High temperature sensitivity of monoterpene emissions from global vegetation

Efstratios Bourtsoukidis (e.bourtsoukidis@cyi.ac.cy)
The Cyprus Institute https://orcid.org/0000-0001-5578-9414

Andrea Pozzer (andrea.pozzer@mpic.de)
Max Planck Institute for Chemistry https://orcid.org/0000-0003-2440-6104

Jonathan Williams (jonathan.williams@mpic.de)
Max Planck Institut für Chemie https://orcid.org/0000-0001-9421-1703

David Makowski (david.makowski@inrae.fr)
INRAE - University of Paris-Saclay https://orcid.org/0000-0001-6385-3703

Josep Penuelas (josep.penuelas@uab.cat)
CSIC, Global Ecology Unit CREAF-CSIC-UAB, Cerdanyola del Vallès 08193, Catalonia, Spain
https://orcid.org/0000-0002-7215-0150

Vaseilios Matthaios (v.matthaios@bham.ac.uk)
University of Birmingham

Georgia Lazoglou (g.lazoglou@cyi.ac.cy)
The Cyprus Institute

Ana Yañez-Serrano (ana.yanez@idaea.csic.es)
Institute of Environmental Assessment and Water Research (IDAEA)

Philippe Ciais (philippe.ciais@lsce.ipsl.fr)
Laboratoire des Sciences du Climat et de l'Environnement https://orcid.org/0000-0001-8560-4943

Jos Lelieveld (Jos.Lelieveld@mpic.de)
Max Planck Institute for Chemistry https://orcid.org/0000-0001-6307-3846

Mihalis Vrekoussis (mvrekous@uni-bremen.de)
University of Bremen https://orcid.org/0000-0001-8292-8352

Nikolaos Daskalakis (daskalakis@uni-bremen.de)
University of Bremen

Jean Sciare (j.sciare@cyi.ac.cy)

Article

Keywords:

DOI: https://doi.org/

Additional Declarations: (Not answered)

High temperature sensitivity of monoterpene emissions from global vegetation

E. Bourtsoukidis[1,2,*], A. Pozzer[2], J. Williams[2,1], D. Makowski[3], J. Peñuelas[4,5], V. N. Matthaios[6,7], G. Lazoglou[1], A. M. Yañez-Serrano[8], J. Lelieveld[2,1], P. Ciais[9,1], M. Vrekoussis[10,1,11], N. Daskalakis[10],and Jean Sciare[1]

[1]Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute; Nicosia, Cyprus.
[2]Max Planck Institute for Chemistry (MPIC); Mainz, Germany.
[3]French National Institute for Agriculture, Food, and Environment (INRAE), Unversity Paris-Saclay, AgroParisTech (UMR MIA); Paris, France.
[4]Center for Research Ecology and Forestry Applications (CREAF); Bellaterra, Barcelona, Spain.
[5]Global Ecology Unit CREAF-CSIC-UAB; Bellaterra, Barcelona, Spain.
[6]School of Geography, Earth and Environmental Sciences, University of Birmingham; Birmingham, UK.
[7]Department of Environmental Health, Harvard T.H. Chan School of Public Health; Boston, MA, USA.
[8]Institute of Environmental Assessment and Water Research (IDAEA); Barcelona, Spain.
[9]Laboratoire des Sciences du Climat et de l'Environnement (LSCE); Gif-sur-Yvette, France.
[10]Institute of Environmental Physics (IUP), University of Bremen; Bremen, Germany.
[11]Center of Marine Environmental Sciences (MARUM), University of Bremen; Bremen, Germany.

Correspondence: e.bourtsoukidis@mpic.de

Abstract

Terrestrial vegetation emits vast amounts of monoterpenes into the atmosphere, influencing ecological interactions and atmospheric chemistry. Global emissions, mostly driven by responses to abiotic stress, are simulated as a function of temperature with a fixed exponential relationship (β coefficient) across forest ecosystems and environmental conditions. We applied meta-analysis algorithms on all published monoterpene emission data and show that this relationship is more intricate and sensitive than previously thought. We find that co-occurring environmental stresses amplify the temperature sensitivity of the emissions that is primarily related to the specific plant functional type (PFT). On average, warmer ecosystems appear more sensitive, indicating the adjustment of plants in response to thermal stress. Implementing a PFT-dependent β in a biogenic emission model, coupled with a chemistry – climate model, demonstrated that atmospheric processes are exceptionally dependent on monoterpene emissions which are subject to amplified variations under rising temperatures.

Introduction

Plants emit large amounts of monoterpenes (MT; $C_{10}H_{16}$) in response to abiotic drivers and biotic interactions, with significant feedbacks to the ecophysiological and biogeochemical processes of terrestrial ecosystems[1–3]. Once released, these Volatile Organic Compound (VOC) species affect atmospheric chemistry and physics due to their high reactivity and propensity to form particles, thereby influencing air quality, atmospheric oxidative capacity and radiative balance-related climate feedbacks[4,5].

Considered as the 'thermometer' of plants[6], monoterpene emissions from global vegetation are fundamentally simulated with an empirical approach that considers temperature as a key regulator of MT emissions by global vegetation[7]:

$$E_{MT} = E_{30}\exp(\beta(T - T_{30}) \quad [1]$$

Here, E_{30} is the emission potential under standardized conditions (30 °C), and β is an empirical coefficient (°C^{-1}) derived from the fit of the regression between measured emissions and temperature (T). This equation remains the best approximation for MTs stored in leaf pools, but an additional light-dependent function, similar to that used for isoprene, must be applied for instantaneously produced MTs[7] (Supplementary Text).

To date, a fixed dependence on temperature (β=0.10 °C^{-1}) is used for all types of vegetation. The purpose of this meta-analysis is to conduct a detailed evaluation of this parameter by collecting and analyzing all experimentally derived β coefficients published (see Supplementary Information references) since the first observation of this relationship in 1980[8]. Moreover, we aim to assess potential refinements of β and investigate the resulting impact on atmospheric chemistry and physics by performing coupled simulations between the most established emission model for biogenic trace gases (MEGAN)[7] and an atmospheric chemistry – climate model (EMAC)[9].

Results

Meta-analysis of monoterpene observations

Screening of all indexed research articles and data sets published during 1980-2020 yielded a collection of 696 β coefficients of regression fits obtained under diverse conditions. The reported values varied widely, ranging from -0.07 to 2.39 °C^{-1}. To identify the factors that explain this variability, we collected an additional 35 parameters that described the experimental procedure, location, time, vegetation state (i.e. the plant functional type and tree characteristics) and the characteristics of MT emission (Supplementary Table 1).

Almost all observations were in the Northern Hemisphere, and the majority of the studies were conducted under field conditions, using branch enclosures and offline chemical analyses for quantifying VOCs using adsorbent tubes (Supplementary Figs. 1-2). The derivation of the rate of emission (enclosures vs eddy covariance) and the techniques of sampling and analyzing VOCs did not differ β amongst the studies (offline vs. online). Attempts to establish clear relationships with the potential drivers were hindered by the complexity of inter-related parameters and in some cases the lack of reported information for each parameter.

