ARTICLE

Check for updates

High temperature sensitivity of monoterpene emissions from global vegetation

Efstratios Bourtsoukidis [1✉], Andrea Pozzer [1,2], Jonathan Williams [1,2], David Makowski [3], Josep Peñuelas [4,5], Vasileios N. Matthaios[6,7], Georgia Lazoglou[1], Ana Maria Yañez-Serrano[8], Jos Lelieveld [1,2], Philippe Ciais[1,9], Mihalis Vrekoussis [1,10,11], Nikos Daskalakis [10] & Jean Sciare[1]

Terrestrial vegetation emits vast amounts of monoterpenes into the atmosphere, influencing ecological interactions and atmospheric chemistry. Global emissions 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 40 years of published monoterpene emission data and show that relationship between emissions and temperature is more sensitive and intricate than previously thought. Considering the entire dataset, a higher temperature sensitivity ($\beta = 0.13 \pm 0.01\,°C^{-1}$) is derived but with a linear increase with the reported coefficients of determination (R^2), indicating that co-occurring environmental factors modify the temperature sensitivity of the emissions that is primarily related to the specific plant functional type (PFT). 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.

[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] Department of Public Health Policy and Systems, University of Liverpool, L69 3GB Liverpool, 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. ✉email: e.bourtsoukidis@cyi.ac.cy

Global vegetation is considerably affected by rising temperatures, yet our understanding of plant response to warming and its implications for the biosphere-atmosphere interactions remains incomplete. Biogenic volatile organic compounds (bVOCs), particularly monoterpenes (MT; $C_{10}H_{16}$), are known to be emitted in response to abiotic drivers such as temperature[1–3]. Due to their essential role in warming-induced responses and emission composition characteristics, MTs have even been attributed the name 'thermometer of plants'[4]. Interlinked to their temperature-dependent emission, the multifaceted roles of MTs in the earth's system extend from communication signals to influencing atmospheric processes[5,6].

Once released into the atmosphere, they become key players in atmospheric chemistry and physics[7,8]. Due to their high reactivity with atmospheric radicals, they are key contributors to the formation and growth of secondary organic aerosols (SOA), indirectly impacting the radiative balance of the atmosphere[9–11]. In the presence of nitrogen oxides, MTs also participate in the formation of tropospheric ozone, a harmful atmospheric pollutant[12]. Their role on atmospheric oxidation and SOA formation make their accurate simulation essential for understanding and predicting the impacts of climate change.

Global emission models employ empirical algorithms to simulate variations in MT emission rates[13]. These algorithms encompass a temperature response mechanism rooted in enzymatic activity and a light response mechanism based on electron transport[14]. Together, they offer a straightforward yet mechanistic approach to MT emission modeling. Generally, rising temperatures stimulate the emission of MTs stored in leaf pools, while both temperature and light govern the de novo production and subsequent release of these compounds[15]. Recent scientific progress has prompted the incorporation of additional parameters into the model, such as leaf age, soil moisture, leaf area index, and CO_2 inhibition (refer to Supplementary information). Despite these advancements, the majority of studies still attributes temperature as the primary driver of MT emissions from global vegetation, commonly represented by the following equation:[13]:

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

Here, E_{30} is the emission potential under standardized conditions (30 °C), and β is an empirical coefficient (in °C^{-1}) derived from the fit of the regression between measured emissions and temperature (T). It should be noted that MT emissions are best described by leaf temperature since leaf and air temperatures can vary considerably among plant species due to the different plant-specific ability to cool down plant tissues[16]. E_{30} is a Plant Functional Type (PFT)-specific parameter that defines the MT emissions at 30 degrees Celsius (or 303.15 K), independent of the temperature responses from the individual plants. These are specified by the slope of the exponent (i.e. the β-coefficient) and are notable indicators of the plant responses to temperature.

To date, global emission models, such as MEGAN[13], use a fixed dependence on temperature ($\beta = 0.10$ °C^{-1}) for all types of vegetation and ecosystems. The purpose of this meta-analysis is to conduct a comprehensive evaluation of this parameter by collecting and analyzing all experimentally derived β coefficients published in the four decades since the first observation of this relationship in 1980[17]. 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)[13] and an atmospheric chemistry–climate model (EMAC)[18].

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 from regression fits obtained under diverse locations (Fig.1). The reported values varied widely, ranging from −0.07 to 2.39 °C^{-1}. To identify the factors that explain this variability, we collected 35 additional 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, including the conclusion of each study (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 Fig. 1). 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), validating the upscaling of β from leaf to canopy levels (Supplementary Fig. 2). In the majority of cases, the reported MT emissions were solely temperature-dependent, accounting for 499 instances or 72% of the total reports. Emissions dependent on both temperature and light were reported in 142 instances (20.4%), while the remaining 55 β-coefficients (7.9%) were derived when authors identified both temperature and another environmental driver as influencing the emissions (Supplementary Fig. 3). The three distinct clusters exhibited the following median values: temperature alone (0.11 ± 0.13 °C^{-1}), temperature and light (0.14 ± 0.12 °C^{-1}), and temperature combined with another environmental driver (0.10 ± 0.05 °C^{-1}) (Supplementary Fig. 4a). The additional environmental drivers reported include relative humidity, soil water content, CO_2 levels, seasonality, transpiration, and photosynthesis. The clearest observed pattern was the increase in β with the coefficient of determination (R^2) obtained from the regression fits between emission and temperature for each experimental dataset. In this context, R^2 serves as a statistical measure that characterizes the goodness of the fit of Eq. (1), elucidating the degree to which the parameterization explains the intrinsic association between temperature and emission rates. Our meta-analysis, based on a standard random-effect statistical model[19], determined the values of β within 95% confidence intervals, across R^2 bins (Fig. 2). A linear relationship was identified between the quality of the 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 indicates a clear underestimation of a globally uniform β.

