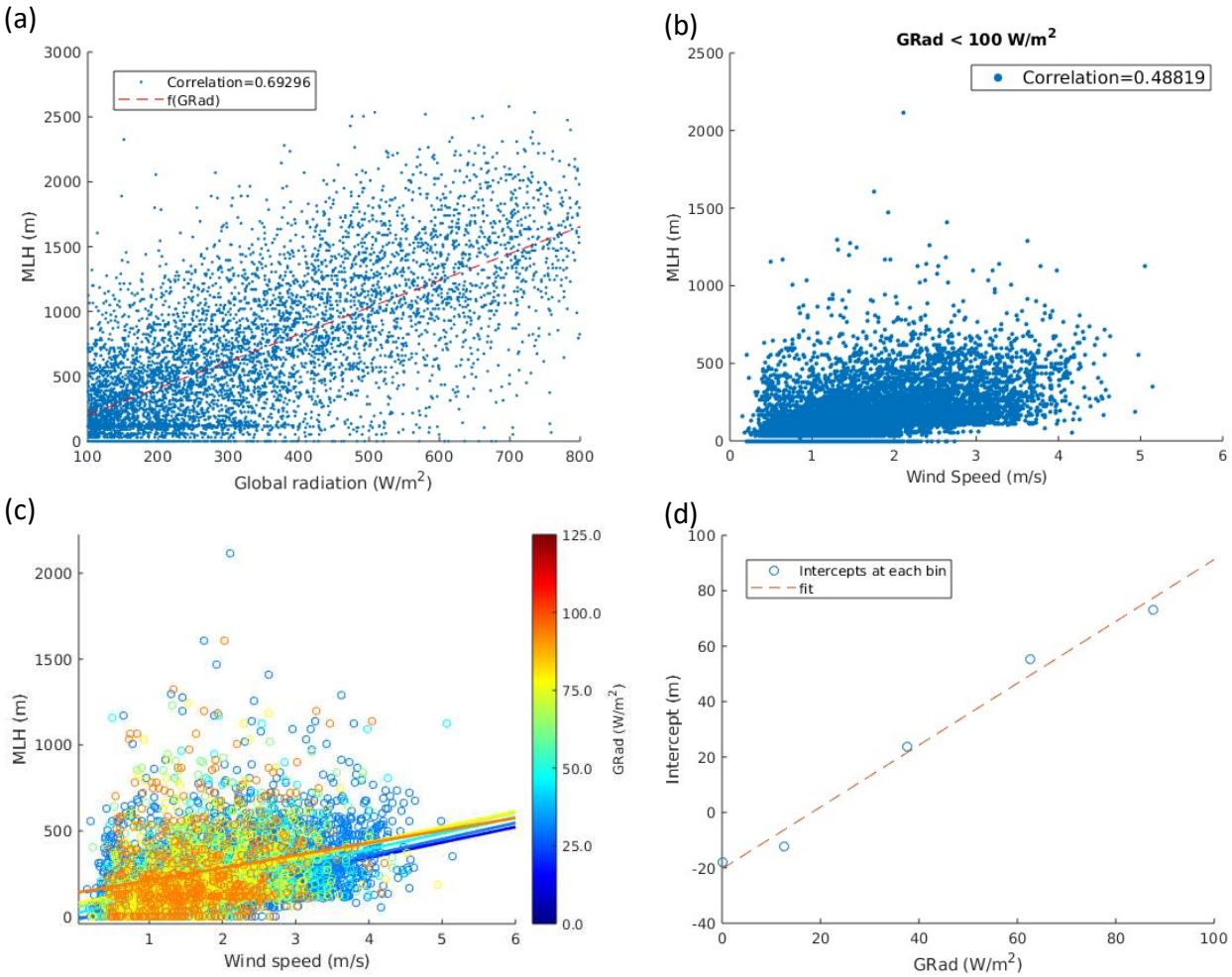

Figure 5. *Deriving the proxy for mixing layer height (MLH) from global radiation (GRad) and wind speed (WS)* ***(a)*** *when radiation is above 100 W/m^2, the MLH follows directly global radiation* ***(b)*** *when radiation is below 100 W/m^2 MLH correlates also with wind speed* ***(c)*** *observed MLH as function of wind speed at different radiation bins with fits for each bin* ***(d)*** *intercepts, from panel (c) recalculated based on forced average slope b_{wind}, as function of radiation. The given correlations represent the values of Pearson's correlation coefficient r.*