The clearest observed pattern was the increase in β with the coefficient of determination (R^2) from the regression fits between emission and temperature. Our meta-analysis, based on a standard random-effect statistical model[10], determined the values of β within 95% confidence intervals, across R^2 bins (Fig. 1). We identified a relationship between the quality of regression fit and β for MT emissions that had larger values as the quality of the fit increased. The majority of the data (ca. 52%) had an R^2 >0.6, and the reported dependencies on temperature were significantly higher than the established value of 0.10 °C^{-1}. This observation indicated a clear underestimation of a globally uniform β. Interestingly, β obtained from field data was systematically higher than β obtained under controlled conditions, suggesting amplified sensitivities to temperature under natural conditions, when environmental stresses co-occur.

A low R^2 indicates a weak to non-existent relationship between MT emissions and temperature, so we chose to disregard the lowest 10% of the reported β (56 values), which represented R^2<0.2. This approach increased the quality of our data set while retaining a substantial number of observations for further statistical analyses. R^2 was not reported in 81 cases, so the overall data that we considered for further analysis are comprised of 559 experimental values.

We applied a machine learning analysis using the feature-importance ranking (FIR) approach, with individual parameters for vegetation, location, and time to better identify the most important parameters that influenced the β values (Fig. 2). The results indicated that the type of ecosystem, particularly the specific plant functional type (PFT; as defined in Guenther et al.[7]), was important to the derived temperature sensitivities. The geographic location (latitude and longitude) remained associated with the vegetation types that grow within the corresponding geographically distributed ecosystems. Our results nonetheless indicated that the responses of MT emissions to temperature could vary dynamically across the life span of plants. Finally, the FIR analysis did not clearly identify seasonality or a relationship between tree height and the coefficient of temperature dependency.

Revised temperature-dependent coefficients (β)

The temperature-dependent coefficients (β) were used for each PFT in MEGAN[7] (Fig. 3). Surprisingly, we did not find any β reported for agricultural ecosystems or grasslands. Almost all PFTs (other than boreal needle leaf deciduous trees and broadleaf deciduous shrubs) had values significantly larger than 0.10 $°C^{-1}$. This finding was particularly important for the widespread boreal needle leaf evergreen forests (β_{NEB} = 0.15 $°C^{-1}$; range 0.12-0.19 $°C^{-1}$) and tropical broadleaf evergreen forests (β_{BETr} = 0.20 $°C^{-1}$; range 0.14-0.3 $°C^{-1}$). Compared to needle leaf trees and broadleaf shrubs, broadleaf trees display higher dependencies on temperature in the release of MTs by their foliage. We note that removing the lowest 10% from our dataset had negligible impact on the revised values (Supplementary Fig. 3). A uniform β was calculated from all 559 data points as 0.13±0.01 $°C^{-1}$, 30% higher than the established value.

A β coefficient for each MT species was derived, taking into account the importance of processes of atmospheric oxidation due to their different rates of reaction with atmospheric radicals[11]. No significant differences were identified across all MTs, even though β was higher for Δ3-carene and trans-β-ocimene (Supplementary Fig. 4).

Summarizing the reconsiderations needed for determining β for the MTs, we concluded that the most sensible approach was to assign a different value for each PFT. Alternatively, a uniform value of 0.13 $°C^{-1}$ could be considered, and potential discrepancies for different monoterpenes may not be disregarded due to insufficient published data.

Model simulations

To comprehensively evaluate the implications of revised MT emission temperature dependencies, we performed coupled simulations employing the most widely used VOC emission model (MEGAN)[7] and a global atmospheric chemistry – climate model (EMAC)[9]. The simulations were performed at 1.8 by 1.8-degree resolution and with hourly outputs for

the original version of the model (β=0.10 °C^{-1}) and three additional scenarios for comparison. First, the MEGAN code was modified (‘‘PFT’’ simulation; Supplementary Code 1) to assign a different β for each PFT. For the PFTs that do not have any measurements, we used the value of 0.13 °C^{-1} that considers the complete dataset (Fig. 3). Second, β was uniformly set to 0.13 °C^{-1}, similar to the current approach (‘‘ALL’’ simulation). Third, β was modified based on β derived for each MT (‘‘MTRP’’ simulation). All three simulations produced significant and similar changes in the rate of emission and atmospheric feedbacks compared to the current version. We focused on the most sensible approach (considering Figs. 2, 3 and Supplementary Figs. 3-4), which was the revision of β across PFTs.

The global annual average of MT emissions was 13% lower (Supplementary Table 2). The most significant changes were found for Siberia and were particularly pronounced during winter. We attributed this large discrepancy to the lower β from broadleaf deciduous shrubs and the generally low MT emissions in this region. Average annual β was higher for tropical ecosystems, particularly for the Amazon rainforest, where MT emissions were locally up to ca. 30% higher. The simulations indicated that MT emissions could even double during the Amazonian dry season despite the similar annual averages (Supplementary Fig. 5). Investigating the daily extremes for each season indicated that MT emissions could occasionally be more than three-fold higher in both tropical and arctic ecosystems (Supplementary Fig. 6).

The most pronounced difference was the diurnal variation in the rates of emission (Supplementary Figs. 7-8). Tropical forests had particularly pronounced diel cycles that could be further amplified on seasonal and daily bases. Maximum differences in the daily standard deviation could increase by more than 300%, demonstrating the large variations in MT emissions introduced by increased dependencies on temperature. We applied Eq. 1 using data for temperature collected at the Amazonian Tall Tower Observatory (ATTO) in 2014 and 2015[12] to demonstrate the dynamics of PFT-dependent β (Supplementary Fig. 9). The increase in β (from 0.10 to 0.20 °C^{-1}) led to slightly lower average emissions. The variations in the standard deviation of the hourly simulated emissions, however, averaged twice as large and were particularly pronounced under extreme conditions such as the El Niño year of 2015.

The exponential relationship between temperature and MT emission is standardized at 30 °C, and we found that significant changes in the simulated emissions occurred at either low or very high temperatures (Fig. 5, Supplementary Fig. 10). We derived β weighted by area coverage for each PFT to illustrate the simplified dynamics of emission for the three dominant ecosystems (Supplementary Table 3). The current models of emission have clearly overestimated MT emissions under 30 °C. In contrast, the current models have underestimated the emissions for temperatures above 30 °C, helping to explain the observed discrepancies in diurnal variation of MT emission in warm environments[13,14].

The emission potential at standard conditions (E_{30}) was another important parameter that defined the rate of MT emissions. We chose to perform the simulations adopting the values in MEGAN because our analysis identified large uncertainties in the potential emissions, which increased further due to the uncertainties in the conversion of units (typically from ng C g(dry

weight)$^{-1}$ h^{-1} to μg C m^{-2} h^{-1}) and the leaf area index derived from the literature and satellite observations (Supplementary Figs. 11-12).