The linear increase of β with the goodness of fit of Eq. (1) (i.e. R^2) reveals that when temperature is the dominant driver, the temperature sensitivity reaches its maximum. In fact, when considering instances with $R^2 > 0.9$, the β-coefficient is 0.17 ± 0.3 °C^{-1}, 70% higher than the value currently considered. Conversely, at mid- and lower R^2 values, β declines, demonstrating a decrease in temperature sensitivity as the emission variance becomes less explainable by temperature. This implies that other environmental drivers regulate the emission rates. Considering the emissions that were reported to be dependent on both temperature and light, β and R^2 followed the same trend. However, the values for the temperature and light (T + L) cluster were consistently higher compared to the temperature-only (T) cluster (Supplementary Fig 4b). An ANOVA test conducted on these two datasets yielded a p-value of 0.017, indicating statistically significant differences in temperature responses when MTs are also emitted as a function of light.
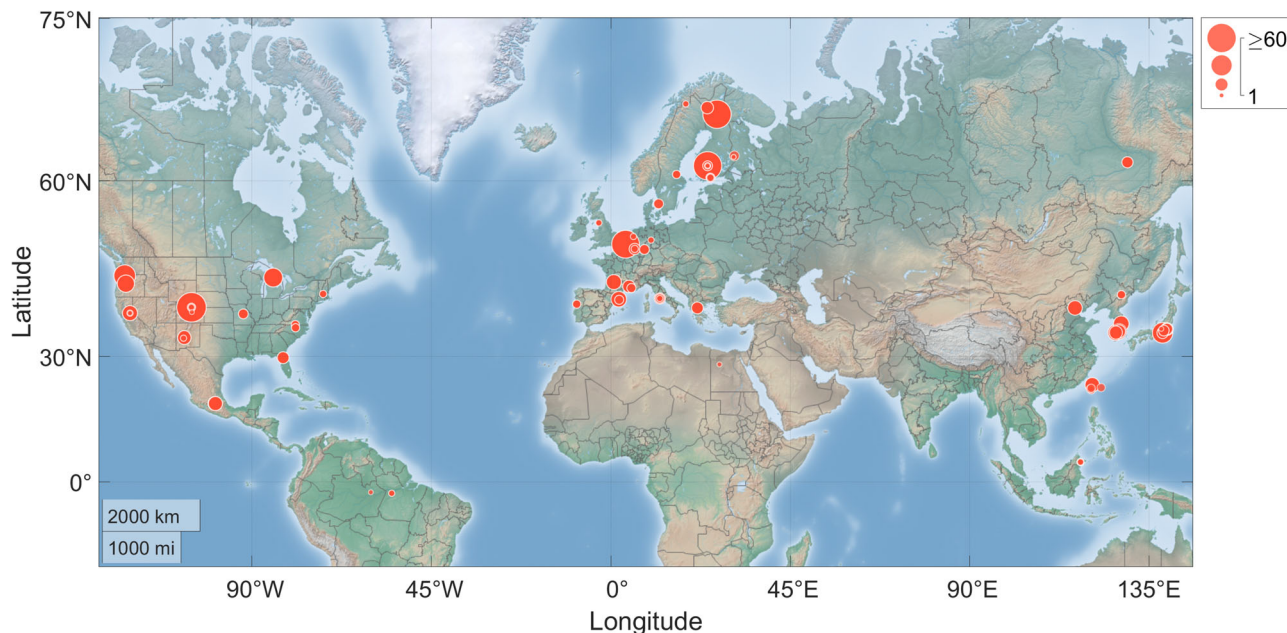
Fig. 1 Locations of experimentally derived β. The size of the circles indicates the number of coefficients obtained from the literature. The studies considered can be found as a separate table in the Supplementary information.
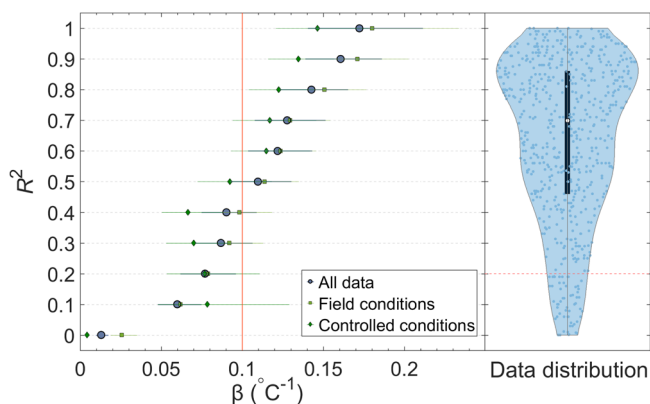

Fig. 2 Groups of coefficients of determination (R^2) for experimentally derived dependencies (β) of MT emissions on temperature. The blue and green error bars indicate 95% confidence intervals. On the right, the distribution of the $N = 696$ values of β extracted from the literature is displayed together 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.

Interestingly, β obtained from field data was systematically higher than β obtained under controlled conditions, suggesting amplified sensitivities to temperature under natural conditions. This discrepancy was further enhanced at high R^2 values, where the MT emissions were primarily linked to temperature. Based on this observation, it can be concluded that laboratory studies may underestimate the temperature sensitivity in real-life conditions.

