

Improved modelling of biogenic emissions in human-disturbed forest edges and urban areas

Received: 9 April 2025

Accepted: 18 August 2025

Published online: 28 August 2025



Yanli Zhang¹, Haofan Ran^{1,2}, Alex Guenther³✉, Qiang Zhang⁴✉, Christian George⁵, Wahid Mellouki⁶, Guoying Sheng¹, Ping'an Peng¹ & Xinming Wang^{1,2}✉

Biogenic volatile organic compounds (BVOCs) are critical to biosphere-atmosphere interactions, profoundly influencing atmospheric chemistry, air quality and climate, yet accurately estimating their emissions across diverse ecosystems remains challenging. Here we introduce GEE-MEGAN, a cloud-native extension of the widely used MEGAN2.1 model, integrating dynamic satellite-derived land cover and vegetation within Google Earth Engine to produce near-real-time BVOC emissions at 10–30 m resolution, enabling fine-scale tracking of emissions in rapidly changing environments. GEE-MEGAN reduces BVOC emission estimates by 31% and decreases root mean square errors by up to 48.6% relative to MEGAN2.1 in human-disturbed forest edges, and reveals summertime BVOC emissions up to 25-fold higher than previous estimates in urban areas such as London, Los Angeles, Paris, and Beijing. By capturing fine-scale landscape heterogeneity and human-driven dynamics, GEE-MEGAN significantly improves BVOC emission estimates, providing crucial insights to the complex interactions among BVOCs, climate, and air quality across both natural and human-modified environments.

Biogenic volatile organic compounds (BVOCs) are critical drivers of biosphere-atmosphere interactions, significantly influencing ozone formation, secondary organic aerosol (SOA) production^{1,2}, and broader atmospheric chemistry^{3,4}. These processes have profound implications for air quality, climate, and human health. However, despite their significance, accurately quantifying BVOC emissions across diverse ecosystems remains a persistent challenge. This complexity stems from the intricate interplay of biological, climatic, and anthropogenic factors^{5,6}, compounded by the dynamic nature of vegetation and environmental conditions.

Existing BVOC models, such as the widely used Model of Emissions of Gases and Aerosols from Nature (MEGAN2.1)⁷, have made

significant strides by providing a comprehensive framework for emission estimation. However, these models often rely on coarse spatial data and static parameters, which limits their ability to capture the heterogeneity of urban and natural landscapes, leading to inconsistent and inaccurate BVOC emission estimates across cities and regions^{8–12}. Such limitations are particularly problematic in rapidly changing environments, such as urban areas^{13–15}, where vegetation types, land use, and greening strategies are highly dynamic^{16–18}. Furthermore, computational constraints hinder the application of high-resolution models over large geographic scales, leaving critical gaps in our understanding of BVOCs' impacts on air quality and climate^{19,20}.

¹State Key Laboratory of Advanced Environmental Technology, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou, China.

²University of Chinese Academy of Sciences, Beijing, China. ³Department of Earth Science System, University of California, Irvine, CA, USA. ⁴Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing, China. ⁵Univ Lyon, Université Claude Bernard Lyon 1, CNRS, IRCELYON, Villeurbanne, France. ⁶Institut de Combustion, Aérothermique, Réactivité Environnement (ICARE), CNRS, Orléans, France. ✉e-mail: alex.guenther@uci.edu; qiangzhang@tsinghua.edu.cn; wangxm@gig.ac.cn