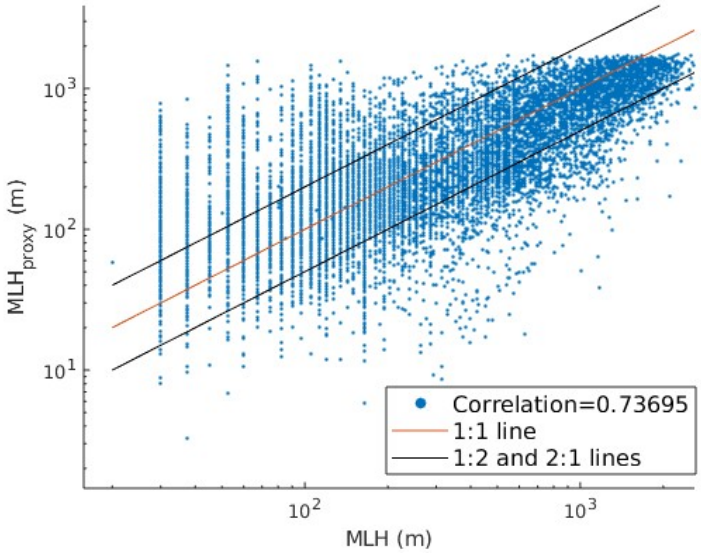

Figure 6. *The created mixing layer height proxy compared to observed mixing layer height. The given correlation represents the value of Pearson's correlation coefficient r.*

Since we did not have other MLH datasets against which to compare our proxy, we cannot discuss its spatial representativeness. For this purpose, long-term measurements of MLH in other polar environments are needed. Furthermore, the impact of snow cover on MLH should be studied and parameterized, e.g. using the existing datasets on Arctic MLHs (e.g. Di Liberto et al., 2012; Mazzola et al., 2016).

3. Gross Primary Production

In the Earth's northerly latitudes, boreal tree species give way to smaller shrubs and sedges as taiga transitions to tundra. Both the taiga and tundra ecosystems, and the transition zone between them, are expected to respond strongly to local and regional climate change, which is amplified in polar regions (Screen and Simmonds, 2010). In line with this, a persistent and long-term greening trend has been noted in the satellite record in latitudes over 50°N (Piao et al. 2019).

The polar greening trend has been noted in satellite retrievals of Leaf Area Index (LAI). This is an estimate of (one sided) leaf area per unit ground area made from reflected radiation in the visible and near-infrared regions of the spectrum. Although obviously related to new plant growth, LAI anomaly is not a direct measure of Gross Primary Production (GPP), which is a seasonally and diurnally dynamic process, and is defined as the amount of carbon dioxide assimilated via photosynthesis per unit time and leaf area. At the local scale GPP can be estimated as the sum of eddy covariance measurements of net ecosystem carbon exchange (NEE), and (modelled) ecosystem respiration (Re):

$$NEE = GPP - Re \quad (3.1)$$

Remote sensing data are required to scale across the landscape, as eddy covariance systems typically have localized footprints on the order of 1 km. The Light Use Efficiency (LUE) paradigm is usually applied

to remotely estimate GPP from remote sensing data:

$$GPP = LUE \times APAR = LUE \times fAPAR \times PAR \quad (3.2)$$

Here PAR is the Photosynthetically Active Radiation (PAR) that drives the photosynthetic reactions, fAPAR is the fraction of the radiation that is absorbed by a plant canopy, and APAR is the total amount of radiation absorbed by the canopy (APAR = fAPAR × PAR). The LUE term, which is also dynamic, represents the operating efficiency of the canopy at a fixed fAPAR (or LAI). In a boreal type forest this term is strongly related to the temperature of the ecosystem, in a Mediterranean ecosystem this term could be dominated by water availability. Remote sensing estimates of GPP (Eq. 3.2) can then be validated or trained with eddy covariance tower observation derived GPP (Eq. 3.1) when orbits or overpasses are aligned with flux footprints (Ryu et al. 2019).