To better understand the key factors influencing β values, we employed machine learning techniques, specifically the Feature Importance Ranking (FIR) tool. This approach evaluates the significance of individual parameters, such as vegetation type, location, and time variables, in the determination of β (Fig. 3). The results suggest that the type of ecosystem, specifically the
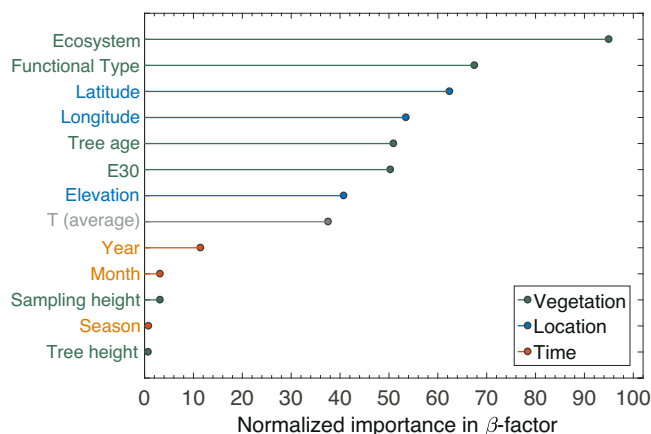

Fig. 3 Feature importance ranking. Indicative contribution of individual parameters of vegetation, location, and time parameters to the variation of the β coefficient. The parameters are explained in Supplementary Table 1.

plant functional type (PFT) as defined by Guenther et al.[13], is a critical factor that affects the temperature sensitivity of MT emissions. The geographic location (latitude and longitude) remained associated with the vegetation types that grow within the corresponding geographically distributed ecosystems. These results, however, suggest that the temperature responses of MT emissions may exhibit dynamic variations throughout the tree's age. Finally, the FIR analysis did not clearly identify seasonality or a relationship between tree or sampling height and the coefficient of temperature dependency. However, a relatively consistent ratio between sampling and tree height has been identified (0.7 ± 0.2), revealing the preference towards sun-exposed parts of the tree (Supplementary Fig. 5). Summarizing, it should be noted that while a clear the relationship between PFT and the β coefficient became evident from the FIR analysis, the role of the other parameters might have underappreciated due to the frequently unreported values.
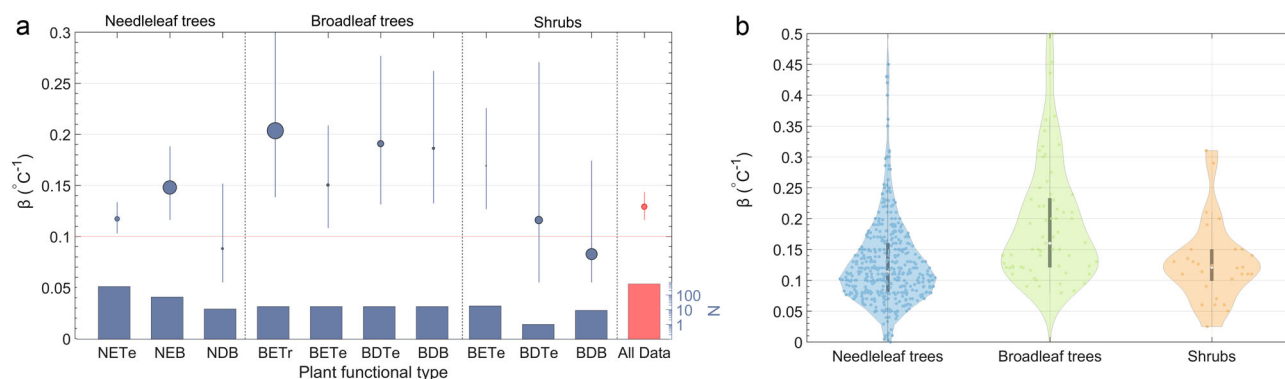
Fig. 4 Plant functional type-dependent β coefficients. Dependence on temperature (β) for the plant functional types (**a**) and wider categories of woody plants (**b**). The error bars in panel a indicate 95% confidence intervals, and the size of the blue bullet points is proportional to the global surface area of each plant functional type. NETe needleleaf evergreen temperate forest, NEB needleleaf evergreen boreal forest, NDB needleleaf deciduous boreal forest, BETr broadleaf evergreen tropical forest, BETe broadleaf evergreen temperate forest/shrubs, BDTe broadleaf deciduous temperate forest/shrubs, BDB broadleaf deciduous boreal forest/shrubs.

Revised temperature-dependent coefficients (β). The statistical model used in this study analyzed the entire dataset consisting of 559 data points with an R^2 value greater than 0.2 (see methods). The results showed a uniform value of $0.13 \pm 0.1\,°C^{-1}$ for β, with different temperature-dependent coefficients for each PFT, as seen in Fig. 4. Removing the lowest 10% of data points had little effect on the revised values, as demonstrated in Supplementary Fig. 6. The analysis showed that, except for boreal needle leaf deciduous trees and broadleaf deciduous shrubs, almost all other PFTs have significantly larger values than $0.10\,°C^{-1}$. This is particularly pronounced for the boreal needle leaf evergreen forests ($\beta_{NEB} = 0.15\,°C^{-1}$; range $0.12–0.19\,°C^{-1}$) and the tropical broadleaf evergreen forests ($\beta_{BETr} = 0.20\,°C^{-1}$; range $0.14–0.3\,°C^{-1}$). Generally, broadleaf trees display higher temperature dependencies in the release of MTs by their foliage compared to needle-leaf trees and broadleaf shrubs. Notably, no β values were reported for agricultural ecosystems or grasslands, highlighting the need for further research to determine the temperature responses for these climate-sensitive ecosystems.

A β coefficient was derived for each of the reported MT species, considering the importance of processes of atmospheric oxidation due to their different rates of reaction with atmospheric radicals. Supplementary Fig. 7 shows that, although β was higher for Δ3-carene and trans-β-ocimene, no significant differences were identified across all MTs. However, it is important to note that the majority of studies (74%) reported the sum of MTs, which limits the ability to draw a clear conclusion on the temperature sensitivity of speciated MTs. This limitation is further exacerbated by the fact that Proton Transfer Reaction instruments measure the cumulative total of all MTs, whereas Gas Chromatography methods only account for selected species in their reported sums.

