Final Report, December 2019

PRO3FILE
Predicting Ozone Fluxes, Impacts and Critical Levels on European Forests

Schaub M[1,2,*], Schönbeck L[1,2], Häni M[1,2], Trotsiuk V[1,2], Cailleret M[1,2]

[1]Swiss Federal Research Institute WSL; [2]SwissForestLab
**Corresponding author: marcus.schaub@wsl.ch*

1. Introduction

The study aims to make use of data from long-term monitoring plots across Europe where ozone concentrations have been measured since 2000, in parallel to forest and vegetation variables. Ozone-related effects and Critical Levels on selected endpoints such as tree growth are derived by quantifying ozone fluxes and applying multiple statistical techniques that consider for confounding abiotic and biotic environmental factors.

Measured (ICP Forests[1], LWF[2], Sanasilva[3], Swiss NFI[4], NCEI[5]) and modeled (Meteotest[6], EMEP[7], ECMWF[8]) data from different national (LWF, Sanasilva, Swiss NFI) and international (ICP Forests, EMEP) networks, are combined to maximize sample size, diversity, temporal and spatial coverage along large gradients of forest types and environmental conditions.

While the data requirements are attached in Appendix 1 and outlined in more detail in the interim report from June 2018, this final report provides information on (i) the materials and methods of the extensive data processing and (ii) the output from empirically assessed dose-response relationships between PODySPEC, AOT40 respectively and radial growth across European forests.

[1] ICP Forests: http://icp-forests.net/
[2] LWF: https://www.wsl.ch/lwf
[3] Sanasilva: https://www.wsl.ch/sanasilva
[4] Swiss NFI: https://www.lfi.ch/index-en.php
[5] NCEI: https://gis.ncdc.noaa.gov/geoportal/catalog/search/resource/details.page?id=gov.noaa.ncdc:C00532
[6] Meteotest: https://meteotest.ch/
[7] EMEP: https://www.emep.int/
[8] ECMWF: http://www.ecmwf.int/en/research/climate-reanalysis/era-20c

2. Materials and methods of data processing

2.1 Data requirements for DO3SE

For the DO_3SE model (Deposition of Ozone for Stomatal Exchange; CLRTAP, 2017), the following parameters are required: air temperature (*Ts_C*), vapor pressure deficit (*VPD*), wind speed (*uh_zR*), precipitation (*precip*), air pressure (*P*) and photosynthetically activate radiation (*PAR*) in gapless hourly time resolution for the growing season from April to September. These data were obtained mainly from the ICP Forests database. *VPD* is not provided by the ICP Forests database hence it was calculated based on *Ts_C* and relative humidity (*RH*) according to Murray (1967) and Monteith & Unsworth (1990).

2.2.1 Hourly meteorological data download from the ICP Forests data base

Following a formal data request[9], hourly meteorological data were downloaded from the ICP Forests database[10] on 19.01.2017, 13.11.2017, 16.02.2018 and 25.09.2018. For QA/QC, the different versions were compared and checked for file size, number of data points, number of years and plots included to detect and resolve systematic errors from the data base (e.g. missing data, wrong plot codes, wrong coordinates, etc.). A cleaned version was then used for further processing and read into R statistical software, version 3.5.2 (R Core Team 2018).

2.2.2 Hourly meteorological data from countries

Since some data from countries/plots/years/parameters seemed to be missing in the ICP Forests database and since the available meteorological data at hourly resolution seemed to pose the bottleneck for the application of DO_3SE in ICP Forests Level II plots, we approached the countries to request the missing original, hourly data for meteorological parameters. On behalf of the ICP Forests Expert Panel on Meteorology, and in close collaboration with the ICP Forests database manager, an official call for meteorological data at hourly resolution has been launched in June 2018, after the ICP Forests Task Force in Riga was informed about the importance of hourly data for ozone flux modeling (Appendix 2).

The call for hourly data resulted in a volume of 3.54 GB resubmitted data. The following countries (country code) re-submitted hourly data: France (1), Belgium (2), Germany (4), UK (6), Austria (14), Norway (55), Bulgaria (63), and Cyprus (66).

2.2.3 Hourly meteorological data validation by the ICP Forests data base manager

The following steps have been performed by the ICP Forests DB manager:

- Range tests (high, low values)
- Completeness tests (measured day per variable on a plot)
- Anomalies and plausibility tests
- Redundancy tests between daily and hourly records

For the detailed script, see Appendix 3.

[9] ICP Forests data request: https://www.icp-forests.org/pdf/ICPF_data_application_form.pdf
[10] ICP Forests data base: http://icp-forests.net/page/level-ii

2.2.4 Hourly meteorological data derived from daily data

For sites where only daily data was available, the data was processed to retrieve hourly data in the following way. Hourly *Ts_C* was modelled using the chillR package (Luedeling, 2019). This package is using plot latitude, day of year (DOY) and minimum (T_{min}) and maximum (T_{max}) temperatures to derive hourly temperature data. T_{min} and T_{max} were estimated from the mean daily temperature. For this, the following model (1) using month of year and mean daily temperature was developed to estimate the daily temperature amplitude between minimum and maximum temperature:

$$\sqrt{T_{ampl}} = a + b * T_{mean} + c * MOY \quad (1)$$

Where T_{ampl} is the amplitude in temperature, i.e. the difference between T_{max} and T_{min}, T_{mean} is the mean daily temperature and MOY is the month of the year (Figure 1).

Hourly values of *PAR* were derived using the sunrise and sunset hours, generated by the chillR package based on plot latitude. *PAR* before sunrise and after sunset were set to 0 and at noon PAR was set to 1. A spline function (Figure 2) was used to interpolate standardized hourly values between 0 and 1. As the area under the graph corresponds to the mean daily *PAR*, the standardized values could be back-transformed to hourly *PAR* values using the mean standardized *PAR* and the measured mean daily *PAR*. If global radiation (*Rg*) was measured instead of *PAR*, it was calculated by applying a factor derived from the linear relationship between all available *PAR* and *Rg* data (forced intercept of 0 between all available data, $R^2 = 0.90$, $p < 0.001$).

Hourly values for *VPD*, air pressure (*P*) and wind speed were not calculated but kept identical to the daily values. Hourly precipitation values were calculated from total precipitation, divided by 24 (hours).

2.3 Hourly meteorological data validation and gap-filling

DO_3SE requires gapless hourly meteorological data from air temperature (*Ts_C*), vapor pressure deficit (*VPD*), wind speed (*uh_zR*), precipitation (*precip*), air pressure (*P*) and photosynthetically activate radiation (*PAR*).

The hourly meteorological data downloaded from the ICP Forests webpage on 25 September 2018, contained 56.8 mio data points. The additional data from the countries and modeled data from daily meteorological data were imported and merged with the already existing data from the previous downloads. In order to increase its quality and completeness, a range of sophisticated routines was developed.

First, data was filtered according to visual check of the data; erroneous data was deleted (only very few sites and parameters affected). After the visual checks, small negative *VPD* and *PAR* values (-20 to 0) where corrected to 0. If P (atmospheric pressure) was not available (only very few ICP Forests sites record this parameter), the data was retrieved from the nearest 0.125° x 0.125° grid point from ERA Interim (Dee *et al*. 2011, downloaded in June 2017). To receive hourly time series for *P*, linear gap-filling of hourly data was applied since ERA Interim data are based on a 6-hour time resolution.

Following, a *four-step step-wise gap-filling procedure* was applied to generate gapless time-series (as required by DO_3SE). The basis of processing was the merged data set in hourly time resolution of all variables (*Ts_C*, *VPD*, *uh_zR*, P, *precip* and *PAR*):

(a) Small gaps (< 14 days) were filled linearly for all time-series (as stated above: *Ts_C*, *VPD*, *uh_zR*, *P* and *PAR*) except for precipitation, where small gaps were filled with zeroes (no precipitation assumed). Gaps larger than 14 days remained gaps after this step. See Fig. 3a1 as an example for this first gap-filling step for all parameters except for precipitation, and Fig. 3a2 for precipitation.

(b) After this first step, for each site and parameter, a linear multivariate site-specific model was built and gaps in the corresponding parameter where filled if the model performed with an adjusted $R^2 > 0.6$ (site-specific modelling). See Fig. 3b as an example for this second gap-filling step.

(c) After the multivariate site-specific modelling, each parameter of every site was compared linearly to the same parameter of all other sites. The highest correlating time series was then used to fill remaining gaps if $R^2 > 0.7$ (cross-site gap filling). See Fig. 3c as an example for this third gap-filling step. This step was not done for *PAR* and *precip*, because these two parameters are too local.

(d) As a last step (after the cross-site gap-filling), the remaining data gaps were completed with filling in an average of a moving window, seven days before and seven days after for every single hour. See Fig. 3d as an example of this fourth and last gap-filling step. This step was not done for *PAR* and *precip*, because these two parameters are too local.

After the step-wise application of the four above outlined steps, data were visually checked and the time series were only kept if the processing succeeded, i.e. the data was gapless for the corresponding years. Alongside all of these steps, all data points were flagged according to their source, to ensure reproducibility and transparency.

2.4 Hourly meteorological and ozone data modelled by EMEP

Grid-based data on a 0.1° x 0.1° resolution were obtained from EMEP and could directly be used in the DO_3SE model. In addition, $FstO_3$ values were provided by EMEP, coming directly from the DO_3SE-model ran by EMEP, enabling the calculation of POD0 and POD1 and thus validation and comparison of the DO_3SE model parameterization in the different categories.

2.5 Hourly meteorological data modelled by ERA5

Grid-based (0.25° x 0.25°) ERA5 hourly data on single levels were downloaded from the climate data storage application program interface[11]. ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables on a 30 km grid and was produced using 4D-Var data assimilation in Integrated Forecast System. ERA5 is the successor of ERA-interim, which was, with lower time and spatial resolution and quality issues, successfully and officially replaced on 31 August 2019. Using the coordinates of the plots where forest growth data was available, the nearest grid cell based on the 30 km grid raster was selected. The hourly meteorological data were combined with the active and passive ozone data from the ICP Forests database.

2.6 Transformation from bi-weekly to hourly ozone data from passive samplers

Empirical functions such as the one developed by Loibl *et al*. (1994) are commonly used to reconstruct hourly ozone data based on aggregated weekly to monthly means. Previous works by Gerosa *et al*. (2007) have shown that this function performed well to estimate AOT40 values at 88 forest monitoring plots in France, Italy, Spain, and Switzerland; even though the atmospheric conditions significantly differ from the sites in the Austrian Alps where it has been calibrated. Similarly, Calatayud *et al*. (2016) tested this method on 24 Spanish sites by mimicking passive measurements through the aggregation of measured hourly data. PODy values calculated using aggregated hourly ozone data and Loibl's function were generally close to the ones calculated using the measured hourly data, but overestimated at sites with high ozone concentration, especially at sites with high PODy values, which can be critical when assessing ozone effects on forest growth.

[11] Url: https://cds.climate.copernicus.eu

Other approaches can generate hourly profiles based on meteorological conditions (e.g., Krupa *et al*. 2003), but their parameters are site-specific as well, and thus cannot be applied under large environmental gradients such as throughout Europe.

In consequence, we developed a new approach that allows the reconstruction of hourly ozone data based on measurements aggregated over longer time-scales (daily to monthly). We used a three-cosine function proposed by Huld *et al*. (2006), which was initially developed to estimate daily temperature profiles for energy systems based on daily temperature means. Considering that the temporal change in $[O_3]$ in a 24-hours period follows dynamics similar to the temperature ones (with or without lag), this function would allow us to mimic the increase in ozone during daytime, which is commonly followed by a decrease during the night. It is able to predict hourly $[O_3]$ values based on daily mean $[O_3]$ and three parameters (see equation below and Figure 4).

```
if (h<=hdawn)              {O3pred = O3mean - O3ampl * cos(pi*(hdawn-h)/(24+hdawn-hpeak))}
if (h<=hpeak & h>hdawn)    {O3pred = O3mean + O3ampl * cos(pi*(hpeak-h)/(hpeak-hdawn))}
if (h>hpeak)               {O3pred = O3mean - O3ampl * cos(pi*(24+hdawn-h)/(24+hdawn-hpeak))}
if (O3pred<0)              {O3pred=0} # Negative values are not allowed
```

where *O3mean* is the daily mean $[O_3]$; *O3ampl* the daily amplitude in $[O_3]$ (difference between the minimum and maximum); *Hdawn* and *Hpeak* the hours when at minimum and maximum $[O_3]$ occur, respectively. Considering that *O3mean* is the only output from passive samplers, we then analyzed how *O3ampl*, *Hdawn* and *Hpeak* change over space and time. For this aim, we used the most accurate data from the category 1 sites (i.e., the data without gaps or that were filled with the first gap-filling procedure only; 21 sites from 5 countries), and we developed linear mixed-effects models to predict them based on site-specific hourly meteorological variables such as the daily amplitude in temperature, the relative humidity, or the hour at sunset, which are known to impact the daily ozone profiles (e.g., Klingberg *et al*. 2012).

The new function better reproduces the daily profiles than the Loibl's function (Figures 4 and 5). It especially doesn't overestimate ozone concentration values during the first part of the day (from 5am to 3pm) and doesn't underestimate it in late afternoon (Figure 5). In consequence, the AOT40 (and POD estimates) were more accurately predicted than with Loibl's function or when assuming a constant ozone concentration throughout the day (Figures 6 and 7).

2.7 Summary of the data collected and cleaned

The resulting numbers from the data call and data validation and gap-filling procedures for plots, countries and plot*years are shown in Table 1.

The different sources of meteorological and ozone data allowed for several categories combining the one and the other source type. Figure 8 reveals the number of covered plots, countries and plot*years for each of the categories when combining the different data sources for meteorology and ozone:

category 1) measured hourly meteo & measured hourly ozone from ICP Forests;

category 2) measured hourly meteo & measured bi-weekly ozone from ICP Forests, transformed into hourly ozone values;

category 3) measured daily meteo & measured hourly ozone from ICP Forests;

category 4) measured daily meteo & measured bi-weekly ozone from ICP Forests, transformed into hourly ozone values;

category 5) modeled hourly meteorological and ozone data from EMEP based on CLRTAP (2017) parameterization for DO_3SE;

category 6) modeled hourly AOT40, POD0 and POD1 output from EMEP, based on parameterization according to Simpson *et al*. (2012);

category 7) modeled hourly meteo from ERA5 & measured hourly ozone from ICP Forests;

category 8) modeled hourly meteo from ERA5 & measured bi-weekly ozone from ICP Forests, transformed into hourly values.

2.8 Vegetation data

From the ICP Forests database, we used the individual tree data for each plot and each survey (dataset named *gr_ipm*) and its associated metadata (datasets *gr_pli*; *si_plt*; *si_sta*). First, we cleaned the data by removing surveys for which (i) the diameter of every tree is identical to the diameter from the previous survey; (ii) there was only a one-year gap with the present survey (in that case the diameter increment is in the same order of magnitude than the measurement error, and many negative diameter increments can occur); (iii) there was no link with the previous diameter measurement (*e.g., change in tree ID*); (iv) the number and size of incoming trees was too high while plot area was assumed to be constant; (v) there was fertilization or the presence of large external disturbances such as windthrows.

At tree level, several modifications were applied to correct – among others – wrong species identity, tree status (dead/living), and obvious typo in diameter data (e.g., +10 or -10 cm between two surveys; or exchange in diameter values between two trees).

Finally, we only selected stands dominated by our species of interest: beech, spruce, scots pine, fir, and oaks. From the initial 860 plots and 3’373 plot*surveys, we ended up with 498 plots and 1’235 plots*surveys. Automatic and manual procedures were also developed under R to screen and clean these data.

3. Data analyses

3.1 DO_3SE parameterization

The gapless hourly data containing air temperature (*Ts_C*), vapour pressure deficit (*VPD*), wind speed (*uh_zR*), precipitation (*precip*), atmospheric pressure (*P*), photosynthetically activate radiation (*PAR*) and ozone concentration (*O_3, ppb*) were supplied to DO_3SE. Only data from 1 April to 30 September were used. In addition, latitude, longitude, elevation and dominant tree species were supplied. The dominant tree species was determined per plot by selecting the species with the highest basal area in that plot, resulting in plots with dominating *Fagus sylvatica*, *Quercus robur*, *Quercus petraea*, *Picea abies*, *Abies alba* and *Pinus sylvestris* trees.

The CLRTAP mapping manual (CLRTAP 2017) was closely followed for species specific parametrization of the ozone stomatal flux model. Considering that most of the *F. sylvatica* and *P. abies* dominated plots in the dataset are growing in the continental region, the parameters for that region were used. *Q. robur* and *Q. petraea* were both grouped under deciduous oak species in the Mediterranean region. The parameters for *P. sylvestris* were obtained from Büker *et al.* (2015).

Initially, we set up similar DO_3SE model parameterizations for measured (ICP Forests) and modelled (EMEP and ERA5) input parameters. Parameterization on measurement heights differed between ICP Forests, ERA5 and EMEP (see Table 2). In addition, the DO_3SE model runs by EMEP (Fig. 8, category 6) were based on parameterizations according to Simpson *et al.* (2012) and not according to CLRTAP (2017) (shown in Table 2). EMEP, i.e. Simpson *et al*. (2012) only distinguishes between coniferous (CF) and deciduous (DF) forests and uses different phenological parameters, as well as different parameters for the growing season, compared to CLRTAP (2017).

3.2 Empirical growth models

To analyze the potential links between radial tree growth and ozone, we followed a phenomenological approach with the non-linear mixed-effects model developed by Rohner *et al*. (2018). This model, which is implemented in the forest model MASSIMO (Thürig *et al*. 2005), is able to predict the individual basal area increment (*BAI*) based on several explanatory variables reflecting individual-tree characteristics, site and stand conditions, climate, nitrogen deposition, and ozone (e.g. Figure 9).

Based on the Swiss NFI data, this model was able to explain a large part of the variance in *BAI* of spruce, pines, beech, fir and oaks in Switzerland. Hence, we believe that this model may be an appropriate approach to assess potential ozone effects on radial growth across Europe.

BAI is simulated as:

$$BAI = e^{b_1*(1-e^{b_2*DBH})} * e^{b_3} + \varepsilon \qquad (1)$$

Where b_1, b_2, and b_3 are the coefficients that have to be estimated and ε the residual error. b_3 is function of n explanatory variables V_1 to V_n according to:

$$b_3 = \beta_O + \beta_1 V_1 + \cdots + \beta_n V_n + b_{plot} \qquad (2)$$

where β_0 is an estimated fixed intercept, $\beta 1$ to βn are the model coefficients estimated for the included explanatory variables, and b_{plot} is a random intercept with the ICP Forests plots as grouping factor, which takes into account the grouped structure of the data (several trees per plot). All numerical variables not distributed around zero (except for *DBH*) were centered and scaled before being included in the models to enhance model convergence.

The explanatory variables (V_n) tested were grouped into three categories:

1) Tree- and stand- related variables: based on the inventory data (dendrometric surveys), we calculated and included tree *DBH*, stand basal area (*BA*), the basal area of larger trees (*BAL*), a stand density index (*SDI*), the mean DBH of the 100 thickest trees per ha (*DDOM*).
2) Site-related variables: site topography was considered with the *Asp* index which considers both, slope and aspects (see Schumacher *et al*. 2004), and soil *pH*. Soil water content is not available at most of the sites, so we used a new open-access database of harmonized soil properties[12] (Batjes 2016) on a 30*30 arc-seconds grid to estimate it. Altitude was not included as potential explanatory variables as it is implicitly considered in the climate variables.
3) Atmospheric variables: we included classical (i) climate variables reflecting both, temperature (mean annual temperature or mean temperature during the growing season

[12] Url: https://www.isric.org/explore/wise-databases

[*gr*; April to September]) and climatic water balance (*WB*; calculated as the precipitations (*P*) - Potential Evapotranspiration (*PET*) during the entire year or during the growing season), (ii) atmospheric nitrogen deposition (*Ndep*), which was estimated using the EMEP model, and (iii) one ozone metric (mean ozone concentration, AOT40, POD0, or POD1).

