BOREAL ENVIRONMENT RESEARCH 31: 45–62
ISSN 1797-2469 (online)

Helsinki 7 August 2026

Profound analysis of sulfuric acid proxies at a boreal forest site in southern Finland

Ilona Jaakkola[1)2)], Lubna Dada[3)], Tuomo Nieminen[1)4)], Nina Sarnela[1)], Tuukka Petäjä[1)], Kai Puolamäki[5)], Veli-Matti Kerminen[1)] and Markku Kulmala[1)]

[1)] *Institute for Atmospheric and Earth System Research / Physics, Faculty of Science, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland*
[2)] *Station for Measuring Ecosystem - Atmosphere Relations II (SMEAR II), University of Helsinki, Korkeakoski, 35500, Finland*
[3)] *Paul Scherrer Institute, Forschungsstrasse 111, 5232 Villigen PSI Switzerland*
[4)] *Department of Physics, Faculty of Science, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland*
[5)] *Department of Computer Science, Faculty of Science, University of Helsinki, P.O. Box 68, 00014 Helsinki, Finland*
**corresponding author's e-mail: ilona.a.jaakkola@helsinki.fi*

Received 13 Nov. 2025, final version received 15 May 2026, accepted 26 May 2026

Jaakkola I., Dada L., Nieminen T., Sarnela N., Petäjä T., Puolamäki K., Kerminen V.-M. & Kulmala M. 2026: Profound analysis of sulfuric acid proxies at a boreal forest site in southern Finland. *Boreal Env. Res.* 31: 45–62.

We analyzed four sulfuric acid proxies, based on known sources and sinks of sulfuric acid (OH, represented with global or UVB radiation, and stabilized Criegee intermediates oxidizing SO2, and condensing on preexisting particles and formation of sulfuric acid dimers). Using data measured between 2012 and 2023 at SMEAR II (Station for Measuring Ecosystem–Atmosphere Relations) in Finland, we show that the best-performing proxy had all four terms. Exclusion of the dimer term did not impact the performance considerably, but the Criegee term was essential for capturing the diurnal cycle of sulfuric acid. UVB and global radiation performed similarly and can therefore both be used in the estimation of sulfuric acid concentration. Obtaining fitting coefficients separately for each season increased the seasonal performance only slightly, implying that a large training dataset is sufficient to develop sulfuric acid proxies even if the measured predictor variables exhibit clear seasonal variation.

Introduction

Sulfuric acid (H_2SO_4) has a low saturation vapor pressure, and it can bind strong intramolecular bonds, which makes it a key component in atmospheric new particle formation (NPF; Weber *et al.* 1996, Kulmala *et al.* 2006, Sihto *et al.* 2006, Kulmala *et al.* 2014b, Lehtipalo *et al.* 2018). NPF occurs commonly on Earth and is a prominent source of atmospheric aerosol particles (Kulmala *et al.* 2004, 2016, Kerminen *et al.* 2018). Particles impact the radiative balance of the Earth by scattering and absorbing solar radiation and by altering cloud properties when acting as cloud condensation nuclei (CCN; Twomey 1977, Albrecht 1989, Ramanathan *et*

Editor in charge of this article: Antti Arola
https://doi.org/10.60910/ber2026.xs5t-v078

al. 2001, Lohmann and Feichter 2005, Kerminen *et al.* 2012, Gryspeerdt *et al.* 2014, Rosenfeld *et al.* 2014, IPCC 2021). Particles hence form an important step in the cooling feedbacks between ecosystems and the atmosphere (Kulmala *et al.* 2014a, Ezhova *et al.* 2018, 2025, Yli-Juuti *et al.* 2021, Artaxo *et al.* 2022). However, they are also involved in the formation of atmospheric haze, leading to reduced air quality and health risks (Laumbach and Kipen 2012, Guo *et al.* 2014, Huang *et al.* 2014, Brauer *et al.* 2016, Zamora *et al.* 2019, Wang *et al.* 2020).

The occurrence and intensity of NPF are sensitive to circumferential conditions, such as the number of preexisting particles and concentrations of precursor gases (Kerminen *et al.* 2018). Therefore, decreased trace gas concentrations resulting from legislation aimed at improving air quality (Kuula *et al.* 2022, Ylivinkka *et al.* 2025) have led to fewer and less intense NPF events in several locations (Hamed *et al.* 2010, Kyrö *et al.* 2014, Saha *et al.* 2018, Shen *et al.* 2022). In some locations, however, NPF frequency has not decreased (Salma *et al.* 2021) even if the NPF intensity decreased (Kalivitis *et al.* 2019). At SMEAR II (Station for Measuring Ecosystem-Atmosphere Relations) in Hyytiälä, Finland, the NPF frequency and particle formation rates have decreased during the period from 1996 to 2019 although particle growth rate did not have a significant trend (Li *et al.* 2024). For the period of 1996–2012, Nieminen *et al.* (2014) showed increasing particle growth and formation rates despite the decreasing trend of precursor vapors, but they concluded that the NPF was favored by temperature and air mass origin. Further, Dada *et al.* (2017) showed that NPF is more likely to occur in temperatures between –21 to 25 °C and clear sky conditions at SMEAR II. However, the mechanisms behind NPF and aerosol–cloud–radiation interactions are not yet fully understood, and therefore it is important to study such changes and their consequences on atmospheric processes (Arneth *et al.* 2009, Kulmala 2015, IPCC 2021, Shen *et al.* 2021, Feng *et al.* 2022, Yan *et al.* 2022).

Sulfuric acid is challenging to measure due to its low atmospheric concentration (approximately $10^5 - 10^7$ molec. cm^{-3}), and both instrument maintenance and data processing require specialized expertise (Junninen *et al.* 2010, Mikkonen *et al.* 2011). Therefore, it is routinely measured only at a few measurement stations, and the available data sets are commonly relatively short, making it challenging to gain a comprehensive understanding of boundary layer processes or to quantify long-term trends. Thus, various sulfuric acid proxies have been developed for atmospheric studies as a substitute for the missing measurement data on sulfuric acid.

The proxies are based on the known source and sink reactions of sulfuric acid. The first proxy, developed using measurement data from SMEAR II, accounted for the formation of sulfuric acid due to hydroxyl radicals (OH) oxidizing sulfur dioxide (SO_2), and loss of sulfuric acid due to condensing on preexisting particles (condensation sink, CS; Petäjä *et al.* 2009). In 2011, Mikkonen *et al.* (2011) introduced statistical proxies using measurement data from several locations in Europe and the USA and concluded that the best performing proxy accounted for global radiation, SO_2 concentration, condensation sink, and relative humidity as predictor variables.

The simple sulfuric acid proxy by Petäjä *et al.* (2009) performs relatively well but still underestimates the sulfuric acid concentration. Therefore, Dada *et al.* (2020) revised the proxy and added new source and sink terms: oxidation of sulfur dioxide by stabilized Criegee intermediates (sCI) and formation of sulfuric acid dimers. When the Criegee term was included, the diurnal cycle of sulfuric acid concentration was correctly estimated at several measurement locations, including Cyprus, Hungary, and China. The formation of hydroxyl radicals requires solar radiation, and as a primary atmospheric oxidant, they are quickly consumed after sunset. Stabilized Criegee intermediates, however, are produced in ozonolysis of alkenes, and are thus available for oxidation both day and nighttime (Mauldin III *et al.* 2012, Boy *et al.* 2013, Sipilä *et al.* 2014, Zhu *et al.* 2026). Therefore, stabilized Criegee intermediates are more significant during dark periods.

All proxies, and their modified versions, have since been tested also in other environments, including China (Lu *et al.* 2019, Yang *et al.* 2021, Guo *et al.* 2026), Amazon (Myers *et al.*

2022), and volcano plumes (Rose *et al.* 2021). Generally, the proxies perform well during daytime and outside of major sources of SO_2, but often local adjustments or Criegee intermediate term are needed for better overall performance. Moreover, when relying on short campaigns, the seasonal variation are not accounted for.

In this work, we further investigate sulfuric acid proxies by Dada *et al.* (2020). We aim to compare the different proxies to determine the necessary terms for a reliable sulfuric acid proxy and, on the other hand, to identify the terms that lead to the best possible proxy for sulfuric acid at SMEAR II. Unlike in Dada *et al.* (2020), we employ both UVB and global radiation as estimates for OH concentration to discern the differences caused by the selected radiation type and determine which one is more suitable for proxies. Furthermore, we also analyze the data seasonally; for example, the widely used proxy by Petäjä *et al.* (2009) was developed using only a short dataset from spring and early summer. Additionally, we estimate the accuracy of the proxies with an independent dataset. While the fitted coefficients derived here are directly applicable to SMEAR II, comparing the relative roles of the proxy source and sink terms is expected to be informative also beyond this site.

Sulfuric acid proxies

We investigated four sulfuric acid proxies introduced in Dada *et al.* (2020) to analyze their performance in estimating sulfuric acid concentration at SMEAR II. Since OH is not measured continuously at SMEAR II, we used solar radiation to represent OH concentration. Earlier studies have shown that both UVB and global radiation correlate well with the OH concentration, and thus they can be used as substitutes for the OH concentration in proxies (Berresheim *et al.* 2002, Rohrer and Berresheim 2006, Petäjä *et al.* 2009, Dada *et al.* 2020). In this study we fitted proxies with both UVB and global radiation to evaluate the disparities between the radiation types when used in proxies to estimate the OH concentration.

Proxy 1 accounts for the formation of sulfuric acid due to OH radicals (the first term on the right-hand side in Eq. 1) and stabilized Criegee intermediates (the second term), and loss due to condensing on preexisting particles (the third term), and formation of molecular dimers (the fourth term):

$$\begin{aligned}\frac{\mathrm{d}[H_2SO_4]}{\mathrm{d}t} = {} & k_{\mathrm{RAD}}\mathrm{RAD}[SO_2] \\ & + k_{\mathrm{sCI}}[O_3][\mathrm{MT}][SO_2] \\ & - \mathrm{CS}[H_2SO_4] - k_{\mathrm{DIM}}[H_2SO_4]^2.\end{aligned} \quad (1)$$

In Eq. 1, k_{RAD}, $k_{\mathrm{sCI,}}$ and k_{DIM} are fitting coefficients, RAD refers to either UVB or global radiation, and MT refers to monoterpene concentration.

Criegee intermediates are formed in the ozonolysis of alkenes (Mauldin III *et al.* 2012, Boy *et al.* 2013, Sipilä *et al.* 2014). At SMEAR II, the most common alkenes are monoterpenes, which are emitted by the surrounding conifers (Rinne *et al.* 2005, Hakola *et al.* 2012, Hellén *et al.* 2018). Hence, we used monoterpene concentration as an estimate of the ambient alkene concentration, following the approach of Dada *et al.* (2020).