Atmospheric implications

The direct implications on the processes of atmospheric oxidation were evaluated by investigating the differences in concentration between model simulations. The hydroxyl radical (OH) is the most important daytime oxidant that reacts strongly with most organic molecules, particularly MTs. Our results showed that using a PFT-dependent β increases OH concentrations globally (Fig. 4). The boreal-forest belt was particularly affected, with ca. 12% higher annual average OH concentrations compared to the base simulation (β=0.10 °C^{-1}). Increased OH concentrations in NH summer were particularly pronounced (up to 20%) (Supplementary Fig. 13). Significant daytime changes in MT emissions in EMAC indicated that OH concentrations could occasionally be even as much as double compared to what is currently simulated (Supplementary Fig. 14).

In contrast to OH, average tropospheric concentrations of ozone (O_3) were only moderately affected as the maximum O_3 concentration change in the summer was within 4 - 5% for the Northern Hemisphere (Supplementary Fig. 15). Regional changes in O_3 concentration on specific days, however, could increase by up to 10% (Supplementary Fig. 16). Monoterpenes contributed both to the production (in the presence of NOx) and chemical removal of tropospheric O_3 by reaction schemes that could differ considerably between locations (e.g. VOC or NOx limited environments[15]). By reducing MT emissions in our updated emission model, O_3 loss was generally minimized, increasing its abundance globally, except for the hot and dry season of the Amazon rainforest when the high ambient MT concentrations remove O_3 at higher rates compared to the base model simulations.

Changes in the dynamics of atmospheric oxidation in forests influenced the yield of secondary organic aerosol (SOA) particles around the globe (Fig 4, Supplementary Fig. 17). The revised simulations indicated lower SOA production over boreal and temperate forests. Seasonal averages nonetheless indicated increased SOA formation over both tropical forests and oceans. Another consistent feature was the increase in SOAs in Siberia of the order of 10% considering the annual averages.

Our results thus underscore that the temperature dependency of MT emissions is a sensitive coefficient that introduces important feedbacks in atmospheric oxidation processes and reaction products. The feedbacks include an amplified seasonality and demonstrate the importance of accurately simulating the temperature effects on MT emissions.

Discussion

As global temperatures are rising, the emissions of monoterpenes from vegetation are expected to increase exponentially, adding greater amounts of these reactive species to the atmosphere. The implications range from changes in crucial ecological functions (e.g. plant defense and signaling)[16,17] to atmospheric composition and eventually radiative transfer and the properties of clouds[4]. Therefore, the accurate simulation of MT emissions is essential for evaluating plant responses and the respective feedbacks and improving future projections.

Various algorithms for simulating MT emissions have been developed accounting for either short-term volatilization (emission-based)[7,8] or long-term production of MT that are linked to photosynthesis (production-based)[18,19]. The emission-based models, such as MEGAN, have been widely applied to describe observations at the scales of leaves, branches, and canopies by simulating global MT emissions from terrestrial vegetation, accounting for multiple parameters such as temperature, light, foliar age, soil moisture concentration, and CO_2 inhibition[7]. Our meta-analysis focused on the effect of temperature and in particular, on the evaluation of the emission-response defining temperature sensitivity (β) of MT emissions.

The β coefficient for the temperature-dependent term in Eq. 1 can vary across seasons[20,21], tree species[22], tree age[23], developmental stage[24], ambient humidity[25], level of photosynthetically active radiation[26], CO_2 abundance[27], soil moisture concentration[28], and other drivers[29]. Both the FIR analysis and simplified correlation analyses on the new data collected in this study demonstrated that the vegetation type was the most important regulator of the sensitivity to temperature of MT emissions from vegetation. On average, plants that grow in warmer ecosystems appear more sensitive, indicating the adjustment of plants in response to thermal stress. Applying a PFT-dependent β helps to explain the frequently overestimated MT emissions in several environments[13,30–33].

Besides β, other parameters also contribute to the emissions of MT by terrestrial vegetation. The potential emission at standard condition (E_{30}) is an important parameter that regulates the rate of emission by forests. Our meta-analysis collected E_{30} values that vary by orders of magnitude for each ecosystem (Supplementary Figs. 11-12) revealing substantial uncertainties in the magnitude of simulated emissions. Furthermore, uncertainties in the leaf area index and the lack of studies over widespread ecosystems including agricultural and grasslands add to the unknowns and shape the direction of future research. Our results demonstrated that co-occurring environmental stresses amplify the temperature responses (β) of MT emissions. The accurate formulation of a parameterization that incorporates diverse and frequently co-occurring drivers of emission is challenging, because plants from different or even the same ecosystems have developed different strategies to cope with environmental stresses. Plants are also living organisms, so they are further subject to interspecific variability of emissions of different MTs[34].

Relating these uncertainties to atmospheric processes, we identified important implications on the mechanisms of atmospheric oxidation and consequently SOA formation. Introducing PFT-dependent β decreased MT emissions under mean temperatures below 30 °C, thereby increasing global OH concentrations. A critical review of tropospheric OH radicals identified significant and unexplained discrepancies between measured and simulated OH concentrations in unpolluted forested regions[35]. By revising MT emissions through their temperature sensitivity, our updated simulations help explain these observed discrepancies, corroborating the strong links between the emissions and atmospheric processes.

With newly discovered biogenic MT sources[12,36–38] and the challenge of modeling co-occurring environmental stresses on the biosphere, our results demonstrate the need for more process-oriented research of biosphere-atmosphere interactions, particularly in tropical, grassland and agricultural ecosystems. As climatic warming intensifies, biogenic VOC

emissions from global forests will play a crucial role in the evaluation of the health of ecosystems, and influence the atmospheric oxidation capacity with implications for the chemical composition, aerosols and climate.

Methods

Experimental design

Searching the Web of Science, selecting results from the WOS, BCIBIOSIS, CCC, DRCI, and RSCI databases, and using the keywords ‘‘monoterpenes’’, ‘‘emissions’’, and/or ‘‘temperature’’ identified 745 peer-reviewed studies published between 1980 and 2020. Screening of these articles (page by page) decreased the number of studies to 84 that reported β derived from regression fits of experimental observations. The data for tropical plants were not sufficient to perform reliable statistical analyses, so three additional studies (which reported only data) were considered and analyzed. The timelines for the data for temperature and MT emission were extracted from published plots[6,39] using a web-based tool (WebPlotDigitizer; https://automeris.io), and data were directly obtained from Yáñez-Serrano et al.[40]. We compiled a data set consisting of 696 values of β. To the best of our knowledge, we accounted for all values reported in the literature. We may, however, have missed some, mainly because they were not appropriately indexed in the literature data bases.

We investigated potential relationships with experimental, geographical, plant-specific, seasonal, and regression-fit variables by extracting all available information for 35 parameters (Supplementary Table 1), vectorising them (i.e. annotated/assigned a number to each character class), and then proceeded with our statistical analyses.

The main conclusion of each study was classified in three categories: those driven by temperature, temperature and light, and temperature and/or other environmental parameters (e.g. soil moisture concentration and relative humidity). The experimental approaches used varied in environmental conditions, determination of the rate of emission, VOC sampling, and chemical analysis (Supplementary Fig. 2). These variables were comprehensively reported in all studies, so the annotation was relatively straightforward. The β data points were annotated using the method that was used to collect the majority of the data only when two sampling methods were used. The regression fits for individual monoterpenes were annotated accordingly when measurements were performed using gas chromatography.