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)[13] and a state-of-the-art global atmospheric chemistry–climate model (EMAC)[18]. The simulations were performed at 1.8 by 1.8-degree resolution and with hourly outputs. In total, we performed four simulations: 1) a

BASE simulation, where MEGAN was used in its standard configuration ($\beta = 0.10\,°C^{-1}$), 2) an ALL simulation, where MEGAN was used in its standard configuration but with β set at $0.13\,°C^{-1}$, 3) a MTRP simulation, where β was based on the monoterpene species (Supplementary Fig. 7), and 4) a PFT simulation, where β in MEGAN varied based on PFT (Fig. 4, Supplementary Code 1). The new simulations resulted in significant and comparable changes in emission rates and atmospheric feedbacks compared to the current model (Supplementary Fig. 8). Based on the results presented in Figs. 3 and 4, as well as Supplementary Figs. 6, 7, we determined that the most appropriate strategy was to revise the temperature sensitivity coefficient (β) across different plant functional types (PFTs).

In the "PFT" simulation, global annual average MT emissions decreased by 13%, with the most pronounced reductions occurring for needle-leaf evergreen boreal trees (44%), broadleaf deciduous boreal trees (44%), and broadleaf evergreen temperate trees (41%) (Fig. 5, Supplementary Table 2). The largest seasonal increase was observed in Siberia during winter, which is attributed to the lower β values from broadleaf deciduous shrubs and generally low MT emissions in this region. It should be noted that the β applied for this PFT is grounded in findings from a single study[20]. Although MT emissions from this ecosystem remain relatively low[21], emerging research indicates that the subarctic tundra might exhibit higher sensitivity to temperature changes in both MT and isoprene emissions[22–24]. Given the accelerated warming of this delicate ecosystem compared to the global average[25], it is crucial to gather more data from the pan-Arctic region to accurately characterize the temperature responses of regional vegetation.

In contrast, tropical ecosystems, especially the Amazon rainforest, had higher annual β values, resulting in local increases of up to 30% in MT emissions. The simulations also showed that during the Amazonian dry season, MT emissions could even double despite similar annual averages (Supplementary Fig. 9). Moreover, the investigation of daily extremes for each season revealed occasional increases of more than three-fold in both tropical and arctic ecosystems (Supplementary Fig. 10). In light of the higher temperature sensitivity of MT emissions, it becomes evident that the increasing frequency and strength of heat waves[26] will markedly amplify the warming-induced MT emissions.

The most pronounced difference between the two simulations was the diurnal variation in the rates of emission (Supplementary Fig. 11, 12). Tropical forests exhibited especially marked diel cycles, which could be further enhanced over seasonal and daily timescales. The maximum differences in daily standard deviation
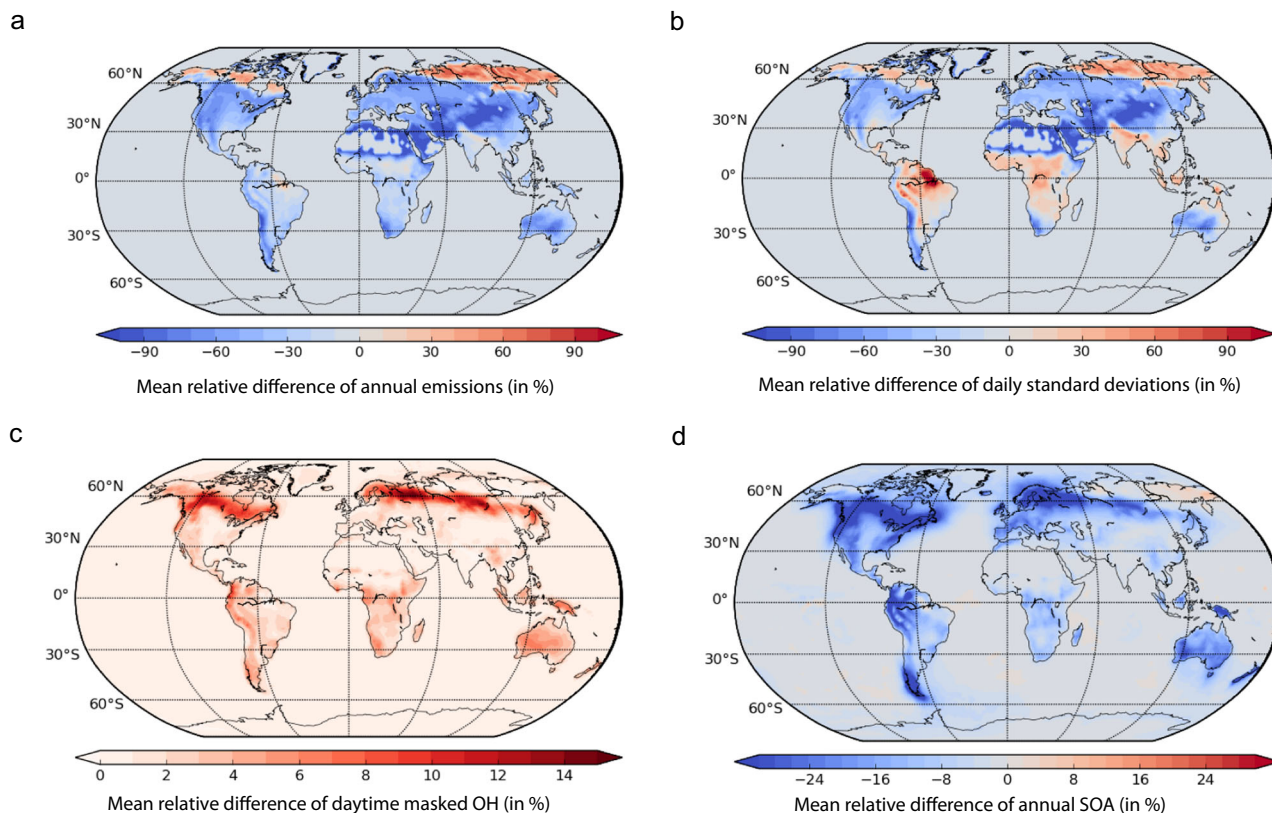
Fig. 5 Model simulations. Annual mean relative differences between the BASE simulation, where MEGAN was used in its standard configuration ($\beta = 0.10$ °C^{-1}), and a PFT-dependent β, based on the median values presented in Fig. 4. The differences are averages of hourly data over a climatic year. **a** Total monoterpene emissions, **b** The standard deviation of daily emissions, **c** Daytime OH, and **d** SOA production.
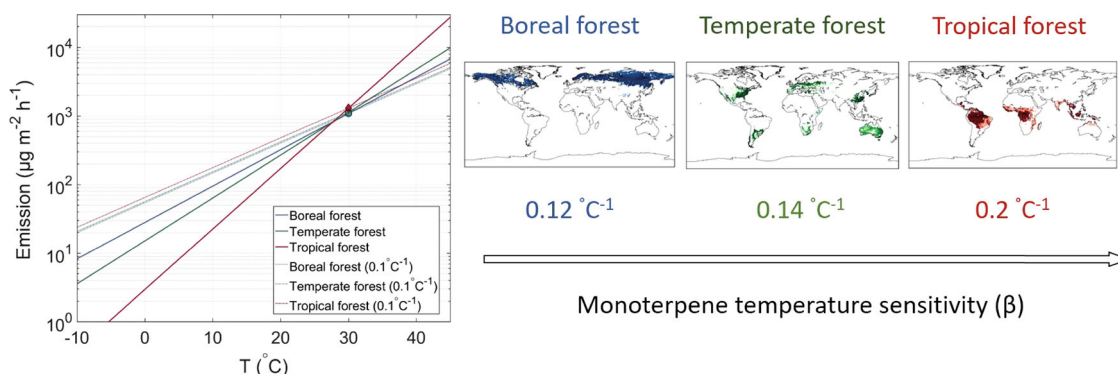