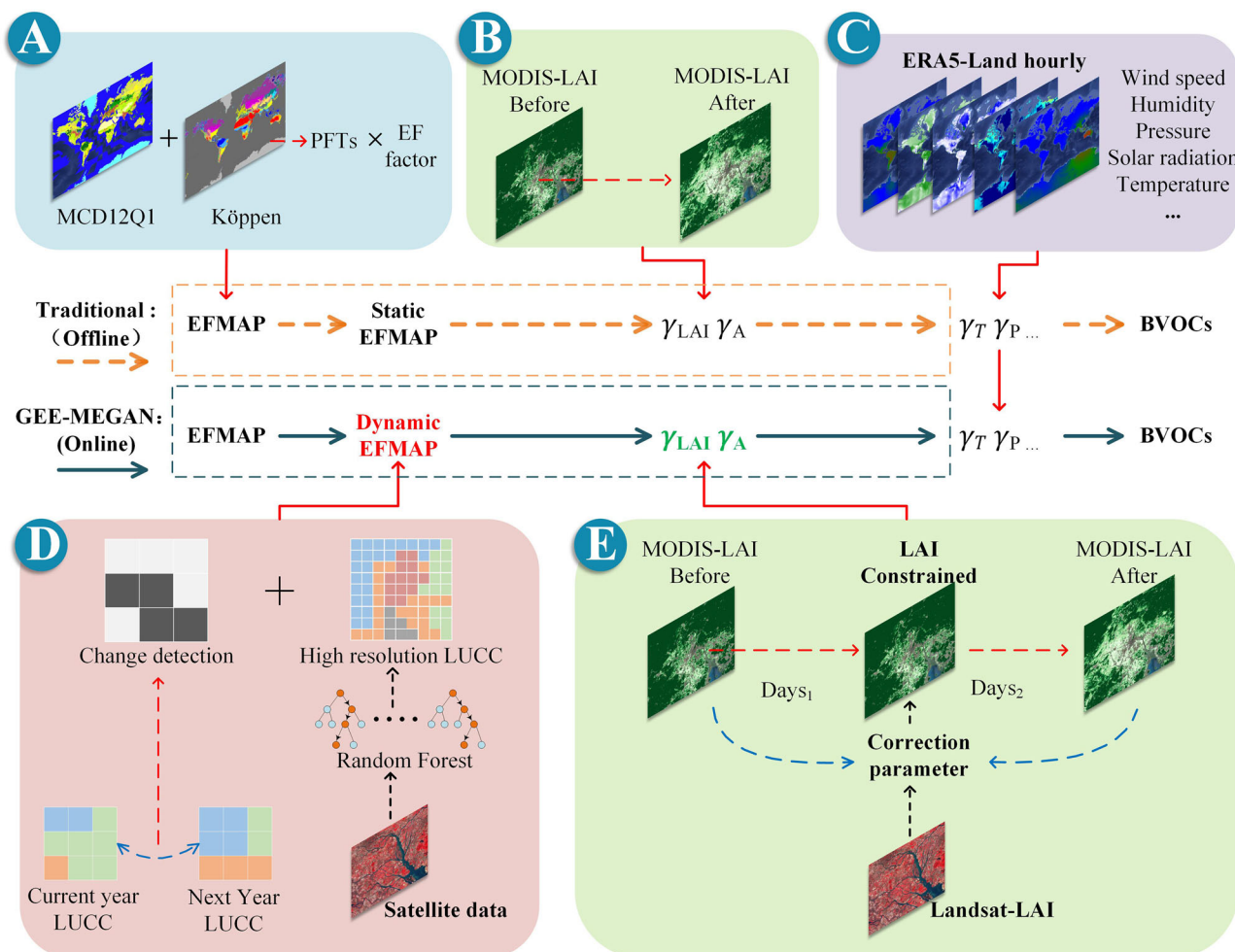


Fig. 1 | Comparison of GEE-MEGAN and traditional biogenic volatile organic compounds (BVOC) calculation processes. Each panel illustrates a set step in the calculation of BVOC emissions. **A** Calculate the standard emission factor grid map (EFMAP) in MEGAN2.1 based on plant functional types (PFTs) and emission factor (EF); **B** Derivation of leaf area index (γ_{LAI}) and environmental activity factor (γ_A) in

MEGAN2.1; **C** Calculation of temperature (γ_T) and light (γ_P) activity factors using meteorological data in MEGAN 2.1; **D** Dynamic standard emission factor grid map construction in GEE-MEGAN, incorporating land use/cover change (LUCC) and wildfire data. **E** Updated calculation of γ_{LAI} and γ_A in GEE-MEGAN using high-resolution LAI data from Landsat.

To address these challenges, we introduce GEE-MEGAN, an innovative modeling framework (Fig. 1 and Fig. S1 in the Supplementary Information) that combines the strengths of MEGAN2.1 with the cloud-based computational power of Google Earth Engine (GEE)^{21,22}. By integrating multi-source remote sensing data (e.g., MODIS, Landsat, Sentinel) and machine learning^{23,24}, GEE-MEGAN dynamically refines critical model inputs (Fig. 1) such as leaf area index (LAI) and plant functional types (PFTs), enabling near-real-time, high-resolution BVOC emission estimates at spatial scales ranging from 10–30 m locally to global domains. These advancements in accuracy, resolution, and computational efficiency position GEE-MEGAN as a transformative tool for advancing research on biosphere-atmosphere interactions, particularly in understanding the complex feedback loops between BVOCs, climate and air quality.

Results and discussion

Better prediction of observed LAI and BVOC fluxes

GEE-MEGAN accurately captured both daily and monthly emission trends at regional scales compared to conventional models at identical spatial resolutions. For example, hourly isoprene emission rates for August 2010 across South America showed that daily average emissions closely aligned with trends derived from the inventory from MEGAN model coupled with the Model for Hydrocarbon emissions by

the CANopy (MOHYCAN) and constraint by Ozone Monitoring Instrument (OMI) Formaldehyde (HCHO) observation (MEGAN2.1-MOHYCAN-OMI, details in Fig. S2a)²⁵, achieving a Spearman Rank Correlation Coefficient of $R = 0.83$ ($p < 0.001$, $N = 31$). Monthly regional emissions also exhibited minimal discrepancies when compared to the MEGAN default and MEGAN2.1 models for South America^{7,25,26}, with differences ranging from 6.4% to 12% (Fig. S2c–h and Table S1). These results validate GEE-MEGAN's consistency with the core dynamics of MEGAN2.1 across large temporal and spatial scales. Notably, GEE-MEGAN consistently produced isoprene emission estimates approximately 50% higher than those derived from OMI-MEGAN and Global Ozone Monitoring Experiment-2 (GOME2) based MEGAN model^{25,27}. This discrepancy is consistent with recent studies suggesting that OMI-MEGAN underestimates emissions due to biases in the OMI formaldehyde (HCHO) column data²⁸, which likely skew emission retrievals downward.