In proxy 2, we excluded the Criegee intermediate source term to evaluate its importance for the sulfuric acid proxy:

$$\begin{aligned}\frac{\mathrm{d}[H_2SO_4]}{\mathrm{d}t} = {} & k_{\mathrm{RAD}}\mathrm{RAD}[SO_2] \\ & - \mathrm{CS}[H_2SO_4] - k_{\mathrm{DIM}}[H_2SO_4]^2.\end{aligned} \quad (2)$$

Proxy 3 excludes the dimer formation term:

$$\begin{aligned}\frac{\mathrm{d}[H_2SO_4]}{\mathrm{d}t} = {} & k_{\mathrm{RAD}}\mathrm{RAD}[SO_2] \\ & + k_{\mathrm{sCI}}[O_3][\mathrm{MT}][SO_2] - \mathrm{CS}[H_2SO_4].\end{aligned} \quad (3)$$

Proxy 4 is similar to the proxy by Petäjä *et al.* (2009), with the exception that the fitting coefficient does not have a power-law dependency on radiation:

$$\begin{aligned}\frac{\mathrm{d}[H_2SO_4]}{\mathrm{d}t} = {} & k_{\mathrm{RAD}}\mathrm{RAD}[SO_2] \\ & - \mathrm{CS}[H_2SO_4].\end{aligned} \quad (4)$$

The power-law dependency on radiation was excluded because it did not improve the results, unlike in Petäjä *et al.* (2009), and because it lacks a theoretical background. The proxies were additionally compared to proxies by Petäjä *et al.* (2009) and Mikkonen *et al.* (2011). Proxies developed with UVB radiation were compared with the UVB proxy by Petäjä *et al.* (2009, eq. 3), and proxies developed with global radiation are compared with the global radiation proxies by Petäjä *et al.* (2009, eq. 4) and Mikkonen *et al.* (2011, eq. 9).

We further tested a proxy with an additional fitting coefficient in the CS term:

$$\frac{\mathrm{d}[\mathrm{H_2SO_4}]}{\mathrm{d}t} = k_{\mathrm{RAD}}\mathrm{RAD}[\mathrm{SO_2}] + k_{\mathrm{sCI}}[\mathrm{O_3}][\mathrm{MT}][\mathrm{SO_2}] - k_{\mathrm{CS}}\mathrm{CS}[\mathrm{H_2SO_4}] - k_{\mathrm{DIM}}[\mathrm{H_2SO_4}]^2. \quad (5)$$

With this, we attempted to estimate the efficiency of the CS, which is lower than anticipated, for example, in the case of sulfuric acid (Tuovinen *et al.* 2020). However, the proxy results in an equation where one fitting coefficient is divided by another (Eq. S5 in the Supplementary Information, SI), which is computationally unstable. Therefore, the results given by this proxy varied significantly between refits and depended on the selected initial starting points of the fitting. Hence, we abandoned this proxy from further analysis. Table 1 summarizes the utilized source and sink terms of the fitted proxies.

Fitting method

We assumed steady-state conditions between the sources and sinks of sulfuric acid and solved the proxies (Eqs. 1–4) for sulfuric acid concentration. The solved forms of the equations are presented in SI (Eq. S1-S5). Fitting was conducted by minimizing the sum of the squared logarithm of the ratio between sulfuric acid concentration given by the proxies and the measured sulfuric acid concentration. We used MATLAB's GlobalSearch algorithm with *fmincon* as the local solver, employing the sequential quadratic programming (SQP) algorithm for constrained nonlinear optimization (Nocedal and Wright 2006). The method gave the optimal fitting coefficients (k_{RAD}, k_{sCI}, and k_{DIM}) for the different proxies. To reduce the amount of bias in the fitting coefficients caused by the optimization, we refitted the data 100 times using the same parameters. We selected the fitting coefficients associated with the smallest value of the loss function.

Confidence intervals for the fitting coefficients were estimated using block bootstrapping. The data was resampled 500 times, and as a block size, we used 24 data points, equivalent to a lag of three days, since we have data from every third hour. This way, we can exclude most of the autocorrelation in atmospheric time series data (Figs. S1 and S2 in SI). We report basic bootstrap confidence intervals of coefficients falling between the 5th and 95th percentiles.

Measurement site, instrumentation, and methods

Site

The measurements were conducted at the rural SMEAR II research station in Hyytiälä, southern Finland (61.51°N, 24.17°E, 181 m a.s.l.; Hari and Kulmala 2005). A 64-year-old Scots pine-dominated forest surrounds the station, with a canopy height of approximately 20 m. The nearest cities are Tampere (50 km south-

Table 1. Table illustrates the utilized source and sink terms of sulfuric acid in each proxy fitted in this paper.

	Source terms		Sink terms	
Proxy 1	$k_{\mathrm{RAD}}\mathrm{RAD}[\mathrm{SO_2}]$	$k_{\mathrm{sCI}}[\mathrm{O_3}][\mathrm{MT}][\mathrm{SO_2}]$	$\mathrm{CS}[\mathrm{H_2SO_4}]$	$k_{\mathrm{DIM}}[\mathrm{H_2SO_4}]^2$
Proxy 2	$k_{\mathrm{RAD}}\mathrm{RAD}[\mathrm{SO_2}]$		$\mathrm{CS}[\mathrm{H_2SO_4}]$	$k_{\mathrm{DIM}}[\mathrm{H_2SO_4}]^2$
Proxy 3	$k_{\mathrm{RAD}}\mathrm{RAD}[\mathrm{SO_2}]$	$k_{\mathrm{sCI}}[\mathrm{O_3}][\mathrm{MT}][\mathrm{SO_2}]$	$\mathrm{CS}[\mathrm{H_2SO_4}]$	
Proxy 4	$k_{\mathrm{RAD}}\mathrm{RAD}[\mathrm{SO_2}]$		$\mathrm{CS}[\mathrm{H_2SO_4}]$	

west; 260 000 inhabitants) and Jyväskylä (90 km northeast; 149 000 inhabitants).

Ambient measurements

We used data measured at SMEAR II during five periods between 2012 and 2023 (22.3.–7.5.2012, 8.4.–23.6.2013, 18.8.2016–17.4.2017, 8.3.2018–28.5.2020, and 23.5.2022–29.5.2023). The data availability, compared to the maximum data availability during these periods, was 57%. The measurement data were averaged to hourly medians. However, monoterpenes, needed for the Criegee intermediate concentration estimation, were measured only every third hour (see Gas measurements section), leading to only eight data points per day.

All data points with zero values were excluded from the determination of fitting coefficients, as they sometimes resulted in unrealistic fitting coefficients for the proxies. This limited available data points from 7 031 to 5 085 (72%).

When fitting proxies without the Criegee source term (proxies 2 and 4), we additionally set minimum thresholds for the incoming solar radiation. The thresholds were set because those proxies are not valid when radiation, producing hydroxyl radicals, is not present (Petäjä *et al.* 2009, Mikkonen *et al.* 2011). The radiation thresholds were 0.015 Wm^{-2} for UVB and 10 Wm^{-2} for global radiation, in accordance with Mikkonen *et al.* (2011). The thresholds represent approximately the 20th percentile of radiation data. The number of data points with a radiation threshold was 3 258, resulting in a data availability of 64% compared to the otherwise used dataset without zero values.

To evaluate the performance of the proxies, these datasets were further divided into separate training and test datasets. To avoid autocorrelation, we draw 12 random one-week-long segments of the data as a test dataset, corresponding roughly to 10% of the entire observation dataset (672 data points, 461 with radiation limits).

When performing seasonal analysis, the seasons were divided as follows: spring included March, April and May (1 698 / 1 201; 361 / 262), summer June, July and August (1 225 / 932; 191 / 150), autumn September, October and November (983 / 502; 168 / 70), and winter December, January and February (379 / 122; 80 / 20). The first two numbers before the semicolon indicate the number of data points in the training dataset during each season, and the latter indicates the number of data points when applying the radiation limit. The numbers after the semicolon refer to the testing dataset, respectively.

Gas measurements

Sulfuric acid was measured with a Chemical Ionization Atmospheric Pressure interface Time-of-Flight spectrometer (CI-APi-ToF; Tofwerk AG, Thun, Switzerland, and Aerodyne Research Inc., USA; Eisele and Tanner 1993, Jokinen *et al.* 2012). The nitrate ion (NO_3^-) was used as a reagent ion, and the data were processed using the *tofTools* package in MATLAB (Junninen *et al.* 2010). The measurements since 2016 were conducted at a height of 35 m above the ground level and before that at ca. 1.5 m above ground. The instrument was calibrated with calibration method where sulfuric acid is produced *in situ* from SO_2 with mercury lamp (Kürten *et al.*, 2012). The concentrations were calculated with unit mass resolution by normalizing the sulfuric acid signals (HSO_4^- and $HNO_3HSO_4^-$) with reagent ion signals (NO_3^-, $HNO_3\,NO_3^-$, and $(HNO_3)_2NO_3^-$) and multiplying them with a calibration factor.

The measurements of sulfur dioxide, ozone, and monoterpene concentrations were conducted at a height of 16.8 m above ground level (Hari and Kulmala 2005). The sulfur dioxide analyzer was Model 43i-TLE (Thermo Fisher Scientific, USA), and the ozone analyzer was Model 49C (Thermo Fisher Scientific, USA). Monoterpenes were measured with a proton transfer reaction quadrupole mass spectrometer (PTR-QMS, Ionicon Analytik GmbH). Monoterpene signal was calibrated regularly using standard gas by Apel–Riemer Environmental Inc. (Taipale *et al.* 2008). Ambient monoterpene concentration was measured only every third hour at SMEAR II, which limits our whole dataset to eight measurement points per day when using hourly median data.

Condensation sink

The condensation sink was calculated from ambient ground-level particle number size distribution data, using the method suggested by Kulmala *et al.* (2012). We utilized integrated data from a twin differential mobility particle sizer (DMPS) and an aerodynamic particle sizer (APS; Aalto *et al.* 2001, Ylivinkka *et al.* 2025). We assumed that the density of particles at SMEAR II is 1.5 g cm^{-3} when converting aerodynamic particle diameters measured by APS to Stoke's diameters (Kannosto *et al.* 2008). The hygroscopic growth of particles was corrected when calculating the condensation sink (Laakso *et al.* 2004).

Radiation

UVB radiation was measured with a Solar SL 501A pyranometer (wavelength range 280–320 nm) and global radiation with a Middleton solar EQ08-S pyranometer (wavelength range 300–3000 nm). Radiation measurements were conducted above the canopy level.

Methods

The coefficient of determination is calculated as:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \quad (6)$$

where y_i are the measured sulfuric acid values, $\hat{y}_i$ proxy values, and $\bar{y}$ is the mean of the measured values. Logarithms of the measured sulfuric acid and proxy values were used in the calculation.

Linear regression between proxy and measured sulfuric acid values is performed on logarithmic scale using the maximum likelihood effective-variance method (Orear 1982).