The year, month, and season of the experiments were assessed. The experiments were typically conducted within one year. If the experiments were conducted over two consecutive years and if the regression fits were applied for all available data, we used the year in which the majority of the observations were collected. If the experiments were conducted in the same month of two different years, then the first year was annotated in the data points. If the year of the experiments was not reported (typical for laboratory experiments), then we annotated the year prior to publication. The month of an experiment was annotated with numbers 1-12, and 13 was used for longer and/or mixed periods across years. Seasonal observations and laboratory experiments were annotated differently. The seasons were annotated accordingly in this classification.

The exact location of the field experiments included the latitude, longitude, and meters above sea level. If the authors reported a specific station instead of the exact coordinates, we used the cited literature to derive this information, but we used Google Earth and derived the coordinates if deriving this information was not possible. Information for the tree species (both Latin and common names) was collected from each publication. Considering the geographical information, we assigned each tree species to its plant functional type (PFT) as defined in the Model of Emissions of Gases and Aerosols from Nature (MEGAN)[7]. Tree age (in years), tree height (in meters), and sampling height (in meters) were rarely reported together, so several data gaps could not be filled. Finally, the minimum, maximum, mean, and range of atmospheric temperatures were filled as reported, but if not reported, the data were collected from the publication figures when possible.

These studies had different scientific objectives, so the collection and classification of all 35 parameters was a strenuous task. All parameters were cross-checked by more than one author of this study to avoid mistakes during collection.

Statistical analysis

Data were analyzed using random-effects models, which are commonly used in meta-analyses[41–43]. A first model without a covariate was fitted to estimate the global mean β across all studies. Another model with functional type as a categorical covariate was then fitted to the data set for analyzing the relationship between β and PFT. Finally, a categorical variable including categories of R^2 covering intervals of 0.1 was defined and included as a covariate in another random-effect model for studying the relationship between β and R^2. This model was fitted to the data collected under either field or controlled conditions for analyzing the sensitivity of β to the type of experiment. A random study effect was included in all fitted models to account for the heterogeneity between studies. The models were fitted using the lmer function of the lme4 R package by restricted maximum likelihood[44]. The accuracy of estimated β was assessed by calculating 95% confidence intervals. All calculations were implemented using R v4.1.2.

Unit conversions for potential emission at standardized conditions (30 °C, E_{30}).

The rates of emission of monoterpenes were typically reported in units of carbon mass per gramme of foliar dry weight per hour (ng C g(dry weight)$^{-1}$ h^{-1}). The biogenic emission factors in MEGAN are in μg m^{-2} h^{-1}, so we used the leaf area index (LAI) and the specific leaf area (SLA) for each point to convert the reported values. SLA was obtained from the TRY database[45] (https://try-db.org, last accessed 8 June 2022). The data sets used were 1-km [46] maps scaled up from trait data measured in situ. The LAI data we used were the monthly climatological data from the ORNL DAAC database[47]. The data for SLA and LAI were extracted from the above data sets, for the longitudes and latitudes of our data points while in case of missing values, the nearest neighbor with valid data was used.

Machine learning

We used machine learning methods to examine the relationship and the importance of β coefficient to the three broad categories (vegetation type, location and time variables) identified in our meta-analysis. As each of these broad categories had sub-divisions, we

examined them using the Featured Importance Ranking (FIR) approach of machine learning. Feature (or variable) importance ranking refers to a task that measures the contributions of individual input features (variables) to the performance of a supervised learning model[48]. Feature importance ranking has become one of the most powerful tools in explainable/interpretable models to facilitate understanding and discovery of key factors in a specific domain[49,50].

Specifically, we used Caret and Cubist packages in R with tuning two hyper-parameters: neighbors (#Instances) and committees (#Committees), which are the ones to most likely have the largest effect on the final performance of the Cubist model[51]. Cubist is a rule-based tree algorithm, where a tree is grown, and the terminal leaves contain regression models. These models are based on the predictors used in previous splits. In these algorithms, the prediction is made using the linear regression model at the terminal node of the tree but is smoothed by taking into account the prediction from another model in the previous node of the tree (which also occurs recursively up the tree). The tree is reduced to a set of rules, which initially are paths from the top of the tree to the bottom[52]. The performance is taken from each combination of the hyper-parameters tuning with the grid search method with cross-validation (CV)[53]. To avoid bias in data selection, we applied 10-fold CV[53,54], while the model's final performance against the test dataset was validated using R^2 and Root Mean Square Error (RMSE). We also tested an ANN and a random forest algorithm but Cubist had the smaller RMSE and the greater R^2 out of the three (Supplementary Fig. 18).

Model simulations

We used the global ECHAM/MESSy Atmospheric chemistry – Climate (EMAC) model, which simulates atmospheric chemical and meteorological processes and interactions with oceans and the biosphere[55,56]. The model uses the second version of the Modular Earth Submodel System (MESSy2) to link multi-institutional computer codes. The core atmospheric model was the 5th generation European Centre/Hamburg general circulation model (ECHAM5), into which updates and improvements in boundary layer, radiation, and convection routines have been introduced[56–58]. Additional descriptions, references, and information for the model are available at https://www.messy-interface.org. We applied EMAC (ECHAM5 version 5.3.02, MESSy version 2.55.0) with a spherical truncation of T63 (corresponding to a quadratic Gaussian grid of approximately 1.8 × 1.8 degrees of latitude and longitude) with 31 levels of vertical hybrid pressure to 10 hPa.

The various submodels represented tropospheric processes and their interactions with oceans, land, and human influences describing emissions, including isotopic composition, radiative processes, atmospheric multiphase chemistry, aerosols, and mechanisms of deposition[57,59]. The set-up used in this simulation was identical to that used by Pozzer et al.[9], where a detailed scheme of the oxidation of volatile organic compounds (VOCs) (the Mainz Organic Mechanism) was coupled to a base set of volatility (ORACLE[60]) to simulate the partitioning of organic gases and aerosols in unprecedented detail for a chemistry – climate model. We only briefly summarized the most important characteristics (see Pozzer et al.[9] for more details).

EMAC simulates gas-phase and heterogeneous chemistry using the MECCA submodel, which accounts for the photochemical oxidation of natural and anthropogenic VOCs[60–62]. Processes of aerosol microphysics and gas/aerosol partitioning were simulated using the GMXe submodel[63]. The distribution of aerosol sizes was described using seven interacting lognormal modes (four hydrophilic and three hydrophobic modes). The composition of aerosols was uniform within each mode (internally mixed) but could vary between modes (externally mixed). The four modes of hydrophilic size (nucleation, Aitken, accumulation, and coarse) encompassed the spectrum of aerosol sizes. The composition of inorganic aerosols was determined using the ISORROPIA-II thermodynamic equilibrium submodel[64], which calculates the gas/liquid/solid equilibrium partitioning of inorganic compounds and water. The components of aeolian dust can exist in the form of mineral salts in the solid phase and ions in the aqueous phase[65]. The composition and atmospheric evolution of organic aerosol compounds were simulated using the ORACLE sub-model, which represents classes of the volatility of organics by their effective saturation concentrations[66]. The biogenic emissions of non-methane VOCs were calculated online using MEGAN[7]. The results of the model for the last decade have been extensively tested against measured data for gases and particles from ground-based networks monitoring air quality and global observations from satellites[67–70].