Remotely, fAPAR is estimated from broadband vegetation indices such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI). Such indices are algebraic combinations of visible and near infrared radiometer channels, and quantify the red-edge region of the electromagnetic spectrum where plants switch from being absorbers to scatterers of incident radiation. Note also that fAPAR is strongly (but asymptotically) related to LAI, which is estimated from space using similar methods to fAPAR. The LUE term can be estimated using a look up table approach using meteorological data and *a priori* knowledge. Alternatively, additional remote sensing data such as Sun-Induced Fluorescence (SIF) or carotenoid sensitive vegetation indices such as the Photochemical Reflectance Index (PRI) or the related Carotenoid Chlorophyll Index (CCI) could potentially be used to estimate LUE remotely and replace the look up table approach. Briefly, SIF is a measurement of emission of radiation by chlorophyll molecules which relates to LUE and APAR. PRI/CCI are related visible band reflectance indices that respond to changes in photosensitive pigments collectively referred to as the carotenoids. These signals relate to changes in photosynthetic energy partitioning and present seasonal dynamics in line with changes in GPP and photosynthesis (Porcar-Castell et al. 2014, Gamon et al. 2016). Recent developments reported strong correlations between SIF and GPP, both on daily timescales, for a large number of biomes in a global analysis (Li et al 2018). Nonetheless, there are a number of uncertainties to be resolved relating primarily to angular and canopy structural effects before either SIF or PRI can be used in an operational capacity to estimate GPP from space.

In addition to SIF and PRI/CCI, there are also further methods targeted at remote sensing of northerly vegetation such as the Plant Phenology Index (PPI) which is designed to scale linearly with LAI and be insensitive to snow and soil background and has shown strong correlations with GPP (Jin et al. 2014). Finally, the timing of the spring recovery of photosynthesis in the boreal forest has also been estimated indirectly from microwave observations of snowmelt and an advancing trend noted in the period 1979-2014 (Pulliainen et al. 2017).

In summary, there are now sufficient remote sensing tools and proxies to detect changes in greenness (NDVI, EVI, PPI --> LAI), interannual GPP (SIF, PRI, CCI --> LUE) and spring timing/phenology (PPI, microwave). The challenge is to put these tools together in a logically consistent and useful manner to detect the response of GPP in boreal and polar vegetation at the site and regional scales. This should be informed by ecophysiological understanding of the generative photosynthetic mechanisms at the leaf and individual scale, and by physically based scaling using radiative transfer theory at the larger

canopy and satellite pixel scales. Laboratory and field measurements provide the evidence with which we can build the theory upon which spaceborne retrievals of GPP depend.

3.1. Remote and optical methods of observing GPP related phenomena

In the Optics of Photosynthesis laboratory (OPL) we have focused on building the mechanistic evidence base for the use of satellite estimates of productivity (e.g. SIF, PRI). Without such work it is very difficult to have confidence in remote estimates of productivity, as satellite observations of vegetation are subject to any number of confounding influences including angular-geometric effects, atmospheric radiative transfer processes, instrumental drift and uncertainties.

Hence, we have used laboratory and field measurements of leaf optical properties to show that seasonal changes in photosynthetic energy partitioning and carotenoid pigment pools drive the dynamics of chlorophyll fluorescence (Porcar-Castell et al. 2011, Zhang et al. 2019), observed remotely as SIF, and the PRI respectively (Porcar-Castell et al. 2012) for boreal species. In addition to laboratory measurements, OPL also maintain monitoring Pulse Amplitude Modulated (MoniPAM) fluorometers, high spectral resolution above canopy spectrometers and broadband optical sensors at the Station for Measurement of Ecosystem Atmosphere Relations II (SMEARII), Hyytiälä, Finland. Data from these systems can be used to scale from our laboratory and leaf level observations to the canopy scale, which is what a remote sensing instrument observes.

Figure 7 shows how PRI measured by a sensor positioned several meters above the top of the canopy at SMEARII, closely follows the dynamics of GPP. Note that the above canopy PRI shown here has been filtered for snow and outliers, which is a significant challenge when using PRI to estimate changes in LUE from spaceborne data. More work is required to be certain that PRI is following GPP via changes in leaf optical properties, rather than as an artefact of the filtering or seasonal changes in solar illumination angles. This work is currently being conducted by the OPL group (Zhang et al. 2020).