GEE-MEGAN utilizes high-resolution Landsat data for LAI estimation, outperforming conventional approaches that rely on coarser datasets, such as MODIS (Fig. 1E vs. Fig. 1B). When compared to observed LAI data (<https://calvalportal.ceos.org/lpv-direct-v2.1>), GEE-MEGAN's Landsat-based LAI estimates reduced the mean absolute errors (MAE) by 12.5% at 500-m resolution and by 11.7% at a 3-km resolution relative to MODIS-based estimates (Fig. S3), and reduced

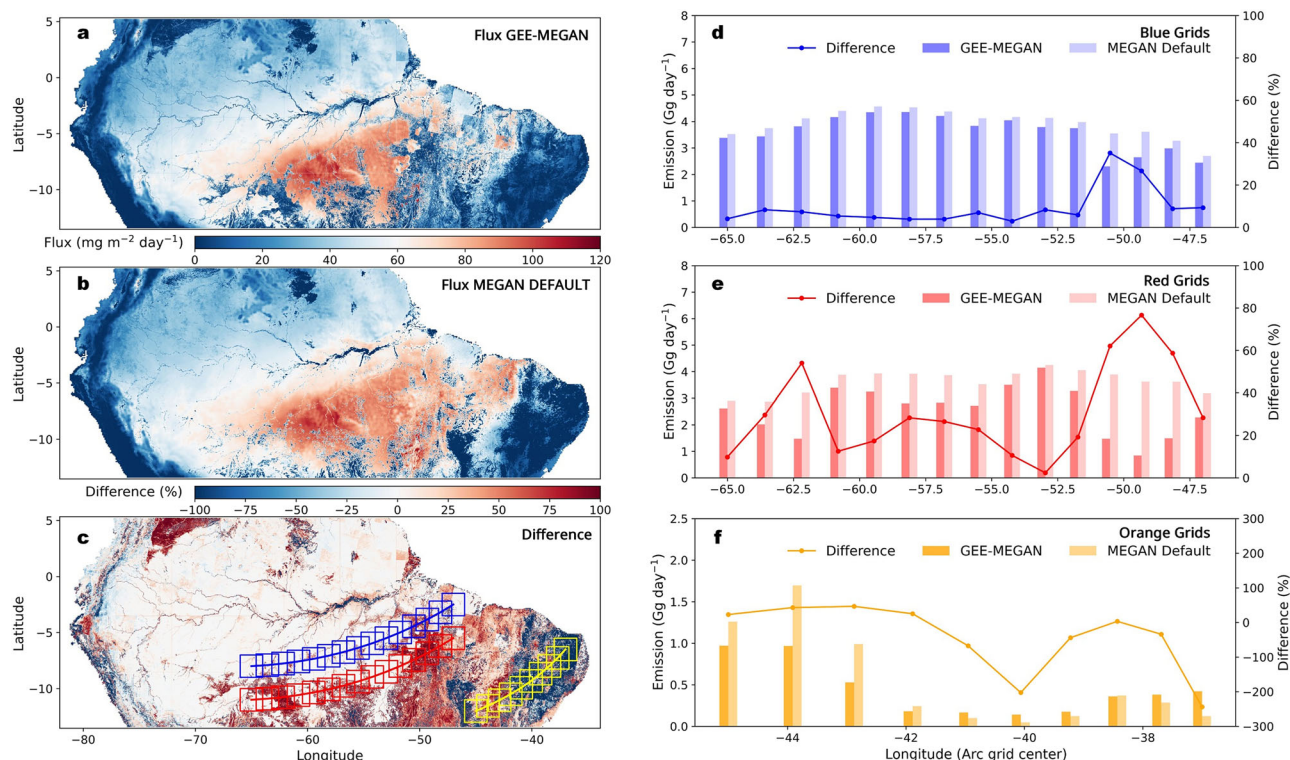


Fig. 2 | Isoprene emission simulations over the Amazon region in August 2019 using GEE-MEGAN and MEGAN-default. **a** Isoprene emission at a 30-m resolution for the Amazon region using the GEE-MEGAN, which incorporated a total of 2,211 Landsat scenes. **b** Simulation results obtained with the MEGAN-default. **c** Relative

difference computed as [(b) – (a)]/(b). **d–f** Total isoprene emissions and their differences for the 2° × 2° (220 km × 220 km) grids, uniformly distributed along the red line (Arc of Deforestation), blue line (inner Amazon edge) and yellow line (with reforestation) as indicated in (c).

the RMSE by 9.1% at 500-m resolution and by 3.6% at a 3-km resolution. Based on these refined LAI inputs, GEE-MEGAN further improves BVOC emission estimates, as evaluated in Table S2. In comparison to observed BVOC fluxes across diverse ecosystems in North America, South America, Europe, and East Asia^{29–33} (Table S2), GEE-MEGAN outperformed MEGAN2.1, reducing RMSE by 21.6–48.6% and improving correlation coefficients (R) from 0.83 to 0.90. The most significant improvements were observed in regions with timely captured dynamic canopy changes.

Enhanced spatial resolution in original and disturbed forests

GEE-MEGAN captures greater spatial heterogeneity in BVOC emissions (Fig. 2a), especially at forest edges affected by human activities, containing emission hotspots and dead-zones, which are often fragmented or averaged out in coarser-resolution datasets^{34,35}. Unlike conventional models that rely on static parameters and coarse spatial data, GEE-MEGAN dynamically integrates climatic norms and land use/cover change (LUCC) data with high-resolution satellite observations (Fig. 1D, Table S3), producing a more accurate and responsive emission factor map (EFMAP) capable of detecting abrupt changes.