The dynamics were not nudged, unlike the study by Pozzer et al.[9], but the model ran freely, forced only by climatological sea-surface temperature (SST) and sea-ice coverage (SIC) obtained from the ERA5 data for 2010-2019, which allowed the model to recreate a climatological meteorology without strong extremes in the forcing (SST and SIC). The meteorology/radiation and the chemistry were also fully decoupled, so all simulations performed developed identical binary meteorology, allowing comparisons between the simulations. In summary, any difference between the concentrations of MTs in the simulations were only due to changes in the emissions and not in different modes of transport.

We performed four simulations: 1) a BASE simulation, where MEGAN was used in its standard configuration (β=0.10 °C^{-1}), 2) an ALL simulation, where MEGAN was used in its standard configuration but with β set at 0.13 °C^{-1}, 3) an MTRP simulation, where β was based on the monoterpene emitted (Supplementary Fig. 3) a PFT simulation, where β in MEGAN varied based on PFT (Fig. 3, Supplementary Code 1).

Data availability

The meta-analysis data are available upon request to all scientists agreeing to the Cyprus Institute protocol (https://doi.org/10.XXXXX/zenodo.XXXXX; to be provided). The software code is available at the supplementary information.

Competing interests

The authors declare that they have no conflict of interest.

Author contributions

E.B. designed the research and drafted the paper. E.B., N.D., J.P., M.V., and A.M.Y-S collected the data. E.B., G.L. D.M., V.N.M., and A.P. analyzed the data. E.B. P.C. J.L., J.P., A.P., J.S and J.W. performed research. All authors contributed to editing the article.

Acknowledgements

We acknowledge the EMME-CARE project from the European Union's Horizon 2020 Research and Innovation Programme (grant agreement No. 856612), as well as matching co-funding by the Government of the Republic of Cyprus. We acknowledge the use of the HPC cluster Aether at the University of Bremen, financed by DFG within the scope of the Excellence Initiative. We acknowledge the support of BmBf project ATTO (01LK1602B).

References

1 1. Vickers, C. E., Gershenzon, J., Lerdau, M. T. & Loreto, F. A unified mechanism of action for volatile isoprenoids in plant abiotic stress. *Nat Chem Biol* **5**, 283–291 (2009).
2. Gershenzon, J. & Dudareva, N. The function of terpene natural products in the natural world. *Nat Chem Biol* **3**, 408–414 (2007).
3. Loreto, F., Dicke, M., Schnitzler, J.-P. & Turlings, T. C. J. Plant volatiles and the environment: Plant volatiles and the environment. *Plant Cell Environ* **37**, 1905–1908 (2014).
4. Riipinen, I. *et al.* The contribution of organics to atmospheric nanoparticle growth. *Nature Geosci* **5**, 453–458 (2012).
5. Yli-Juuti, T. *et al.* Significance of the organic aerosol driven climate feedback in the boreal area. *Nat Commun* **12**, 5637 (2021).
6. Jardine, K. J. *et al.* Monoterpene ' *thermometer* ' of tropical forest-atmosphere response to climate warming: Monoterpene ' *thermometer* '. *Plant, Cell & Environment* **40**, 441–452 (2017).
7. Guenther, A. B. *et al.* The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geosci. Model Dev.* **5**, 1471–1492 (2012).
8. Tingey, D. T., Manning, M., Grothaus, L. C. & Burns, W. F. Influence of Light and Temperature on Monoterpene Emission Rates from Slash Pine. *Plant Physiol.* **65**, 797–801 (1980).
9. Pozzer, A. *et al.* Simulation of organics in the atmosphere: evaluation of EMACv2.54 with the Mainz Organic Mechanism (MOM) coupled to the ORACLE (v1.0) submodel. *Geosci. Model Dev.* **15**, 2673–2710 (2022).
10. Makowski, D. *et al. Systematic review of meta-analyses to assess the impacts of farming practices - A methodological framework*. https://osf.io/byuw9 (2021) doi:10.31219/osf.io/byuw9.
11. Stewart, D. J. *et al.* The kinetics of the gas-phase reactions of selected monoterpenes and cyclo-alkenes with ozone and the NO3 radical. *Atmospheric Environment* **70**, 227–235 (2013).
12. Bourtsoukidis, E. *et al.* Strong sesquiterpene emissions from Amazonian soils. *Nat Commun* **9**, 2226 (2018).
13. Fares, S. *et al.* Observations of Diurnal to Weekly Variations of Monoterpene-Dominated Fluxes of Volatile Organic Compounds from Mediterranean Forests: Implications for Regional Modeling. *Environ. Sci. Technol.* **47**, 11073–11082 (2013).
14. Sarkar, C. *et al.* PTR-TOF-MS eddy covariance measurements of isoprene and monoterpene fluxes from an eastern Amazonian rainforest. *Atmos. Chem. Phys.* **20**, 7179–7191 (2020).
15. Atkinson, R. Atmospheric chemistry of VOCs and NOx. *Atmospheric Environment* **34**, 2063–2101 (2000).
16. Peñuelas, J. & Staudt, M. BVOCs and global change. *Trends in Plant Science* **15**, 133–144 (2010).
17. Yáñez-Serrano, A. M. *et al.* Amazonian biogenic volatile organic compounds under global change. *Glob Change Biol* **26**, 4722–4751 (2020).
18. Niinemets, U., Seufert, G., Steinbrecher, R. & Tenhunen, J. D. A model coupling foliar monoterpene emissions to leaf photosynthetic characteristics in Mediterranean evergreen Quercus species. *New Phytol* **153**, 257–275 (2002).
19. Grote, R. *et al.* Process-based modelling of isoprenoid emissions from evergreen leaves of Quercus ilex (L.). *Atmospheric Environment* **40**, 152–165 (2006).