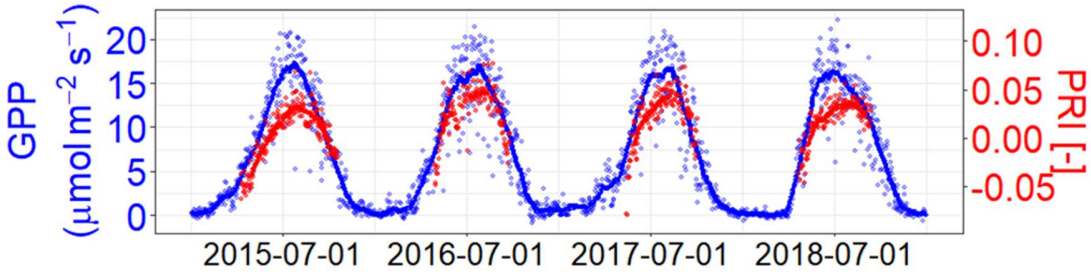

Figure 7. *Above canopy PRI and eddy covariance GPP from SMEARII. GPP and some optical data available on Smart-SMEAR API (Junninen et al. 2009).*

A focus on understanding scaling is needed before we can be fully confident in using either SIF or PRI as a proxy for photosynthesis at larger scales. This is because canopy optical properties and vegetation indices can be thought of as a convolution of a canopy structure signal with the leaf optical property of interest, which conveys the photosynthetic information. The estimation of photosynthesis from space is therefore an inverse problem where we deconvolve the signal of interest from the effect of canopy

structure. Hence a solid understanding of structure is required, which has a seasonal component due to the changing angles of illumination and scattering over the season. To this end, OPL has developed realistic 3D canopy simulations of SIF which show how architectural complexity (e.g. understory growth) can strongly influence remote sensing observations (Liu et al. 2019).

3.2. Future prospects for space based remote sensing of GPP and harmonization of signals

At the SMEARII field station, there is solid evidence that leaf optical properties track photosynthetic seasonality and some evidence that above canopy signals from tower, drone and satellite systems could be used as potential proxies of GPP. However more work is needed on understanding the scaling of leaf optical properties to canopy scale signals and satellite pixels before PRI and SIF can be used in an operational sense. As an alternative method that could potentially be used operationally as a ready proxy of GPP, the PPI has also shown good performance in tracking GPP at SMEARII and other northerly sites (Jin et al 2014). The advantage of the PPI is that it is well tested, and has decent spatial-temporal coverage. The disadvantage is that, unlike PRI and SIF, PPI is only indirectly related to photosynthetic energy usage, and therefore follows subtle changes in leaf area or greenness rather than photosynthesis per se. On longer timescales of weeks to months, these phenological changes are correlated with GPP, even in evergreen forests. Regardless of the index used, more fieldwork is needed in tundra and transitional landscapes, as presumably the benefits of PRI, SIF and PPI apply readily to such locations. Only then, we can properly understand the impacts of global greening and/or browning on the polar carbon cycle and GPP dynamics.

3.3. Spectral characterization and low-tech proxies for Arctic tundra vegetation

To inform upcoming hyperspectral satellite missions, we performed a detailed in-situ spectral characterization of low-Arctic tundra to identify the most discriminative wavelengths for differentiating spectrally similar vegetation communities (Beamish et al. 2018a). Using in-situ and simulated hyperspectral data, our multi-temporal approach revealed that for hyperspectral data, reflectance in blue and red wavebands collected in the senescent phase when colour differences are maximized, are best for differentiating spectrally similar vegetation communities (Beamish et al., 2017). This information can and should be used to inform how and when we interpret and current broadband data and future hyperspectral data when estimating vegetation variables.