For every species and data category, we selected the best and most parsimonious models based on the Akaike Information Criteria (AIC). During this procedure, we also excluded variables with high inter-correlation (also called “competing variables”; e.g., *SDI* and stand *BA*). Interactions among variables and potential non-linear effects of the climatic, site, stand, and atmospheric variables were not yet tested.

4. Results

4.1 DO3SE outputs

AOT40, POD0 & POD1 for categories 1-8 were assessed for at least 221 plot*years based on measured data, and 278’122 plot*years based on modelled EMEP data. Category 2 contained the least plot*years, due to the relative low amount of both, hourly meteorological data and bi-weekly ozone data, whereas hourly meteorological data from ICP Forests, still seems to pose the main bottle neck. Bi-weekly ozone data from passive samplers allowed a spatio-temporal coverage of 1’094 plot*years, but there seems to be little overlap in the amount of plot*years of ozone and those of meteorological data (Figure 8).

For AOT40, measured data show obvious larger variation than modelled EMEP data (Figure 10). The order of magnitude is quite similar between measured and modelled data, although for areas dominated by *F. sylvatica* or *P. abies,* often the measured ozone levels (AOT40) seem to be doubled with respect to modelled EMEP data. For POD0/1, the most obvious difference is the order of magnitude difference between modelled POD0/1 output from EMEP – based on Simpson *et al*. 2012 (category 6) and the modelled output using CLRTAP (2017) parameterization (categories 1-5). The EMEP output consistently shows lower POD values than the outputs using the CLRTAP (2017) parameterization.

There is no clear difference between uptake of ozone in deciduous or evergreen species. Often, *P. sylvestris* and *Quercus* species show slightly higher values for yearly ozone uptake than other species (Figure 10). When visualizing modelled AOT and POD on a spatial scale (Figures 11-16), it becomes clearer that the spatial resolution of measured hourly meteorological data is the most limiting of all. Whilst category 1 (hourly meteo, hourly ozone) mostly cover continental Europe (Figure 11), combining hourly meteorological data with bi-weekly ozone provides more data on Mediterranean parts of Europe (Figure 12). The highest spatial resolution, obtained with EMEP modelled data (Figures 15 and 16), shows a clear north-south gradient that is harder to distinguish with the relatively low spatial resolution in categories 1-4. When looking at category 5 (Figure 15), we find that even when AOT40 values are relatively low, the stomatal uptake indicated by POD0/1 shows fairly high values scattered over Europe, reflecting the differences in the dominance of tree species in different sites and their differences in stomatal uptake (Figure 15).

Some remarkable differences were observed between categories using the CLRTAP (2017) parameterization for DO_3SE (categories 1-5, 7 and 8) and the outcomes from the EMEP DO_3SE model (category 6), using parameterization for CF and DF according to Simpson *et al*. (2012). We therefore decided to study these differences more precisely by using the comparison between category 5 (EMEP data, with parameterization based on CLRTAP (2017) and category 6 (EMEP outputs, based on DO_3SE parameterization from Simpson *et al.* (2012). Figure 17a shows the differences in POD1 for the two different parameterizations. Although POD1 for *F. sylvatica* seems to be corresponding between the two different approaches, for all other species, POD1 from using DO_3SE with the CLRTAP (2017) parameterization were largely overestimating the predictions from EMEP. To test whether these differences originate from errors in data and/or our DO_3SE model, or from the species-specific parameterization, we used the parameterization as described in Simpson *et al*. (2012) to reproduce the outcomes from the EMEP model. When applying the CF/DF parameterization (Simpson *et al*. 2012) to our DO_3SE model (instead of CLRATP 2017 parameterization), the EMEP output can be fairly well reproduced (see Figure 17b). It is remarkable that the differences in POD1 between both approaches are that big, considering recent arguments that more species-specific parameterizations for gas exchange and growing season (which was not even included in the parameterization used in this project now) is needed to better predict species specific responses to ozone.

4.2 Output from empirical growth models

Using meteorological and ozone data from atmospherical models is a classical procedure in ozone risk assessment studies (e.g., de Vries *et al*. 2017 with EMEP; Paoletti *et al*. 2018 with CHIMERE) as it allows the consideration of all study plots despite the potential lack of atmospheric stations there, and the use of a harmonized dataset. We therefore considered the EMEP approach (category 6) as our reference. In that case, the mixed-effect models developed to predict tree-level basal area increment of the 6 species of interest based on the variables mentioned above showed a good predictive accuracy, with R^2 ranging between 0.52 and 0.67 (Table 3), confirming our choices of following the modeling procedure from Rohner *et al*. (2018) and of the explanatory variables.

When using EMEP data (category 6) and mean ozone concentrations, the models showed that:

- trees which experienced high relative competition intensity (high stand *BA* and *DDOM*) had lower *BAI*. Unexpectedly, this was not observed for Scots pine (and to a lower extent for spruce) although this species is a pioneer and less shade-tolerant than the others. This may be due to the fact that the productive stands are also the ones with higher *DDOM*.
- As expected, higher climatic water balance is related with higher tree growth.
- All tree species located in warm sites had higher growth rates than at colder sites; except for oak for which a high mean temperature may be detrimental.
- Nitrogen deposition had a positive effect on tree growth, except for oak and fir. The lack of linear relationships for both species is not due to sampling artefacts (their distribution in nitrogen deposition among plots is similar to the ones for the other species), but may be caused by non-linear and species-specific threshold-based effects of nitrogen deposition.
- There is no apparent effect of the stand topography, soil pH, and soil water content on individual tree growth.
- Finally, mean ozone concentrations show a negative relationship with tree *BAI* for beech and oak only, while there was no significant effect for the coniferous species.

All of these relationships show that the drivers of tree *BAI* differ among the studied tree species, which may be due to both, (i) differences in species-specific environmental requirements (e.g., drought tolerance), and (ii) differences in the representativeness of the study stands within the species distribution. However, they are globally consistent, and are ecologically sound, which confirms the validity of the data and modeling approach used.

Focusing on the effect of ozone, which is the main driver of interest in the present project, we first observed that applying the parameterizations, based on either CLRATP (2017) or Simpson *et al*. (2012) for DO_3SE and EMEP ozone deposition and uptake schemes may alter the AOT40 and POD estimates (with strictly identical input meteorological and ozone data; see Figure 17), and consequently the ozone-growth relationships (see categories 5 *vs*. 6). On the one hand, no significant changes were observed for spruce (Figure 18), and for beech, the sign of the relationships did not change and the p-values differ slightly (Figure 19). On the other side, for oak, Scots pine and fir, the relationships calculated with POD0 and POD1 became more significant with DO_3SE (category 5) than with EMEP (category 6): it became more positive for oak (Figure 20) and pine (Figure 21) and more negative for fir (Figure 22). This increase in significance is due to the larger range (variability among plots and years) in POD estimated by DO_3SE than when POD is calculated by EMEP (Figure 17A).

Second, the relationships may also change with the ozone metric used (mean ozone concentration, AOT40, POD0, or POD1) for a given data category. For instance, with the category 6 data, there was a strong negative relationship between mean ozone concentration and beech *BAI*, which was less significant when using AOT40, and became non-significant with the POD estimates (Figure 19). Similar patterns were observed for the other categories and for the other species (Figures 18 to 22).

Third, relationships may also vary with the number of plots/trees/observations considered in the statistical model. In order to analyze potential sampling-related artefacts due to the decrease in the number of plots*years for which measured data is available (categories 1 to 4), we also built statistical models for these plots using EMEP data as input (crosses in Figures 18 to 22) and compared them with the statistical models built with all plots (category 6). Such comparisons showed that different relationships could be obtained according to the sample size and representativeness (i.e., according to the number and characteristics of the plots considered). For instance, the negative relationships observed between PODy and spruce *BAI* considering all plots became non-significant, and even positive when the number of plots was lower (Figure 18). For this species, no consistent ozone-growth relationship was obtained among all categories (Figure 18), while for beech it was more consistent (always negative; Figure 19).

Finally, the relationships also changed substantially with the source of the input meteorological and ozone data, i.e., if they are modelled or measured (crosses *vs*. dots for categories 1 to 4 in Figures 18 to 22), and if the measured bi-weekly ozone data are derived from passive samplers (categories 2 and 4) or from hourly ozone data (categories 1 and 3). The most evident example arises for category 4, for which the relationships strongly differed from the EMEP ones (e.g., for beech but also valid for the other species).

5. Conclusions

DO_3SE outputs

For PODy and AOT40, we could not find consistence patterns between the different categories and tree species. Such a model's variability could arise from various sources of uncertainties and data quality. In particular, for the temporal and spatial coverage, measured hourly meteorological data from ICP Forests seem to pose the main bottle neck for assessing ozone-growth relationships across European forests. The low numbers of 28 plots from 10 countries and 221 plots*years limit our comparison of the ozone-growth relationships obtained from the different categories. This low spatio-temporal coverage is partly due to the low quality of national data. More efforts should be invested into data aggregation, gap-filling and validation of hourly meteorological data from ICP Forests to allow dose-response relationships on a denser spatio-temporal coverage.

We found significant differences in POD0/1 between the categories 1-5, 7, 8 and the category 6. We therefore applied the Simpson *et al*. (2012) parameterization (instead of CLRTAP 2017) to our

DO_3SE model and could demonstrate that the differences originate from the differing parameterizations (see Figure 17). This suggests that further investigations are needed to harmonize the calibration between the CLRTAP (2017) and the EMEP approach to better assess species-specific responses to ozone, POD0/1 respectively.

Output from empirical growth models

A consistent output from the statistical models developed to analyze the potential link between tree growth and ozone is that the relationships are highly unstable and that the various uncertainties behind them should deserve more attention by us, but also by the overall community. To our knowledge, no study previously tested the robustness of inventory-based ozone-growth responses according to (1) the characteristics of the sampling (e.g., sample size), (2) the source and quality of the meteorological and ozone data used as input (e.g., modelled *vs*. measured data; passive *vs*. active ozone measurements), and (3) the modeling approach used (including both ozone deposition and uptake model; cf. EMEP *vs*. DO_3SE; and statistical growth model). Our analysis is not yet completed: the statistical models have to be improved to better consider (i) the effect of soil water availability using drought indices calculated from soil water balance models; (ii) potential interactions among explanatory variables; (iii) and non-linear relationships. A similar exercise will be also done with the Swiss NFI data. Thus, at this point we do not want to over-interpret the relationships we obtained, and especially the differences in ozone sensitivity among species as they may change with further analyses. However, the current stage of the analyses clearly showed that these various sources of uncertainty must be appropriately considered in future small- and large-scale ozone risk assessment studies. A statistical decomposition of all sources of variability in the ozone-growth relationships would also be needed to detect the predominant ones, and to highlight research priorities. Our study provides a valuable basis for stake holders and helps the decision process towards (i) longer-term data, *vs*. (ii) more plots, *vs*. (iii) more active (hourly) measurements (and neglect modelled data) to obtain robust ozone-growth responses.

Tables and Figures

Table 1. Temporal and spatial coverage of meteorological, ozone and growth data, after the application of the above outlined routines.

	Hourly meteorological data	**Daily meteorological data**	**ERA5/EMEP meteorological data**	**Hourly O_3 active data**	**Bi-weekly O_3 passive data**	**Vegetation data**
# of plots	28	149	5’349	57	191	498
# of countries	10	19	26	12	19	26
# of plots*years	221	1’263	85’584	531	1’094	4’704

Table 2. Parameterizations of the DO_3SE model for both, ICPF and ERA5 input data (according to POD_1SPEC parameterization in the CLRTAP mapping manual (CLRTAP, 2017) and EMEP data (according to the technical description by Simpson *et al*. (2012).

		ICPF/ERA5 (*CLRTAP 2017*)				EMEP (*Simpson et al. 2012*)	
		FSyl	**QRob**	**PAbi-AAlb**	**PSyl**	**CF**	**DF**
Measured height (m)	*Wind*	4.6	4.6	4.6	4.6	20	20
	O_3	2	2	2	2	20	20
Canopy height		25	20	20	20	20[1]	20[1]
Growing season	*SGS50*	105 (1.5)	105 (1.5)	0	0	0	100 (1.5)
	EGS50	297 (-2.00)	297 (-2.00)	366	366	366	307 (-2.00)
LAI parameters	LAI_{min}	0	0	4	4	5	0
	LAI_{max}	4	4	4	4	5	4
	L_S *(days)*	30	30	30	30	1	20
	L_E *(days)*	30	30	30	30	1	30
	g_{max}	155	265	130	180	140	150
	f_{min}	0.13	0.13	0.16	0.1	0.1	0.1
f_{phen} factors	φ_a	0	0.3	0	0.8	0.8	0
	φ_b	1	1	1	1	0.8	1
	φ_d	1	1	1	1	0.8	1
	φ_e	0.4	0.3	0	0.8	0.8	0
	φ_1 *(days)*	20	15	0	40	1	20
	φ_4 *(days)*	20	20	0	40	1	30
	f_{light}	0.006	0.006	0.01	0.006	0.006	0.006
Temp.	T_{min}	5	0	0	0	0	0
	T_{opt}	16	22	14	20	18	20
	T_{max}	33	35	35	36	36	35
VPD	D_{max}	1	1.1	0.5	0.6	0.5	1
	D_{min}	3.1	3.1	3	2.8	3	3.25

SGS50, EGS50 are start and end of growing seasons (day number) at 50° N. Values in parentheses are rates of change in days (e.g. Δ SGS/Δ latitude) with latitude. For example, SGS for DF occurs later at the rate of 1.5 days per degree latitude on moving north, or earlier when moving south; L_s and L_E are days from start growing season to reach maximum LAI, and from end growing season to reach minimum LAI; boreal forests north of 60° N, height is reduced by 5% per degree extra latitude, down to a minimum of 6 m for 74° N and above. LAI_{min}, LAI_{max} for EMEP, is reduced in the same proportion.

Table 3. Summary of the mixed-effect models fitted to tree-level Basal Area Increment for the category 6 plots (i.e., using meteorological and ozone variables from EMEP) and mean ozone concentration as ozone metric.

	Fagus sylvatica	***Quercus rob/pet***	***Pinus sylvestris***	***Picea abies***	***Abies alba***
Nb observations	38'579	10'948	40'792	33'711	8'181
Nb plots	169	117	151	150	64
AIC	357'516	104'189	364'265	324'726	77'472
R^2	0.65	0.59	0.52	0.57	0.67
stand BA	**-0.14**	**-0.08**	**-0.07**	**-0.16**	**-0.24**
DDOM	**-0.10**	**-0.18**	**0.24**	0.03	**-0.19**
grWB	**0.16**	**0.30**	**0.49**	**0.37**	**1.73**
grT	ns	ns	ns	ns	**1.98**
meanT	**0.13**	**-0.21**	**0.47**	**0.54**	ns
Ndep	**0.19**	ns	**0.46**	**0.65**	ns
SWC		ns	ns	ns	ns
Asp	ns	ns	ns	ns	ns
pH	ns	ns	ns	ns	ns
Mean $[O_3]$	**-0.09**	-0.06	ns	ns	ns

Stand BA: stand Basal Area (m^2/ha); *DDOM*: mean Diameter of the Dominant trees (cm); *grWB*: water balance during the growing season (mm); *grT*: mean temperature during the growing season (°C); *meanT*: mean annual temperature (°C); *Ndep*: stand-scale Nitrogen deposition as simulated by EMEP (mg/m^2; 50*50 km grid); *SWC*: soil water content estimated from the WISE database; *Asp*: slope and aspect variable; *pH*: soil pH; *mean $[O_3]$*: mean ozone concentration. Bolded values indicate significant effects at $p<0.01$; otherwise $p<0.05$.

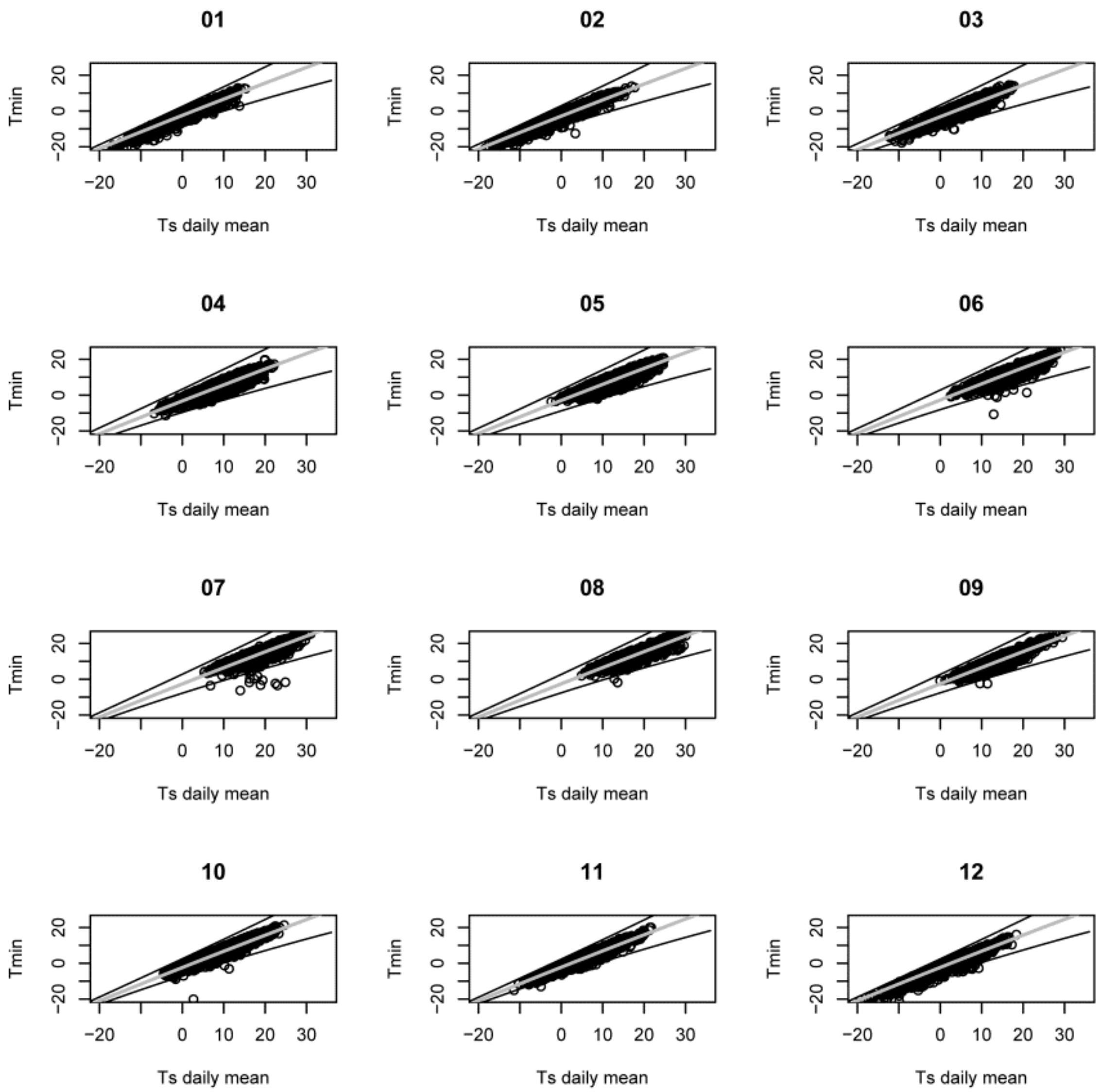

Figure 1. Modelled T_{min} *vs*. measured mean daily temperature (C°) for each month of the year. Black dots show measured data, grey lines show the modelled values and black lines show the credibility intervals of the modelled data.