20. Bourtsoukidis, E., Williams, J., Kesselmeier, J., Jacobi, S. & Bonn, B. From emissions to ambient mixing ratios: online seasonal field measurements of volatile organic compounds over a Norway spruce-dominated forest in central Germany. *Atmos. Chem. Phys.* **14**, 6495–6510 (2014).
21. Tarvainen, V. *et al.* Temperature and light dependence of the VOC emissions of Scots pine. *Atmos. Chem. Phys.* **5**, 989–998 (2005).
22. Fares, S. *et al.* Biogenic emissions from Citrus species in California. *Atmospheric Environment* **45**, 4557–4568 (2011).
23. Lim, J.-H. *et al.* Seasonal variations of monoterpene emissions from Pinus densiflora in East Asia. *Chemosphere* **73**, 470–478 (2008).
24. Hakola, H., Rinne, J. & Laurila, T. The hydrocarbon emission rates of tea-leafed willow (Salix phylicifolia), silver birch (Betula pendula) and European aspen (Populus tremula). *Atmospheric Environment* **32**, 1825–1833 (1998).
25. Schade, G. W., Goldstein, A. H. & Lamanna, M. S. Are monoterpene emissions influenced by humidity? *Geophys. Res. Lett.* **26**, 2187–2190 (1999).
26. Staudt, M. *et al.* Seasonal and diurnal patterns of monoterpene emissions from Pinus pinea (L.) under field conditions. *Atmospheric Environment* **31**, 145–156 (1997).
27. Laffineur, Q. *et al.* Impact of diffuse light on isoprene and monoterpene emissions from a mixed temperate forest. *Atmospheric Environment* **74**, 385–392 (2013).
28. Seco, R. *et al.* Ecosystem-scale volatile organic compound fluxes during an extreme drought in a broadleaf temperate forest of the Missouri Ozarks (central USA). *Glob Change Biol* **21**, 3657–3674 (2015).
29. Filella, I., Wilkinson, M. J., Llusià, J., Hewitt, C. N. & Peñuelas, J. Volatile organic compounds emissions in Norway spruce (Picea abies) in response to temperature changes. *Physiol Plant* **130**, 58–66 (2007).
30. McKinney, K. A., Lee, B. H., Vasta, A., Pho, T. V. & Munger, J. W. Emissions of isoprenoids and oxygenated biogenic volatile organic compounds from a New England mixed forest. *Atmos. Chem. Phys.* **11**, 4807–4831 (2011).
31. Geron, C. D., Daly, R. W., Arnts, R. R., Guenther, A. B. & Mowry, Fred. L. Canopy level emissions of 2-methyl-3-buten-2-ol, monoterpenes, and sesquiterpenes from an experimental Pinus taeda plantation. *Science of The Total Environment* **565**, 730–741 (2016).
32. Tang, J. *et al.* Challenges in modelling isoprene and monoterpene emission dynamics of Arctic plants: a case study from a subarctic tundra heath. *Biogeosciences* **13**, 6651–6667 (2016).
33. Warneke, C. *et al.* Biogenic emission measurement and inventories determination of biogenic emissions in the eastern United States and Texas and comparison with biogenic emission inventories. *J. Geophys. Res.* **115**, D00F18 (2010).
34. Bäck, J. *et al.* Chemodiversity of a Scots pine stand and implications for terpene air concentrations. *Biogeosciences* **9**, 689–702 (2012).
35. Stone, D., Whalley, L. K. & Heard, D. E. Tropospheric OH and HO2 radicals: field measurements and model comparisons. *Chem. Soc. Rev.* **41**, 6348 (2012).
36. Vanhatalo, A. *et al.* Scots Pine Stems as Dynamic Sources of Monoterpene and Methanol Emissions. *Front. For. Glob. Change* **2**, 95 (2020).
37. Edtbauer, A. *et al.* Cryptogamic organisms are a substantial source and sink for volatile organic compounds in the Amazon region. *Commun Earth Environ* **2**, 258 (2021).
38. Mäki, M., Aalto, J., Hellén, H., Pihlatie, M. & Bäck, J. Interannual and Seasonal Dynamics of Volatile Organic Compound Fluxes From the Boreal Forest Floor. *Front. Plant Sci.* **10**, 191 (2019).
39. Langford, B. *et al.* Fluxes and concentrations of volatile organic compounds from a South-East Asian tropical rainforest. *Atmos. Chem. Phys.* **10**, 8391–8412 (2010).
40. Yáñez-Serrano, A. M. *et al.* Heat stress increases the use of cytosolic pyruvate for isoprene biosynthesis. *Journal of Experimental Botany* **70**, 5827–5838 (2019).
41. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis*. (John Wiley & Sons, Ltd, 2009). doi:10.1002/9780470743386.

42. Hedges, L. V., Gurevitch, J. & Curtis, P. S. THE META-ANALYSIS OF RESPONSE RATIOS IN EXPERIMENTAL ECOLOGY. *Ecology* **80**, 1150–1156 (1999).
43. Makowski, D., Piraux, F. & Brun, F. *From Experimental Network to Meta-analysis: Methods and Applications with R for Agronomic and Environmental Sciences*. (Springer Netherlands, 2019). doi:10.1007/978-94-024-1696-1.
44. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using **lme4**. *J. Stat. Soft.* **67**, (2015).
45. Moreno-Martínez, Á. *et al.* A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment* **218**, 69–88 (2018).
46. Camps-Valls, G. *et al.* Global trait maps at 1km resolution. (2020) doi:10.17871/TRY.60.
47. MAO, J. & YAN, B. Global Monthly Mean Leaf Area Index Climatology, 1981-2015. 6.581285 MB (2019) doi:10.3334/ORNLDAAC/1653.
48. Samek, W., Wiegand, T. & Müller, K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *arXiv:1708.08296 [cs, stat]* (2017).
49. Matthaios, V. N. *et al.* Quantifying factors affecting contributions of roadway exhaust and non-exhaust emissions to ambient PM10–2.5 and PM2.5–0.2 particles. *Science of The Total Environment* **835**, 155368 (2022).
50. Matthaios, V. N. *et al.* Factors Influencing Classroom Exposures to Fine Particles, Black Carbon, and Nitrogen Dioxide in Inner-City Schools and Their Implications for Indoor Air Quality. *Environ Health Perspect* **130**, 047005 (2022).
51. Zhou, J. *et al.* Random Forests and Cubist Algorithms for Predicting Shear Strengths of Rockfill Materials. *Applied Sciences* **9**, 1621 (2019).
52. Wang, Y. & Witten, I. *Induction of model trees for predicting continuous classes*. (University of Economics, Faculty of Informatics and Statistics, 1997).
53. Kuhn, M. & Johnson, K. *Applied Predictive Modeling*. (Springer, 2013).
54. Noi, P., Degener, J. & Kappas, M. Comparison of Multiple Linear Regression, Cubist Regression, and Random Forest Algorithms to Estimate Daily Air Surface Temperature from Dynamic Combinations of MODIS LST Data. *Remote Sensing* **9**, 398 (2017).
55. Jöckel, P. *et al.* The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere. *Atmos. Chem. Phys.* **6**, 5067–5104 (2006).
56. Jöckel, P. *et al.* Development cycle 2 of the Modular Earth Submodel System (MESSy2). *Geosci. Model Dev.* **3**, 717–752 (2010).
57. Tost, H. *et al.* Global cloud and precipitation chemistry and wet deposition: tropospheric model simulations with ECHAM5/MESSy1. *Atmos. Chem. Phys.* **7**, 2733–2757 (2007).
58. Ouwersloot, H. G., Pozzer, A., Steil, B., Tost, H. & Lelieveld, J. Revision of the convective transport module CVTRANS 2.4 in the EMAC atmospheric chemistry–climate model. *Geosci. Model Dev.* **8**, 2435–2445 (2015).
59. Pozzer, A. *et al.* Observed and simulated global distribution and budget of atmospheric C2-C5 alkanes. *Atmos. Chem. Phys.* **10**, 4403–4422 (2010).
60. Tsimpidi, A. P., Karydis, V. A., Pozzer, A., Pandis, S. N. & Lelieveld, J. ORACLE (v1.0): module to simulate the organic aerosol composition and evolution in the atmosphere. *Geosci. Model Dev.* **7**, 3153–3172 (2014).
61. Sander, R. *et al.* The atmospheric chemistry box model CAABA/MECCA-3.0. *Geosci. Model Dev.* **4**, 373–380 (2011).
62. Lelieveld, J., Gromov, S., Pozzer, A. & Taraborrelli, D. Global tropospheric hydroxyl distribution, budget and reactivity. *Atmos. Chem. Phys.* **16**, 12477–12493 (2016).
63. Pringle, K. J. *et al.* Description and evaluation of GMXe: a new aerosol submodel for global simulations (v1). *Geosci. Model Dev.* **3**, 391–412 (2010).
64. Fountoukis, C. & Nenes, A. ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+–Ca2+–Mg2+–NH4+–Na+–SO42−–NO3−–Cl−–H2O aerosols. *Atmos. Chem. Phys.* **7**, 4639–4659 (2007).