Results and discussion

Fitting coefficients and confidence intervals

Utilizing the training dataset, we fitted four sulfuric acid proxies (Eqs. 1–4), using both UVB and global radiation, against the measured sulfuric acid data from SMEAR II to obtain fitting coefficients (k_{RAD}, k_{sCI}, and k_{DIM}). We refitted the data 100 times to assess the sensitivity of the fitted coefficients to the optimization procedure. We selected the coefficients corresponding to the lowest loss function (sum of squared logarithms of the ratio between proxy values and measured sulfuric acid concentrations) values (Table 2).

Table 2. Fitting coefficients for different proxies (Eqs. 1–4), confidence intervals given by block bootstrapping, and coefficients of determination (R2), using both UVB and global radiation (GlobRad). Values are obtained using the training dataset.

	k_{RAD} (m^2 W^{-1} s^{-1})	k_{sCI} ($\times 10^{-29}$ cm^6 s^{-1})	k_{DIM} ($\times 10^{-9}$ cm^3 s^{-1})	R^2
UVB				
Proxy 1	6.030 (4.4–8.9)$\times 10^{-6}$	5.693 (4.9–6.8)	4.350 (2.1–9.1)	0.58
Proxy 2	14.12 (11–24)$\times 10^{-6}$	–	10.99 (6.6–24)	0.45
Proxy 3	2.540 (2.2–2.8)$\times 10^{-6}$	5.209 (4.6–5.9)	–	0.55
Proxy 4	4.602 (4.0–5.3)$\times 10^{-6}$	–	–	0.26
GlobRad				
Proxy 1	4.884 (3.7–8.3)$\times 10^{-9}$	5.026 (4.3–6.2)	4.531 (2.1–12)	0.60
Proxy 2	7.387 (5.9–10)$\times 10^{-9}$	–	5.740 (3.9–10)	0.52
Proxy 3	2.079 (1.8–2.4)$\times 10^{-9}$	4.763 (4.2–5.4)	–	0.56
Proxy 4	3.362 (2.8–3.8)$\times 10^{-9}$	–	–	0.36

Histograms of the coefficient distributions are shown in the SI (Figs. S3–S9); for visualization, these histograms include the 75 fits associated with the lowest loss function values, although all 100 fits were included in the numerical evaluation. Although the repeated fits show some spread in the fitted coefficients, the resulting proxy performance among the fits with lowest loss function values remains very similar, indicating that the main conclusions are robust to the exact choice among the best-performing fits.

To further visualize the range of optimal fitting coefficients, we plotted the relation between the minimized loss function and different fitting coefficient values (Fig. S10–S11 in SI). The repeated fits indicate that the loss function contains broad regions of low loss function values rather than a single sharply defined optimum. In practice, several coefficient combinations gave very similar loss values and very similar proxy performance. Therefore, the fitted coefficients should be interpreted as effective fitting parameters rather than as uniquely identifiable physical quantities. However, as the selected fitting coefficients (Table 2) were found in the area of the lowest loss function values, we gained confidence that they were the best-performing coefficients for these proxies.

Confidence intervals for the fitting coefficients were obtained by subjecting the data to block bootstrapping 500 times. The 5th and 95th percentiles of fitting coefficients from the resampled data are shown in brackets (Table 2). Additionally, we plotted the fitting coefficient values from the block bootstrap resamples against each other and colored the plots with loss function values (Fig. S12–S15 in SI). These figures illustrate the relationships between the coefficients and the loss function values. For example, for proxy 1 with UVB radiation, k_{RAD} and k_{DIM} had a rather linear relationship on a logarithmic scale, and the loss function values were uniformly distributed (Fig. S12b in SI). Hence, the coefficients cause a similar level of uncertainty to the proxy. In k_{sCI} against k_{RAD} (Fig. S12a in SI), the data points were more scattered, suggesting that knowing optimal values for k_{RAD} and k_{DIM} does not help in finding an optimal value for k_{sCI}. The uncertainty in k_{sCI} is also visible when comparing the resampled fitting coefficients of proxies 2 and 3 (Figs. S13 and S14 in SI).

Proxy 1 with global radiation produced a few much larger coefficient values (Fig. S12c,d in SI). These coefficients were also associated with high loss function values, which indicated that they were not optimal fitting coefficients for the proxy.

For proxy 4, the loss function values were plotted against its single fitting coefficient (Fig. S15 in SI). The coefficients given in Table 2 for proxy 4 are found in the middle of the fitting coefficient range, giving confidence in their optimality.

Comparison with measured sulfuric acid

Proxies with the Criegee source term (proxies 1 and 3) had the highest coefficient of determination with measured sulfuric acid (R^2 about 0.6; Figs. 1 and S16 in SI), although the data points deviated from the 1:1 line. Surprisingly, the simplest proxy (proxy 4) with a radiation limit yielded values that aligned with the 1:1 line. The R^2 values for proxy 4, however, were lower (0.26 and 0.36 for UVB and global radiation-derived proxies, respectively).

The diurnal variation of the proxies and measured sulfuric acid values highlights the importance of the Criegee source term for the performance of the proxies (Fig. 2). Proxies with the Criegee term (proxies 1 and 3) performed well, with the values of diurnal variation always falling between 25th and 75th percentiles of the measured sulfuric acid concentration. In contrast, the proxies without the Criegee term did not reach the correct nighttime values of sulfuric acid, as their only source term during nighttime drops to zero (see Fig. 7b; Mikkonen *et al.* 2011). Although proxy 2 did not produce the correct nighttime cycle of sulfuric acid, it performed well during daytime (approx. 05:00–17:00; Fig. 2). However, out of the proxies not including the Criegee term (excluding proxies 1 and 3), the proxy by Petäjä *et al.* (2009) produced overall the best performing diurnal cycle of sulfuric acid, particularly when using global radiation, although it mostly underestimated the measured sulfuric acid concentration. The better nighttime

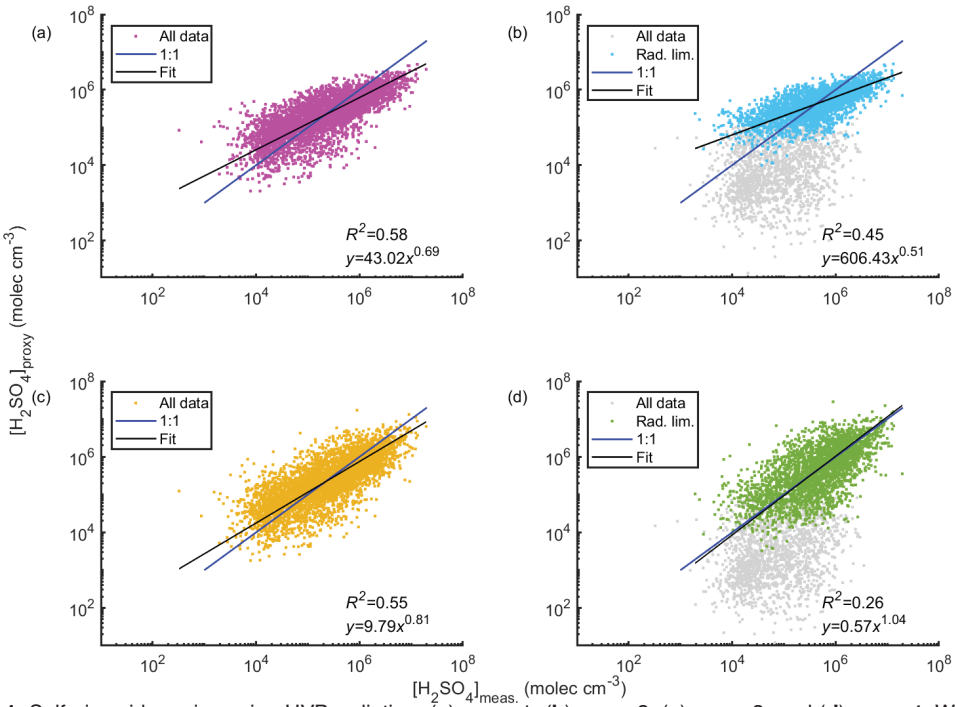

Fig. 1. Sulfuric acid proxies using UVB radiation. (**a**) proxy 1, (**b**) proxy 2, (**c**) proxy 3, and (**d**) proxy 4. We applied radiation limits for proxies 2 and 4 (colored markers). Without radiation limits (gray markers), the R^2 values were –0.32 and –1.16, and fitted lines $y = 3\times10^{-4}x^{1.60}$ and $y = 2\times10^{-7}x^{2.18}$ for proxy 2 and proxy 4, respectively (fitted lines for all data are not shown).

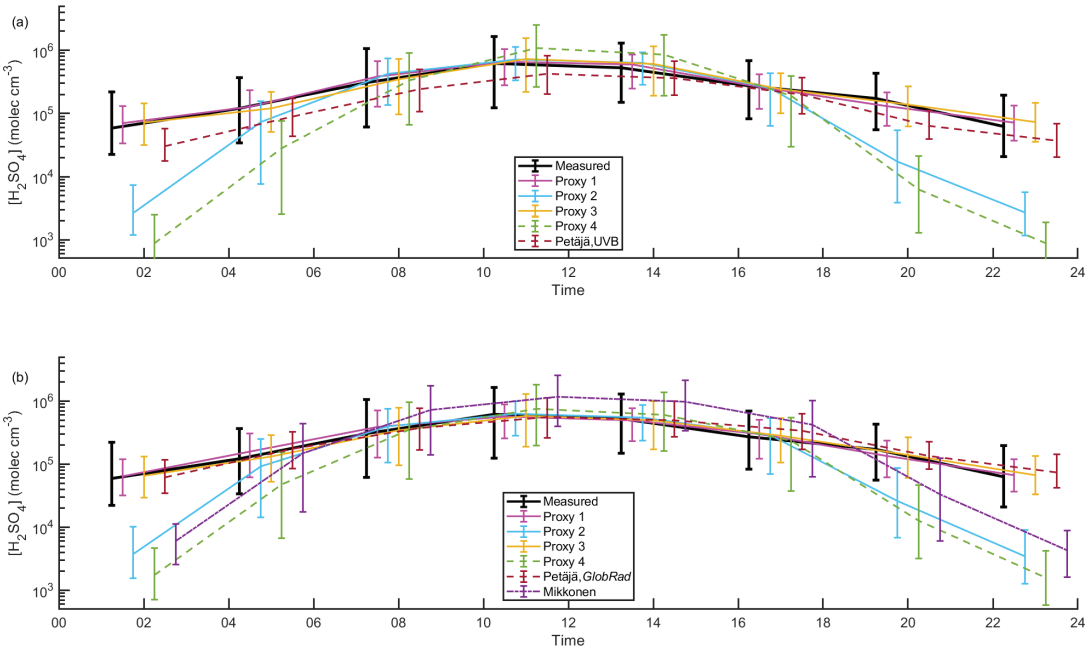

Fig. 2. Hourly averaged diurnal variation of sulfuric acid concentration applying different proxies with (**a**) UVB radiation and (**b**) global radiation. Bars indicate the 25th and 75th percentiles of the data. Petäjä refers to the sulfuric acid proxy by Petäjä *et al.* (2009), and Mikkonen to the sulfuric acid proxy by Mikkonen *et al.* (2011). Notice that due to the time resolution of monoterpene measurements, we have data points only every third hour.

performance of the proxy by Petäjä *et al.* (2009) may be attributed to the fitting coefficient, which exhibited a power-law dependence on radiation intensity.