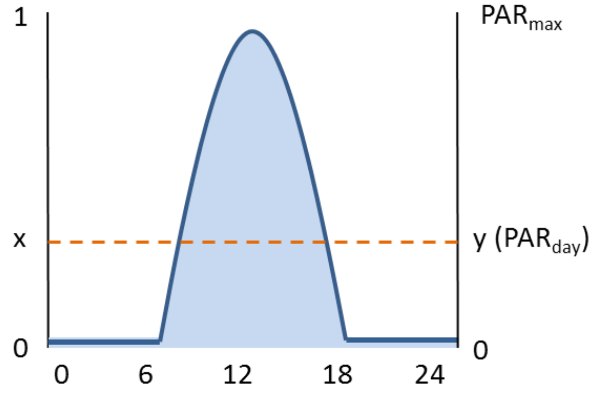

Figure 2. Scaling of daily *PAR* to hourly *PAR* using daily average and daytime hours.

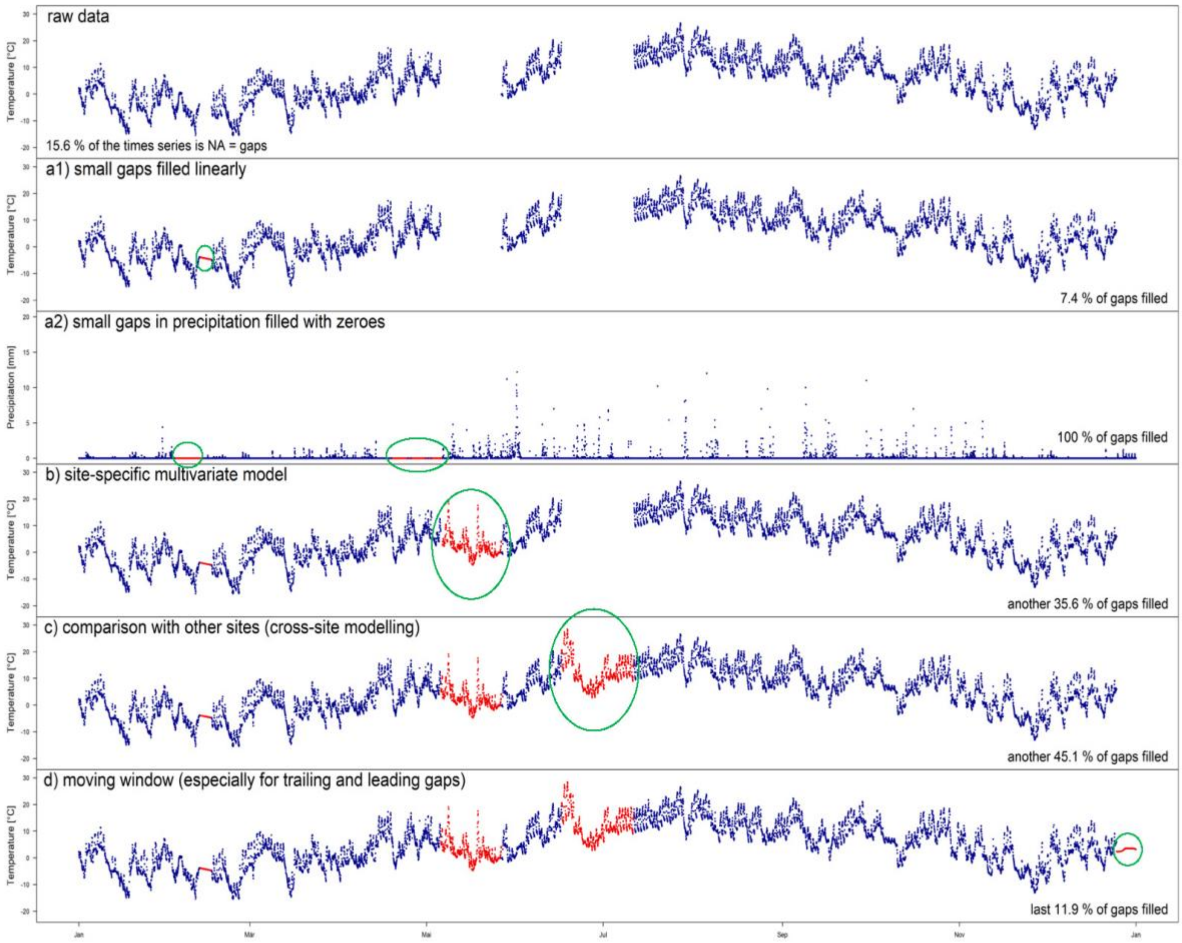

Figure 3. Exemplary time series as processed with the *four-step step-wise gap-filling procedure*. The first "raw data" subplot shows that there are several gaps for a temperature time series of a randomly chosen year and site. (a1) displays that small gaps (< 14 days) are filled linearly. This has been done for all parameters except for precipitation. (a2) shows a randomly chosen precipitation time series where small gaps (<14 days) are filled with zeroes. (b) shows the site-specific multivariate model. In this case, parts of the temperature time series could be reconstructed with the best model found ($temp$ = 0.926*VPD + 0.404*uh_zR + 1.280*$precip$ + 2.389*P – 208.098; R^2 = 0.6277; $p < 0.001$). (c) shows the step where the corresponding temperature time series is linearly compared to all other sites. The best one (in this case a neighboring site) delivered $R^2 > 0.95$ ($p < 0.001$) and hence could be used to reconstruct another gap in the time series. The last step (d) filled the last part of gaps in this case leading to a gapless time series for temperature for this corresponding site and year.

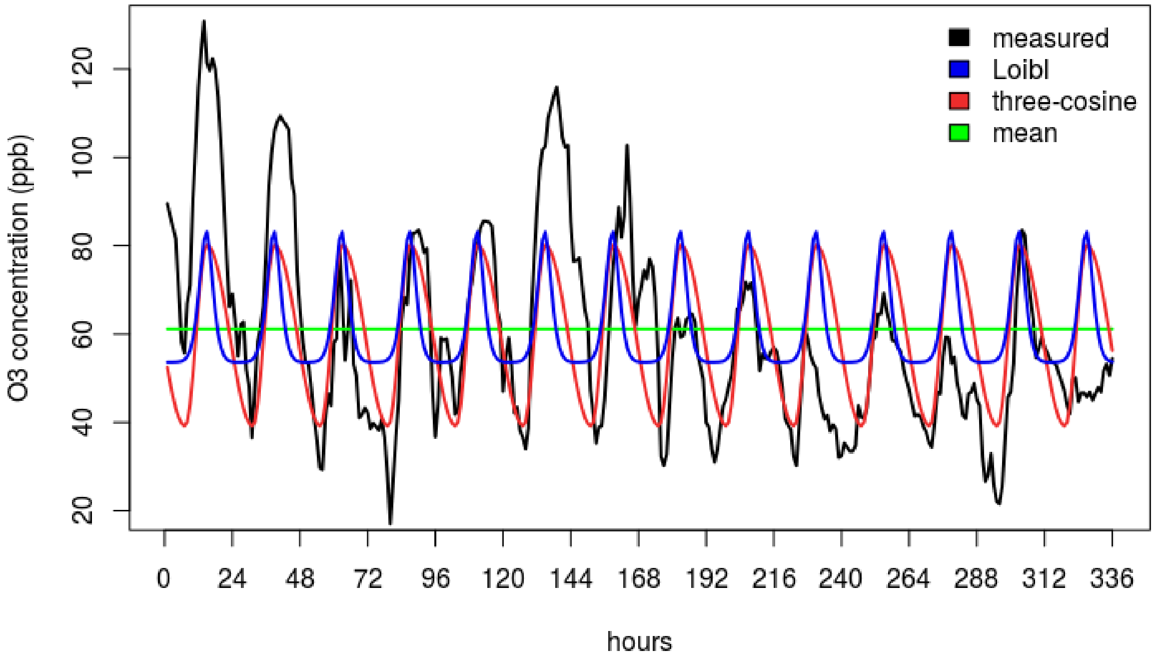

Figure 4. Examples of 2-weeks ozone concentration profiles (in ppb). Measured hourly data from active measurements (in black) are indicated together with the mean weekly values (green), and the hourly values predicted by the three-cosine function (red) and by Loibl's function (blue).

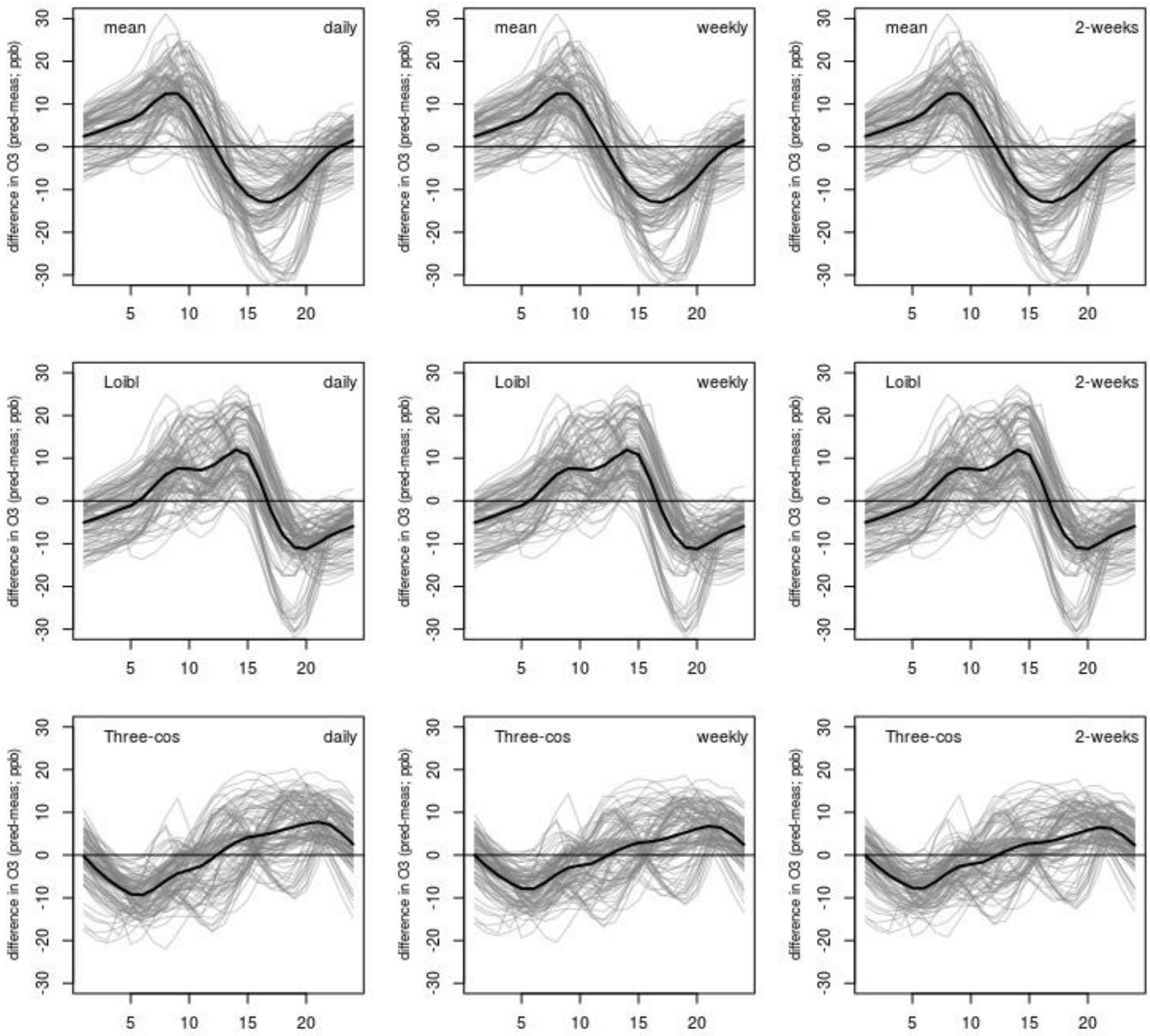

Figure 5. Differences between the daily profiles predicted using the aggregated data at daily (left), weekly (center), and 2-weeks (right) resolution, and the measured ones. Different approaches were used to predict the daily profiles, from top to bottom: constant mean values, Loibl's function, and the three-cosine function whose parameters were estimated based on the mixed-models. Grey lines: data for one specific year at a given site. Black line: average across sites and years.

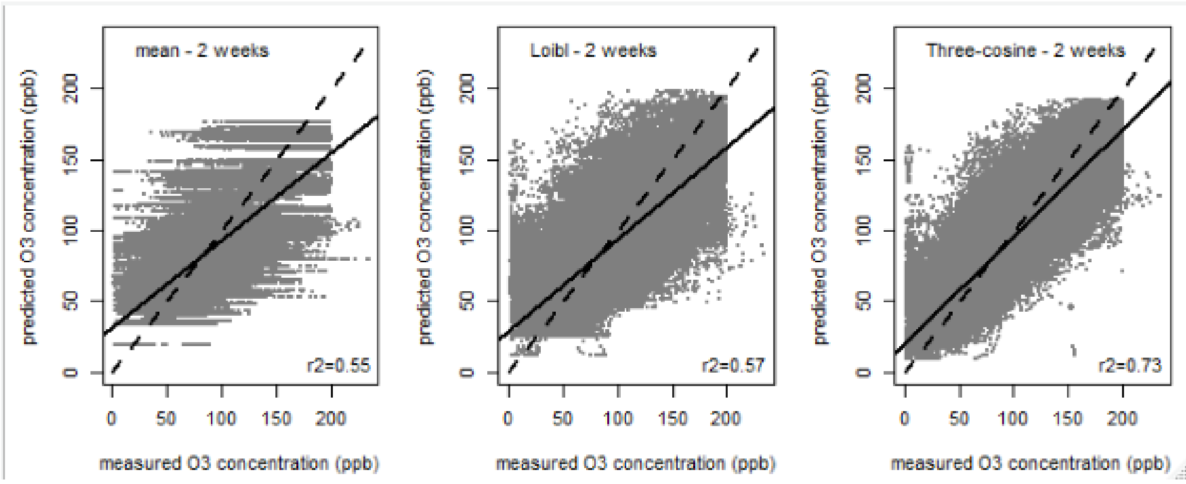

Figure 6. Relationship between the ozone concentration hourly data measured and predicted by the mean, Loibl *et al*. 1994, and three-cosine approaches. The results are shown only for the 2-weeks period, but are similar for the other periods (see Figure 5).

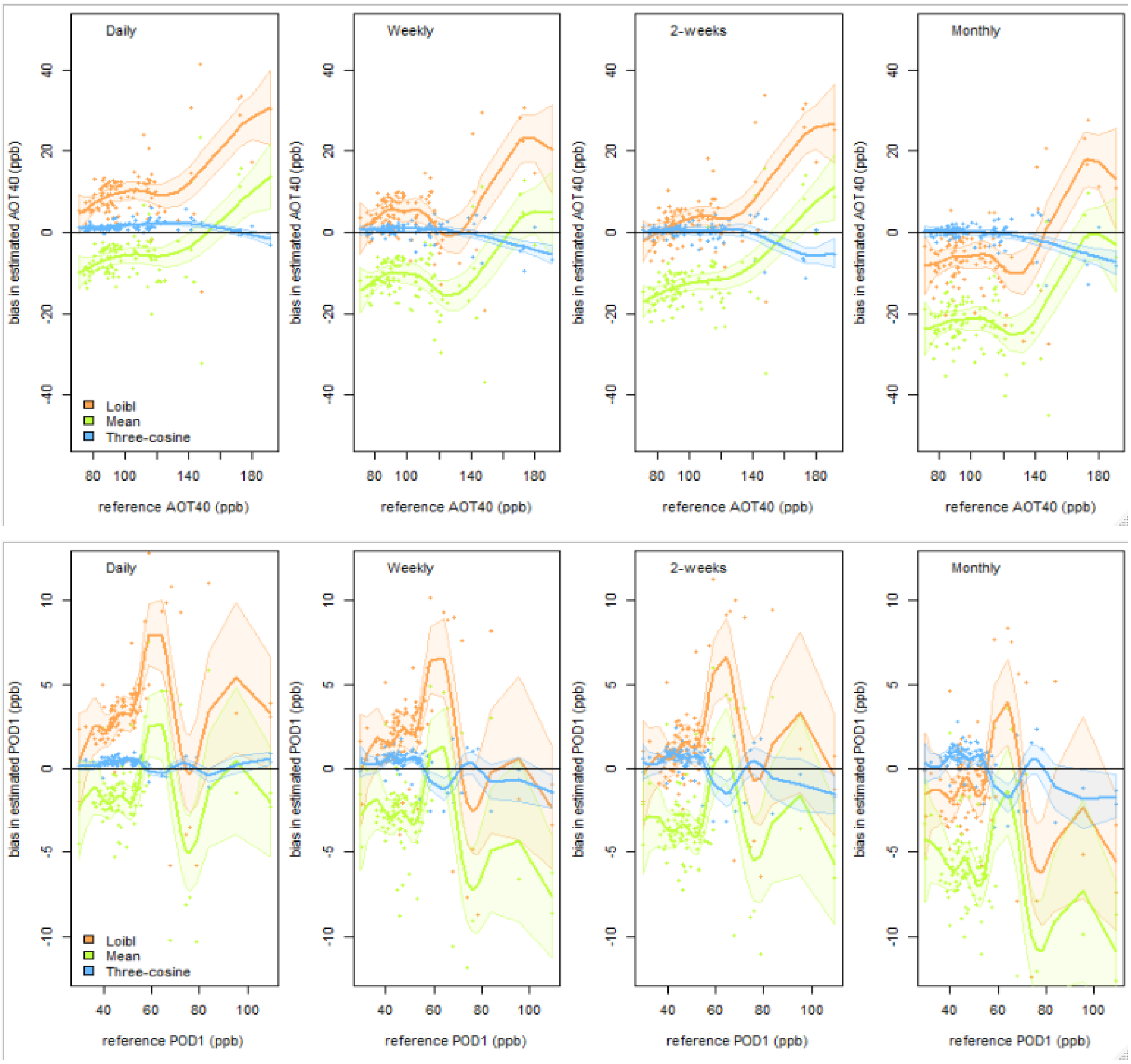

Figure 7. Bias in AOT40 (*top*) and POD1 (*bottom*) estimates calculated using aggregated ozone data and calculated using measured hourly ozone data (reference). Green: mean values (constant ozone within the time-window of interest); Orange: Loibl's function; Blue: three-cosine function with parameters estimated by the mixed-models. Negative values indicate that the aggregated approach underestimates AOT40. Results were shown for beech, but similar results were obtained for the other species. Generalized additive models were used to highlight mean trends and confidence intervals.