In the Amazon rainforest, a critical BVOC source contributing over a quarter of global isoprene emission³⁶, GEE-MEGAN reveals substantial discrepancies in BVOC emissions when compared to the MEGAN default with LUCC data from MODIS. Coarser models often fail to account for critical impacts of forest loss, introducing potential biases of Amazonian BVOC emissions. During the high-BVOC-emitting month of August in 2019, for example, GEE-MEGAN estimated 13% lower Amazonian isoprene emissions than the MEGAN default (Fig. 2a, b). While the discrepancies of 3.0–8.3% occurred deep inside the forest, they were especially pronounced at forest edges by forest degradation or deforestation (Fig. 2c). In these areas, the MEGAN default model overestimated emissions by up to 77% (average 31%) in the

fifteen 2° × 2° grid cells along the Arc of Deforestation^{35,37} (red line in Fig. 2c; e), while the overestimation became smaller, ranging 3–35% (average 9%) in the fifteen 2° × 2° grids closer to the forest interior (blue line in Fig. 2c; d). Moreover, the MEGAN default model underestimated isoprene emission by up to 250% in some southeastern Brazil grids (yellow line in Fig. 2c; f), where recent reforestation or post-fire vegetation recovery occurred^{38,39}.

In paired 2° × 2° grids along the Arc of Deforestation (red grids) and approximately ~300 km into the inner forest (blue grids, Fig. 2c), the MEGAN default model showed an average isoprene emission difference of 6% (maximum 23%). In contrast, GEE-MEGAN estimated an average difference of 35% on average (maximum 68%), demonstrating its strength in capturing emission heterogeneity. Given that each grid cell represents an area of nearly 50,000 km², such large under- or overestimation of reactive BVOC emissions can significantly misrepresent their roles and broader environmental impacts, especially at local and regional scales.

High-resolution BVOC emissions modeling in urban areas

GEE-MEGAN effectively captures the spatial variability of urban green spaces, enabling high-resolution BVOC emission modeling at 10–30 m resolution, resampled to 100 m. This ability is showcased in Fig. 3, where BVOC emissions for July 2019 in London, Paris, Los Angeles, and Beijing are quantified at finer urban scales by GEE-MEGAN, compared to the coarser 500-m resolution of the MEGAN default model. In Paris, the model identified key hotspots, and further evaluation using the BOXMOX chemical box model showed that GEE-MEGAN emissions achieved better agreement with observed isoprene concentrations than the default MEGAN model under identical ERA5 meteorological conditions⁹ (Fig. S4, Note S1 and Table S4). This result provides strong evidence that enhanced spatial detail in emission inputs improves model-observation consistency.