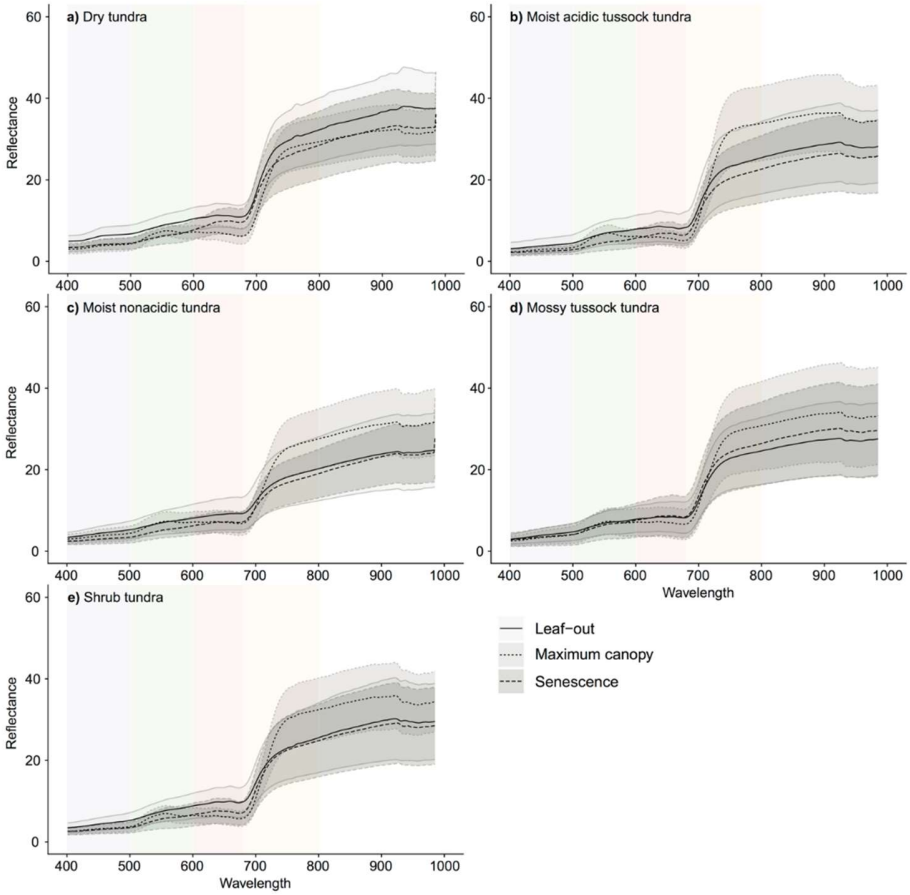

Figure 8. *Detailed spectral characterization of five dominant tundra vegetation communities. The extreme colour differences between vegetation community types and overall lower variability in spectral reflectance at the senesence phase makes it more stable for differentiating spectrally similar vegetation communities.*

Further work looked at the relationship between in-situ field spectroscopy and digital photographs to develop easy, low-tech proxies to validate remote sensing data in remote and logistically challenging field sites (Beamish et al., 2018b). Simple RGB indices showed strong relationships to pigment-driven indices and can be used to estimate and track changes in pigment concentrations in low Arctic vegetation acting as a key validation tool for hyperspectral remote sensing (Beamish et al., 2018c). In particular we found simple normalized red green ratios (nRG) showed strong relationships with the Plant Senescence Reflectance Index (PSRI) which is sensitive to changes in chlorophyll and carotenoid changes induced by senescence (Figure 9).

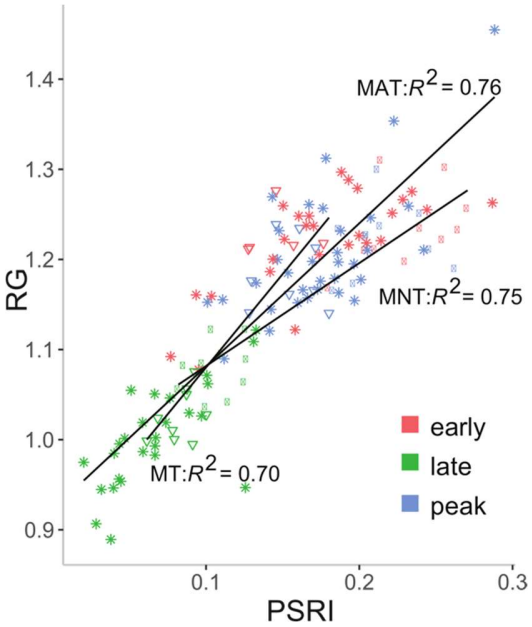

Figure 9. *Simple RGB ratios showed strong correaltions with pigment-driven spectral indices making them suitable as proxies for validating hyperspectral remote sensing*

4. Summary

We developed proxies for condensation sink and for mixing layer height. Furthermore, we explored a concept for the proxy for gross primary production. The CS proxy was verified against measurements available in Finland at SMEAR II station in Hyytiälä, SMEAR I station in Värriö and Pallas station, as well as against data obtained from Villum station in northern Greenland. The proxies captured adequately the observed CS in Finnish stations, but failed to reproduce the observations in Villum. MLH proxy was verified against measurements from Hyytiälä. The datasets for observations and proxy values in Finland are provided in separate files in MS Excel format, with explanation of the sheets and columns below. We did not proceed towards spatially more representative proxies, due to lack of data and resources. However, these proxies can be extended to other areas and modified to utilize satellite retrievals in the future.