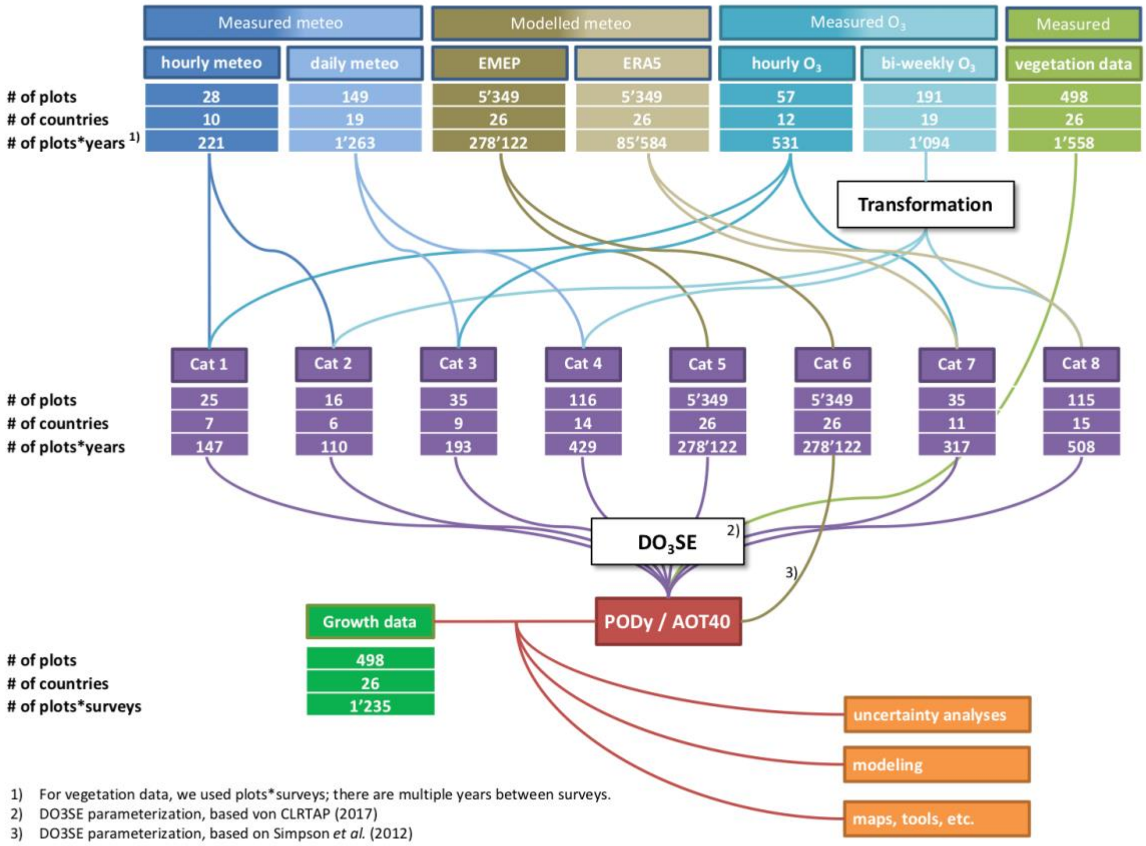

Figure 8. Number of plots, countries and plot*years for each combination of meteorological and ozone data sources with available and gapless, measured and modeled data, resulting in 8 different categories.

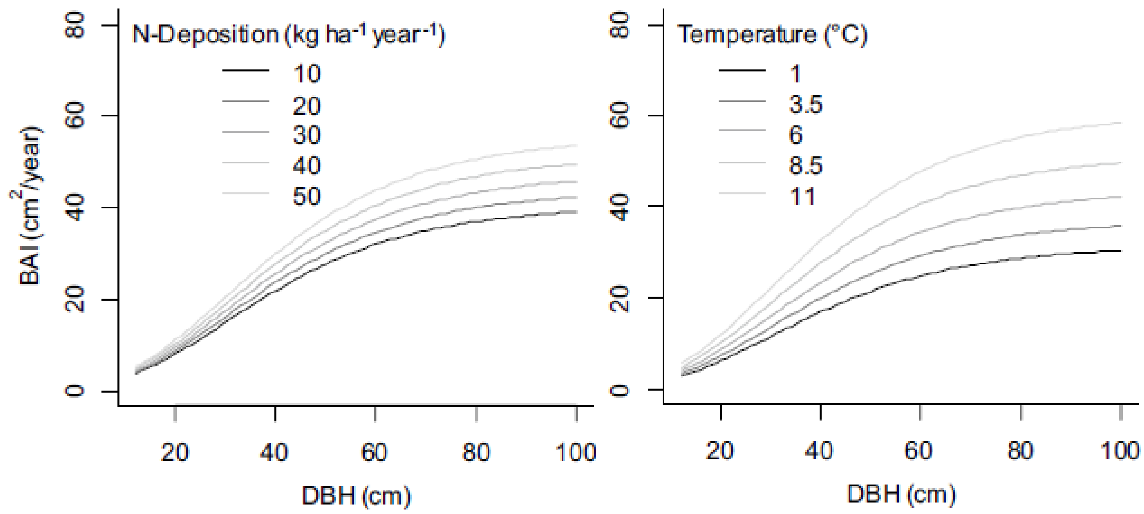

Figure 9. Example of non-linear relationships between tree diameter at breast height of 1.3 m (*DBH*) and basal area increment (*BAI*) as simulated by the mixed-model considering various levels of nitrogen deposition and mean annual temperature (from Rohner *et al*. 2018; example for beech).

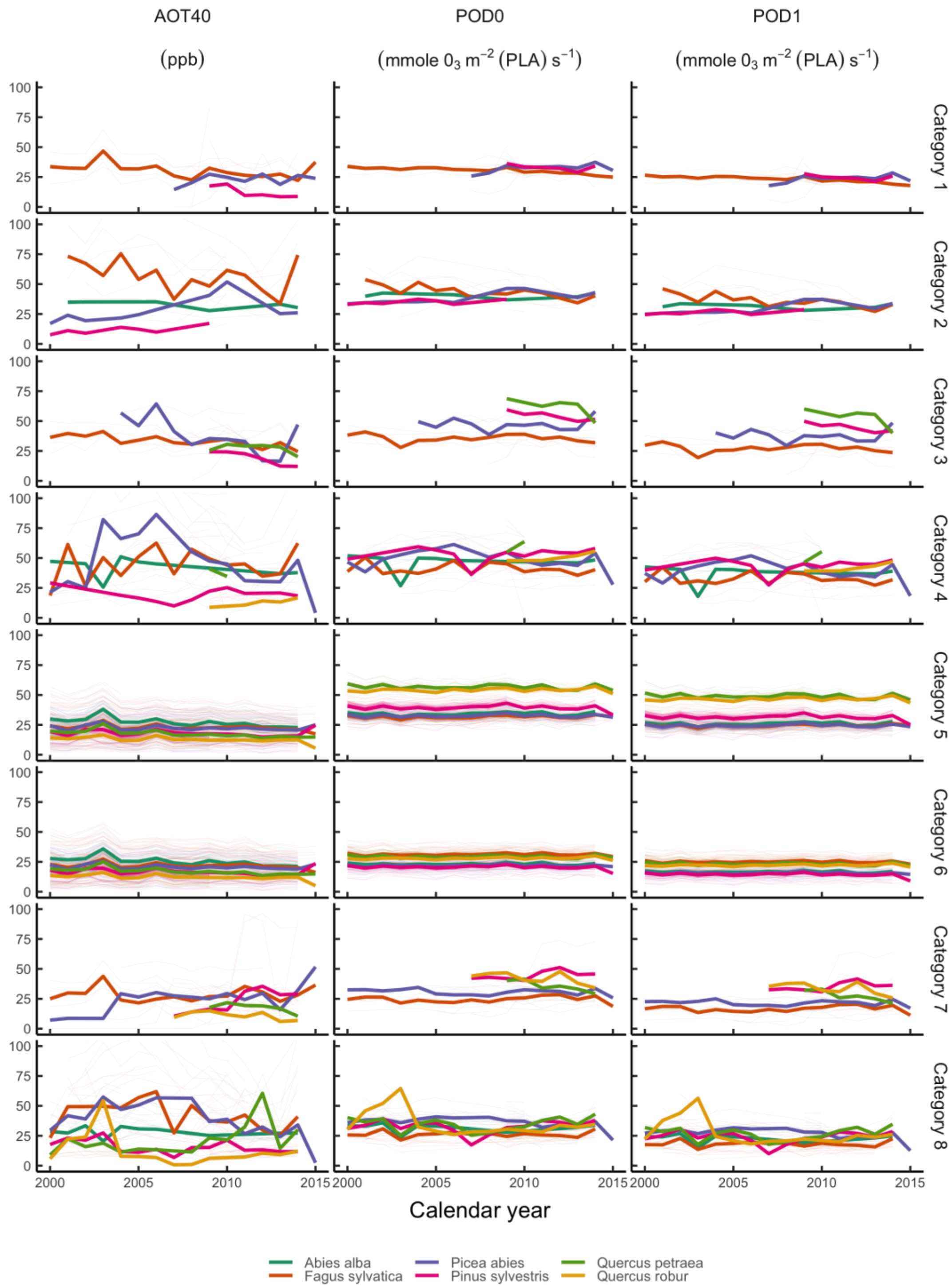

Figure 10. Time series for AOT40, POD0 and POD1 from DO$_3$SE with measured hourly meteorological and measured hourly ozone input data from ICP Forests (category 1), measured hourly meteorological and measured bi-weekly ozone data from ICP Forests, transformed into hourly values (category 2), measured daily meteorological and hourly ozone input data from ICP Forests (category 3), measured daily meteorological and bi-weekly ozone data from ICP Forests, transformed into hourly ozone values (category 4), modeled hourly meteorological and ozone input data from EMEP (category 5), modeled hourly AOT40, POD0 and POD1 output from EMEP (category 6), modeled hourly meteorological input data from ERA5 and measured hourly ozone input data from ICP Forests (category 7), and modeled hourly meteorological input data from ERA5 and measured bi-weekly ozone data from ICP Forests, transformed into hourly values (category 7).

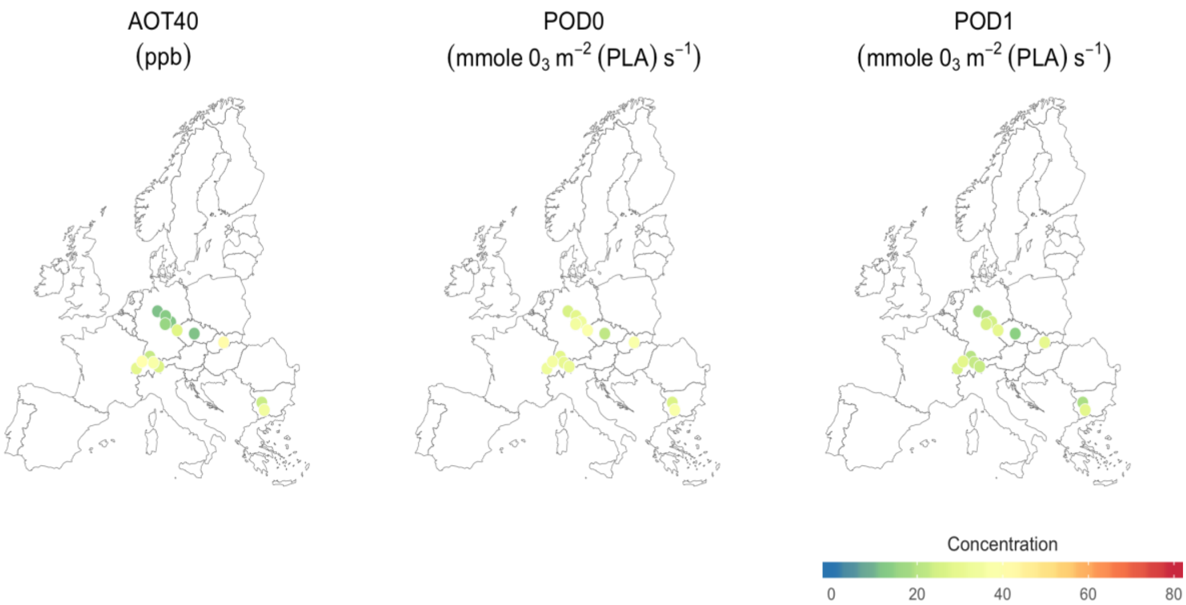

Figure 11. Spatial distribution of mean 2000-2015 values for AOT40, POD0 and POD1 based on measured hourly meteorological and hourly ozone input data from ICP Forests (Figure 8, category 1) for DO_3SE.

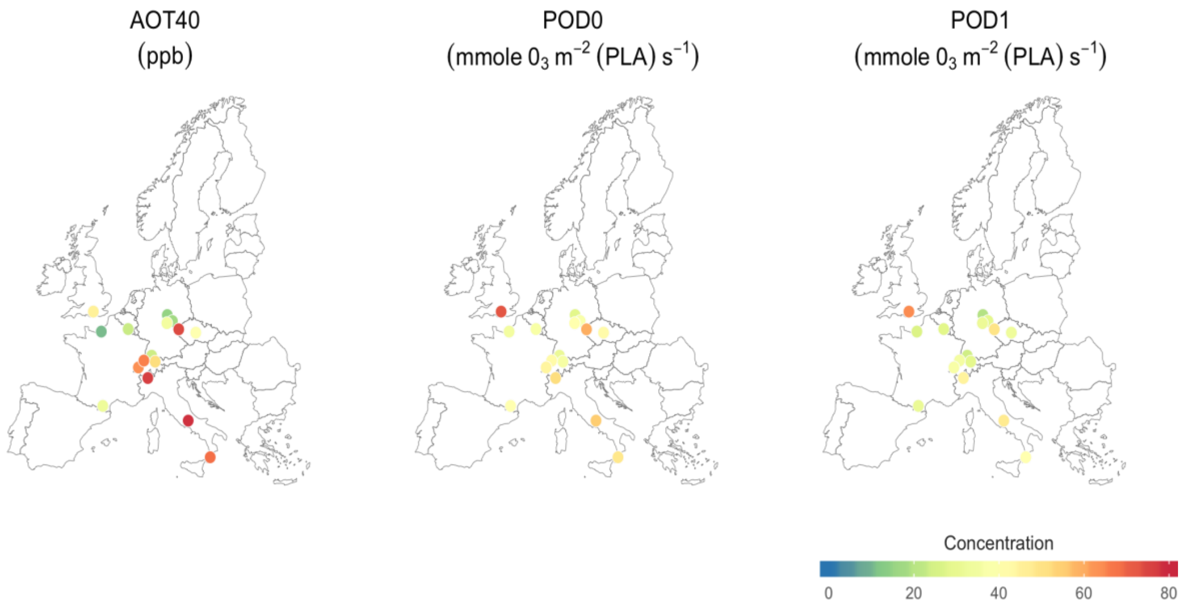

Figure 12. Spatial distribution of mean 2000-2015 values for AOT40, POD0 and POD1 based on measured hourly meteorological and bi-weekly ozone input data from ICP Forests (Figure 8, category 2) for DO_3SE.

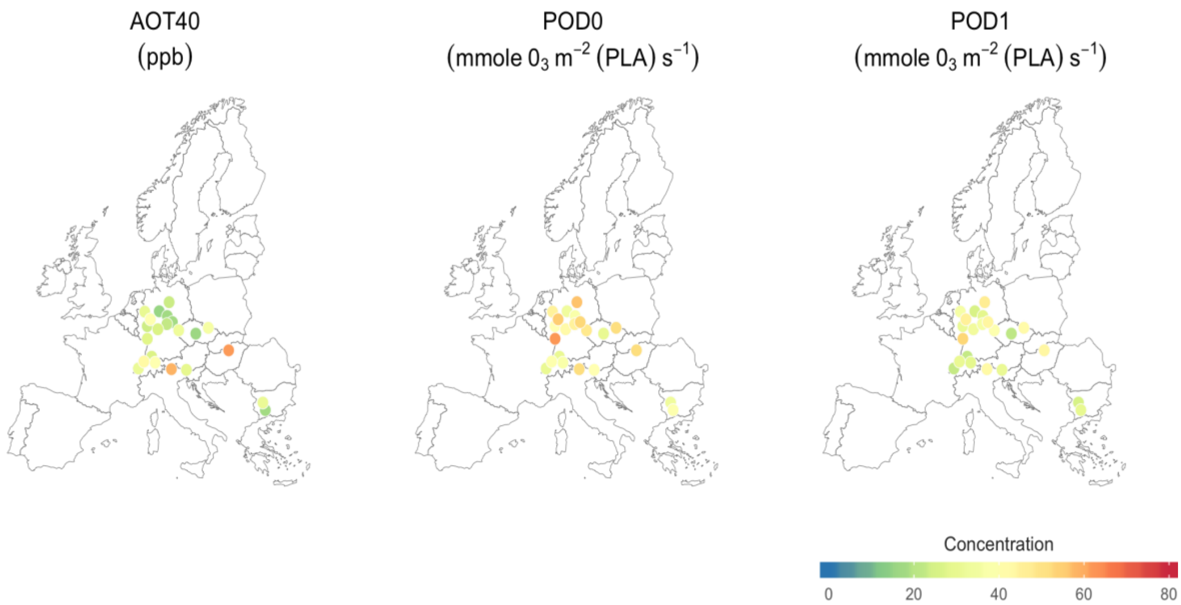

Figure 13. Spatial distribution of mean 2000-2015 values for AOT40, POD0 and POD1 based on measured daily input data (Figure 8, category 3) for DO_3SE.

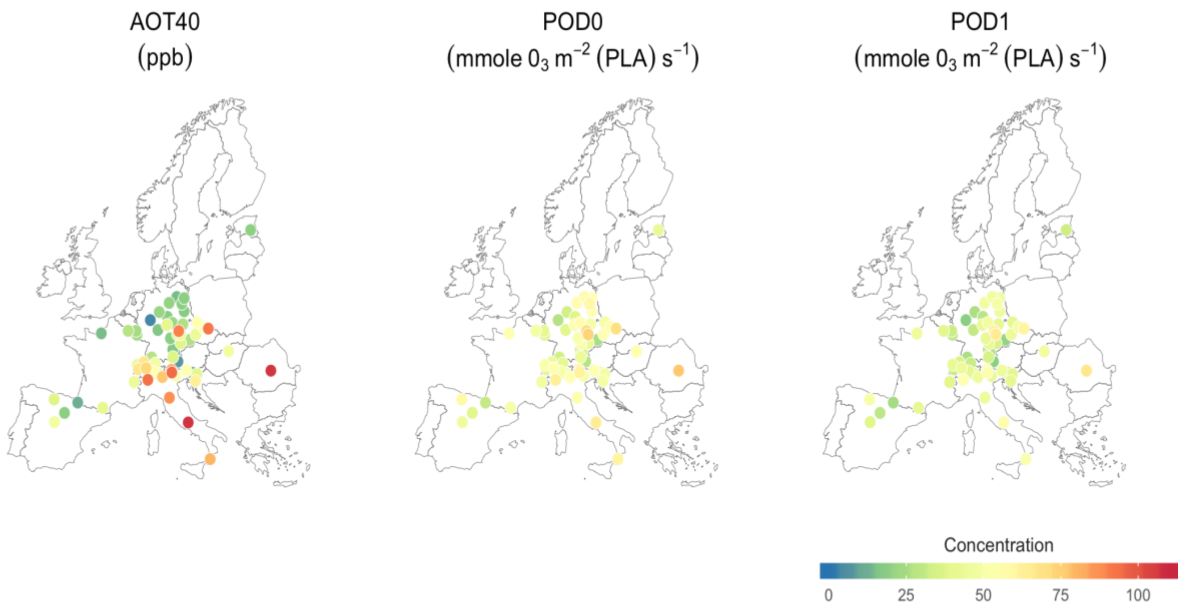

Figure 14. Spatial distribution of mean 2000-2015 values for AOT40, POD0 and POD1 based on measured daily meteorological and bi-weekly ozone input data from ICP Forests (Figure 8, category 4) for DO_3SE.