65. Karydis, V. A., Tsimpidi, A. P., Pozzer, A., Astitha, M. & Lelieveld, J. Effects of mineral dust on global atmospheric nitrate concentrations. *Atmos. Chem. Phys.* **16**, 1491–1509 (2016).
66. Tsimpidi, A. P., Karydis, V. A., Pandis, S. N. & Lelieveld, J. Global combustion sources of organic aerosols: model comparison with 84 AMS factor-analysis data sets. *Atmos. Chem. Phys.* **16**, 8939–8962 (2016).
67. Pozzer, A. *et al.* Distributions and regional budgets of aerosols and their precursors simulated with the EMAC chemistry-climate model. *Atmos. Chem. Phys.* **12**, 961–987 (2012).
68. Pozzer, A. *et al.* Effects of business-as-usual anthropogenic emissions on air quality. *Atmos. Chem. Phys.* **12**, 6915–6937 (2012).
69. Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D. & Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **525**, 367–371 (2015).
70. Karydis, V. A. *et al.* Global impact of mineral dust on cloud droplet number concentration. *Atmos. Chem. Phys.* **17**, 5601–5621 (2017).

Figures

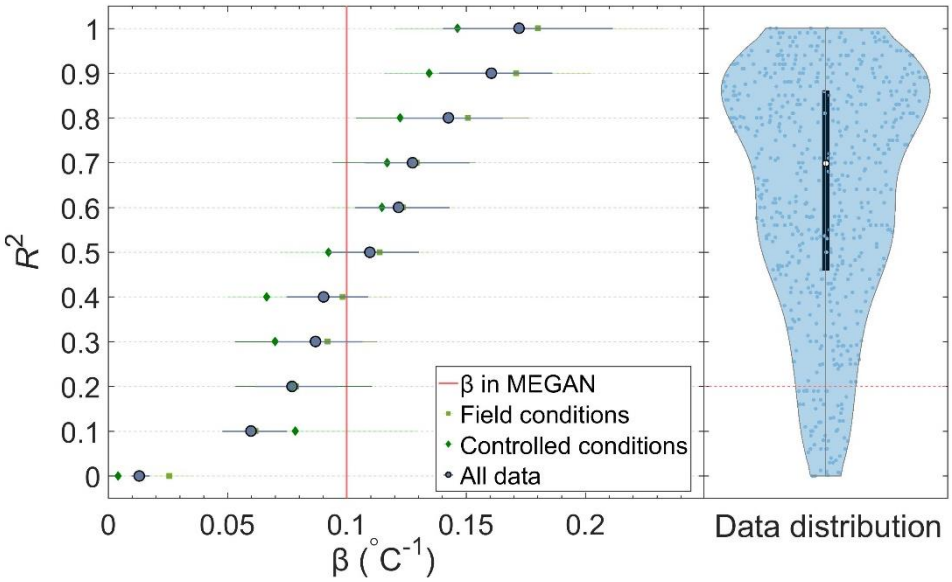
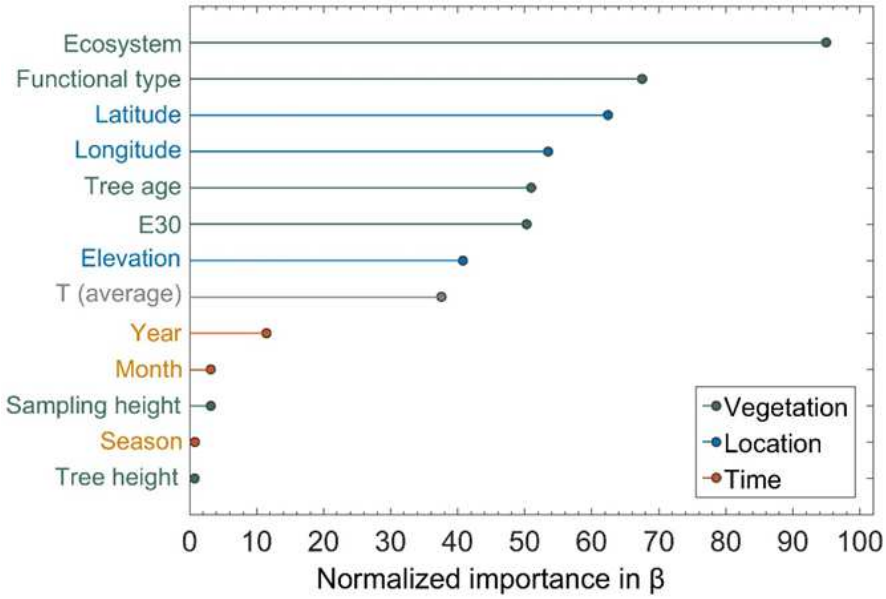


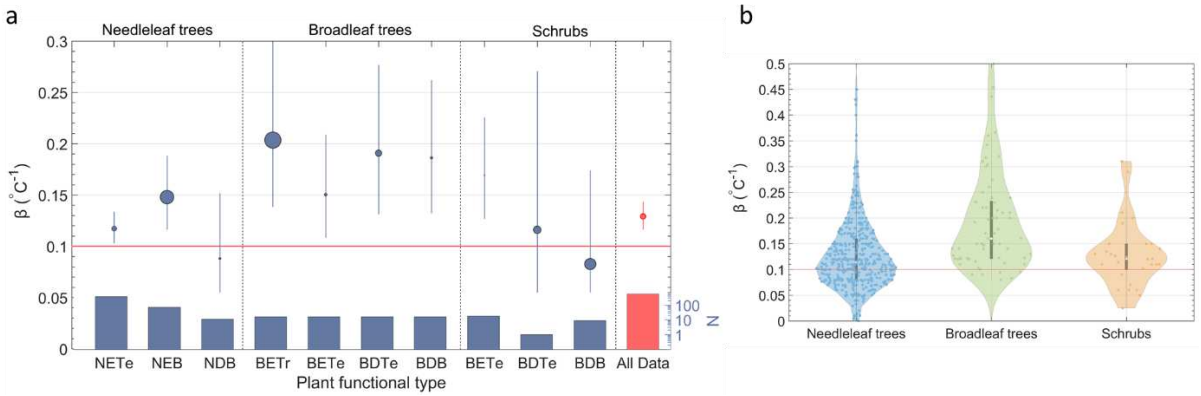
Figure 1. Groups of coefficients of determination (R^2) for experimentally derived dependencies (β) of MT emissions on temperature. The blue and green error bars indicate 95% confidence intervals. On the right, the distribution of the N=696 values of β extracted from the literature is displayed together with a boxplot that illustrates their median (white circle) and 25th and 75th percentiles (lower and upper bounds of the box), respectively. Data below the red dashed line (on the right) were obtained from regressions with poor goodness of fit (R^2<0.2) and were excluded from further analysis.