Proxies 1 and 3 performed similarly using both UVB and global radiation-derived proxies; however, with other proxies, global radiation appeared to be slightly more favorable for the better performance. We used all the data in the diurnal cycle figures without limiting the radiation intensity in any of the proxies.

Residuals

We noticed that the proxy values deviated from the 1:1 line (Figs. 1 and S16 in SI), which is caused by the underestimation of the largest sulfuric acid values and overestimation of the smallest values. It appears that the proxies tended to smooth the observed concentrations, although they otherwise accurately predicted the temporal dynamics. To understand how each measured predictor variable relates to the performance of the proxies and whether the under- and overestimation could be explained by the behavior of the predictor variables, we plotted residuals (proxy values subtracted from the linear model) against the utilized predictor variables (Fig. S17 in SI).

The residuals were mainly independent of SO_2, O_3, and MT concentrations, as well as CS, indicating that these variables did not systematically cause under- or overestimation of the proxy values. Instead, we observed that the under- and overestimation of sulfuric acid concentration was greatest when the radiation level was low (Fig. S17f in SI). Additionally, a slight overestimation trend was associated with increasing radiation intensity; however, adding an exponential fitting coefficient, as in Petäjä *et al.* (2009), did not improve the performance of the proxies. Moreover, although the largest under- and overestimation is related to low radiation intensities, further limiting the radiation data did not lead to improved performance of the proxies. It is also worth noting that values close to the limit of detection of the instruments may cause uncertainty to the proxies even though it is not explicitly visible in Fig. S17 in SI. The different proxies had somewhat similar dependence on the utilized predictor variables, as proxy 1 with UVB radiation, shown in Fig. S17 in SI.

Test dataset

We evaluated the performance of the proxies with an independent test dataset. The calculated R^2 values for the test dataset were lower than those for the training dataset (in all cases, $R^2 < 0.45$), but the results otherwise aligned with those from the training dataset (Fig. 3 and S18–S19 in SI). Figures also show that the data points are more scattered with low sulfuric acid values, which may be related to the uncertainty of the data below the limit of detection, which generally is about 4×10^4 molec. cm^{-3} (Jokinen *et al.* 2012). The uncertainty of the measured sulfuric acid concentration introduces an additional source of uncertainty in the proxy evaluation. Our laboratory calibrations indicate an uncertainty of approximately 25–33% in the sulfuric acid calibration factor (unpublished results), which is similar to that published by Kürten *et al.* (2012) (30–35%). This systematic uncertainty is considered to be the same for all concentration levels, but the relative contribution of signal variability and background subtraction becomes larger at low sulfuric acid concentrations, leading to increased scatter in comparisons between measured and proxy-derived concentrations. In the Cosmics Leaving Outdoor Droplets (CLOUD) experiments, a more conservative uncertainty factor of approximately 1.5 has sometimes been adopted when comparing measurements with independent production estimates (Dunne *et al.* 2016).

For the proxy evaluation presented here, the impact of this uncertainty is expected to be limited. Systematic uncertainties in the calibration factor primarily affect the absolute concentration scale, while the proxy analysis focuses on the ability of the proxy to reproduce the temporal variability of sulfuric acid concentrations. A constant calibration bias would therefore primarily influence the absolute concentration scale and the slope of the regression between measured and proxy-derived concentrations, while increased uncertainty at low concentrations may also increase the scatter in the comparison.

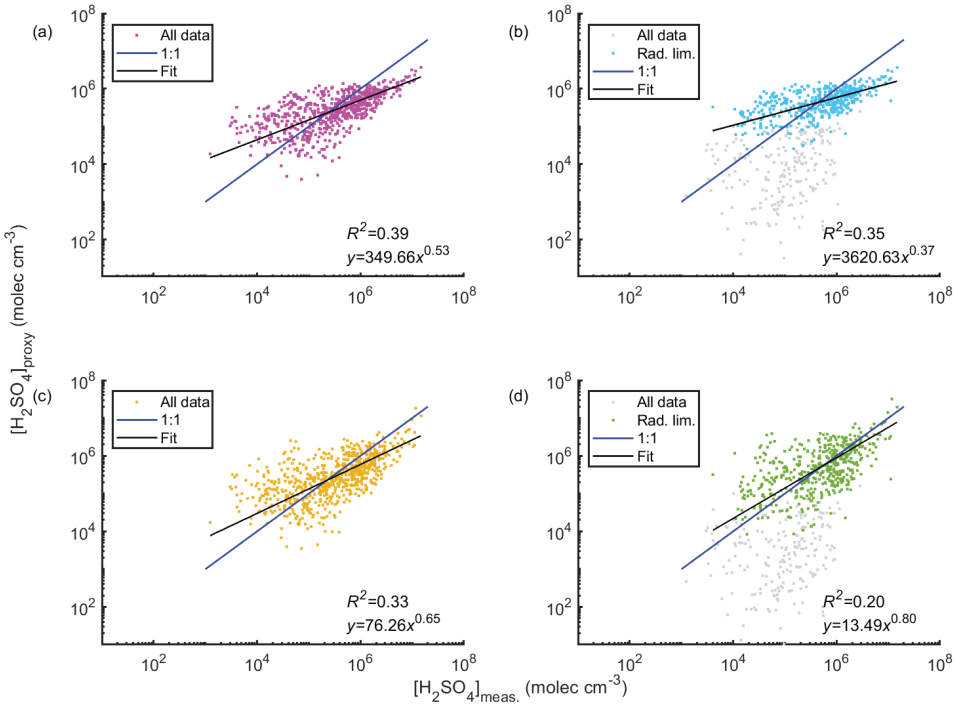

Fig. 3. Sulfuric acid proxies with test dataset using UVB radiation. (**a**) proxy 1, (**b**) proxy 2, (**c**) proxy 3, and (**d**) proxy 4. We applied radiation limits for proxy 2 and proxy 4 (colored markers). Without the radiation limits (gray markers), the R^2 values were –0.57 and –1.37, and fitted lines $y = 9\times10^{-5}x^{1.68}$ and $y = 1\times10^{-8}x^{2.38}$ for proxy 2 and proxy 4, respectively (fitted lines for all data not shown).

We also compared proxy 1 with the global radiation proxy corrected for the hygroscopic growth provided in Dada et al. (2020), using the test dataset. We noticed that they had similar coefficients of determination with the measured sulfuric acid concentration (Fig. S20 and S21 in SI). However, the proxy by Dada et al. (2020) tended to overestimate nighttime concentrations of sulfuric acid. In contrast, during the daytime, the values fell within the 25th and 75th percentiles of the measured sulfuric acid (Fig. S20).

Similar behavior was expected from the proxies, as they are essentially identical. The only differences are the slightly different training datasets (the dataset is longer in this study, and we use CS derived from the combined DMPS and APS dataset) and how the fitting coefficients were derived. In this paper, the coefficients were derived using the global optimization function fmincon of MATLAB, whereas Dada et al. (2020) used the fminsearch function of MATLAB. fminsearch uses the Nelder–Mead simplex method to find a local minimum in an unconstrained problem (Lagarias et al. 1998). Regarding proxies 2 and 4, the fitting coefficients deviate more significantly from those presented in Dada et al. (2020), which is likely due to the radiation thresholds used in this study.

Finally, we assessed the sensitivity of the test dataset selection by evaluating the performance of the proxies separately for each of the 12 one-week-long test segments. We summarize this variability using the standard error of the mean of the segment-wise MSE values, together with the median and interquartile range of the MSEs and the relative ranking of the proxies across the segments (Tables S1 and S2 in SI). The medians and 25th and 75th percentiles of the MSEs (Table S1 in SI), and a table of the relative performance order between the proxies (Table S2 in SI) show that there is variability in the performance across the test data segments. However, the results indicate that the proxies perform mostly similarly between the different test data segments as

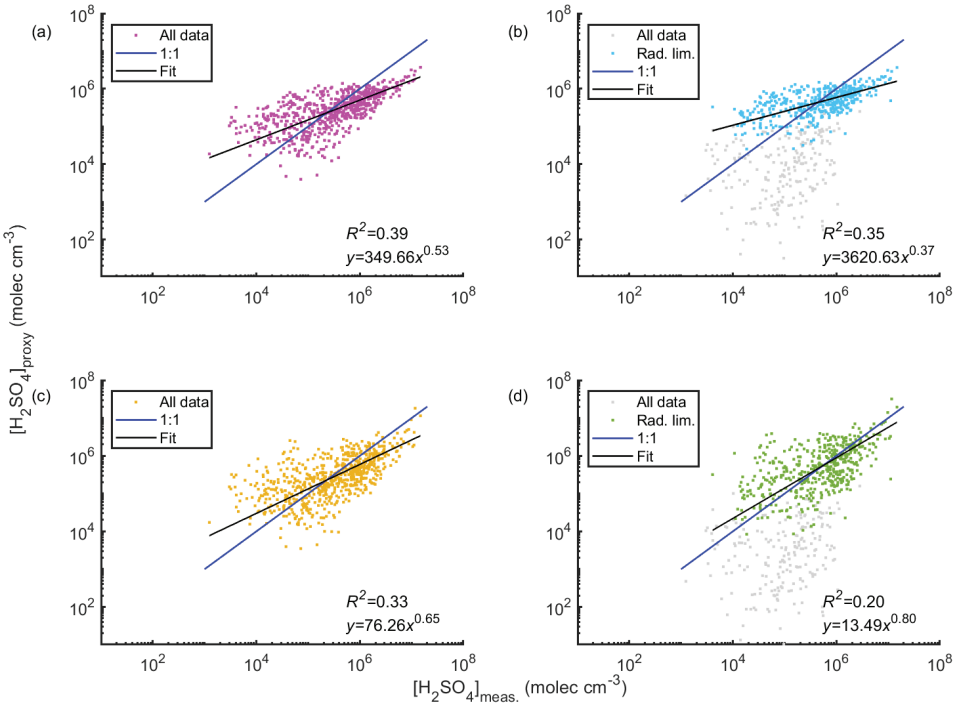

Fig. 4. Proxy 1 using UVB radiation divided seasonally: (**a**) spring, (**b**) summer, (**c**) autumn, and (**d**) winter.

they do with the full test dataset. Although the exact ranking among the best-performing proxies is not fully stable across segments, the main conclusions remain unchanged, and the clearly worse proxy formulations can still be distinguished from the better-performing group.

Seasonal variation

To understand how seasonality of the measured predictor variables affects the proxies, we (1) plotted the data seasonally using the fitting coefficients gained with the whole training dataset (Fig. 4, Table 2), and (2) refitted the training dataset separately for each season (Fig. 5 and Table S3 in SI). The R^2 values were lower for the seasonally divided data compared to the entire training dataset with the exception that proxies 2 and 4 (proxies without the Criegee term) produced higher R^2 values in spring and winter (Fig. 1, 4, and S22–S28 in SI). The lowest R^2 values were mostly related to summertime data, which may also be attributed to the lower variability in sulfuric acid concentration during summer.