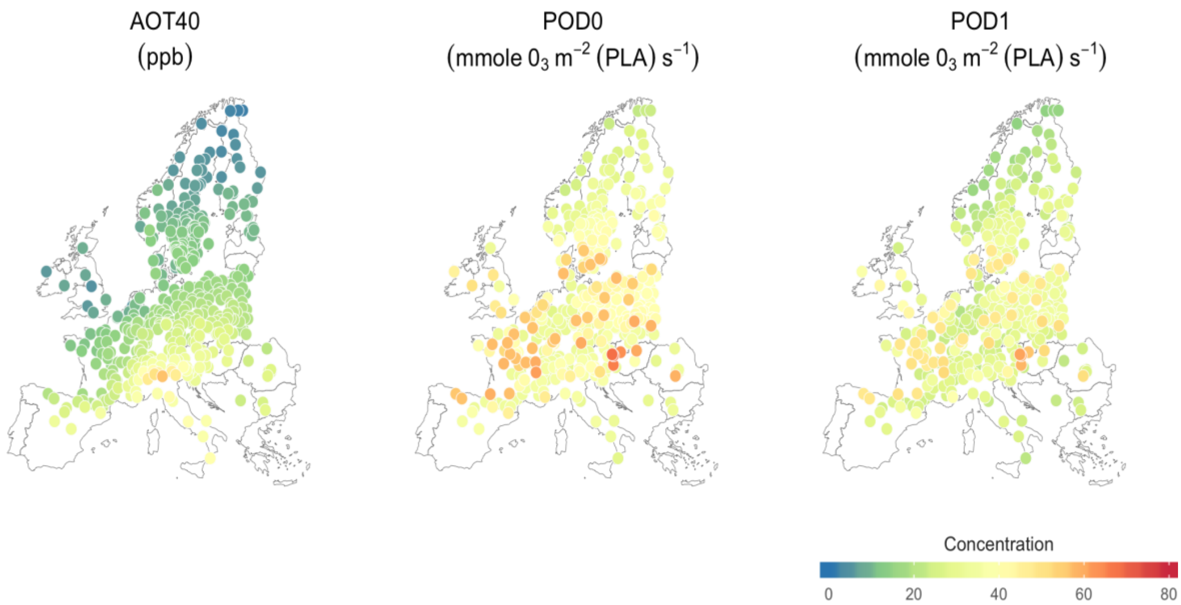

Figure 15. Spatial distribution of mean 2000-2015 values for AOT40, POD0 and POD1 based on modeled input data from EMEP for DO_3SE and CLRTAP (2017) parameterization (Figure 8, category 5).

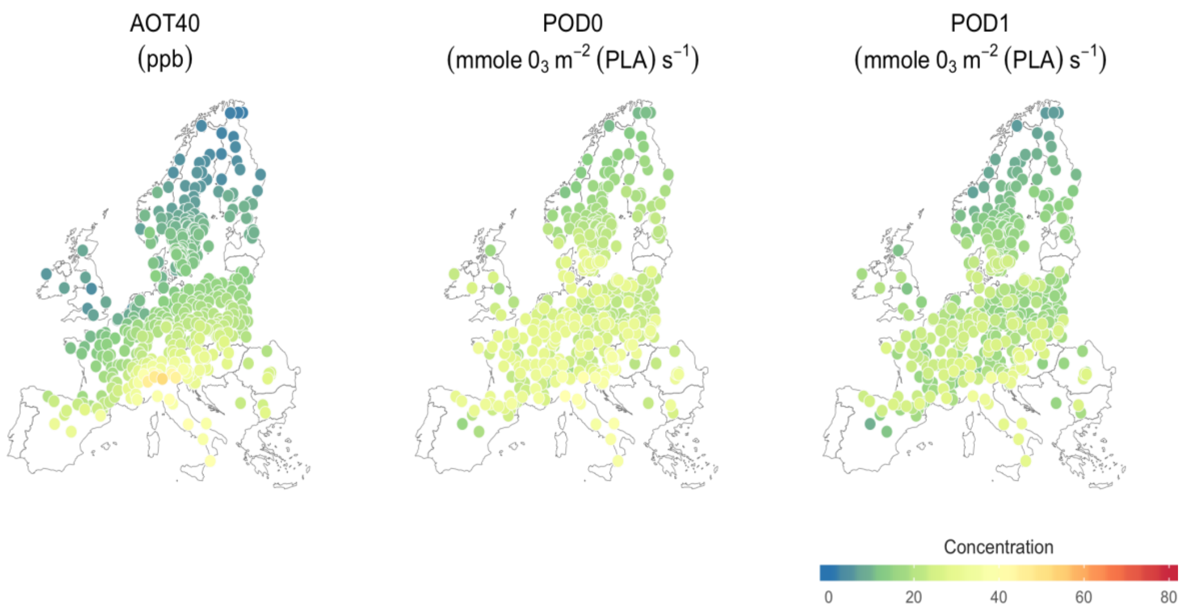

Figure 16. Spatial distribution of mean 2000-2015 values for AOT40, POD0 and POD1 based on modeled output data from EMEP and based on the parameterization from Simpson *et al*. (2012) (Figure 8, category 6).

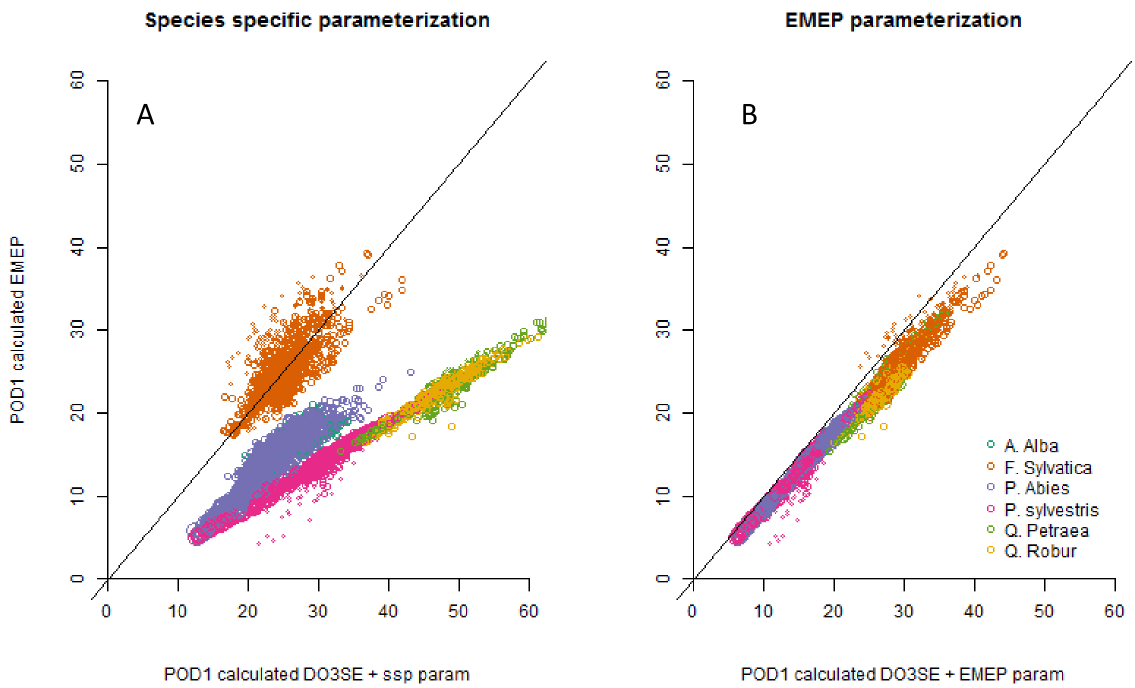

Figure 17. Comparison of POD1 based on species specific parameterization according to the Mapping Manual (CLRATP 2017) and based on parameterizations for coniferous and deciduous forests as used in the EMEP model according to Simpson *et al*. (2012). The y-axis shows POD1 as calculated in category 6, i.e. summing up the $FstO_3$ modelled by EMEP. The x-axis shows POD1 as calculated in category 5, i.e. applying DO_3SE with input variables from EMEP. In Figure A, the DO_3SE parameterization according to CLRATP (2017) and in Figure B the DO_3SE parameterization for CF (coniferous; *A. alba*, *P. abies* and *P. sylvestris*) and DF (deciduous; *F. sylvatica*, *Q. petraea* and *Q. robur*) according to Simpson *et al*. (2012) was used. Differences between both parameterizations are described in Table 2.

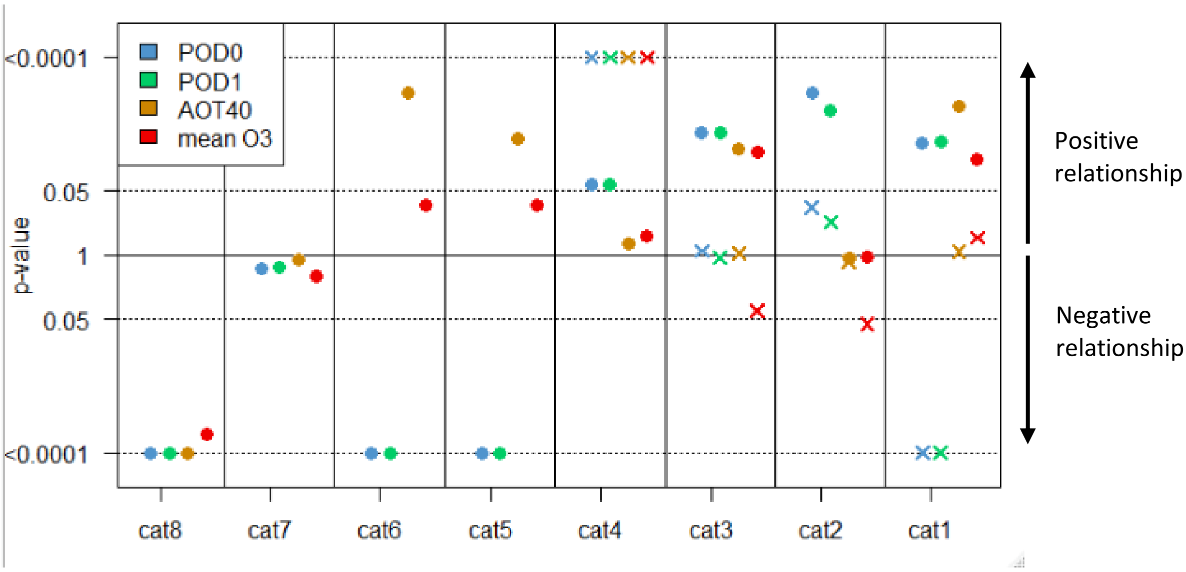

Figure 18. Sign (positive/negative) and significance (p-values in a log-scale) of the relationship between various ozone metrics (mean ozone concentration; AOT40; POD0 and POD1) and individual tree *BAI* of *P. abies* estimated from the mixed-effects models according to various samplings (categories) and sources of input data. *Category 4 to 1*: here, the ozone metrics were calculated using both EMEP data and EMEP model (*crosses*) or using measured data and the DO_3SE model (full dots) at the plots where ozone and meteorological measurements were available (36, 14, 5, and 9 plots for the categories 4 to 1). Note that the relationships calculated between mean $[O_3]$ and tree BAI calculated for the categories 6 and 5 are identical as both use ozone data from EMEP. In that case, the AOT40-based relationship may differ as the phenological module differs.

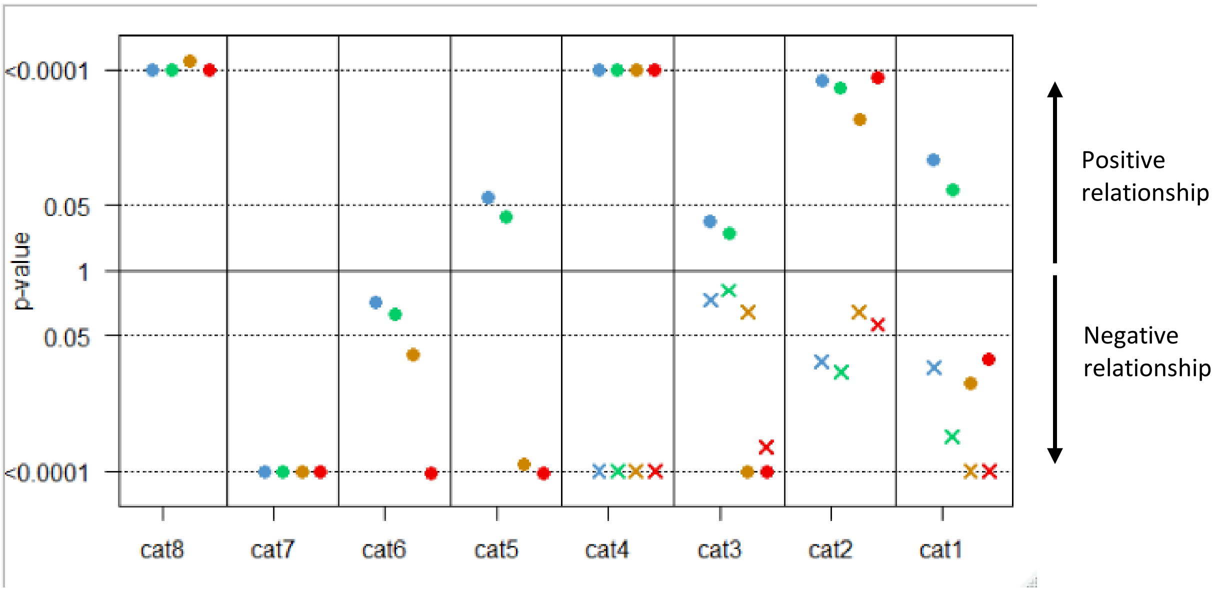

Figure 19. Sign (positive/negative) and significance (p-values in a log-scale) of the relationship between various ozone metrics (mean ozone concentration; AOT40; POD0 and POD1) and individual tree BAI of *F. sylvatica* estimated from the mixed-effects models according to various samplings and sources of input data. The number of plots considered in the categories 6 to 1 are 150, 150, 35, 13, 4, and 10.

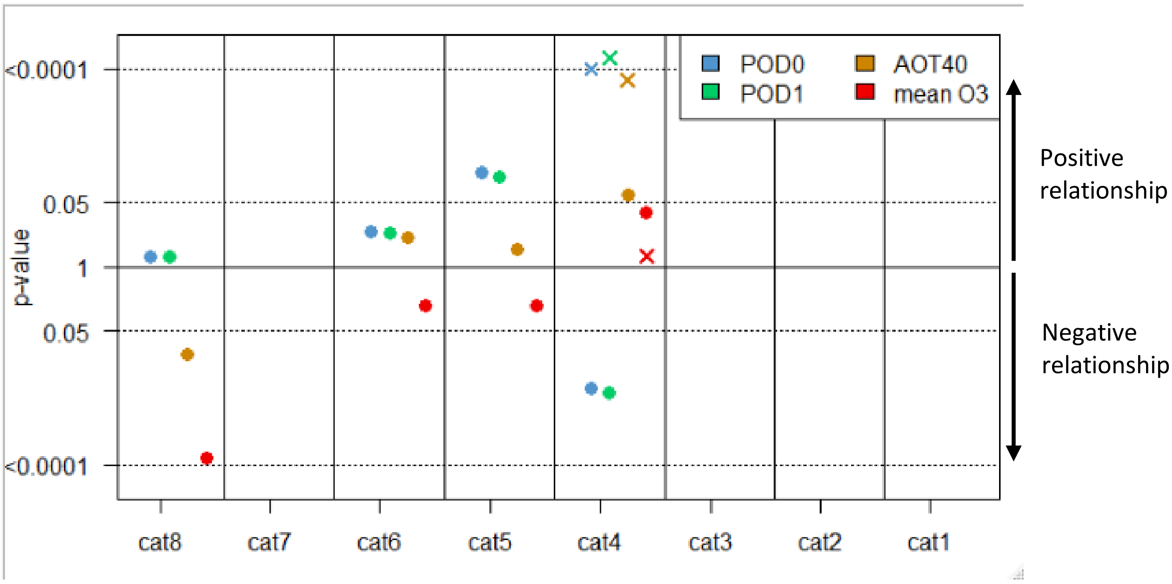

Figure 20. Sign (positive/negative) and significance (p-values in a log-scale) of the relationship between various ozone metrics (mean ozone concentration; AOT40; POD0 and POD1) and individual tree *BAI* of *Quercus sp.* estimated from the mixed-effects models according to various samplings and sources of input data. The number of plots considered in the categories 6 to 4 are 117, 117, and 14. Note that none relationship was shown for the categories 1,2, 3, and 7 because of a too low number of plots (lower than 3) or due to the lack of convergence of the statistical models.

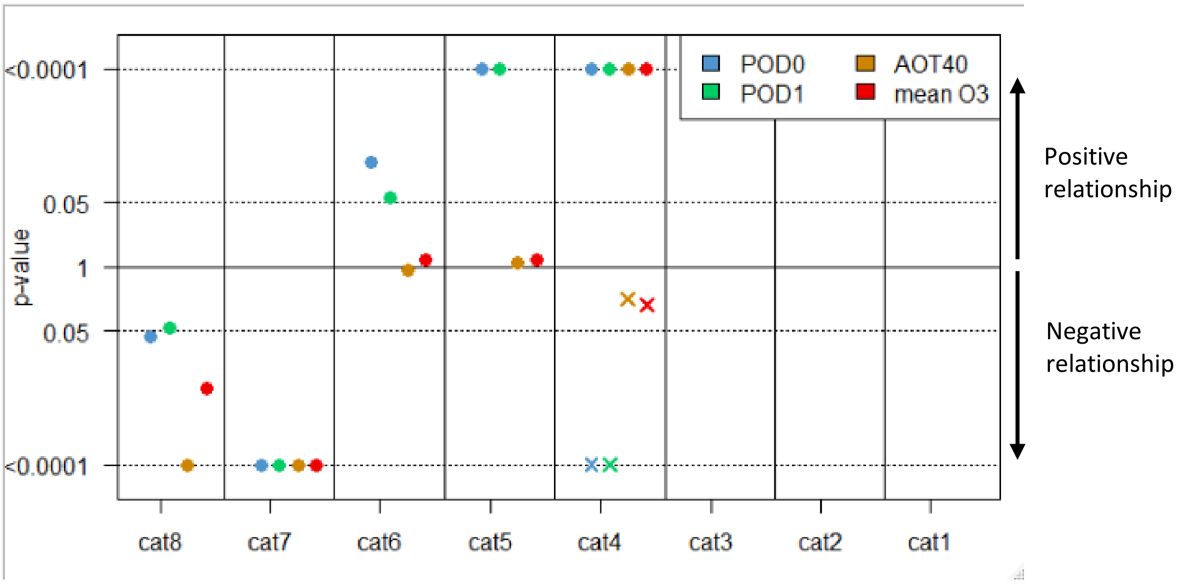

Figure 21. Sign (positive/negative) and significance (p-values in a log-scale) of the relationship between various ozone metrics (mean ozone concentration; AOT40; POD0 and POD1) and individual tree *BAI* of *P. sylvestris* estimated from the mixed-effects models according to various samplings and sources of input data. The number of plots considered in the categories 6 to 4 are 151, 151, and 17. Note that none relationship was shown for the categories 1 to 3 because of a too low number of plots (lower than 3) or due to the lack of convergence of the statistical models.