For GPP, we did not formulate a proxy, but examined different parameters that can be retrieved from satellite data and can be applied for such a proxy. We additionally explored a multitemporal approach to differentiating tundra vegetation communities with hyperspectral data and found the late senescent phase superior for both in-situ and simulated hyperspectral data. As a next step, we looked at the use of low-tech digital cameras to act as proxies for validating hyperspectral data and found promising results with simple RGB ratios.

5. Proxy data sets

Condensation sink:
Filename: iCUPE_D411a_CS_proxy.xlsx
Results for each station are on separate sheets named after the station.
Columns in order: date (YYYY-MM-DD), observed CS (s^{-1}), proxy CS (s^{-1})

Mixing layer height:
Filename: iCUPE_D411b_MLH_proxy.xlsx
Columns in order: time (YYYY-MM-DD hh:mm:ss), observed MLH (m), proxy MLH (m

References

Aalto, P., Hämeri, K., Becker, E., Weber, R., Salm, J., Mäkelä, J. M., Hoell, C., O'Dowd, C., Karlsson, H., Hansson, H.-C., Väkevä, M., Koponen, I., Buzorius, G., and Kulmala, M. (2001). Physical characterisation of aerosol particles during nucleation events, Tellus, 53B, 344–358.

Beamish, Alison Leslie (2018a): Plot level spectroscopy measurements of low-Arctic vegetation, Toolik Research Station and Imnavait Vegetation Grid, Toolik Lake, Alaska. PANGAEA, https://doi.org/10.1594/PANGAEA.885801,

Beamish, Alison Leslie (2018b): Ground-based digital camera (RGB), spectroscopy, and pigment data from Toolik Vegetation Grid, Toolik Lake, Alaska. PANGAEA, https://doi.org/10.1594/PANGAEA.886340,

Beamish, A. L., Coops, N. C., Hermosilla, T., Chabrillat, S., and Heim, B.: Monitoring pigment-driven vegetation changes in a low-Arctic tundra ecosystem using digital cameras (2018c), Ecosphere, 9, https://doi.org/10.1002/ecs2.2123, 2018

Beamish, A. L., Coops, N., Chabrillat, S., and Heim, B.: A Phenological Approach to Spectral Differentiation of Low-Arctic Tundra Vegetation Communities, North Slope, Alaska (2017), Remote Sens., 9, https://doi.org/10.3390/rs9111200, 2017.

Di Liberto, L., Angelini, F., Pietroni, I., Cairo, F., Di Donfrancesco, G., Viola, A., Argentini, S., Fierli, F., Gobbi, G., Maturilli, M., Neuber, R. and Snels, M. (2012). Estimate of the Arctic Convective Boundary Layer Height from Lidar Observations: A Case Study, Advances in Meteorology, vol. 2012, Article ID 851927, 9 pages, https://doi.org/10.1155/2012/851927.

Gamon, J.A., Huemmrich, K.F., Wong, C.Y., Ensminger, I., Garrity, S., Hollinger, D.Y., Noormets, A. and Peñuelas, J. (2016). A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers. Proceedings of the National Academy of Sciences, 113(46), pp.13087-13092.

Jin, H. and Eklundh, L.(2014). A physically based vegetation index for improved monitoring of plant phenology. Remote Sensing of Environment, 152, pp.512-525.

Junninen, H., Lauri, A., Keronen, P., Aalto, P., Hiltunen, V., Hari, P., Kulmala, M. (2009). Smart-SMEAR: on-line data exploration and visualization tool for SMEAR stations. Boreal Environment Research 14, 447–457.

Khalil, M.A.K. and Rasmussen R.A. (1990). The global cycle of carbon monoxide: Trends and mass balance, Chemosphere, 20, 1–2, 227-242, https://doi.org/10.1016/0045-6535(90)90098-E.

Kulmala, M., Maso, M., Mäkelä, J., Pirjola, L., Väkevä, M., Aalto, P., Miikkulainen, P., Hämeri, K., & O'dowd, C. (2016). On the formation, growth and composition of nucleation mode particles. *Tellus. Series B, Chemical and Physical Meteorology*, *53*(4), 479–490. https://doi.org/10.3402/tellusb.v53i4.16622.