Fitting data seasonally increased the R^2 values on average by about 20% compared to seasonally divided data (Table S3 in SI). The coefficients for proxies 3 and 4 (proxies without the dimer term) varied by up to 50% between seasons, whereas in proxies 1 and 2, the differences reached almost 100%. For proxy 1, all the coefficients were higher in autumn and winter compared to spring and summer, while for proxy 2 the coefficients were generally higher in summer.

The differences in the fitted coefficients likely reflect both seasonal differences in the data and the fact that the fitted coefficients are essentially effective parameters rather than direct physical rate constants. Because the source and sink terms are partly interdependent and there are broad regions with low loss function values, different coefficient combinations can lead to similar proxy performance. Seasonal variation in measured predictor variables may nevertheless contribute to the observed differences between seasons as in autumn and winter, all predictor

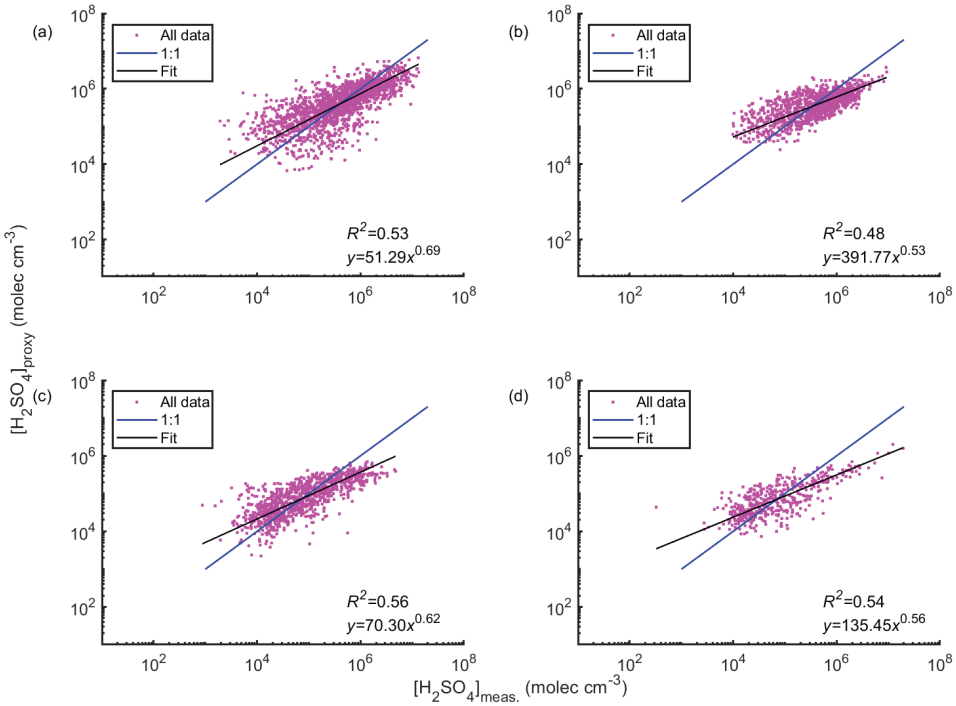

Fig. 5. Proxy 1 using UVB radiation fitted seasonally: (**a**) spring, (**b**) summer, (**c**) autumn, and (**d**) winter.

variables have lower average values than in spring and summer. Additionally, the radiation and Criegee terms exhibit strong seasonal variability (Fig. 7a), and hence proxies 1 and 2 might also show stronger seasonal variability in their fitting coefficients.

Note, however, that due to the distribution of the measurement data, we had fewer autumn and winter data points compared to other seasons, which might at least partially impact our results. The division of data seasonally did not affect the performance of the proxies in producing the diurnal cycle (Fig. S29 and S30 in SI).

Term contributions

We analyzed how the fitted source and sink terms contribute within each proxy by first calculating, for each proxy separately, the median value of each individual source and sink term over the analyzed dataset. Each source-term median was then expressed relative to the sum of the source-term medians, and each sink-term median relative to the sum of the sink-term medians (Fig. 6). These quantities are intended as a descriptive summary of the relative importance of the fitted proxy terms.

As the diurnal cycle already indicated, the dimer term (purple bar in Fig. 6) is a less important sink term than the CS term (yellow bar in Fig. 6) in proxy 1. Interestingly, however, the relative contribution of the Criegee term (red bar in Fig. 6) was only slightly higher, and even smaller in the case of global radiation, than the contribution of dimer term, although the Criegee term was essential in producing the diurnal cycle correctly.

In proxies 1 and 3, the radiation term (blue bar in Fig. 6) in global radiation derived proxies has a larger contribution than in UVB derived proxies. This might explain why proxies derived with global radiation seem to perform better. For proxy 2 without the Criegee term, the contribution of the dimer term became considerably larger than the CS term (Fig. 6). This behavior along with the fact that global minima could not be found for the fitting may suggest that the proxies with the gained fitting coefficients do not necessarily represent the actual strength of the

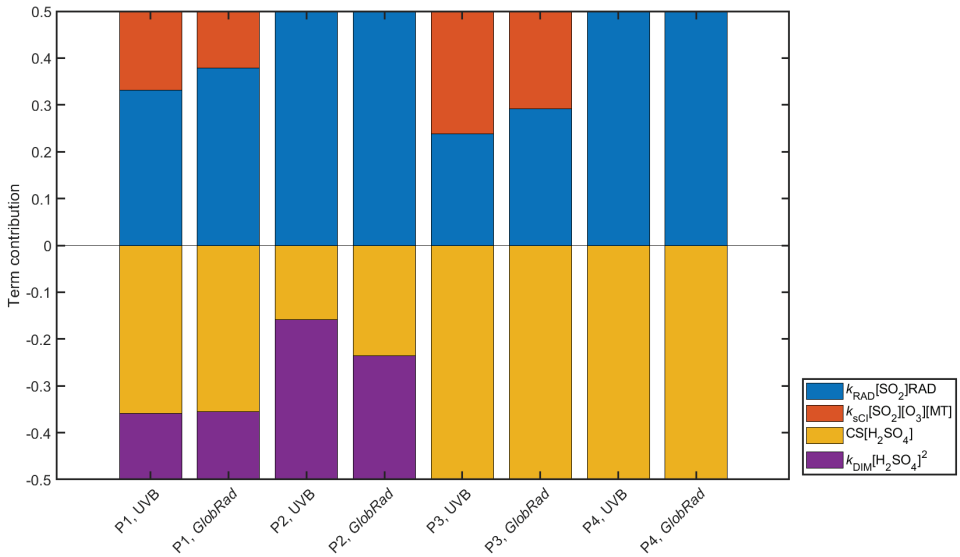

Fig. 6. Term contribution of different sulfuric acid proxies. Positive values refer to the contribution of source terms of sulfuric acid and negative values to the sink terms.

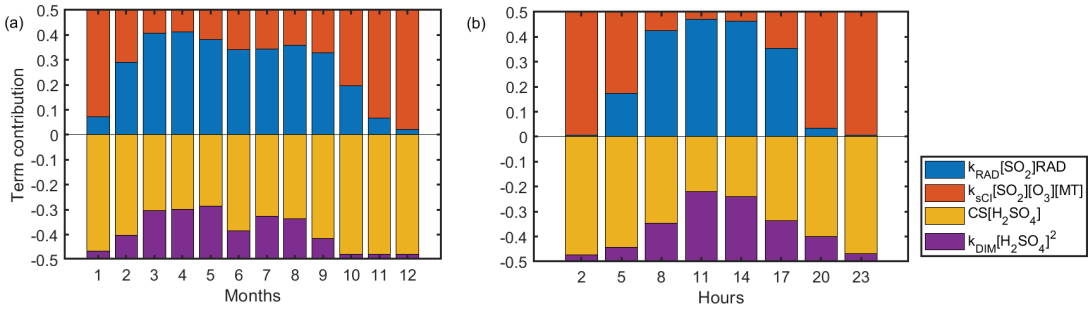

Fig. 7. (**a**) annual and (**b**) diurnal term contribution of proxy 1 with UVB radiation. Positive values refer to the contribution of source terms of sulfuric acid and negative values to the sink terms.

different source and sink terms but in a broader way formation and losses of sulfuric acid that linearly and non-linearly dependent on various parameters. As an example, OH can also be formed in recycling of OH from HO2 (Lelieveld et al. 2016), which is not currently accounted in the proxies.

Monthly term contributions (Fig. 7a) show that the radiation term peaks in early spring and has a slight secondary maximum in early autumn. Hence, the Criegee term has the highest values in winter, although MT concentration is low in winter at SMEAR II. The dimer term peaks in summertime, with June showing a smaller contribution.

Diurnal variation in term contribution shows, as expected, that the radiation and dimer terms peak in daytime due to the available radiation and condensable gases, and the Criegee and CS terms have their maximal contribution during night hours (Fig. 7b). Monthly and diurnal term contributions of other proxies follow the term contributions presented in Fig. 7.

Conclusions and summary

Our analysis of four sulfuric acid proxies revealed that the best-performing proxies at the SMEAR II site account for two source terms of sulfuric acid: formation due to the oxidation of sulfur dioxide by OH radicals and stabilized Criegee intermediates, as well as two sink terms: losses due to the formation of molecular dimers and a condensa-

tion sink (proxy 1). However, proxies accounting only for the loss due to condensation sink (proxy 3) performed relatively well, as already noted also by Dada et al. (2020) and Yang et al. (2021).

At nighttime, proxies without the Criegee term underestimate sulfuric acid concentration. However, during daylight hours (approximately 05:00–17:00 in summer and 08:00–17:00 in winter), the formation of sulfuric acid is dominated by OH oxidizing sulfur dioxide. Therefore, if only daytime data of sulfuric acid is required, a simpler proxy without the Criegee term can be utilized. Studies of volcano plumes (Rose et al. 2021) and sulfuric acid formation in Beijing, China, (Lu et al. 2019) showed also that the simple proxies can perform relatively well, but that the best performance was reached when using local fitting coefficients. At SMEAR II, a proxy that includes the dimer term but excludes the Criegee term (proxy 2) performed better than the simple proxy with only the condensation sink and radiation terms (proxy 4), but among the proxies excluding the Criegee term, the best diurnal cycle is achieved by the proxy proposed by Petäjä et al. (2009), especially with global radiation.

Seasonal analysis revealed that the differences in proxy performance between seasons were relatively minor and may be explained by the data distribution and fewer data points in autumn and winter. Additionally, fitting seasonally divided data improved the results only slightly, which may be due to the relatively wide ranges of the lowest loss function values, resulting in a situation where a slight change in the fitting coefficient value does not significantly impact the results. On the other hand, this can also explain why the widely used proxy by Petäjä et al. (2009) performs relatively well, even though it was derived using only a few months of spring and early summertime data from SMEAR II. The numerical values of the fitted coefficients are site-specific, but the overall comparison of proxy structure and the importance of including the Criegee term are expected to be of broader relevance.