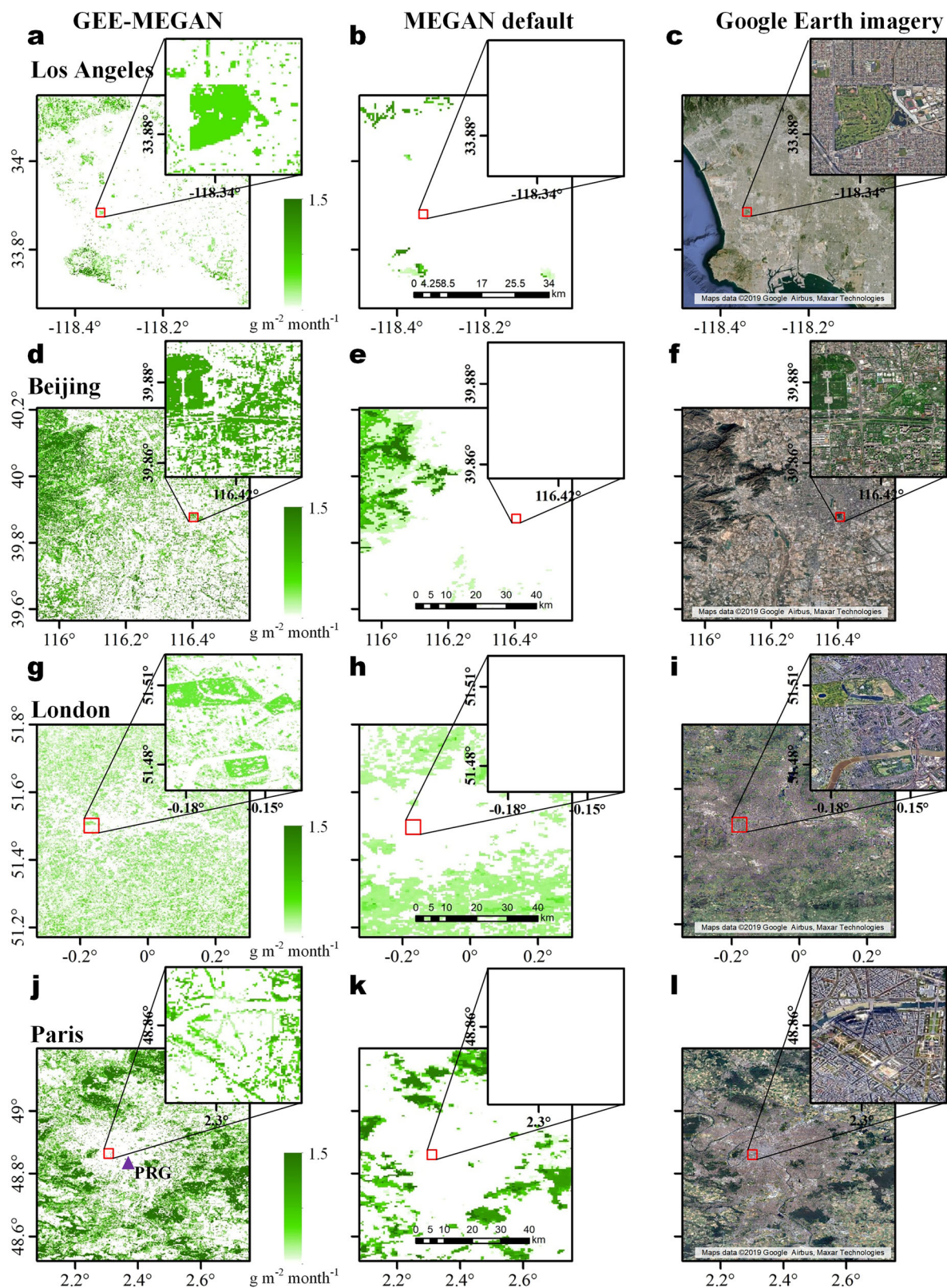


Fig. 3 | Spatial distribution of isoprene emissions in Paris, Los Angeles, London, and Beijing simulations for July 2019. a, d, g, j Monthly mean isoprene emissions modeled by GEE-MEGAN, resampled to 100 m resolution; each panel includes an upper-right inset showing a magnified view of a representative urban sub-region. **b, e, h, k** Corresponding 500 m resolution isoprene fluxes estimated by the default MEGAN model, plotted over the same spatial extents and with identical inset areas

for comparison. **c, f, i, l** Google Earth satellite imagery for the same regions (Maps data ©2019 Google). All panels in each city share a common geographic frame. The Paris Rive Gauche (PRG) site, marked in (j), is located on the seventh floor of the Lamarck B building at Université Paris Cité (30 m above ground level) and is equipped with Proton-Transfer-Reaction Mass Spectrometry (PTR-MS) for ambient isoprene monitoring.

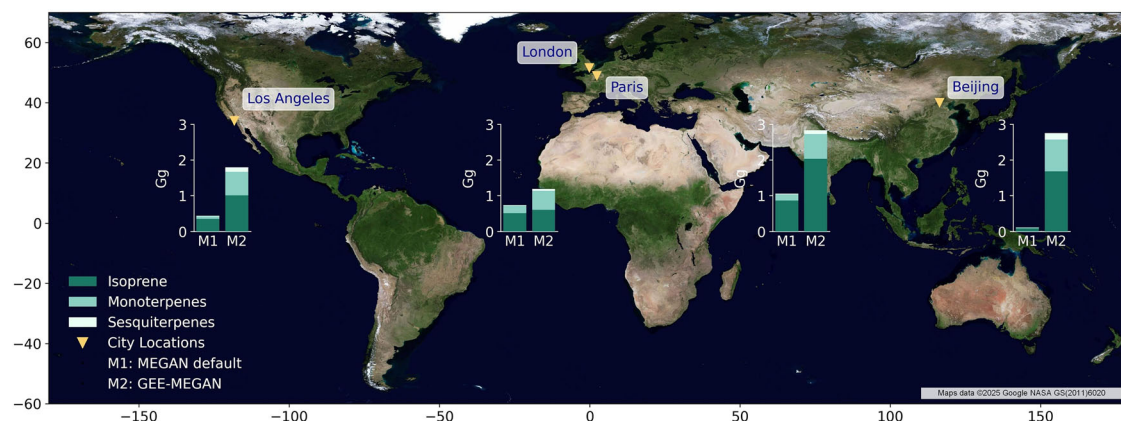


Fig. 4 | Biogenic volatile organic compounds (BVOC) compositions in Paris, Los Angeles, London, and Beijing as modeled by MEGAN default and GEE-MEGAN. The base map is derived from Google Earth remote sensing satellite imagery (Maps data ©2025 Google). The bar depict the emissions of isoprene, monoterpenes, and sesquiterpenes for each city, focus on regions with nighttime light percentages

above 80% (nighttime light data from the National Polar-orbiting Partnership, Visible Infrared Imaging Radiometer Suite). This threshold was used to define urban boundaries consistently across cities. For each city, the same spatial extent was used for both models to ensure direct comparison.

The high spatial heterogeneity captured by GEE-MEGAN, particularly in emission hotspots and high-value zones that are often fragmented or averaged in coarser-resolution models, consistently leads to higher emission estimates. For cities like London, Los Angeles, Paris, and Beijing, GEE-MEGAN estimated BVOC emissions for July 2019 to be 1.6–25 times higher than those predicted by the MEGAN default (Fig. 4). In Beijing, the larger discrepancy arises from fragmented vegetation being poorly resolved by lower-resolution land-cover inputs. GEE-MEGAN identifies an additional 2800.54 km² of vegetated area and a mean patch size of 0.67 km² within the same domain (Table S5), highlighting the critical role of spatial vegetation detail in driving urban BVOC emissions. This is primarily due to GEE-MEGAN's incorporation of fine-scale vegetation components, such as isolated trees or smaller trees and shrubs^{40,41}, which are often overlooked in coarser models (Table S6).

A more specific comparison of BVOC emissions in New York on July 2, 2018 revealed that GEE-MEGAN estimated total emissions at 27 tons, nearly 30% higher than the 21 tons predicted by NYC-MEGAN¹². This discrepancy arises from GEE-MEGAN's superior ability to account for the continuous urban-to-suburban coverage of green spaces, addressing limitations in vegetation coverage used in NYC-MEGAN.