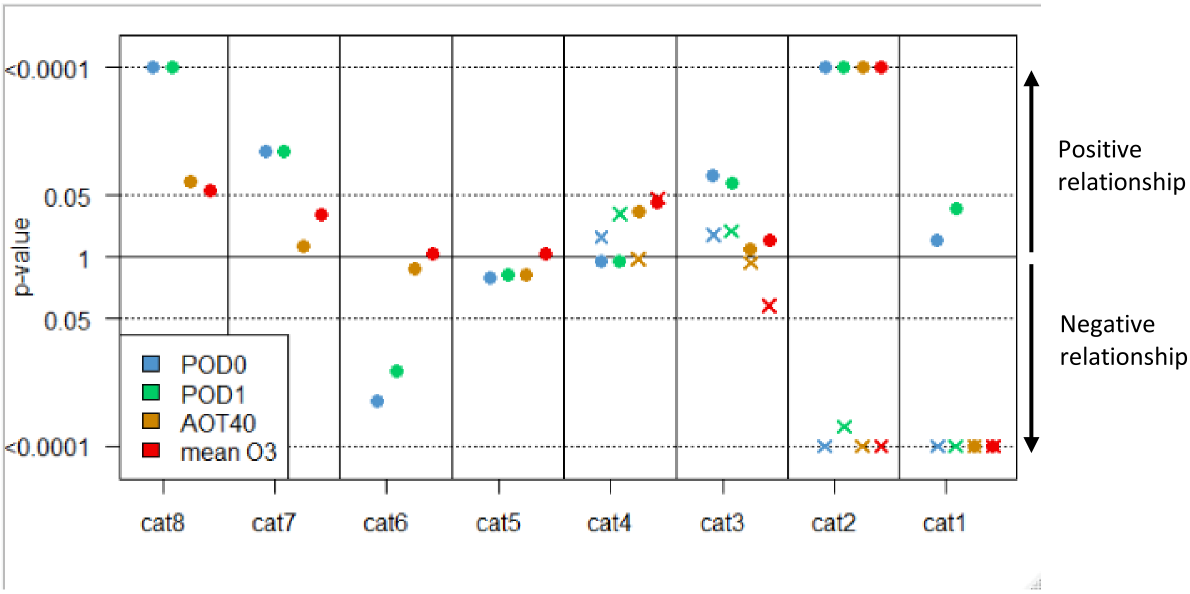

Figure 22. Sign (positive/negative) and significance (p-values in a log-scale) of the relationship between various ozone metrics (mean ozone concentration; AOT40; POD0 and POD1) and individual tree *BAI* of *A. alba* estimated from the mixed-effects models according to various samplings and sources of input data. The number of plots considered in the categories 6 to 4 are 151, 151, and 17

Manuscripts and outputs

1) Cailleret M, Ferretti M, Gessler A, Rigling A, Schaub M (2018) Ozone effects on European forest growth – towards an integrative approach. *Journal of Ecology*. doi:10.1111/1365-2745.12941
2) Schaub M, Haeni M, Calatayud V, Ferretti M, Gottardini E (2018) ICP Forests. Ozone concentrations are decreasing but exposure remains high in European forests. ICP Forests Brief, 3. 6 p. doi:10.3220/ICP1525258743000
3) Yu K, Smith WK, Trugman AT, Condit R, Hubbell SP, Sardans J, Peng C, Zhu K, Peñuelas J, Cailleret M[13], Levanic T, Gessler A, Schaub M, Ferretti M, Anderegg WRL (2019) Pervasive decreases in living vegetation carbon turnover time across forest climate zones. *PNAS*. doi: 10.1073/pnas.1821387116
4) Schaub M, Schönbeck L, Cailleret M, ICP Forests Experts, Haeni M (in prep) Ozone data paper. *Nature Data*.
5) Cailleret M, Haeni M, Schönbeck L, Schaub M (in prep) DO_3SE uncertainty analysis.
6) Cailleret M, Haeni M, Calatayud V, Schaub M (in prep) Refining approaches for calculating stomatal ozone fluxes from passive samplers.
7) Schaub M, Haeni M, Schönbeck L, Trotsiuk V, Cailleret M (in prep) PRO3FILE synthesis paper.
8) Ferretti M, Haeni M, Cailleret M, Gottardini E, Schaub M (in prep) The fingerprint of tropospheric ozone on the vegetation in Europe.

Presentations

1) Cailleret M, Haeni M, Bugmann H, Büker P, Emberson L, Rigling A, Simpson D, Solly E, Waldner P, Gessler A, Schaub M (2017) Long-term impacts of ozone and climate change on forest dynamics in Switzerland. IUFRO 125th Anniversary Congress. Freiburg.
2) Cailleret M, Haeni M, Gessler A, Ferretti M, Prescher AK, Rigling A, Simpson D, Schaub M (2018) Sensitivity of mean ozone concentration, AOT40, and POD1 estimates to different sources for ozone, climate, and vegetation data. 7th ICP Forests Scientific Conference, Riga, Latvia.
3) Schaub M, Haeni M, Etzold S, Waldner P, Büker P, Emberson L, Ferretti M, Gottardini E, Rohner B, Mina M, ICP Forests Experts, Gessler A, Rigling A, Cailleret M (2018) Predicting Ozone Fluxes, Impacts and Critical Levels on European Forests (PRO3FILE). ICP Forests Combined EP, Zvolen, Slovakia.
4) Cailleret M, Haeni M, Schönbeck L, Gessler A, Schaub M (2018) Dose-response relationships between ozone fluxes and tree radial growth across Europe - preliminary results. 8th ICP Forests Scientific Conference, Ankara, Turkey.
5) Cailleret M *et al*. (2019) Introduction to discussion: ozone effects on European forest growth - towards an integrative approach. Joint ICP Forests and ICP Vegetation Expert Workshop 12 April 2019, Birmensdorf, Switzerland.
6) Schaub M, Ferretti M, Haeni M, Cailleret M, Gottardini E (2019) Ozone and foliar symptoms in Europe. Joint ICP Forests and ICP Vegetation Expert Workshop 12 April 2019, Birmensdorf, Switzerland.

[13] M. Cailleret cleaned, gap-filled and aggregated the ICP Forests growth data and provided them to Kailiang Yu for this publication.

7) Schaub M, Haeni M, Schönbeck L, Etzold S, Waldner P, Büker P, Emberson L, Ferretti M, Gottardini E, Rohner B, Mina M, ICP Forests Experts, Gessler A, Rigling A, Cailleret M (2019) Predicting Ozone Fluxes, Impacts and Critical Levels on European Forests (PRO3FILE). Joint ICP Forests and ICP Vegetation Expert Workshop 12 April 2019, Birmensdorf, Switzerland.

8) Schaub M, Cailleret M, Gottardini E, Häni M, Ferretti M (2019) Ozone concentrations and foliar symptoms on European forests. XXV IUFRO World Congress, Curitiba, Brazil.

9) Cailleret M, Haeni M, Ferretti M, Schönbeck L, Simpson D, ICP Forests Experts, Schaub M (2019) Main sources of uncertainty in large-scale assessments of ozone effects on forest growth. XXV IUFRO World Congress, Curitiba, Brazil.

10) Haeni M, Cailleret M, Raspe S, Kumar S, Gessler A, Hunziker S, Schaub M (2019) Twenty years of hourly meteorological data from the in- and outside of 200 forest plots across Europe. XXV IUFRO World Congress, Curitiba, Brazil.

References

Batjes NH (2016) Harmonised soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. *Geoderma* 269:61-68.

Büker P, Feng Z, Uddling J, Briolat A, Alonso R, Braun S, Elvira S, Gerosa G, Karlsson PE, Le Thiec D, Marzuoli R, Mills G, Oksanen E, Wieser G, Wilkinson M, Emberson LD (2015) New flux based dose–response relationships for ozone for European forest tree species. *Environmental Pollution* 206:163-174. doi:10.1016/j.envpol.2015.06.033

Calatayud V, Diéguez JJ, Sicard P, Schaub M, De Marco A (2016) Testing approaches for calculating stomatal ozone fluxes from passive samplers. *The Science of the total environment* 572, 56–67.

CLRTAP, Mills G, Harmens H, Hayes F, Pleijel H, Büker P, González I, Alonso R ,Bender J, Bergmann E, Bermejo V, Braun S, Danielsson H, Gerosa G, Grünhage L, Karlsson P, Marzuoli R, Schaub M, Simpson D (2017). MAPPING CRITICAL LEVELS FOR VEGETATION Revised Chapter 3 of the Manual on Methodologies and Criteria for Modelling and Mapping Critical Loads and Levels and Air Pollution Effects, Risks and Trends. 66 p. https://www.unece.org/fileadmin/DAM/env/documents/2017/AIR/EMEP/Final__new_Chapter_3_v2__August_2017_.pdf

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars AC, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J, Park B, Peubey C, de Rosnay P, Tavolato C, Thépaut Jand Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.* 137: 553-597. doi:10.1002/qj.828

de Vries W, Posch M, Simpson D, Reinds, GJ (2017) Modelling long-term impacts of changes in climate, nitrogen deposition and ozone exposure on carbon sequestration of European forest ecosystems. *Science of the Total Environment* 605:1097-1116.

Gerosa G, Ferretti M, Bussotti F, Rocchini D (2007) Estimates of ozone AOT40 from passive sampling in forest sites in South-Western Europe. *Environmental Pollution* 145:629-635.

Huld T, Šúri m, Dunlop E, Micale F (2006) Estimating average daytime and daily temperature profiles within Europe. *Environmental Modelling & Software* 21:1650-1661. 10.1016/j.envsoft.2005.07.010.

Klingberg J, Karlsson PE, Pihl Karlsson G, Hu Y, Chen D, Pleijel H (2012) Variation in ozone exposure in the landscape of southern Sweden with consideration of topography and coastal climate. *Atmospheric Environment* 47:252-260.

Krupa S, Nosal M, Ferdinand JA, Stevenson RE, Skelly JM (2003) A multi-variate statistical model integrating passive sampler and meteorology data to predict the frequency distributions of hourly ambient ozone (O_3) concentrations. *Environmental Pollution*, 124:173-178.

Loibl W, Winiwarter W, Kopsca A, Zufger J, Baumann R (1994) Estimating the spatial distribution of ozone concentrations in complex terrain. *Atmospheric Environment*, 28:2557-2566.

Luedeling E (2019) chillR: Statistical methods for phenology analysis in temperate fruit trees. R package version 0.70.17 https://CRAN.R-prject.org/package=chillR.

Murray FW (1967) On the computation of saturation vapor pressure. *Journal of Applied Meteorology* 6:203-204.

Monteith JL, Unsworth MH (1990) Principles of environmental physics. Arnold.

Paoletti E, De Marco A, Anav A, Gasparini P, Pompei E (2018) Five-year volume growth of European beech does not respond to ozone pollution in Italy. *Environmental Science and Pollution Research* 25:8233-8239.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rohner B, Waldner P, Lischke H, Ferretti M, Thürig E (2018) Predicting individual-tree growth of central European tree species as a function of site, stand, management, nutrient, and climate effects. *European Journal of Forest Research* 137:29-44. doi: 10.1007/s10342-017-1087-7.

Schumacher S (2004). The role of large-scale disturbances and climate for the dynamics of forested landscapes in the European Alps (Doctoral dissertation, ETH Zurich).

Simpson D, Benedictow A, Berge H, Bergström R, Emberson LD, Fagerli H, Flechard CR, Hayman GD, Gauss M, Jonson JE, *et al.* (2012) The EMEP MSC-W chemical transport model - technical description. *Atmospheric Chemistry and Physics* 12:7825–7865.

Thürig E, Kaufmann E, Frisullo R, Bugmann H (2005) Evaluation of the growth function of an empirical forest scenario model. *Forest ecology and management* 204:53-68.

Appendix

Appendix 1. Overview of the required 2000-2015 input variables and parameters with the respective model applications and data sources, where ADC stands for 'additional data from the countries'.

No	Variables and Parameters	Application	Source I	Source II
	Atmospheric variables			
1	Temperature (*Ts_C*)	DO_3SE, MS	ICPF Meteo & ADC	Meteotest, EMEP
2	Vapor Pressure Deficit (*VPD*) [1]	DO_3SE	ICPF Meteo & ADC	
3	Measured wind speed (*uh_zR*)	DO_3SE	ICPF Meteo & ADC	Meteotest, EMEP
4	Precipitation (*precip*)	DO_3SE, MS	ICPF Meteo & ADC	Meteotest, EMEP
5	Pressure (*P*)	DO_3SE	ICPF Meteo & ADC	Meteotest, EMEP
6	Photosynthetic active radiation (*PAR*)	DO_3SE, MS	ICPF Meteo & ADC	Meteotest, EMEP
7	Global radiation (*R*)	DO_3SE, MS	ICPF Meteo & ADC	Meteotest, EMEP
8	CO2 concentration	DO_3SE, MS		www.co2.earth/monthly-co2
9	Nitrogen deposition	MS	ICPF Deposition	Meteotest, EMEP
10	Measured O3 density (*O3_zR*) - active	DO_3SE, MS	ICPF Ozone & ADC	Meteotest, EMEP, ECMWF, NCEI
11	Measured O3 density (*O3_zR*) - passive	DO_3SE, MS	ICPF Ozone	
12	Phytotoxic O3 dose above y (PODY)	MS		DO3SE, Meteotest, EMEP
13	AOT40	MS	ICPF Ozone & ADC	Meteotest, EMEP
	Measurement data			
14	O_3 measurement height	DO_3SE	ICPF Ozone & ADC	
15	O_3 measurement canopy height	DO_3SE	ICPF Ozone & ADC	
16	Wind speed measurem. height	DO_3SE	ICPF Ozone & ADC	
17	Wind speed measurem. canopy height	DO_3SE	ICPF Ozone & ADC	
	Site properties			
18	Latitude	DO_3SE	ICPF plot infos	Swiss NFI
19	Longitude	DO_3SE	ICPF plot infos	Swiss NFI
20	Elevation	DO_3SE, MS	ICPF plot infos	Swiss NFI
21	Topography (slope and aspect)	MS	ICPF plot infos	Swiss NFI
22	Soil Water Content	MS	ICPF Soil	
23	Soil texture	DO_3SE	ICPF Soil	CH Bodeneignungskarte
24	Soil pH	MS	ICPF Soil	
25	Management type		ICPF plot infos	
	Tree & vegetation characteristics			
26	Tree diameter (*DBH*)	MS	ICPF Growth	Swiss NFI
27	Tree height	DO_3SE, MS	ICPF Growth	Swiss NFI
28	Species identity	DO_3SE, MS	ICPF Growth	Swiss NFI
29	Crown transparency	MS	ICPF Crown	
30	Tree vitality	MS	ICPF Growth	Swiss NFI
31	Leaf Area Index (*LAI*)	DO_3SE	ICPF Pheno	DBH-LAI allometric relationships
32	Leaf nutrients concentration	MS	ICPF Foliage	
33	Root depth (root)	DO_3SE	ICPF Soil	
34	Start of growing season (*SGS*)	DO_3SE	ICPF Pheno	Pheno functions
35	End of growing season (*EGS*)	DO_3SE	ICPF Pheno	Pheno functions

[1]) Calculated from air temperature and relative humidity.

Appendix 2. Script for data validation executed by the ICP Forests DB manager.

```
-- +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
MEH resubmission

-- +++++++++++++++
- the file "fmd_mm_meh.csv" contains all resubmissions

-- +++++++++++++++
- range test executed on the file are described here:
		https://icp-forests.org/documentation/Tests/MM.html#MEH
		warnings have been marked with "q_flag" = 10
		errors have been marked with "q_flag" = 20
		All results have been checked:
		SELECT * FROM fmd_mm_meh WHERE code_country = 6 AND code_variable = 'AT' AND q_flag = 10 ; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 6 AND code_variable = 'RH' AND q_flag = 10 ; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 14 AND code_variable = 'WS' AND q_flag = 10 ; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 66 AND code_variable = 'AT' AND q_flag = 10 ; -- hourly min incorrect
		SELECT * FROM fmd_mm_meh WHERE code_country = 66 AND code_variable = 'ST' AND q_flag = 10 ; -- some hourly min/max values incorrect
		SELECT * FROM fmd_mm_meh WHERE code_country = 66 AND code_variable = 'SR' AND q_flag = 10 ; -- some hourly mean values incorrect
		SELECT * FROM fmd_mm_meh WHERE code_country = 1 AND code_variable = 'SR' AND q_flag = 10 ; -- high values
		SELECT * FROM fmd_mm_meh WHERE code_country = 14 AND code_variable = 'AT' AND q_flag = 10 ; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 1 AND code_variable = 'AT' AND q_flag = 10 ; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 55 AND code_variable = 'AT' AND q_flag = 10 ; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 55 AND code_variable = 'RH' AND q_flag = 10 ; -- low values / some negative - incorrect
		SELECT * FROM fmd_mm_meh WHERE code_country = 4 AND code_variable = 'ST' AND q_flag = 10  LIMIT 1000; -- (very) high values
		SELECT * FROM fmd_mm_meh WHERE code_country = 14 AND code_variable = 'RH' AND q_flag = 10  LIMIT 1000; -- OK (low values)
		SELECT * FROM fmd_mm_meh WHERE code_country = 66 AND code_variable = 'RH' AND q_flag = 10  LIMIT 1000; -- OK (low values)
		SELECT * FROM fmd_mm_meh WHERE code_country = 2 AND code_variable = 'AT' AND q_flag = 10; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 2 AND code_variable = 'WS' AND q_flag = 10; -- OK (most values = 0 / some values very high)
		SELECT * FROM fmd_mm_meh WHERE code_country = 1 AND code_variable = 'RH' AND q_flag = 10  LIMIT 1000; -- OK (no min/max values - mean very low)
		SELECT * FROM fmd_mm_meh WHERE code_country = 4 AND code_variable = 'RH' AND q_flag = 10  LIMIT 1000; -- OK (low values)
		SELECT * FROM fmd_mm_meh WHERE code_country = 4 AND code_variable = 'AT' AND q_flag = 10  LIMIT 1000; -- OK (low values)
		SELECT * FROM fmd_mm_meh WHERE code_country = 2 AND code_variable = 'AP' AND q_flag = 10  LIMIT 1000; -- OK (low values)
		SELECT * FROM fmd_mm_meh WHERE code_country = 55 AND code_variable = 'AP' AND q_flag = 10  LIMIT 1000; -- very low values (around 950)
		SELECT * FROM fmd_mm_meh WHERE code_country = 2 AND code_variable = 'SR' AND q_flag = 10  LIMIT 1000; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 14 AND code_variable = 'SR' AND q_flag = 10  LIMIT 1000; -- OK
		SELECT * FROM fmd_mm_meh WHERE code_country = 4 AND code_variable = 'WS' AND q_flag = 10  LIMIT 1000; -- min values might be max value??
		SELECT * FROM fmd_mm_meh WHERE code_country = 4 AND code_variable = 'SR' AND q_flag = 10  LIMIT 1000; -- some high and low values / some negative values -
incorrect
		SELECT * FROM fmd_mm_meh WHERE code_country = 63 AND code_variable = 'SR' AND q_flag = 10 ; -- most values negative - incorrect (25 rows)
		SELECT * FROM fmd_mm_meh WHERE code_country = 63 AND code_variable = 'AT' AND q_flag = 10 ; -- some values very low or high / difference between min and
max not realistic in some cases
		SELECT * FROM fmd_mm_meh WHERE code_country = 63 AND code_variable = 'RH' AND q_flag = 10  LIMIT 1000; -- very low values (some negative)
		-- Germany Saarland - has been submitted late
		SELECT * FROM fmd_mm_meh WHERE partner_code = 3504 AND code_variable = 'SR' AND q_flag = 10 ; -- OK (high value)
		SELECT * FROM fmd_mm_meh WHERE partner_code = 3504 AND code_variable = 'AT' AND q_flag = 10 ; -- OK (low values)
		SELECT * FROM fmd_mm_meh WHERE partner_code = 3504 AND code_variable = 'RH' AND q_flag = 10 ; -- Low values < 20%

-- +++++++++++++++
- the file submissions.csv contains an overview of all re-submissions
		included in the file "fmd_mm_meh.csv"

		for submissions "plm_new" = 1
			the PLM has been submitted

		for "plm_old" = 1
			PLM and MEM valid for re-submitted MEH were already in the database

		for "plm_exist" = 1
			PLM from other years exist in the database

		for "reconstruct" = 1
			the PLM could be potentially created from former PLM files
			"vertical_position", "instrument" and "location" are unique in former PLM files

-- +++++++++++++++
- the file submission_days.csv includes a list of measured days per variable on a plot (completeness overview)
		! It has not been tested if all hours have been measured per day !

-- +++++++++++++++
- the files in the directory "database" contain a complete download
		of the Meteo Survey from the ICP Forests database
		(without the specific resubmissions from our call)

-- +++++++++++++++
Anomalies and Plausibility in in re-submission

- some re-submitted data includes already data for 2018
		"survey_year" = 2018 :
			SURVEY_year  | code_country          | code_plot
			________________________________________
			2018                 | 6                                      | 716
			2018                 | 66                           | 101
			2018                 | 66                           | 102
			2018                 | 6                                      | 920

- several cases have been submitted with 25 hourly records per day and site
			survey_year    | code_country          | code_plot    | instrument_seq_nr | date_observation    | code_variable          | hours per day
			_______________________________________________________________________________________________________________________
			2011                 | 2                                      | 11                                     | 37
				| "2011-12-31"                  | "WD"                        | "25"
			2011                 | 2                                      | 11                                     | 38
				| "2011-12-31"                  | "WS"                        | "25"
```