We observed that proxies without the Criegee term (proxies 2 and 4) tended to perform slightly better when derived using global radiation than UVB radiation. Despite these differences, both UVB and global radiation-derived proxies performed similarly. Hence, for example, the larger uncertainty related to the Criegee term fitting coefficient did not seem to cause detectable uncertainty for the proxy, as the UVB-derived proxy performed well despite its larger Criegee term contribution. Therefore, both radiation types can be used to derive sulfuric acid proxies.

Relatively considerable uncertainty is associated with different sulfuric acid proxies, which stems from both measurement uncertainties and difficulties in finding representative fitting coefficients by minimizing the difference between proxy values and measured sulfuric acid. The proxies, which account for sulfuric acid formation due to stabilized Criegee radicals oxidizing sulfur dioxide, are better in estimating the nighttime sulfuric acid concentrations correctly. Nevertheless, the proxies seem to smooth the sulfuric acid concentration by overestimating the low, and by underestimating the high, measured sulfuric acid concentrations.

Acknowledgements: The work was supported by the ACCC Flagship funded by the Research Council of Finland (337549) and other Research Council of Finland grants (364226). We thank ACTRIS for aerosol and condensable vapors measurements at SMEAR II. The PIs as well as scientific and technical staff of the SMEAR II station are acknowledged.

Supplementary Information: The supplementary information related to this article is available online at: https://doi.org/10.60910/ber2026.xs5t-v078

References

Aalto, P., Hämeri, K., Becker, E., Weber, R., Salm, J., Mäkelä, J. M., Hoell, C., O'Dowd, C. D., Hansson, H.-C., Väkevä, M., Koponen, I., Buzorius, G. & Kulmala, M. 2001. Physical characterization of aerosol particles during nucleation events. *Tellus B: Chemical and Physical Meteorology*, 53(4):344–358.

Albrecht, B. A. 1989. Aerosols, cloud microphysics, and fractional cloudiness. *Science*, 245(4923):1227–1230.

Arneth, A., Unger, N., Kulmala, M. & Andreae, M. O. 2009. Clean the air, heat the planet? *Science*, 326(5953):672–673.

Artaxo, P., Hansson, H.-C., Andreae, M., Bäck, J., Alves, E., Barbosa, H., Bender, F., Bourtsoukidis, E., Carbone, S., Chi, J., Decesari, S., Despres, V., Ditas, F., Ezhova, E., Fuzzi, S., Hasselquist, N., Heintzenberg, J., Holanda, B., Guenther, A., Hakola, H., Heikkinen, L., Kerminen,

V.-M., Kontkanen, J., Krejci, R., Kulmala, M., Lavric, J., de Leeuw, G., Lehtipalo, K., Machado, L., McFiggans, G., Franco, M., Meller, B., Morais, F., Mohr, C., Morgan, W., Nilsson, M., Peichl, M., Petäjä, T., Praß, M., Pöhlker, C., Pöhlker, M., Pöschl, U., von Randow, C., Riipinen, I., Rinner, J., Rizzo, L., Rosenfeld, D., Silva Dias, M., Sogecheva, L., Stier, P., Swietlicki, E., Sörgel, M., Tunved, P., Virkkula, A., Wang, J., Weber, B., Yáñez-Serrano, A. M., Zieger, P., Mikhailov, E., Backes Meller, B., Gonçalves Morais, F., Smith, J. N. & Kesselmeier, J. 2022. Tropical and boreal forest – atmosphere interactions: a review. *Tellus B: Chemical and Physical Meteorology*, 74:24–163.

Berresheim, H., Elste, T., Tremmel, H. G., Allen, A. G., Hansson, H.-C., Rosman, K., Dal Maso, M., Mäkelä, J. M., Kulmala, M. & O'Dowd, C. D. 2002. Gas-aerosol relationships of H2SO4, MSA, and OH: observations in the coastal marine boundary layer at Mace Head, Ireland. Journal of Geophysical Research: Atmospheres, 107 (D19), 8100, doi:10.1029/2000JD000229

Boy, M., Mogensen, D., Smolander, S., Zhou, L., Nieminen, T., Paasonen, P., Plass-Dülmer, C., Sipilä, M., Petäjä, T., Mauldin, L., Berresheim, H. & Kulmala, M. 2013. Oxidation of SO2 by stabilized Criegee intermediate (sCI) radicals as a crucial source for atmospheric sulfuric acid concentrations. *Atmospheric Chemistry and Physics*, 13(7):3865–3879.

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R. V., Dentener, F., Dingenen, R. v., Estep, K., Amini, H., Apte, J. S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P. K., Knibbs, L. D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J. L. T., Shaddick, G., Anderson, H. R., Vos, T., Forouzanfar, M. H., Burnett, R. T. & Cohen, A. 2016. Ambient air pollution exposure estimation for the Global Burden of Disease 2013. *Environmental Science & Technology*, 50(1):79–88.

Dada, L., Paasonen, P., Nieminen, T., Buenrostro Mazon, S., Kontkanen, J., Peräkylä, O., Lehtipalo, K., Hussein, T., Petäjä, T., Kerminen, V.-M., Bäck, J. & Kulmala, M. 2017. Long-term analysis of clear-sky new particle formation events and nonevents in Hyytiälä. *Atmospheric Chemistry and Physics*, 17:6227–6241, https://doi.org/10.5194/acp-17-6227-2017.

Dada, L., Ylivinkka, I., Baalbaki, R., Li, C., Guo, Y., Yan, C., Yao, L., Sarnela, N., Jokinen, T., Daellenbach, K. R., Yin, R., Deng, C., Chu, B., Nieminen, T., Wang, Y., Lin, Z., Thakur, R. C., Kontkanen, J., Stolzenburg, D., Sipilä, M., Hussein, T., Paasonen, P., Bianchi, F., Salma, I., Weidinger, T., Pikridas, M., Sciare, J., Jiang, J., Liu, Y., Petäjä, T., Kerminen, V.-M. & Kulmala, M. 2020. Sources and sinks driving sulfuric acid concentrations in contrasting environments: implications on proxy calculations. *Atmospheric Chemistry and Physics*, 20(20):11747–11766.

Dunne, E. M., Gordon, H., Kürten, A., Almeida, J., Duplissy, J., Williamson, C., Ortega, I. K., Pringle, K. J., Adamov, A., Baltensperger, U., Barmet, P., Benduhn, F., Bianchi, F., Breitenlechner, M., Clarke, A., Curtius, J., Dommen, J., Donahue, N. M., Ehrhart, S., Flagan, R. C., Franchin, A., Guida, R., Hakala, J., Hansel, A., Heinritzi, M., Jokinen, T., Kangasluoma, J., Kirkby, J., Kulmala, M., Kupc, A., Lawler, M. J., Lehtipalo, K., Makhmutov, V., Mann, G., Mathot, S., Merikanto, J., Miettinen, P., Nenes, A., Onnela, A., Rap, A., Reddington, C. L. S., Riccobono, F., Richards, N. A. D., Rissanen, M. P., Rondo, L., Sarnela, N., Schobesberger, S., Sengupta, K., Simon, M., Sipilä, M., Smith, J. N., Stozhkov, Y., Tomé, A., Tröstl, J., Wagner, P. E., Wimmer, D., Winkler, P. M., Worsnop, D. R. & Carslaw, K. S. 2016. Global atmospheric particle formation from CERN CLOUD measurements. *Science*, 354(6316):1119–1124, https://doi.org/10.1126/science.aaf2649.

Eisele, F. & Tanner, D. 1993. Measurement of the gas phase concentration of H2SO4 and methane sulfonic acid and estimates of H2SO4 production and loss in the atmosphere. *Journal of Geophysical Research: Atmospheres*, 98(D5):9001–9010.

Ezhova, E., Ylivinkka, I., Kuusk, J., Komsaare, K., Vana, M., Krasnova, A., Noe, S., Arshinov, M., Belan, B., Park, S.-B., Lavrič, J. V., Heimann, M., Petäjä, T., Vesala, T., Mammarella, I., Kolari, P., Bäck, J., Rannik, Ü., Kerminen, V.-M. & Kulmala, M. 2018. Direct effect of aerosols on solar radiation and gross primary production in boreal and hemiboreal forests. *Atmospheric Chemistry and Physics*, 18(24):17863–17881.

Ezhova, E., Laanti, T., Lintunen, A., Kolari, P., Nieminen, T., Mammarella, I., Heljanko, K. & Kulmala, M. 2025. Explainable machine learning for modeling of net ecosystem exchange in boreal forests. *Biogeosciences*, 22:257–288, https://doi.org/10.5194/bg-22-257-2025.

Feng, Z., Zheng, F., Liu, Y., Fan, X., Yan, C., Zhang, Y., Daellenbach, K. R., Bianchi, F., Petäjä, T., Kulmala, M. & Bao, X. 2022. Evolution of organic carbon during COVID-19 lockdown period: possible contribution of nocturnal chemistry. *Science of the Total Environment*, 808:152191.

Gryspeerdt, E., Stier, P. & Partridge, D. G. 2014. Satellite observations of cloud regime development: the role of aerosol processes. *Atmospheric Chemistry and Physics*, 14(3):1141–1158.

Guo, S., Hu, M., Zamora, M. L., Peng, J., Shang, D., Zheng, J., Du, Z., Wu, Z., Shao, M., Zeng, L., Molina, M. J. & Zhang, R. 2014. Elucidating severe urban haze formation in China. *Proceedings of the National Academy of Sciences*, 111(49):17373–17378.

Guo, Y., Yan, C., Li, C., Deng, C., Zhang, Y., Zhou, Y., Zheng, H., Jiang, Y., Chen, X., Ma, W., Sarnela, N., Lin, Z., Hua, C., Fan, X., Zheng, F., Feng, Z., Wang, Z., Zhang, Y., Jiang, J., Zhao, B., Kulmala, M. & Liu, Y. 2026. Measurement report: three-year characteristics of sulfuric acid in urban Beijing and derivation of daytime sulfuric acid proxies applicable to inland sites. *Atmospheric Chemistry and Physics*, 26:3467–3487, https://doi.org/10.5194/acp-26-3467-2026.

Hakola, H., Hellén, H., Hemmilä, M., Rinne, J. & Kulmala, M. 2012. In situ measurements of volatile organic compounds in a boreal forest. *Atmospheric Chemistry and Physics*, 12(23):11665–11678.

Hamed, A., Birmili, W., Joutsensaari, J., Mikkonen, S., Asmi,

A., Wehner, B., Spindler, G., Jaatinen, A., Wiedensohler, A., Korhonen, H., Lehtinen, K. E. J. & Laaksonen, A. 2010. Changes in the production rate of secondary aerosol particles in Central Europe in view of decreasing SO2 emissions between 1996 and 2006. *Atmospheric Chemistry and Physics*, 10(3):1071–1091.