GEE-MEGAN provides a more accurate representation of the roles of BVOC emissions in urban atmospheric chemistry. In West Los Angeles, while the MEGAN default model estimated zero BVOC emissions for most of the 251 communities in August 2019 (Fig. 5a), using compound-specific OH reactivity (OHR) coefficients and SOA yield factors²⁰, GEE-MEGAN identified BVOC emissions that were, on average, about 23.0 times higher (Fig. 5b), along with 23.4 times greater OHR and 25.1 times higher SOA formation potential (SOAFP) (Fig. S5). More recent state-of-the-art airborne flux measurements in Los Angeles during summer 2021^{19,20} revealed that biogenic terpenoids accounted for ~15% of VOC emissions, but contributed over 50% of OHR and ozone/SOA formation potentials. In comparison, GEE-MEGAN estimated these BVOCs contributed to 3.5% of VOCs, 26.3% of OHR, and 29.4% of SOAFP, while the MEGAN default model severely underestimated these contributions (1% of VOCs, ~15% of OHR and SOAFP; Fig. S6). Interestingly, significant gaps still exist between observations and GEE-MEGAN estimates. This could be due to biased emission factors or it may reflect the substantial contributions of volatile chemical products (VCPs) to terpenoid emissions in Los Angeles. This suggests that more accurate BVOC emission estimates also enable better quantification of the contribution of temperature-

sensitive VCPs⁴² to total terpenoid emissions, and help disentangle their temperature-driven variability²⁰.

Implications for air quality and climate

GEE-MEGAN provides a solid foundation for investigating the complex interplay between BVOCs and climate by offering near-real-time high-resolution emission modeling. BVOC emissions from forests are the largest source of reactive organic gases released into the atmosphere^{43,44}, profoundly influencing radiative ozone (a potent greenhouse gas in the lower atmosphere) formation, aerosol growth, cloud formation, and atmospheric oxidative capacity, which controls the lifetime of methane (CH₄) and thereby radiative forcing^{4,45–47}. GEE-MEGAN captures BVOC emissions more accurately and efficiently across heterogeneous landscapes, particularly at forest edges where human disturbance is more pronounced. While GEE-MEGAN and the default MEGAN model yield similar isoprene emissions in pristine forests, large discrepancies emerge at forest edges, such as along the Brazilian Amazon's Arc of Deforestation³⁵. For example, during the high-BVOC-emitting month of August 2019, emission discrepancies are more pronounced in areas affected by forest degradation or deforestation (Fig. 2c), highlighting the need for high-resolution emission estimates, especially along global forest edges. Although pathways by which BVOCs influence climate are more complex than currently recognized⁴⁸, emission estimation with high spatiotemporal resolution as provided by GEE-MEGAN is a prerequisite for better delineating human perturbations to atmospheric compositions and climatic impacts of BVOCs at local, regional and even global scales.

GEE-MEGAN's high-resolution, near-real-time emission modeling capabilities can help refine climate predictions by providing insights into the strong and complex feedbacks between climate and BVOC³⁶. As temperature-driven BVOC emissions increase under global warming^{20,49}, GEE-MEGAN can offer critical insights into how ecosystems respond to and feedback on climate change^{43,50–52} by providing more accurate and timely estimates of BVOC emissions driven by both natural and anthropogenic forces, such as deforestation and afforestation. Additionally, GEE-MEGAN supports the assessments of climate resilience strategies^{53–55}, such as urban greening and afforestation/reforestation initiatives^{56–58}, ensuring that these efforts balance the cooling benefits of increased vegetation with minimization of BVOC-related warming and pollution. This dual focus on environmental and climatic impacts optimizes sustainability measures, helping mitigate air quality degradation while maximizing carbon sequestration and heat mitigation benefits.