2011	2	11	39	"2011-12-31"	"AT"	"25"
2011	2	11	40	"2011-12-31"	"RH"	"25"
2011	2	11	41	"2011-12-31"	"AP"	"25"
2011	2	11	42	"2011-12-31"	"PR"	"25"
2011	2	11	43	"2011-12-31"	"SR"	"25"

- there are plausibility issues in the re-submission

2574769 improbable cases were identified being "hourly_mean" = "hourly_max" AND "hourly_min" <> "hourly_max"

1819465 improbable cases were identified being "hourly_mean" = "hourly_min" AND "hourly_min" <> "hourly_max"

following countries and plots are concerned:

code_country	code_plot
1	3
1	6
1	12
1	16
1	17
1	36
1	37
1	41
1	44
1	59
1	63
1	69
1	74
1	75
1	80
1	84
1	86
1	90
1	93
1	98
1	100
2	11
2	21
4	101
4	301
4	302
4	303
4	304
4	305
4	307
4	308
4	603
4	605
4	607
4	609
4	611
4	613
4	901
4	902
4	903
4	904
4	905
4	906
4	907
4	908
4	909
4	910
4	911
4	912
4	913
4	914
4	915
4	916
4	917
4	918
4	919
4	920
4	921
4	922
4	923
4	1201
4	1202
4	1203
4	1502
4	1605
4	1606
4	1607
14	9
14	16
63	1
63	2
63	3
66	101
66	102

Redundancies between database and re-submission

- several cases have been submitted with 730 daily records per year and site (values differ from each other):

survey_year	code_country	code_plot	instrument_seq_nr	hourly	code_variable	days per year
2013	55	19	1	0	"AT"	"730"
2013	55	19	1	0	"RH"	"730"
2013	55	19	1	0	"WD"	"730"
2013	55	19	1	0	"WS"	"730"
2013	55	19	1	3	"AT"	"730"
2013	55	19	1	3	"RH"	"730"
2013	55	19	1	3	"WD"	"730"
2013	55	19	1	3	"WS"	"730"
2013	55	19	1	6	"AT"	"730"
2013	55	19	1	6	"RH"	"730"
2013	55	19	1	6	"WD"	"730"
2013	55	19	1	6	"WS"	"730"
2013	55	19	1	9	"AT"	"730"
2013	55	19	1	9	"RH"	"730"
2013	55	19	1	9	"WD"	"730"
2013	55	19	1	9	"WS"	"730"
2013	55	19	1	12	"AT"	"730"
2013	55	19	1	12	"RH"	"730"
2013	55	19	1	12	"WD"	"730"
2013	55	19	1	12	"WS"	"730"
2013	55	19	1	15	"AT"	"730"
2013	55	19	1	15	"RH"	"730"
2013	55	19	1	15	"WD"	"730"
2013	55	19	1	15	"WS"	"730"
2013	55	19	1	18	"AT"	"730"
2013	55	19	1	18	"RH"	"730"
2013	55	19	1	18	"WD"	"730"
2013	55	19	1	18	"WS"	"730"
2013	55	19	1	21	"AT"	"730"
2013	55	19	1	21	"RH"	"730"
2013	55	19	1	21	"WD"	"730"
2013	55	19	1	21	"WS"	"730"

- there are possibly some redundancies between the database and the re-submission data

799867 cases for the join on the identifiers: "survey_year", "code_country", "code_plot", "instrument_seq_nr", "code_variable", "date_observation", "hourly":

```
PSEUDOCODE
SELECT survey_year, code_country, code_plot, instrument_seq_nr, code_variable, date_observation, hourly, hourly_mean FROM RESUBMISSION.fmd_mm_meh
intersect
SELECT survey_year, code_country, code_plot, instrument_seq_nr, code_variable, date_observation, hourly, hourly_mean FROM DATABASE.fmd_mm_meh;
```

survey_year	partner_code
2011	102
2012	3504
2013	3504
2014	3504
2015	2904

326807 cases with identical values:

```
PSEUDOCODE
SELECT (attributes as above) + ("hourly_mean", "hourly_min", "hourly_max", "hourly_completeness") FROM RESUBMISSION.fmd_mm_meh
intersect
SELECT (attributes as above) + ("hourly_mean", "hourly_min", "hourly_max", "hourly_completeness") FROM DATABASE.fmd_mm_meh
```

survey_year	partner_code
2011	102
2015	2904

-- +++++++++++++++

Additional Issues

- Austria (14) no data for 2002 re-submitted
- France (1) 1994 not complete
- France(1) differing number of datasets per year (between 31 and 65 thousand)
- Cyprus (66) differing number of datasets per year(15: 29256 / 16: 89613 / 17: 168513 / 18: 90721)
- Flanders(102) 2011 less datasets than other years
- Germany Thuringia (3704): data measured in 30 minutes intervals have been aggregated to hourly data by PCC
- Germany NW-FVA (3004): "meh files do not include data of following plots, which have already been provided to the project: 604, 606, 608, 610, 1501

Appendix 3: Script for running DO$_3$SE in Linux

```
require(insol, jsonlite, RcppCNPy, RJSONIO)

#######
# (1) Methods and static parameters ---------------------------------------------------------------------------------------
#######

I_accumulation_period_method = "growing season"        # Period of O3 uptake
I_Astart = NULL; I_Aend = NULL

I_LAI_method = "estimate total"        # "input" (Input all LAI values); "input total" (Input total LAI, split according to fLAI); "estimate total" (Estimate total LAI from primary land cover, split according to fLAI)
I_SAI_method = "estimate total"        # "input" (Input all SAI values); "input total" (Input total SAI, split according to fLAI); "estimate total" (Estimate total SAI from primary land cover, split according to fLAI)
I_fLAI = 1.0                           # Distribution of LAI/SAI in canopy (default: uniform distribution)

I_SAI_methodseason = "forest"          # "LAI" (SAI = LAI); "forest" (Forest method: SAI = LAI + 1); "wheat"
I_gmorph = 1.0                         # Sun/shade morphology factor
I_method="multiplicative"              # stomatal condutance method: "multiplicative" (Jarvis) or "photosynthesis" (hybrid multiplicative/Farquhar-based)

I_f_phen_method = "simple day PLF"     # "disabled" (f_phen supplied or =1); "simple day PLF" (single hump: 3 values, 2 slopes);  "complex day PLF" (double hump: 5 values, 4 slopes);
I_f_phen_b = 1.0; I_f_phen_c = 1.0; I_f_phen_d = 1.0
I_f_phen_2 = 200; I_f_phen_3 = 200;    # used in "complex day PLF"
I_f_phen_limA = 0; I_f_phen_limB = 0;  # start and end of soil water limitation

I_leaf_f_phen_method = "f_phen"        # "disabled" (leaf_f_phen supplied or =1); "f_phen" (leaf_f_phen=f_phen); "day PLF" (single hump between Astart and Aend)
I_leaf_f_phen_a = 0.0; I_leaf_f_phen_b = 1.0; I_leaf_f_phen_c = 0.0; I_leaf_f_phen_1 = 15; I_leaf_f_phen_2 = 30

I_f_light_method = "enabled"           # "enabled" (f_light and leaf_f_light calculated)
I_cosA = 0.5                           # A = mean leaf inclination (0.5 = 60 degrees)

I_f_temp_method = "default"            # "default" (Normal bell-shaped function over T_min -> T_opt -> T_max); "square high" (similar but with straight lines)

I_f_VPD_method = "linear"              # "linear" (between VPD_max and VPD_min); "log": Simple, unparameterised, logarithmic relationship

I_fSWP_exp_curve = "temperate"         # "custom" (fSWP_exp_a and fSWP_exp_b supplied); "temperate" (a = 0.355, b = -0.706); "mediterranean" (a = 0.619, b = -1.024)
I_fSWP_exp_a = NULL; I_fSWP_exp_b = NULL

I_f_O3_method = "disabled"             # "disabled" (f_O3 supplied or =1); "wheat" or "potato" (specific methods)
I_VPD_crit=1000                        # Critical daily VPD threshold above which stomatal conductance will stop increasing (kPa).

# Parameters of the photosynthesis-based stomatal conducance model
I_g_sto_0=30000                        # Closed stomatal conductance (umol m-2 s-1)
I_m=16.83                              # Species-specific sensitivity to An (dimensionless)
I_V_cmax_25=30; I_J_max_25=60          # Vcmax and Jmax
I_Tleaf_method="EB"                    # "input" (Tleaf_C supplied); "ambient" (Use ambient air temperature); "Nikolov"; EB"; "de Boeck"
I_O3_method="martin2000"               # "disabled"; "martin 2000"
I_K_z=24                               # Coefficient for ozone damage (dimensionless)
I_F_0=37                               # Threshold flux for ozone damage (nmol O3 m-2 PLA s-1)

I_Y=1                                  # POD threshold (nmol O3 m-2 s-1)

I_albedo=0.12
I_Rsoil=100                            # soil resistance (s/m)
I_OTC=FALSE                            # Is open top chamber experiment?

I_soil_texture="loam"; I_b=NULL; I_FC=NULL; I_SWP_AE=NULL; I_Ksat=NULL
I_root=12; I_source="P-M"

I_CO2_method="constant"                # "constant" (Use CO2_constant value); "input" (CO2 concentration supplied)
I_CO2_constant=391                     # Constant CO2 concentration (ppm)
I_PWP=-4.0; I_ASW_FC=NULL; I_initial_SWC=NULL


#######
# (2) Phenological module ---------------------------------------------------------------------------------------------------
#######

# Options in the current version of DO3SE: "constant" (provided); all year" (1-365); "forest latitude" (from EMEP)
# Species-specific phenological models for estimating SGS and EGS -> only for deciduous species; Oak vs. beech

phenofunction=function(species, method, lat, long, Temp)     # T (Daily mean temperature) <- vector[365]
{

 I_growing_season_method=I_SGS=I_EGS=NULL

 if (species %in% c("PAbi", "PSyl","AAlb"))
 {
   I_growing_season_method="all year"
 }

 if (species %in% c("FSyl", "QRob"))
 {
   if (method=="forest latitude") {I_growing_season_method = "forest latitude"}
   if (method=="constant")
   {
     I_growing_season_method = "constant"

     Ssen=0  # Colouring state for Senescence [arbitrary unit]
     Rf=0    # Date when the accumulation of forcing starts
```

```
    if (species == "FSyl")
    {
     # Flushing from Vitasse et al. 2011; Sigmoid fucntion (lower bias)
     t0=49     # 48.1 in the paper
     alpha=-0.2; beta=-4.5
     Fstar=59.8
     # Senescence from Delpierre et al. 2009 AFM
     Pstart=12.5
     Tb=25
     y=2; x=2
     Ycrit=5160
    }

    if (species == "QRob")
    {
     # Flushing from Vitasse et al. 2011; Sigmoid fucntion (lower bias)
     t0=52     # 51.4 in the paper
     alpha=-0.29; beta=14.16
     Fstar=12.2
     # Senescence from Delpierre et al. 2009 AFM
     Pstart=14.5
     Tb=26.5
     y=0; x=2
     Ycrit=10178
    }

    # Flushing
    for (d in t0:200) # day of the year
    {
     Rf = Rf + (1 / (1+exp(alpha*(Temp[d]-beta))))
     if (Rf >= Fstar) {tf=d; break}
    }

    # Senescence
    for (d in 200:365)
    {
     P<-as.numeric(daylength(lat, long, d, 1)[,3])

     fPd=P/Pstart
     if (P<Pstart)
     {
      if (Temp[d]<Tb)  {Rsen=((Tb-Temp[d])^x) * (fPd^y)}
      if (Temp[d]>=Tb) {Rsen=0}
      Ssen=Ssen+Rsen
     }
     if (P>Pstart)
     {
      Ssen=0
     }
     if (Ssen>=Ycrit) {Y90=d; break}
    }

    I_SGS=tf
    I_EGS=Y90
  }
 } # End condition deciduous

 return(c(I_growing_season_method, I_SGS, I_EGS))
}
```