Hari, P. & Kulmala, M. 2005. Station for measuring ecosystem-atmosphere relations (SMEAR II). *Boreal Environment Research*, 10(5):315–322.

Hellén, H., Praplan, A. P., Tykkä, T., Ylivinkka, I., Vakkari, V., Bäck, J., Petäjä, T., Kulmala, M. & Hakola, H. 2018. Long-term measurements of volatile organic compounds highlight the importance of sesquiterpenes for the atmospheric chemistry of a boreal forest. *Atmospheric Chemistry and Physics*, 18(19):13839–13863.

Huang, R.-J., Zhang, Y., Bozzetti, C., Ho, K.-F., Cao, J.-J., Han, Y., Daellenbach, K. R., Slowik, J. G., Platt, S. M., Canonaco, F., Zotter, P., Wolf, R., Pieber, S. M., Bruns, E. A., Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbaszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., El Haddad, I. & Prévôt, A. S. H. 2014. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature*, 514(7521):218–222.

IPCC 2021. *Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jokinen, T., Sipilä, M., Junninen, H., Ehn, M., Lönn, G., Hakala, J., Petäjä, T., Mauldin III, R. L., Kulmala, M. & Worsnop, D. R. 2012. Atmospheric sulphuric acid and neutral cluster measurements using CI-APi-TOF. *Atmospheric Chemistry and Physics*, 12(9):4117–4125.

Junninen, H., Ehn, M., Petäjä, T., Luosujärvi, L., Kotiaho, T., Kostiainen, R., Rohner, U., Gonin, M., Fuhrer, K., Kulmala, M. & Worsnop, D. R. 2010. A high-resolution mass spectrometer to measure atmospheric ion composition. *Atmospheric Measurement Techniques*, 3(4):1039–1053.

Kalivitis, N., Kerminen, V.-M., Kouvarakis, G., Stavroulas, I., Tzitzikalaki, E., Kalkavouras, P., Daskalakis, N., Myriokefalitakis, S., Bougiatioti, A., Manninen, H. E., Roldin, P., Petäjä, T., Boy, M., Kulmala, M., Kanakidou, M. & Mihalopoulos, N. 2019. Formation and growth of atmospheric nanoparticles in the eastern Mediterranean: results from long-term measurements and process simulations. *Atmospheric Chemistry and Physics*, 19:2671–2686, https://doi.org/10.5194/acp-19-2671-2019.

Kannosto, J., Virtanen, A., Lemmetty, M., Mäkelä, J. M., Keskinen, J., Junninen, H., Hussein, T., Aalto, P. & Kulmala, M. 2008. Mode resolved density of atmospheric aerosol particles. *Atmospheric Chemistry and Physics*, 8(17):5327–5337.

Kerminen, V.-M., Chen, X., Vakkari, V., Petäjä, T., Kulmala, M. & Bianchi, F. 2018. Atmospheric new particle formation and growth: review of field observations. *Environmental Research Letters*, 13(10):103003.

Kerminen, V.-M., Paramonov, M., Anttila, T., Riipinen, I., Fountoukis, C., Korhonen, H., Asmi, E., Laakso, L., Lihavainen, H., Swietlicki, E., Svenningsson, B., Asmi, A., Pandis, S. N., Kulmala, M. & Petäjä, T. 2012. Cloud condensation nuclei production associated with atmospheric nucleation: a synthesis based on existing literature and new results. *Atmospheric Chemistry and Physics*, 12(24):12037–12059.

Kulmala, M. 2015. Atmospheric chemistry: China's choking cocktail. *Nature*, 526(7574):497–499.

Kulmala, M., Lehtinen, K. E. J. & Laaksonen, A. 2006. Cluster activation theory as an explanation of the linear dependence between formation rate of 3 nm particles and sulphuric acid concentration. *Atmospheric Chemistry and Physics*, 6(3):787–793.

Kulmala, M., Luoma, K., Virkkula, A., Petäjä, T., Paasonen, P., Kerminen, V.-M., Nie, W., Qi, X., Shen, Y., Chi, X. & Ding, A. 2016. On the mode-segregated aerosol particle number concentration load: contributions of primary and secondary particles in Hyytiälä and Nanjing. *Boreal Environment Research*, 21:319–331.

Kulmala, M., Nieminen, T., Nikandrova, A., Lehtipalo, K., Manninen, H. E., Kajos, M. K., Kolari, P., Lauri, A., Petäjä, T., Krejci, R., Hansson, H.-C., Swietlicki, E., Lindroth, A., Christensen, T. R., Arneth, A., Hari, P., Bäck, J., Vesala, T. & Kerminen, V.-M. 2014a. CO2-induced terrestrial climate feedback mechanism: from carbon sink to aerosol source and back. *Boreal Environment Research*, 19(suppl. B):122–131.

Kulmala, M., Petäjä, T., Ehn, M., Thornton, J., Sipilä, M., Worsnop, D. & Kerminen, V.-M. 2014b. Chemistry of atmospheric nucleation: on the recent advances on precursor characterization and atmospheric cluster composition in connection with atmospheric new particle formation. *Annual Review of Physical Chemistry*, 65:21–37.

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

Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V.-M., Birmili, W. & McMurry, P. 2004. Formation and growth rates of ultrafine atmospheric particles: a review of observations. *Journal of Aerosol Science*, 35(2):143–176.

Kürten, A., Rondo, L., Ehrhart, S. & Curtius, J. 2012. Calibration of a chemical ionization mass spectrometer for the measurement of gaseous sulfuric acid. *The Journal of Physical Chemistry A*, 116(24):6375–6386, https://doi.org/10.1021/jp212123n.

Kuula, J., Timonen, H., Niemi, J. V., Manninen, H. E., Rönkkö, T., Hussein, T., Fung, P. L., Tarkoma, S., Laakso, M., Saukko, E., Ovaska, A., Kulmala, M., Karppinen, A., Johansson, L. & Petäjä, T. 2022. Opinion: insights into updating Ambient Air Quality Directive 2008/50/EC. *Atmospheric Chemistry and Physics*, 22:4801–4808, https://doi.org/10.5194/acp-22-4801-2022.

Kyrö, E.-M., Väänänen, R., Kerminen, V.-M., Virkkula, A., Petäjä, T., Asmi, A., Dal Maso, M., Nieminen, T.,

Juhola, S., Shcherbinin, A., Riipinen, I., Lehtipalo, K., Keronen, P., Aalto, P. P., Hari, P. & Kulmala, M. 2014. Trends in new particle formation in eastern Lapland, Finland: effect of decreasing sulfur emissions from Kola Peninsula. *Atmospheric Chemistry and Physics*, 14(9):4383–4396.

Laakso, L., Petäjä, T., Lehtinen, K. E. J., 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.

Lagarias, J. C., Reeds, J. A., Wright, M. H. & Wright, P. E. 1998. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147.

Laumbach, R. J. & Kipen, H. M. 2012. Respiratory health effects of air pollution: update on biomass smoke and traffic pollution. *Journal of Allergy and Clinical Immunology*, 129(1):3–11.

Lehtipalo, K., Yan, C., Dada, L., Bianchi, F., Xiao, M., Wagner, R., Stolzenburg, D., Ahonen, L. R., Amorim, A., Baccarini, A., Bauer, P. S., Baumgartner, B., Bergen, A., Bernhammer, A.-K., Breitenlechner, M., Brilke, S., Buchholz, A., Buenrostro Mazon, S., Chen, D., Chen, X., Dias, A., Dommen, J., Draper, D. C., Duplissy, J., Ehn, M., Finkenzeller, H., Fischer, L., Frege, C., Fuchs, C., Garmash, O., Gordon, H., Hakala, J., He, X., Heikkinen, L., Heinritzi, M., Helm, J. C., Hofbauer, V., Hoyle, C. R., Jokinen, T., Kangasluoma, J., Kerminen, V.-M., Kim, C., Kirkby, J., Kontkanen, J., Kürten, A., Lawler, M. J., Mai, H., Mathot, S., Mauldin III, R. L., Molteni, U., Nichman, L., Nie, W., Nieminen, T., Ojdanic, A., Onnela, A., Passananti, M., Petäjä, T., Piel, F., Pospisilova, V., Quéléver, L. L. J., Rissanen, M. P., Rose, C., Sarnela, N., Schallhart, S., Schuchmann, S., Sengupta, K., Simon, M., Sipilä, M., Tauber, C., Tomé, A., Tröstl, J., Väisänen, O., Vogel, A. L., Volkamer, R., Wagner, A. C., Wang, M., Weitz, L., Wimmer, D., Ye, P., Ylisirniö, A., Zha, Q., Carslaw, K. S., Curtius, J., Donahue, N. M., Flagan, R. C., Hansel, A., Riipinen, I., Virtanen, A., Winkler, P. M., Baltensperger, U., Kulmala, M. & Worsnop, D. R. 2018. Multicomponent new particle formation from sulfuric acid, ammonia, and biogenic vapors. *Science Advances*, 4(12):eaau5363.

Lelieveld, J., Gromov, S., Pozzer, A. & Taraborrelli, D. 2016. Global tropospheric hydroxyl distribution, budget and reactivity. *Atmospheric Chemistry and Physics*, 16:12477–12493, https://doi.org/10.5194/acp-16-12477-2016.

Li, X., Li, H., Yao, L., Stolzenburg, D., Sarnela, N., Vettikkat, L., de Jonge, R. W., Baalbaki, R., Uusitalo, H., Kontkanen, J., Lehtipalo, K., Daellenbach, K. R., Jokinen, T., Aalto, J., Keronen, P., Schobesberger, S., Nieminen, T., Petäjä, T., Kerminen, V.-M., Bianchi, F., Kulmala, M. & Dada, L. 2024. Over 20 years of observations in the boreal forest reveal a decreasing trend of atmospheric new particle formation. *Boreal Environment Research*, 29:25–52.

Lohmann, U. & Feichter, J. 2005. Global indirect aerosol effects: a review. *Atmospheric Chemistry and Physics*, 5(3):715–737.

Lu, Y., Yan, C., Fu, Y., Chen, Y., Liu, Y., Yang, G., Wang, Y., Bianchi, F., Chu, B., Zhou, Y., Yin, R., Baalbaki, R., Garmash, O., Deng, C., Wang, W., Liu, Y., Petäjä, T., Kerminen, V.-M., Jiang, J., Kulmala, M. & Wang, L. 2019. A proxy for atmospheric daytime gaseous sulfuric acid concentration in urban Beijing. *Atmospheric Chemistry and Physics*, 19:1971–1983, https://doi.org/10.5194/acp-19-1971-2019.

Mauldin III, R., Berndt, T., Sipilä, M., Paasonen, P., Petäjä, T., Kim, S., Kurtén, T., Stratmann, F., Kerminen, V.-M. & Kulmala, M. 2012. A new atmospherically relevant oxidant of sulphur dioxide. *Nature*, 488(7410):193–196.