Kulmala, M., Kerminen, V., Petäjä, T., Ding, A., & Wang, L. (2017). Atmospheric gas-to-particle conversion: why NPF events are observed in megacities? *Faraday Discussions*, *200*, 271–288. https://doi.org/10.1039/c6fd00257a

Kulmala, M., Petäjä, T., Nieminen, T., Sipilä, M., Manninen, H., Lehtipalo, K., Dal Maso, M., Aalto, P., Junninen, H., Paasonen, P., Riipinen, I., Lehtinen, K., Laaksonen, A., & Kerminen, V. (2012). Measurement of the nucleation of atmospheric aerosol particles. *Nature Protocols*, *7*(9), 1651–1667, https://doi.org/10.1038/nprot.2012.091

Laakso, L., Petäjä, T., Lehtinen, K., Kulmala, M., Paatero, J., Hõrrak, U., Tammet, H., & Joutsensaari, J. (2004). Ion production rate in a boreal forest based on ion, particle and radiation measurements. *Atmospheric Chemistry and Physics*, *4*(7), 1933–1943. https://doi.org/10.5194/acp-4-1933-2004

Li, X., Xiao, J., He, B, Arain, M.A., Beringer, J., Desai, A.R., Emmel, C., Hollinger, D.Y., Krasnova, A., Mammarella, I., Noe, S.M., Serrano Ortiz, P., Rey-Sanchez, A.C., Rocha, A.V., and Varlagin, A. (2018). Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global analysis based on OCO-2 and flux tower observations. *Glob Change Biol*. 24, 3990–4008. https://doi.org/10.1111/gcb.14297

Liu, W., Atherton, J., Mõttus, M., Gastellu-Etchegorry, J.P., Malenovský, Z., Raumonen, P., Åkerblom, M., Mäkipää, R. and Porcar-Castell, A. (2019). Simulating solar-induced chlorophyll fluorescence in a boreal forest stand reconstructed from terrestrial laser scanning measurements. Remote Sensing of Environment, 232, p.111274.

Mazzola, M., Tampieri, F., Viola, A., Lanconelli, C. and Choi, T. (2016), Stable boundary layer vertical scales in the Arctic: observations and analyses at Ny-Ålesund, Svalbard. Q.J.R. Meteorol. Soc., 142: 1250-1258, https://doi.org/10.1002/qj.2727.

Paasonen, P., Asmi, A., Petäjä, T., Kajos, M. K., Äijälä, M., Junninen, H., Holst, T., Abbatt, J. P. D., Arneth, A., Birmili, W., van der Gon, H. D., Hamed, A., Hoffer, A., Laakso, L., Laaksonen, A., Leaitch, W. R., Plass-Dülmer, C., Pryor, S. C., Räisänen, P., Swietlicki, E., Wiedensohler, A., Worsnop, D. R., Kerminen, V-M. & Kulmala, M. (2013) Warming-induced increase in aerosol number concentration likely to moderate climate change, Nature Geoscience, 6, 6, 438-442.

Paasonen, P., Peltola, M., Kontkanen, J., Junninen, H., Kerminen, V.-M., & Kulmala, M. (2018) Comprehensive analysis of particle growth rates from nucleation mode to cloud condensation nuclei in Boreal forest, Atmos. Chem. Phys., 18, 12085-12103.

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., Bjerke, J.W., Chen, A., Ciais, P., Tømmervik, H. and Nemani, R.R. (2019). Characteristics, drivers and feedbacks of global greening. Nature Reviews Earth & Environment, pp.1-14.

Porcar-Castell, A., Garcia-Plazaola, J.I., Nichol, C.J., Kolari, P., Olascoaga, B., Kuusinen, N., Fernández-Marín, B., Pulkkinen, M., Juurola, E. and Nikinmaa, E. (2012). Physiology of the seasonal relationship between the photochemical reflectance index and photosynthetic light use efficiency. Oecologia, 170(2), pp.313-323.

Porcar-Castell, A. (2011). A high-resolution portrait of the annual dynamics of photochemical and non-photochemical quenching in needles of Pinus sylvestris. Physiologia Plantarum, 143(2), pp.139-153.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Van der Tol, C., Flexas, J., Pfündel, E.E., Moreno, J., Frankenberg, C. and Berry, J.A. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. Journal of experimental botany, 65(15), pp.4065-4095.