Fig. 6 Monoterpene emissions over temperature gradients for the three main ecosystems. a Emissions using the standardized potential emission at 30 °C and β in MEGAN v2.1 (dashed lines) and using β derived from this study (panel **b** and Supplementary Table 2). The maps in the right panel represent the global coverage of boreal (blue), temperate (green), and tropical (red) forests.

exceeded 300%, highlighting the substantial variability in MT emissions due to temperature dependencies. We applied Eq. (1) using data for temperature collected at the Amazon Tall Tower Observatory (ATTO) in 2014 and 2015[27] to demonstrate the dynamics of PFT-dependent β (Supplementary Fig. 13). The increase in β (from 0.10 to 0.20 °C^{-1}) led to slightly lower average emissions because the average temperature is lower than 30°C (see Fig. 6). 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. Such amplified diurnal variations help to explain the discrepancies in the standard deviation of the

emissions between the model and observations in tropical ecosystems[28].

The relationship between temperature and MT emission follows an exponential function that is standardized at 30 °C (Eq. (1)). Revisions of the β values according to their PFT resulted in substantial changes in the simulated emissions occurred at either low or very high temperatures (Fig. 6, Supplementary Fig. 14). 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). As shown in Fig. 6a and Supplementary Fig. 14, the current models of emission have clearly overestimated MT emissions under 30 °C.

5

In contrast, the current models have underestimated the emissions for temperatures above 30 °C, which may help to explain the observed discrepancies in diurnal variation of MT emission in warm environments[28,29].

The parameter E_{30}, representing the emission potential at standard conditions, plays a crucial role in determining 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 Fig. 15, 16).

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. The model simulations showed that using a PFT-dependent β increases OH concentrations globally (Fig. 5). 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. 17). Substantial 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. 18). It is therefore concluded that increasing the temperature sensitivity of MT emissions will result in a global increase of the atmospheric oxidative capacity of OH, demonstrating the important role of MTs in atmospheric chemistry. The implications of an enhanced OH abundance over the northern latitudes include a reduction in the atmospheric lifetime of methane, which is prolifically emitted in these regions[30].

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. 19). Regional changes in O_3 concentration on specific days, however, could increase by up to 10% (Supplementary Fig. 20). 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[31]). 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 over forests influenced the yield of secondary organic aerosol particles around the globe (Fig. 5, Supplementary Fig. 21). 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.

The study's results underscore the crucial role of the temperature dependency of MT emissions in shaping atmospheric oxidation processes and reaction products. The model simulations demonstrated that this sensitive coefficient plays an important role in atmospheric processes that are amplified seasonally and have a significant impact on the formation of secondary organic aerosol particles globally. Given the higher temperature sensitivities presented here, the rising global temperatures will amplify MT emissions. Consequently, this increase in MTs will lead to a corresponding rise in the production of SOA, thereby enhancing the radiative cooling effects. These findings emphasize the urgent need for accurate simulations of the temperature effects on MT emissions to improve our understanding of atmospheric chemistry and physics, and ultimately, to better predict the impacts of climate change.

Discussion

As global temperatures continue to rise and extreme heat events become more frequent[26], the temperature sensitivity of forests is emerging as a critical issue for understanding the impacts of climate change on forest ecosystems and atmospheric chemistry. One important aspect of this sensitivity is the emissions of bVOCs, such as monoterpenes, which will increase with rising temperatures. The stronger emissions resulting from rising temperatures can have far-reaching and uncertain consequences for the biosphere, potentially disrupting its delicate balance and feedback mechanisms between ecology, atmospheric chemistry and climate[6]. Therefore, the accurate simulation of MT emissions is essential for evaluating plant responses and the respective feedbacks and improving future projections.