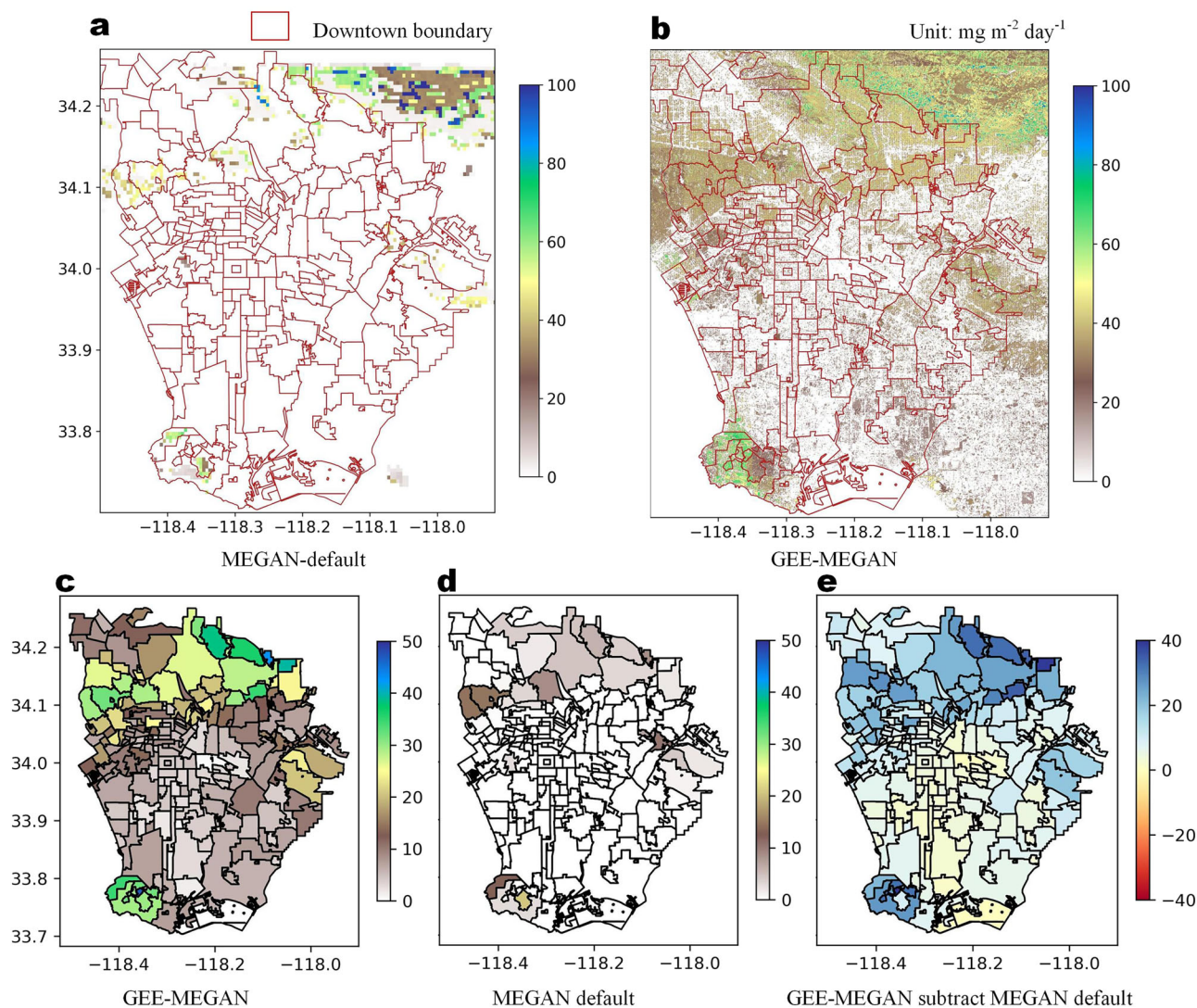


Fig. 5 | Comparisons of biogenic volatile organic compounds (BVOC) emissions and community-level spatial distributions modeled by MEGAN default and GEE-MEGAN for West Los Angeles during August 2019. a Total BVOC emissions modeled by MEGAN default at a resolution of 500 m. **b** Total BVOC emissions modeled by GEE-MEGAN at a finer resolution of 30 m. Red boundary lines indicate

community boundaries within West Los Angeles (<https://lacounty.maps.arcgis.com/home/index.html>). **c** Community-average daily BVOC emissions modeled by GEE-MEGAN. **d** Community-average daily BVOC emissions modeled by MEGAN default. **e** Differences in community-average daily emissions between GEE-MEGAN and MEGAN default, calculated as (c) minus (d).

GEE-MEGAN also equips urban planners and even non-experts with powerful tools to assess the air quality impacts of urban greening initiatives. As BVOCs increasingly influence urban atmospheric chemistry due to declining anthropogenic emissions and expanding greening spaces^{20,42,59}, accurately capturing their emissions in urban areas, however, becomes challenging because of complex vegetation patterns and dynamic land-use changes. By resolving emissions at finer spatial scale, down to the street and community level, GEE-MEGAN substantially outperforms the 500-m resolution MEGAN default model. As illustrated by the examples from Los Angeles discussed above, GEE-MEGAN provides a more precise understanding of the contributions of biogenic emissions to total urban terpenoid emissions, which is essential for evaluating their roles in air pollution, including ozone and SOA formation^{19,20,42}. This ability enables more nuanced exposure assessments by accounting for interactions between biogenic and anthropogenic sources across diverse finer-scale microenvironments in densely populated areas^{58,59}. Additionally, GEE-MEGAN's near-real-time modeling capability supports adaptive air-quality management in response to events such as heat waves^{16,60–62} or rapid urbanization⁶³, ensuring that green infrastructure projects

maximize environmental benefits while minimizing unintended air pollution consequences.

Despite its advancements, GEE-MEGAN does not yet fully account for complex environmental factors influencing BVOC emissions⁶⁴, such as soil moisture, nitrogen availability, evapotranspiration, and biotic stresses. Integrating these factors through GEE's multisource, multi-scale data capabilities and advanced machine learning techniques could further refine its predictions. Although ERA5-Land hourly meteorology (~11 km) was used in this study to ensure global applicability and consistency, our results indicate that improving the spatial representation of vegetation yields substantial gains in BVOC emission accuracy. Nevertheless, incorporating high-resolution meteorological reanalysis products or AI-driven weather prediction models^{65,66} could further enhance GEE-MEGAN's capacity to simulate urban-scale processes, offering deeper insights into BVOC-related dynamics in air quality and climate change.

In summary, GEE-MEGAN bridges critical gaps in BVOC emission modeling by combining scalability, precision, and real-time capabilities. These features make it a transformative tool for improving our understanding of roles played by biogenic emissions in a rapidly

urbanizing and warming world, addressing both air quality and climate challenges. By providing actionable insights for sustainable urban planning and climate mitigation, GEE-MEGAN sets a benchmark in biosphere-atmosphere interaction modeling.