Pulliainen, J., Aurela, M., Laurila, T., Aalto, T., Takala, M., Salminen, M., Kulmala, M., Barr, A., Heimann, M., Lindroth, A. and Laaksonen, A. (2017). Early snowmelt significantly enhances boreal springtime carbon uptake. Proceedings of the National Academy of Sciences, 114(42), pp.11081-11086.

Ryu, Y., Berry, J.A. and Baldocchi, D.D. (2019). What is global photosynthesis? History, uncertainties and opportunities. Remote sensing of environment, 223, pp.95-114.

Screen, J., Simmonds, I. The central role of diminishing sea ice in recent Arctic temperature amplification. Nature 464, 1334–1337 (2010). https://doi.org/10.1038/nature09051.

Seibert, P. (2000). Review and intercomparison of operational methods for the determination of the mixing height. Atmospheric Environment (1994), 34(7), 1001–1027. https://doi.org/10.1016/s1352-2310(99)00349-0

Vakkari, V., O'Connor, E., Nisantzi, A., Mamouri, R., & Hadjimitsis, D. (2015). Low-level mixing height detection in coastal locations with a scanning Doppler lidar. *Atmospheric Measurement Techniques*, *8*(4), 1875–1885. https://doi.org/10.5194/amt-8-1875-2015.

Vakkari, V. (2018). Mixing layer heigth measurements from Hyytiälä (2016-2018) with HALO photonics scanning doppler lidar. Unpublished raw data.

Vakkari, V., Manninen, A., O'Connor, E., Schween, J., van Zyl, P., & Marinou, E. (2019). A novel post-processing algorithm for Halo Doppler lidars. *Atmospheric Measurement Techniques*, *12*(2), 839–852., https://doi.org/10.5194/amt-12-839-2019.

Wiedensohler, A., Birmili, W., Nowak, A., Sonntag, A., Weinhold, K., Merkel, M., Wehner, B., Tuch, T., Pfeifer, S., Fiebig, M., Fjäraa, A. M., Asmi, E., Sellegri, K., Depuy, R., Venzac, H., Villani, P., Laj, P., Aalto, P., Ogren, J. A., Swietlicki, E., Williams, P., Roldin, P., Quincey, P., Hüglin, C., Fierz-Schmidhauser, R., Gysel, M., Weingartner, E., Riccobono, F., Santos, S., Grüning, C., Faloon, K., Beddows, D., Harrison, R., Monahan, C., Jennings, S. G., amp, apos, Dowd, C. D., Marinoni, A., Horn, H. G., Keck, L., Jiang, J., Scheckman, J., McMurry, P. H., Deng, Z., Zhao, C. S., Moerman, M., Henzing, B., de Leeuw, G., Löschau, G., and Bastian, S. (2012). Mobility particle size spectrometers: Harmonization of technical standards and data structure to facilitate high quality long-term observations of atmospheric particle number size distributions, Atmospheric Measurement Techniques, 5, 657-685, https://doi.org/10.5194/amt-5-657-2012.

Zhang, C., Atherton, J., Peñuelas, J., Filella, I., Kolari, P., Aalto, J., Ruhanen, H., Bäck, J. and Porcar-Castell, A. (2019). Do all chlorophyll fluorescence emission wavelengths capture the spring recovery of photosynthesis in boreal evergreen foliage?. Plant, cell & environment, 42(12), pp.3264-3279.

Zhang, C., Atherton, J., Rajewicz, P., Riikonen, A., Kolari, P., Fernández-Marín, B., Garcia-Plazaola, J. I., and Porcar-Castell, A. (2020). Seasonal dynamics of spectral vegetation indices at leaf, ecosystem and satellite scales for a boreal evergreen coniferous forest , EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-18337, https://doi.org/10.5194/egusphere-egu2020-18337.

Zhou, J., Swietlicki, E., Berg, O. H., Aalto, P. P., Hämeri, K., Nilsson, E. D., and Leck, C. (2001). Hygroscopic properties of aerosol particles over the central arctic ocean during summer, Journal of Geophysical Research: Atmospheres, 106, 32111-32123, https://doi.org/10.1029/2000jd900426.