To date, various algorithms have been developed to simulate MT emissions, which account for either short-term volatilization (emission-based)[13,17] or long-term production of MT linked to photosynthesis (production-based)[32,33]. Emission-based models, such as MEGAN, are commonly used to describe observations at different scales, from leaves and branches to canopies, by simulating global MT emissions from terrestrial vegetation. These models account for multiple parameters, including temperature, light, foliar age, soil moisture, and CO_2 inhibition. In this meta-analysis, we focused on the effect of temperature, particularly in evaluating the warming emission-response that defines temperature sensitivity (β) of MT emissions.

The β coefficient for the temperature-dependent term in Eq. (1) is shown to vary across seasons[34,35], tree species[36], tree age[37], developmental stage[38], ambient humidity[39], level of photosynthetically active radiation[40], CO_2 abundance[41], soil moisture[42], and other drivers[43]. 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 warming and applying a PFT-dependent β helps to explain the frequently overestimated MT emissions in several environments[29,44–47].

Using a PFT-dependent β is a considerable step forward in understanding the sensitivity of MT emissions to temperature. However, given the complexity of nature, it is still necessary to acknowledge the remaining uncertainties in estimating and modeling the β for MT emissions. For example, our findings reveal that the sensitivity of MT emissions to temperature is also influenced by co-occurring environmental drivers and is different for laboratory and field conditions. On the same lines, it has been recently shown that the β coefficient can significantly increase under heat stress conditions[48]. Creating an accurate parameterization that considers the diverse and frequently co-occurring drivers of emissions is a challenging task since plants have developed distinct coping mechanisms to deal with environmental stressors. It has been demonstrated that even plants of the same species and identical growing conditions can exhibit large chemodiversity in their MT emissions, adding to the complexity of plant responses[49].

Besides the temperature sensitivity parameter (β), other parameters also play a role in the emissions of MTs by terrestrial

vegetation. One critical parameter is the potential emission at standard conditions (E_{30}), which regulates the rate of emission by forests. Our meta-analysis collected E_{30} values from different ecosystems (see Supplementary Figs. 15, 16) and found that they vary greatly by orders of magnitude, revealing substantial uncertainties in the magnitude of simulated emissions. Given that both β and E_{30} define the emission rate of MTs from global vegetation, further research is necessary to validate the current approach of implementing constant E_{30} values in emission models. For example, although CO_2 concentrations can impact both β and E_{30} values, limited data in the reviewed literature hindered a conclusive assessment. Moreover, uncertainties in the leaf area index and the lack of studies in widespread ecosystems, including agricultural and grasslands, add to the unknowns and shape the direction of future research.

The estimation of MT emissions has significant implications for atmospheric processes, specifically for the mechanisms of atmospheric oxidation and subsequent SOA formation. Our study demonstrates that increasing β leads to a decrease in MT emissions under mean temperatures below 30 °C, resulting in an increase in global OH concentrations. This observation is especially relevant given the existing discrepancies between measured and simulated OH concentrations in unpolluted forested regions[50]. Our updated simulations, which incorporate the PFT-based temperature sensitivity of MT emissions, provide an explanation for these discrepancies and underscore the strong link between MT emissions and atmospheric processes. These findings not only enhance our understanding of the complex mechanisms of atmospheric oxidation but also highlight the critical importance of accurately estimating MT emissions in simulating the behavior of atmospheric processes.

With newly discovered biogenic MT sources[27,51–54] and the challenge of modeling co-occurring environmental drivers on the biosphere, our study highlights the need for more process-oriented research of biosphere-atmosphere interactions, particularly in tropical, pan-Arctic, grassland, and agricultural ecosystems. As the effects of climate change intensify, biogenic VOC emissions from global vegetation will play a crucial role in evaluating the health of ecosystems and influencing 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 available for tropical plants were insufficient for conducting reliable statistical analyses. To address this limitation, we included three additional studies that reported relevant data. The timelines for temperature and MT emission were extracted from published plots in Jardine et. al[4], and Langford et al.[55] using a web-based tool (WebPlotDigitizer; https://automeris.io). Data from the third study were directly obtained from Yáñez-Serrano et al.[56] 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. The studies used in this meta-analysis are listed in the Supplementary Information (Supplementary References list).

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.

We categorized the main conclusions of each study into three groups based on the factors driving monoterpene emissions: temperature alone, temperature and light, and temperature in combination with other environmental factors such as soil moisture and relative humidity. The experimental techniques employed in these studies varied, including different environmental conditions, methods for determining emission rates, VOC sampling, and chemical analyses (Supplementary Fig. 2). However, these variables were comprehensively reported in all studies, making the annotation process relatively straightforward. We annotated the β data points based on the method that was used to collect the majority of the data, unless there were two different sampling methods used. In such cases, we used the annotation method that corresponded to the primary sampling method. Additionally, when gas chromatography was used for measuring individual monoterpenes, we annotated the regression fits accordingly.

The year, month, and season of the experiments were assessed. Typically, the experiments were conducted within a single year. In cases where the experiments were carried out over two consecutive years and 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, we annotated the year of the first observation. For experiments where the year was not reported (typical for laboratory experiments), 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)[13]. 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. Considering that a low R^2 indicates a weak to non-existent relationship between MT emissions and temperature, 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 analysis are comprised of 559 experimental values.

The remaining data were analyzed using random-effects models, which are commonly used in meta-analyses[57–59]. 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[60]. 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, E₃₀). 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[61] (https://try-db.org, last accessed 8 June 2022). The data sets used were 1-km[62] maps scaled up from trait data measured in situ. The LAI data we used were the monthly climatological data from the ORNL DAAC database[63]. 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[64], and effectively addresses inter-correlated parameters. 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[65,66].