#######
 # (3) Generates Configuration file --
#######

```
conf <- function(site,growing_season_method,SGS,EGS,accumulation_period_method,Astart,Aend,
            LAI_a,LAI_b,LAI_c,LAI_d,LAI_1,LAI_2,
            SAI_methodseason,fmin,gmax,gmorph,method,f_phen_method,
            f_phen_limA,f_phen_limB,f_phen_a,f_phen_b,
            f_phen_c,f_phen_d,f_phen_e,f_phen_1,
            f_phen_2,f_phen_3,f_phen_4,leaf_f_phen_method,
            leaf_f_phen_a,leaf_f_phen_b,leaf_f_phen_c,leaf_f_phen_1,
            leaf_f_phen_2,f_light_method,f_lightfac,
            cosA,f_temp_method,T_min,T_opt,T_max,
            f_VPD_method,VPD_min,VPD_max,
            f_SW_method,SWP_min,SWP_max,
            fSWP_exp_curve,fSWP_exp_a,fSWP_exp_b,
            f_O3_method,VPD_crit,g_sto_0,m,
            V_cmax_25,J_max_25,Tleaf_method,O3_method,
            K_z,F_0,height,Lm,Y,height_method,
            LAI_method,SAI_method,fLAI,
            lat,long,elev,albedo,Rsoil,OTC,z_u,
            z_O3,h_u,h_O3,CO2_method,CO2_constant,
            soil_texture,b,FC,SWP_AE,Ksat,root,
            PWP,ASW_FC,source,initial_SWC)
 {
 foo <- RJSONIO::fromJSON("data/configs/config.json")
 head(foo)

 foo$LC_config$LCs[[1]]$name <- ifelse(is.null(site),'null',site)
 foo$LC_config$LCs[[1]]$season$growing_season_method <- ifelse(is.null(growing_season_method),'null',growing_season_method)
 foo$LC_config$LCs[[1]]$season$SGS <- ifelse(is.null(SGS),'null',SGS)
 foo$LC_config$LCs[[1]]$season$EGS <- ifelse(is.null(EGS),'null',EGS)
 foo$LC_config$LCs[[1]]$season$accumulation_period_method <- ifelse(is.null(accumulation_period_method),'null',accumulation_period_method)
 foo$LC_config$LCs[[1]]$season$Astart <- ifelse(is.null(Astart),'null',Astart)
 foo$LC_config$LCs[[1]]$season$Aend <- ifelse(is.null(Aend),'null',Aend)
 foo$LC_config$LCs[[1]]$season$LAI_a <- ifelse(is.null(LAI_a),'null',LAI_a)
```

```
foo$LC_config$LCs[[1]]$season$LAI_b <- ifelse(is.null(LAI_b),'null',LAI_b)
foo$LC_config$LCs[[1]]$season$LAI_c <- ifelse(is.null(LAI_c),'null',LAI_c)
foo$LC_config$LCs[[1]]$season$LAI_d <- ifelse(is.null(LAI_d),'null',LAI_d)
foo$LC_config$LCs[[1]]$season$LAI_1 <- ifelse(is.null(LAI_1),'null',LAI_1)
foo$LC_config$LCs[[1]]$season$LAI_2 <- ifelse(is.null(LAI_2),'null',LAI_2)
foo$LC_config$LCs[[1]]$season$SAI_method <- ifelse(is.null(SAI_method),'null',SAI_methodseason)
foo$LC_config$LCs[[1]]$gsto$fmin <- ifelse(is.null(fmin),'null',fmin)
foo$LC_config$LCs[[1]]$gsto$gmax <- ifelse(is.null(gmax),'null',gmax)
foo$LC_config$LCs[[1]]$gsto$gmorph <- ifelse(is.null(gmorph),'null',gmorph)
foo$LC_config$LCs[[1]]$gsto$method <- ifelse(is.null(method),'null',method)
foo$LC_config$LCs[[1]]$gsto$f_phen_method <- ifelse(is.null(f_phen_method),'null',f_phen_method)
foo$LC_config$LCs[[1]]$gsto$f_phen_limA <- ifelse(is.null(f_phen_limA),'null',f_phen_limA)
foo$LC_config$LCs[[1]]$gsto$f_phen_limB <- ifelse(is.null(f_phen_limB),'null',f_phen_limB)
foo$LC_config$LCs[[1]]$gsto$f_phen_a <- ifelse(is.null(f_phen_a),'null',f_phen_a)
foo$LC_config$LCs[[1]]$gsto$f_phen_b <- ifelse(is.null(f_phen_b),'null',f_phen_b)
foo$LC_config$LCs[[1]]$gsto$f_phen_c <- ifelse(is.null(f_phen_c),'null',f_phen_c)
foo$LC_config$LCs[[1]]$gsto$f_phen_d <- ifelse(is.null(f_phen_d),'null',f_phen_d)
foo$LC_config$LCs[[1]]$gsto$f_phen_e <- ifelse(is.null(f_phen_e),'null',f_phen_e)
foo$LC_config$LCs[[1]]$gsto$f_phen_1 <- ifelse(is.null(f_phen_1),'null',f_phen_1)
foo$LC_config$LCs[[1]]$gsto$f_phen_2 <- ifelse(is.null(f_phen_2),'null',f_phen_2)
foo$LC_config$LCs[[1]]$gsto$f_phen_3 <- ifelse(is.null(f_phen_3),'null',f_phen_3)
foo$LC_config$LCs[[1]]$gsto$f_phen_4 <- ifelse(is.null(f_phen_4),'null',f_phen_4)
foo$LC_config$LCs[[1]]$gsto$leaf_f_phen_method <- ifelse(is.null(leaf_f_phen_method),'null',leaf_f_phen_method)
foo$LC_config$LCs[[1]]$gsto$leaf_f_phen_a <- ifelse(is.null(leaf_f_phen_a),'null',leaf_f_phen_a)
foo$LC_config$LCs[[1]]$gsto$leaf_f_phen_b <- ifelse(is.null(leaf_f_phen_b),'null',leaf_f_phen_b)
foo$LC_config$LCs[[1]]$gsto$leaf_f_phen_c <- ifelse(is.null(leaf_f_phen_c),'null',leaf_f_phen_c)
foo$LC_config$LCs[[1]]$gsto$leaf_f_phen_1 <- ifelse(is.null(leaf_f_phen_1),'null',leaf_f_phen_1)
foo$LC_config$LCs[[1]]$gsto$leaf_f_phen_2 <- ifelse(is.null(leaf_f_phen_2),'null',leaf_f_phen_2)
foo$LC_config$LCs[[1]]$gsto$f_light_method <- ifelse(is.null(f_light_method),'null',f_light_method)
foo$LC_config$LCs[[1]]$gsto$f_lightfac <- ifelse(is.null(f_lightfac),'null',f_lightfac)
foo$LC_config$LCs[[1]]$gsto$cosA <- ifelse(is.null(cosA),'null',cosA)
foo$LC_config$LCs[[1]]$gsto$f_temp_method <- ifelse(is.null(f_temp_method),'null',f_temp_method)
foo$LC_config$LCs[[1]]$gsto$T_min <- ifelse(is.null(T_min),'null',T_min)
foo$LC_config$LCs[[1]]$gsto$T_opt <- ifelse(is.null(T_opt),'null',T_opt)
foo$LC_config$LCs[[1]]$gsto$T_max <- ifelse(is.null(T_max),'null',T_max)
foo$LC_config$LCs[[1]]$gsto$f_VPD_method <- ifelse(is.null(f_VPD_method),'null',f_VPD_method)
foo$LC_config$LCs[[1]]$gsto$VPD_min <- ifelse(is.null(VPD_min),'null',VPD_min)
foo$LC_config$LCs[[1]]$gsto$VPD_max <- ifelse(is.null(VPD_max),'null',VPD_max)
foo$LC_config$LCs[[1]]$gsto$f_SW_method <- ifelse(is.null(f_SW_method),'null',f_SW_method)
foo$LC_config$LCs[[1]]$gsto$SWP_min <- ifelse(is.null(SWP_min),'null',SWP_min)
foo$LC_config$LCs[[1]]$gsto$SWP_max <- ifelse(is.null(SWP_max),'null',SWP_max)
foo$LC_config$LCs[[1]]$gsto$fSWP_exp_curve <- ifelse(is.null(fSWP_exp_curve),'null',fSWP_exp_curve)
foo$LC_config$LCs[[1]]$gsto$fSWP_exp_a <- ifelse(is.null(fSWP_exp_a),'null',fSWP_exp_a)
foo$LC_config$LCs[[1]]$gsto$fSWP_exp_b <- ifelse(is.null(fSWP_exp_b),'null',fSWP_exp_b)
foo$LC_config$LCs[[1]]$gsto$f_O3_method <- ifelse(is.null(f_O3_method),'null',f_O3_method)
foo$LC_config$LCs[[1]]$gsto$VPD_crit <- ifelse(is.null(VPD_crit),'null',VPD_crit)
foo$LC_config$LCs[[1]]$pn_gsto$g_sto_0 <- ifelse(is.null(g_sto_0),'null',g_sto_0)
foo$LC_config$LCs[[1]]$pn_gsto$m <- ifelse(is.null(m),'null',m)
foo$LC_config$LCs[[1]]$pn_gsto$V_cmax_25 <- ifelse(is.null(V_cmax_25),'null',V_cmax_25)
foo$LC_config$LCs[[1]]$pn_gsto$J_max_25 <- ifelse(is.null(J_max_25),'null',J_max_25)
foo$LC_config$LCs[[1]]$pn_gsto$Tleaf_method <- ifelse(is.null(Tleaf_method),'null',Tleaf_method)
foo$LC_config$LCs[[1]]$pn_gsto$O3_method <- ifelse(is.null(O3_method),'null',O3_method)
foo$LC_config$LCs[[1]]$pn_gsto$K_z <- ifelse(is.null(K_z),'null',K_z)
foo$LC_config$LCs[[1]]$pn_gsto$F_0 <- ifelse(is.null(F_0),'null',F_0)
foo$LC_config$LCs[[1]]$height <- ifelse(is.null(height),'null',height)
foo$LC_config$LCs[[1]]$Lm <- ifelse(is.null(Lm),'null',Lm)
foo$LC_config$LCs[[1]]$Y <- ifelse(is.null(Y),'null',Y)
foo$LC_config$height_method <- ifelse(is.null(height_method),'null',height_method)
foo$LC_config$LAI_method <- ifelse(is.null(LAI_method),'null',LAI_method)
foo$LC_config$SAI_method <- ifelse(is.null(SAI_method),'null',SAI_method)
foo$location$albedo <- ifelse(is.null(albedo),'null',albedo)
foo$location$Rsoil <- ifelse(is.null(Rsoil),'null',Rsoil)
foo$location$OTC <- ifelse(is.null(OTC),'null',OTC)
foo$location$z_u <- ifelse(is.null(z_u),'null',z_u)
foo$location$z_O3 <- ifelse(is.null(z_O3),'null',z_O3)
foo$location$h_u <- ifelse(is.null(h_u),'null',h_u)
foo$location$h_O3 <- ifelse(is.null(h_O3),'null',h_O3)
foo$met_config$CO2_method <- ifelse(is.null(CO2_method),'null',CO2_method)
foo$met_config$CO2_constant <- ifelse(is.null(CO2_constant),'null',CO2_constant)
foo$SMD_config$soil_texture <- ifelse(is.null(soil_texture),'null',soil_texture)
foo$SMD_config$soil$b <- ifelse(is.null(b),'null',b)
foo$SMD_config$soil$FC <- ifelse(is.null(FC),'null',FC)
foo$SMD_config$soil$SWP_AE <- ifelse(is.null(SWP_AE),'null',SWP_AE)
foo$SMD_config$soil$Ksat <- ifelse(is.null(Ksat),'null',Ksat)
foo$SMD_config$root <- ifelse(is.null(root),'null',root)
foo$SMD_config$PWP <- ifelse(is.null(PWP),'null',PWP)
foo$SMD_config$ASW_FC <- ifelse(is.null(ASW_FC),'null',ASW_FC)
foo$SMD_config$source <- ifelse(is.null(source),'null',source)
foo$SMD_config$initial_SWC <- ifelse(is.null(initial_SWC),'null',initial_SWC)

foo$location$lat <- lat
foo$location$long <- long
foo$location$elev <- elev

foo$LC_config$fLAI[[1]] <- ifelse(is.null(fLAI),'null',paste0('fffff',fLAI,'ddddd'))

ex <- RJSONIO::toJSON(foo,pretty=TRUE)
ex <- gsub('"null"',"null",ex)
ex <- gsub("\\t","    ",ex)
ex <- gsub('" : ','": ',ex)
ex <- gsub(',\n',', \n',ex)
ex <- gsub('"fffff',"[",ex)
ex <- gsub('ddddd"',"]",ex)

ss <- randomstring(n=1,length=25)
ss <- paste0("data/configs/",ss,".json")
write(ex,file=ss)
```

```
  return(ss)
}

#######
# (4) DO3SE run ---------------------------------------------------------------------------------------------------
#######

proc <- function(co,site,year) {
  if(length(co)==0) {stop("Error, please provide a config file!")}
  if(length(site)==0) {stop("Error, please provide a site!")}
  if(length(year)==0) {stop("Error, please provide a year!")}

  file <- paste0("data/meteo/",site,"-",year,".csv")
  outcsv <- paste0("output/",site,"-",year,"out.csv")
  if(!file.exists(file)) {stop("Error, the file does not exist!")}

  out <- paste0("output/",site,"-",year,"out.npy")
  system(paste0("python -m pyDO3SE.cli run_csv ",co," ",file," --output-format=npy>",out),wait=TRUE)
  file.remove(co)

  system(paste0("python DO3SE-model/pyDO3SE/utils.py ",out," ",outcsv))
  bar <- read.csv(outcsv,na.strings=c(-999))[,-1]
  file.remove(out)
  return(bar)
}

#######
# (5) Main overall function ------------------------------------------------------------------------------------
#######

runDO3SE=function(site,year,param,LAI,maxH,species,phenomethod,lat,long,elev)
{

# Parameterisation -> EMEP (modelled) vs measured data

if (param=="measured")
 {
   I_height_method="constant"          # "input" (Input canopy height); "constant" (Use constant height from primary land cover)
   I_z_u=4.6                           # Measurement height for windspeed (m)-> from LWF (or 25, from earlier function-script)
   I_z_O3=2                            # Measurement height for O3 (m)
   I_h_u=0.5                           # Canopy height for windspeed measurement, default = target canopy (m)
   I_h_O3=0.5                          # Canopy height for O3 measurement, default = target canopy (m)
   I_f_SW_method="fSWP exp"            # "disabled": (f_SW supplied or =1); "fSWP exp" (Use fSWP exponential curve); "fSWP linear" (Use linear fSWP
                                       function between SWP_min and SWP_max); "fLWP exp" (Use fSWP exponential curve, but with LWP instead of SWP);
                                       "fPAW" (Use fPAW relationship)
 }

if (param=="EMEP")
 {
   I_height_method="constant"        # "input" (Input canopy height); "constant" (Use constant height from primary land cover)
   I_z_u=20                  # Measurement height for windspeed (m)
   I_z_O3=20                  # Measurement height for O3 (m)
   I_h_u=20               # Canopy height for windspeed measurement, default = target canopy (m)
   I_h_O3=20              # Canopy height for O3 measurement, default = target canopy (m)
   I_f_SW_method="fSWP exp"        # "disabled": (f_SW supplied or =1); "fSWP exp" (Use fSWP exponential curve); "fSWP linear" (Use linear fSWP function between SWP_min and
SWP_max); "fLWP exp" (Use fSWP exponential curve, but with LWP instead of SWP); "fPAW" (Use fPAW relationship)
 }

# Species-specific parameters-------------------------------------------------------------------------------------------

if (species=="FSyl") #& param=="measured")
 {
   I_fmin=0.13               # Minimum stomatal conductance (fraction)
   I_gmax=155                # Maximum stomatal conductance (mmol O3 m-2 s-1)
   I_f_lightfac=0.006        # Coefficient factor between PPFD and the factor responsible of the reduction in st. conductance beacuase of low light availability
   I_T_min=5; I_T_opt=16; I_T_max=33
   I_VPD_min=3.1; I_VPD_max=1                # (kPa)
   I_Lm=0.07                         # leaf dimension (m)
   I_SWP_min=-1.25; I_SWP_max=-0.05          # SWP (MPa) for minimum and maximum gsto
   I_f_phen_a=0; I_f_phen_e=0.4
   I_f_phen_1=20; I_f_phen_4=20              # used in "simple day PLF"
 }

if (species=="QRob")
 {
   I_fmin=0.13               # Minimum stomatal conductance (fraction)
   I_gmax=265                # Maximum stomatal conductance (mmol O3 m-2 s-1)
   I_f_lightfac=0.006        # Coefficient factor between PPFD and the factor responsible of the reduction in st. conductance beacuase of low light availability
   I_T_min=0; I_T_opt=22; I_T_max=35
   I_VPD_min=3.1; I_VPD_max=1.1              # (kPa)
   I_Lm=0.042                        # leaf dimension (m)
   I_SWP_min=-2; I_SWP_max=-1          # SWP (MPa) for minimum and maximum gsto (mediterranean broadleaf)
   I_f_phen_a=0.3; I_f_phen_e=0.3      # was zero, I made it to 0.3 as in the MM (LS)
   I_f_phen_1=15; I_f_phen_4=20              # used in "simple day PLF"
 }

if (species=="QuPe")  ## same as Q robur?
 {
   I_fmin=0.13               # Minimum stomatal conductance (fraction)
   I_gmax=265                # Maximum stomatal conductance (mmol O3 m-2 s-1)
   I_f_lightfac=0.006        # Coefficient factor between PPFD and the factor responsible of the reduction in st. conductance beacuase of low light availability
   I_T_min=0; I_T_opt=22; I_T_max=35
   I_VPD_min=3.1; I_VPD_max=1.1              # (kPa)
   I_Lm=0.042                        # leaf dimension (m)
   I_SWP_min=-2; I_SWP_max=-1          # SWP (MPa) for minimum and maximum gsto (mediterranean broadleaf)
   I_f_phen_a=0.3; I_f_phen_e=0.3      # was zero, I made it to 0.3 as in the MM (LS)
   I_f_phen_1=15; I_f_phen_4=20              # used in "simple day PLF"
```

```
 }

if (species=="PAbi")
 {
  I_fmin=0.16                  # Minimum stomatal conductance (fraction)
  I_gmax=130                   # Maximum stomatal conductance (mmol O3 m-2 s-1)
  I_f_lightfac=0.01            # Coefficient factor between PPFD and the factor responsible of the reduction in st. conductance beacuase of low light availability
  I_T_min=0; I_T_opt=14; I_T_max=35
  I_VPD_min=3; I_VPD_max=0.5                   # (kPa)
  I_Lm=0.008                            # leaf dimension (m)
  I_SWP_min=-0.5; I_SWP_max=-0.05            # SWP (MPa) for minimum and maximum gsto
  I_f_phen_a=0; I_f_phen_e=0.0
  I_f_phen_1=0; I_f_phen_4=0                # used in "simple day PLF"
 }

if (species=="AbAl")
 {
  I_fmin=0.16                  # Minimum stomatal conductance (fraction)
  I_gmax=130                   # Maximum stomatal conductance (mmol O3 m-2 s-1)
  I_f_lightfac=0.01            # Coefficient factor between PPFD and the factor responsible of the reduction in st. conductance beacuase of low light availability
  I_T_min=0; I_T_opt=14; I_T_max=35
  I_VPD_min=3; I_VPD_max=0.5                   # (kPa)
  I_Lm=0.008                            # leaf dimension (m)
  I_SWP_min=-0.5; I_SWP_max=-0.05            # SWP (MPa) for minimum and maximum gsto
  I_f_phen_a=0; I_f_phen_e=0.0
  I_f_phen_1=0; I_f_phen_4=0                # used in "simple day PLF"
 }

if (species=="PSyl") #Bueker 2015
 {
  I_fmin=0.1                   # Minimum stomatal conductance (fraction)
  I_gmax=180                   # Maximum stomatal conductance (mmol O3 m-2 s-1)
  I_f_lightfac=0.006           # Coefficient factor between PPFD and the factor responsible of the reduction in st. conductance beacuase of low light availability
  I_T_min=0; I_T_opt=20; I_T_max=36
  I_VPD_min=2.8; I_VPD_max=0.6                  # (kPa)
  I_Lm=0.008                            # leaf dimension (m)
  I_SWP_min=-1.5; I_SWP_max=-0.7     # SWP (MPa) for minimum and maximum gsto
  I_f_phen_a=0.8; I_f_phen_e=0.8
  I_f_phen_1=40; I_f_phen_4=40                # used in "simple day PLF"
 }

# Phenological module ----------------------------------------------------------------------------------------------------------

clim=read.csv(paste0("data/meteo/",site,"-",year,".csv"),sep=",")
Temp=aggregate(clim,by=list(clim$dd), FUN=mean)$Ts_C
if (length(Temp)<300) {Temp=c(rep(4,91),Temp,rep(4,91))} # if climate data at the beginning and end of the year are missing
pheno=phenofunction(species, phenomethod, lat, long, Temp)

# Main run DO3SE function -----------------------------------------------------------------------------------------------------

co <- conf(site=site,
        growing_season_method=pheno[1],SGS=as.numeric(pheno[2]),EGS=as.numeric(pheno[3]),
        accumulation_period_method=I_accumulation_period_method,Astart=I_Astart,Aend=I_Aend,
        height_method=I_height_method,
        LAI_method=I_LAI_method,SAI_method=I_SAI_method,fLAI=I_fLAI,
        SAI_methodseason=I_SAI_methodseason,
        fmin=I_fmin,gmax=I_gmax,gmorph=I_gmorph,
        method=I_method,
        f_phen_method=I_f_phen_method,
        f_phen_limA=I_f_phen_limA, f_phen_limB=I_f_phen_limB,
        f_phen_a=I_f_phen_a, f_phen_b=I_f_phen_b, f_phen_c=I_f_phen_c, f_phen_d=I_f_phen_d, f_phen_e=I_f_phen_e,
        f_phen_1=I_f_phen_1, f_phen_2=I_f_phen_2, f_phen_3=I_f_phen_3, f_phen_4=I_f_phen_4,
        leaf_f_phen_method=I_leaf_f_phen_method,
        leaf_f_phen_a=I_leaf_f_phen_a, leaf_f_phen_b=I_leaf_f_phen_b, leaf_f_phen_c=I_leaf_f_phen_c,
        leaf_f_phen_1=I_leaf_f_phen_1, leaf_f_phen_2=I_leaf_f_phen_2,
        f_light_method=I_f_light_method, f_lightfac=I_f_lightfac, cosA=I_cosA,
        f_temp_method=I_f_temp_method, T_min=I_T_min, T_opt=I_T_opt, T_max=I_T_max,
        f_VPD_method=I_f_VPD_method, VPD_min=I_VPD_min, VPD_max=I_VPD_max,
        f_SW_method=I_f_SW_method,
        SWP_min=I_SWP_min, SWP_max=I_SWP_max,
        fSWP_exp_curve=I_fSWP_exp_curve, fSWP_exp_a=I_fSWP_exp_a, fSWP_exp_b=I_fSWP_exp_b,
        f_O3_method=I_f_O3_method, VPD_crit=I_VPD_crit, g_sto_0=I_g_sto_0, m=I_m,
        V_cmax_25=I_V_cmax_25, J_max_25=I_J_max_25,
        Tleaf_method=I_Tleaf_method, O3_method=I_O3_method,
        K_z=I_K_z, F_0=I_F_0, Lm=I_Lm,
        Y=I_Y,
        albedo=I_albedo, Rsoil=I_Rsoil,
        OTC=I_OTC, z_u=I_z_u, z_O3=I_z_O3,
        soil_texture=I_soil_texture, b=I_b, FC=I_FC, SWP_AE=I_SWP_AE, Ksat=I_Ksat,
        root=I_root,
        source=I_source,
        h_u=I_h_u, h_O3=I_h_O3,
        CO2_method=I_CO2_method, CO2_constant=I_CO2_constant,
        PWP=I_PWP, ASW_FC=I_ASW_FC, initial_SWC=I_initial_SWC,

      # Site-dependent variables that have to be read in another input file
         lat=lat,
         long=long,
         elev=elev,
         height=maxH,
         LAI_a=0.0,LAI_b=LAI,LAI_c=LAI,LAI_d=0.0,LAI_1=30,LAI_2=30) #LAI_1: 15 or 30?

 # do the main processing
 foo <- proc(co,site, year)

 POD0=tail(foo$POD_0,1)
 PODy= tail(foo$POD_y,1)
```

```
  AOT40= tail(foo$AOT_40,1)

  return(c(POD0,PODy,AOT40))
}


#######
# (6) Minor functions ----------------------------------------------------------------------------------------------
#######
randomstring <- function(n=1,length=12) {
  randomString <- c(1:n)
  for(i in 1:n) {
    randomString[i] <- paste(sample(c(0:9, letters, LETTERS),
                             length, replace=TRUE),collapse="")
  }
  return(randomString)}
```