Mikkonen, S., Romakkaniemi, S., Smith, J. N., Korhonen, H., Petäjä, T., Plass-Duelmer, C., Boy, M., McMurry, P. H., Lehtinen, K. E. J., Joutsensaari, J., Hamed, A., Mauldin III, R. L., Birmili, W., Spindler, G., Arnold, F., Kulmala, M. & Laaksonen, A. 2011. A statistical proxy for sulphuric acid concentration. *Atmospheric Chemistry and Physics*, 11(21):11319–11334.

Myers, D. C., Kim, S., Sjostedt, S., Guenther, A. B., Seco, R., Vega Bustillos, O., Tota, J., Souza, R. A. F. & Smith, J. N. 2022. Sulfuric acid in the Amazon basin: measurements and evaluation of existing sulfuric acid proxies. *Atmospheric Chemistry and Physics*, 22:10061–10076, https://doi.org/10.5194/acp-22-10061-2022.

Nieminen, T., Asmi, A., Dal Maso, M., Aalto, P., Keronen, P., Petäjä, T., Kulmala, M. & Kerminen, V.-M. 2014. Trends in atmospheric new-particle formation: 16 years of observations in a boreal-forest environment. *Boreal Environment Research*, 19(suppl. B):191–214.

Nocedal, J. & Wright, S. J. 2006. Sequential quadratic programming. In: *Numerical optimization*, Springer New York, New York, NY, pp. 529–562.

Orear, J. 1982. Least squares when both variables have uncertainties. American Journal of Physics, 50:912–916, https://doi.org/10.1119/1.12972.

Petäjä, T., Mauldin III, R., Kosciuch, E., McGrath, J., Nieminen, T., Paasonen, P., Boy, M., Adamov, A., Kotiaho, T. & Kulmala, M. 2009. Sulfuric acid and OH concentrations in a boreal forest site. *Atmospheric Chemistry and Physics*, 9(19):7435–7448.

Ramanathan, V., Crutzen, P. J., Kiehl, J. T. & Rosenfeld, D. 2001. Aerosols, climate, and the hydrological cycle. *Science*, 294(5549):2119–2124.

Rinne, J., Ruuskanen, T. M., Reissell, A., Taipale, R., Hakola, H. & Kulmala, M. 2005. On-line PTR-MS measurements of atmospheric concentrations of volatile organic compounds in a European boreal forest ecosystem. *Boreal Environment Research*, 10(5):425–436.

Rohrer, F. & Berresheim, H. 2006. Strong correlation between levels of tropospheric hydroxyl radicals and solar ultraviolet radiation. *Nature*, 442(7099):184–187.

Rose, C., Rissanen, M. P., Iyer, S., Duplissy, J., Yan, C., Nowak, J. B., Colomb, A., Dupuy, R., He, X.-C., Lampilahti, J., Tham, Y. J., Wimmer, D., Metzger, J.-M., Tulet, P., Brioude, J., Planche, C., Kulmala, M. & Sellegri, K. 2021. Investigation of several proxies to estimate sulfuric acid concentration under volcanic plume conditions. *Atmospheric Chemistry and Physics*, 21:4541–4560,

https://doi.org/10.5194/acp-21-4541-2021.

Rosenfeld, D., Andreae, M. O., Asmi, A., Chin, M., de Leeuw, G., Donovan, D. P., Kahn, R., Kinne, S., Kivekäs, N., Kulmala, M., Lau, W., Schmidt, K. S., Suni, T., Wagner, T., Wild, M. & Quaas, J. 2014. Global observations of aerosol-cloud-precipitation-climate interactions. *Reviews of Geophysics*, 52(4):750–808.

Saha, P. K., Robinson, E. S., Shah, R. U., Zimmerman, N., Apte, J. S., Robinson, A. L. & Presto, A. A. 2018. Reduced ultrafine particle concentration in urban air: changes in nucleation and anthropogenic emissions. *Environmental Science & Technology*, 52(12):6798–6806.

Salma, I., Thén, W., Aalto, P., Kerminen, V.-M., Kern, A., Barcza, Z., Petäjä, T. & Kulmala, M. 2021. Influence of vegetation on occurrence and time distributions of regional new aerosol particle formation and growth. *Atmospheric Chemistry and Physics*, 21:2861–2880, https://doi.org/10.5194/acp-21-2861-2021.

Shen, J., Bigi, A., Marinoni, A., Lampilahti, J., Kontkanen, J., Ciarelli, G., Putaud, J. P., Nieminen, T., Kulmala, M., Lehtipalo, K. & Bianchi, F. 2021. Emerging investigator series: COVID-19 lockdown effects on aerosol particle size distributions in northern Italy. *Environmental Science: Atmospheres*, 1(5):214–227.

Shen, X., Sun, J., Ma, Q., Zhang, Y., Zhong, J., Yue, Y., Xia, C., Hu, X., Zhang, S. & Zhang, X. 2022. Long-term trend of new particle formation events in the Yangtze River Delta, China and its influencing factors: 7-year dataset analysis. *Science of the Total Environment*, 807:150783, https://doi.org/10.1016/j.scitotenv.2021.150783.

Sihto, S.-L., Kulmala, M., Kerminen, V.-M., Dal Maso, M., Petäjä, T., Riipinen, I., Korhonen, H., Arnold, F., Janson, R., Boy, M., Laaksonen, A. & Lehtinen, K. E. J. 2006. Atmospheric sulphuric acid and aerosol formation: implications from atmospheric measurements for nucleation and early growth mechanisms. *Atmospheric Chemistry and Physics*, 6(12):4079–4091.

Sipilä, M., Jokinen, T., Berndt, T., Richters, S., Makkonen, R., Donahue, N. M., Mauldin III, R. L., Kurtén, T., Paasonen, P., Sarnela, N., Ehn, M., Junninen, H., Rissanen, M. P., Thornton, J., Stratmann, F., Herrmann, H., Worsnop, D. R., Kulmala, M., Kerminen, V.-M. & Petäjä, T. 2014. Reactivity of stabilized Criegee intermediates (sCIs) from isoprene and monoterpene ozonolysis toward SO2 and organic acids. *Atmospheric Chemistry and Physics*, 14(22):12143–12153.

Taipale, R., Ruuskanen, T. M., Rinne, J., Kajos, M. K., Hakola, H., Pohja, T. & Kulmala, M. 2008. Technical note: quantitative long-term measurements of VOC concentrations by PTR-MS – measurement, calibration, and volume mixing ratio calculation methods. *Atmospheric Chemistry and Physics*, 8(22):6681–6698.

Tuovinen, S., Kontkanen, J., Jiang, J. & Kulmala, M. 2020. Investigating the effectiveness of condensation sink based on heterogeneous nucleation theory. *Journal of Aerosol Science*, 149:105613.

Twomey, S. 1977. The influence of pollution on the shortwave albedo of clouds. *Journal of the Atmospheric Sciences*, 34(7):1149–1152.

Wang, Y., Yu, M., Wang, Y., Tang, G., Song, T., Zhou, P., Liu, Z., Hu, B., Ji, D., Wang, L., Zhu, X., Yan, C., Ehn, M., Gao, W., Pan, Y., Xin, J., Sun, Y., Kerminen, V.-M., Kulmala, M. & Petäjä, T. 2020. Rapid formation of intense haze episodes via aerosol–boundary layer feedback in Beijing. *Atmospheric Chemistry and Physics*, 20(1):45–53.

Weber, R., Marti, J., McMurry, P. H., Eisele, F., Tanner, D. & Jefferson, A. 1996. Measured atmospheric new particle formation rates: implications for nucleation mechanisms. *Chemical Engineering Communications*, 151(1):53–64.

Yan, C., Shen, Y., Stolzenburg, D., Dada, L., Qi, X., Hakala, S., Sundström, A.-M., Guo, Y., Lipponen, A., Kokkonen, T. V., Kontkanen, J., Cai, R., Cai, J., Chan, T., Chen, L., Chu, B., Deng, C., Du, W., Fan, X., He, X.-C., Kangasluoma, J., Kujansuu, J., Kurppa, M., Li, C., Li, Y., Lin, Z., Liu, Y., Liu, Y., Lu, Y., Nie, W., Pulliainen, J., Qiao, X., Wang, Y., Wen, Y., Wu, Y., Yang, G., Yao, L., Yin, R., Zhang, G., Zhang, S., Zheng, F., Zhou, Y., Arola, A., Tamminen, J., Paasonen, P., Sun, Y., Wang, L., Donahue, N. M., Liu, Y., Bianchi, F., Daellenbach, K. R., Worsnop, D. R., Kerminen, V.-M., Petäjä, T., Ding, A., Jiang, J. & Kulmala, M. 2022. The effect of COVID-19 restrictions on atmospheric new particle formation in Beijing. *Atmospheric Chemistry and Physics*, 22(18):12207–12220.

Yang, L., Nie, W., Liu, Y., Xu, Z., Xiao, M., Qi, X., Li, Y., Wang, R., Zou, J., Paasonen, P., Yan, C., Xu, Z., Wang, J., Zhou, C., Yuan, J., Sun, J., Chi, X., Kerminen, V.-M., Kulmala, M. & Ding, A. 2021. Toward building a physical proxy for gas-phase sulfuric acid concentration based on its budget analysis in polluted Yangtze River Delta, East China. *Environmental Science & Technology*, 55(10):6665–6676, https://doi.org/10.1021/acs.est.1c00738.

Yli-Juuti, T., Mielonen, T., Heikkinen, L., Arola, A., Ehn, M., Isokääntä, S., Keskinen, H.-M., Kulmala, M., Laakso, A., Lipponen, A., Luoma, K., Mikkonen, S., Nieminen, T., Paasonen, P., Petäjä, T., Romakkaniemi, S., Tonttila, J., Kokkola, H. & Virtanen, A. 2021. Significance of the organic aerosol driven climate feedback in the boreal area. *Nature Communications*, 12(1):1–9.

Ylivinkka, I., Keskinen, H.-M., Ahonen, L. R., Heikkinen, L., Aalto, P. P., Nieminen, T., Lehtipalo, K., Luoma, K., Banerji, S., Aalto, J., Levula, J., Kesti, J., Ezhova, E., Kulmala, M. & Petäjä, T. 2025. Long-term PM trends at boreal forest site in southern Finland from three different measurement techniques. *Aerosol Research*, 3:503–520, https://doi.org/10.5194/ar-3-503-2025.

Zamora, M. L., Peng, J., Hu, M., Guo, S., Marrero-Ortiz, W., Shang, D., Zheng, J., Du, Z., Wu, Z. & Zhang, R. 2019. Wintertime aerosol properties in Beijing. *Atmospheric Chemistry and Physics*, 19(22):14329–14338.

Zhu, Y., Chen, Q., He, L., Shang, C., Jiang, L., Guan, D., Yao, G. & Guo, W. 2026. Revisiting the critical role of stabilized Criegee intermediates (sCIs) in sulfuric acid formation: coupling mechanistic updates with interpretable machine learning. *EGUsphere* [preprint], https://doi.org/10.5194/egusphere-2025-6276.