587

588

589 **Figure 2.** Feature importance ranking. Indicative contribution of individual parameters of
590 vegetation, location, and time parameters to the variation of the β coefficient.



591

592

593 **Figure 3.** Plant functional type-dependent β coefficients. Dependence on temperature (β) for
594 the plant functional types (A) and wider categories of woody plants (B). The error bars in
595 panel a indicate 95% confidence intervals, and the size of the blue bullet points is proportional
596 to the global surface area of each plant functional type. NETe, needleleaf evergreen temperate
597 forest; NEB, needleleaf evergreen boreal forest; NDB, needleleaf deciduous boreal forest;
598 BETr, broadleaf evergreen tropical forest; BETe, broadleaf evergreen temperate forest/shrubs;
599 BDTe, broadleaf deciduous temperate forest/shrubs; BDB, broadleaf deciduous boreal
600 forest/shrubs.

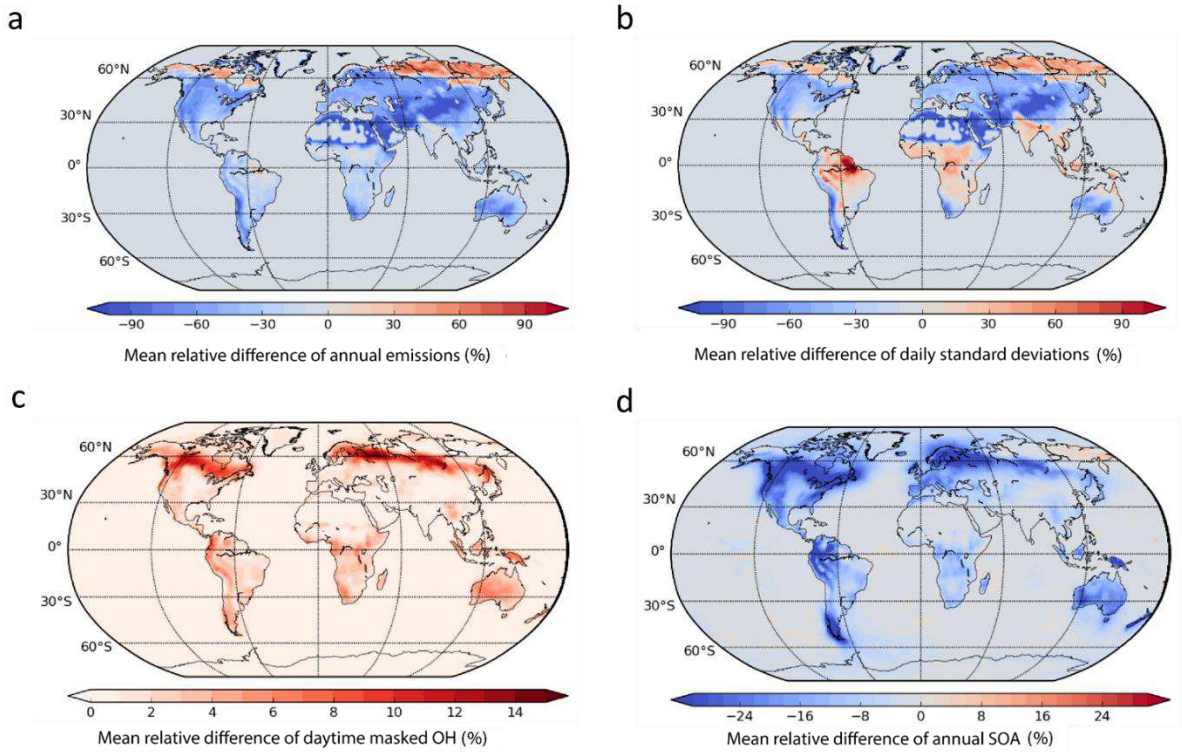


Figure 4. Model simulations. Annual mean relative differences (β-factor = 0.1 °C^{-1} and β-factor from Fig. 3 based on all data) derived from hourly data over a climatic year. A. Total monoterpene emissions, B. Standard deviation of daily emissions, C. Daytime OH, and D. SOA production**.**

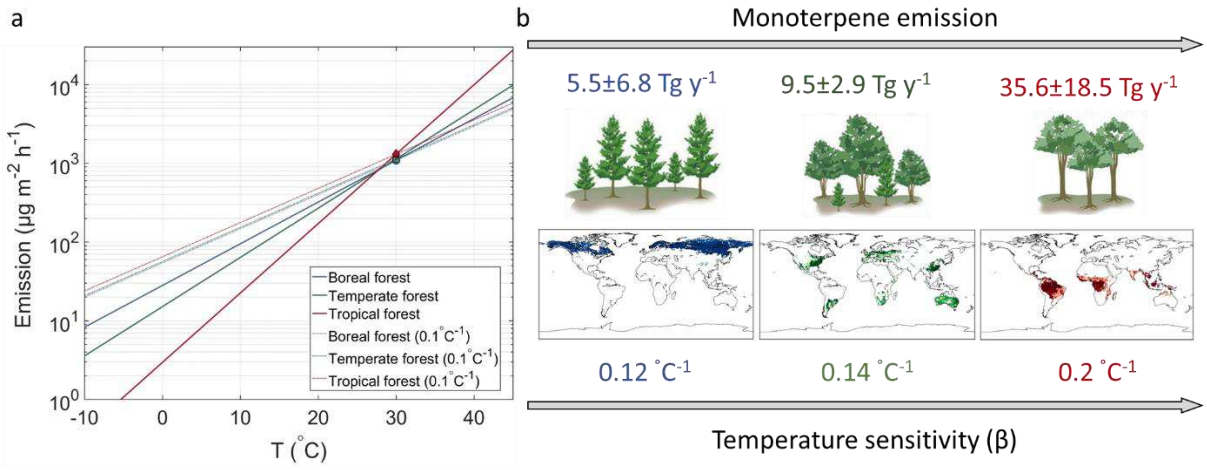


Figure 5. Monoterpene emissions over temperature gradients for the three main ecosystems. A. Emissions using the standardized potential emission at 30 °C and β in MEGAN v2.1 (dashed lines) and using β derived from this study (panel b and Supplementary Table 2). The maps in panel B represent the global coverage of boreal (blue), temperate (green), and tropical (red) forests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Bourtsoukidisetal.Supplement.pdf
- Bourtsoukidisetal.Supplement.pdf