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[67]. 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[68]. The performance is taken from each combination of the hyper-parameters tuning with the grid search method with cross-validation (CV)[69]. To avoid bias in data selection, we applied 10-fold CV[69,70], 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. 22).

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[71,72]. 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[72–74]. 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[73,75]. The set-up used in this simulation was identical to that used by Pozzer et al.[18], 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[76]) 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.[18] 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[76–78]. Processes of aerosol microphysics and gas/aerosol partitioning were simulated using the GMXe submodel[79]. 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[80], 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[81]. The composition and atmospheric evolution of organic aerosol compounds were simulated using the ORACLE submodel, which represents classes of the volatility of organics by their effective saturation concentrations[82]. The biogenic emissions of non-methane VOCs were calculated online using MEGAN[13]. 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[83–86].

The dynamics were not nudged, unlike the study by Pozzer et al.[18], 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.

Data availability

Code availability
The code used for including a plant functional type-dependent temperature sensitivity can be found in the Supplementary Information.

References
1. Loreto, F. & Schnitzler, J.-P. Abiotic stresses and induced BVOCs. *Trends Plant Sci.* **15**, 154–166 (2010).
2. Niinemets, Ü. Mild versus severe stress and BVOCs: thresholds, priming and consequences. *Trends Plant Sci.* **15**, 145–153 (2010).
3. Loreto, F., Pinelli, P., Manes, F. & Kollist, H. Impact of ozone on monoterpene emissions and evidence for an isoprene-like antioxidant action of monoterpenes emitted by Quercus ilex leaves. *Tree Physiol.* **24**, 361–367 (2004).
4. Jardine, K. J. et al. Monoterpene ' thermometer' of tropical forest-atmosphere response to climate warming: Monoterpene ' thermometer'. *Plant Cell Environ.* **40**, 441–452 (2017).
5. Rosenkranz, M., Chen, Y., Zhu, P. & Vlot, A. C. Volatile terpenes – mediators of plant-to-plant communication. *Plant J.* **108**, 617–631 (2021).
6. Peñuelas, J. & Staudt, M. BVOCs and global change. *Trends Plant Sci.* **15**, 133–144 (2010).
7. Riipinen, I. et al. The contribution of organics to atmospheric nanoparticle growth. *Nat. Geosci* **5**, 453–458 (2012).
8. Junninen, H. et al. Terpene emissions from boreal wetlands can initiate stronger atmospheric new particle formation than boreal forests. *Commun. Earth. Environ.* **3**, 93 (2022).
9. Kavouras, I. G., Mihalopoulos, N. & Stephanou, E. G. Formation of atmospheric particles from organic acids produced by forests. *Nature* **395**, 683–686 (1998).
10. Zhang, H. et al. Monoterpenes are the largest source of summertime organic aerosol in the southeastern United States. *Proc. Natl. Acad. Sci. USA* **115**, 2038–2043 (2018).
11. Petäjä, T. et al. Influence of biogenic emissions from boreal forests on aerosol–cloud interactions. *Nat. Geosci.* **15**, 42–47 (2022).
12. Benjamin, M. T. & Winer, A. M. Estimating the ozone-forming potential of urban trees and shrubs. *Atmos. Environ.* **32**, 53–68 (1998).
13. 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).
14. Guenther, A. et al. Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys.* **6**, 3181–3210 (2006).
15. Staudt, M. & Lhoutellier, L. Monoterpene and sesquiterpene emissions from Quercus coccifera exhibit interacting responses to light and temperature. *Biogeosciences* **8**, 2757–2771 (2011).
16. Blonder, B. & Michaletz, S. T. A model for leaf temperature decoupling from air temperature. *Agri. Forest Meteorol.* **262**, 354–360 (2018).
17. 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).
18. 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).
19. Makowski, D. et al. *Systematic review of meta-analyses to assess the impacts of farming practices - A methodological framework.* https://osf.io/byuw9. https://doi.org/10.31219/osf.io/byuw9 (2021).
20. Hellén, H. et al. Sesquiterpenes dominate monoterpenes in northern wetland emissions. *Atmos. Chem. Phys.* **20**, 7021–7034 (2020).
21. Angot, H. et al. Biogenic volatile organic compound ambient mixing ratios and emission rates in the Alaskan Arctic tundra. *Biogeosciences* **17**, 6219–6236 (2020).
22. Li, T., Baggesen, N., Seco, R. & Rinnan, R. Seasonal and diel patterns of biogenic volatile organic compound fluxes in a subarctic tundra. *Atmos. Environ.* **292**, 119430 (2023).
23. Seco, R. et al. Strong isoprene emission response to temperature in tundra vegetation. *Proc. Natl. Acad. Sci. USA.* **119**, e2118014119 (2022).
24. Rinnan, R. et al. Separating direct and indirect effects of rising temperatures on biogenic volatile emissions in the Arctic. *Proc. Natl. Acad. Sci. USA.* **117**, 32476–32483 (2020).
25. Rantanen, M. et al. The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth Environ.* **3**, 168 (2022).
26. Perkins-Kirkpatrick, S. E. & Lewis, S. C. Increasing trends in regional heatwaves. *Nat. Commun.* **11**, 3357 (2020).
27. Bourtsoukidis, E. et al. Strong sesquiterpene emissions from Amazonian soils. *Nat. Commun.* **9**, 2226 (2018).
28. 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).
29. 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).
30. Saunois, M. et al. The Global Methane Budget 2000–2017. *Earth Syst. Sci. Data* **12**, 1561–1623 (2020).
31. Atkinson, R. Atmospheric chemistry of VOCs and NOx. *Atmos. Environ.* **34**, 2063–2101 (2000).
32. 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).
33. Grote, R. et al. Process-based modelling of isoprenoid emissions from evergreen leaves of Quercus ilex (L.). *Atmos. Environ.* **40**, 152–165 (2006).
34. 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).
35. Tarvainen, V. et al. Temperature and light dependence of the VOC emissions of Scots pine. *Atmos. Chem. Phys.* **5**, 989–998 (2005).
36. Fares, S. et al. Biogenic emissions from Citrus species in California. *Atmos. Environ.* **45**, 4557–4568 (2011).
37. Lim, J.-H. et al. Seasonal variations of monoterpene emissions from Pinus densiflora in East Asia. *Chemosphere* **73**, 470–478 (2008).
38. 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). *Atmos. Environ.* **32**, 1825–1833 (1998).
39. Schade, G. W., Goldstein, A. H. & Lamanna, M. S. Are monoterpene emissions influenced by humidity? *Geophys. Res. Lett.* **26**, 2187–2190 (1999).
40. Staudt, M. et al. Seasonal and diurnal patterns of monoterpene emissions from Pinus pinea (L.) under field conditions. *Atmos. Environ.* **31**, 145–156 (1997).
41. Laffineur, Q. et al. Impact of diffuse light on isoprene and monoterpene emissions from a mixed temperate forest. *Atmos. Environ.* **74**, 385–392 (2013).
42. 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).
43. 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).
44. 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).
45. 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. *Sci. Total Environ.* **565**, 730–741 (2016).
46. 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).
47. 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).
48. Nagalingam, S., Seco, R., Kim, S. & Guenther, A. Heat stress strongly induces monoterpene emissions in some plants with specialized terpenoid storage structures. *Agri. Forest Meteorol.* **333**, 109400 (2023).
49. Bäck, J. et al. Chemodiversity of a Scots pine stand and implications for terpene air concentrations. *Biogeosciences* **9**, 689–702 (2012).
50. Stone, D., Whalley, L. K., Heard, D. E. & Tropospheric, O. H. and HO2 radicals: field measurements and model comparisons. *Chem. Soc. Rev.* **41**, 6348 (2012).

51. Vanhatalo, A. et al. Scots Pine Stems as Dynamic Sources of Monoterpene and Methanol Emissions. *Front. For. Glob. Change* **2**, 95 (2020).
52. 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).
53. 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).
54. Byron, J. et al. Chiral monoterpenes reveal forest emission mechanisms and drought responses. *Nature* **609**, 307–312 (2022).
55. 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).
56. Yáñez-Serrano, A. M. et al. Heat stress increases the use of cytosolic pyruvate for isoprene biosynthesis. *J. Exp. Botany* **70**, 5827–5838 (2019).
57. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis.* (John Wiley & Sons, Ltd, 2009). https://doi.org/10.1002/9780470743386.
58. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The Meta-Analysis Of Response Ratios In Experimental Ecology. *Ecology* **80**, 1150–1156 (1999).
59. 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). https://doi.org/10.1007/978-94-024-1696-1.
60. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Soft.* **67**, (2015).
61. Moreno-Martínez, Á. et al. A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sens. Environ.t* **218**, 69–88 (2018).
62. Camps-Valls, G. *et al.* Global trait maps at 1km resolution. https://doi.org/10.17871/TRY.60 (2020)
63. MAO, J. & YAN, B. Global Monthly Mean Leaf Area Index Climatology, 1981-2015. 6.581285 MB https://doi.org/10.3334/ORNLDAAC/1653 (2019).
64. Samek, W., Wiegand, T. & Müller, K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *arXiv*:1708.08296 [cs, stat] (2017).
65. 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. *Sci. Total Environ.* **835**, 155368 (2022).
66. 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).
67. Zhou, J. et al. Random Forests and Cubist Algorithms for Predicting Shear Strengths of Rockfill Materials. *Appl. Sci.* **9**, 1621 (2019).
68. Wang, Y. & Witten, I. *Induction of model trees for predicting continuous classes.* (University of Economics, Faculty of Informatics and Statistics, 1997).
69. Kuhn, M. & Johnson, K. *Applied Predictive Modeling.* (Springer, 2013).
70. 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 Sens.* **9**, 398 (2017).
71. 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).
72. Jöckel, P. et al. Development cycle 2 of the Modular Earth Submodel System (MESSy2). *Geosci. Model Dev.* **3**, 717–752 (2010).
73. 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).
74. 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).
75. Pozzer, A. et al. Observed and simulated global distribution and budget of atmospheric C2-C5 alkanes. *Atmos. Chem. Phys.* **10**, 4403–4422 (2010).
76. Tsimpidi, A. P., Karydis, V. A., Pozzer, A. & Lelieveld, J. ORACLE (v1.0): module to simulate the organic aerosol composition and evolution in the atmosphere. *Geosci. Model Dev.* **7**, 3153–3172 (2014).
77. Sander, R. et al. The atmospheric chemistry box model CAABA/MECCA-3.0. *Geosci. Model Dev.* **4**, 373–380 (2011).
78. Lelieveld, J., Gromov, S., Pozzer, A. & Taraborrelli, D. Global tropospheric hydroxyl distribution, budget and reactivity. *Atmos. Chem. Phys.* **16**, 12477–12493 (2016).
79. 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).
80. 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).
81. 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).
82. 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).
83. 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).
84. Pozzer, A. et al. Effects of business-as-usual anthropogenic emissions on air quality. *Atmos. Chem. Phys.* **12**, 6915–6937 (2012).
85. 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).
86. Karydis, V. A. et al. Global impact of mineral dust on cloud droplet number concentration. *Atmos. Chem. Phys.* **17**, 5601–5621 (2017).

Acknowledgements

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., A.P., P.C., J.L., J.P., J.S and J.W. performed research. All authors contributed to editing the article.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-01175-9.

Correspondence and requests for materials should be addressed to Efstratios Bourtsoukidis.

Peer review information *Communications Earth & Environment* thanks Andrea Ghirardo, Thomas Holst, Tao Li, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Clare Davis. A peer review file is available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.