Methods

GEE-MEGAN model

The GEE-MEGAN framework integrates the MEGAN2.1⁷ biogenic emission model with the computational capabilities of GEE, providing a flexible and scalable platform for estimating BVOC emissions across diverse spatial and temporal scales and bridging the gap between high-resolution emission modeling and large-scale atmospheric analyses.

Data source module. This module aggregates inputs from a range of data sources to enhance model accuracy and adaptability: Fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis (ERA5) climate data, MODIS-derived LAI, and reconstructed PFTs are used for large-scale analysis, and higher-resolution data such as Landsat imagery are employed for improving LAI precision, particularly in urban and regional studies (see Table S3 and Note S2). In addition to ERA5, GEE-MEGAN can incorporate other meteorological datasets such as Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2), the Climate Forecast System Version 2 (CFSv2), or the Global Land Data Assimilation System (GLDAS-2.1), and also support user-defined inputs such as WRF model output or CESM projections (Table S7).

Preprocessing module. To ensure consistency and reliability, this module processes the input through: (1) Temporal filtering: Aligning data to the required timeframes for simulations; (2) Spatial reprojection: Standardizing data projections to maintain spatial consistency; and (3) Resampling: Adjusting spatial resolution to match the scale of the analysis. Further details are provided in Note S3.

BVOC online calculation module. This module implements the core MEGAN2.1 computational framework within GEE, including canopy energy balance (Fig. S7) for simulating canopy conditions to estimate BVOC emissions, environmental correction factors for adjusting emissions based on environmental variables such as temperature and light, and standardized EF_s⁷ for applying standardized emission rates for various vegetation types. Despite improving GEE-MEGAN computational efficiency and enabling near-real-time simulations, GEE-MEGAN produces emission activity factors and outputs consistent with MEGAN2.1 when using identical inputs (see Fig. S8 and Note S4).

Output and post-processing module. This module supports both online and local data processing, allowing data export for integration into atmospheric chemistry models. Outputs are available in formats compatible with subsequent analyses, with additional details outlined in Note S5.

Scalability and customization. GEE-MEGAN provides three operational scales tailored to different applications: local scale (10–30 m) that utilizes high-resolution remote sensing and machine learning (e.g., Random Forest algorithm) to refine land use and cover change classification (LUCC), enhancing accuracy in urban and regional estimates, regional scale (500 m) that balances resolution and computational efficiency for broader geographic analyses, and global scale (500–5000 m) that employs coarser datasets to support global-scale modeling of BVOC emissions (see Table S3).

Box modeling

To evaluate isoprene concentrations and validate GEE-MEGAN estimates, atmospheric simulations were performed using the BOXMOX

model⁶⁷, based on the Model for Ozone and Related Chemical Tracers (MOZART-4) chemical mechanism. The simulations focused on the Paris Rive Gauche site in southeastern Paris, where Proton-Transfer-Reaction Mass Spectrometry measurements of ambient isoprene concentrations were available for comparison with GEE-MEGAN outputs⁹. Key inputs for the BOXMOX simulations included meteorological data (2-m air temperature, boundary layer height, and photolysis rates derived from the ERA5 reanalysis data) and anthropogenic emissions (pollutant data such as nitrogen oxides, VOCs and other species obtained from Copernicus Atmosphere Monitoring Service).

Refinement of PFTs, emission factor maps (EFMAP) and LAI in GEE-MEGAN

PFTs (Table S8) in GEE-MEGAN were derived using a combination of the Köppen-Geiger climate classification map⁶⁸ (Table S9) and the MODIS MCD12 land cover dataset⁶⁹ (Fig. S9 and Table S10), following established methodologies used in the evolution of the Community Land Model (CLM)⁷⁰. This integration leverages climate reanalysis data and MODIS land use categories to ensure accurate and consistent PFTs mapping. Detailed derivation approaches are provided in Table S11 and Note S6, with comparison of multiple PFTs classifications shown in Fig. S10.

For local modeling at 30 m and 10 m resolution, high-resolution datasets from Landsat and Sentinel satellites were used, with a supervised machine learning algorithm, specifically a random forest algorithm (ee.Classifier.smileRandomForest), to update near-real-time land cover mapping (Note S7 and S8). The random forest was configured with 50 decision trees, and the maximum depth and other parameters were set to default values. Training labels were generated using a combination of datasets, including MODIS MCD12⁶⁹, GLC_FCS30D⁷¹, Global Sentinel-2 10 m Canopy Height⁷², EAS 10 m Annual Land Cover⁷³, and Finer Resolution Observation and Monitoring of Global Land Cover 10 m (FROM-GLC10) datasets⁷⁴. Each dataset was initially mapped to the Annual International Geosphere-Biosphere Programme classification⁶⁹, and a hybrid sampling approach was employed to generate training labels. This approach combines spatially uniform sampling, which selected 5000 points randomly across the target area, and stratified sampling, which selected 1000 points from each LUCC classification category.

The predictor variables used in the random forest model were derived from Landsat and Sentinel satellite bands (visible, infrared, and near-infrared) as well as from calculated indices, including the normalized difference vegetation index, enhanced vegetation index, and soil adjusted vegetation index⁷⁵. Historical data were filtered to create training labels and corresponding attributes, which were then used to infer land use classification based on near-real-time remote sensing data. Finally, the predicted LUCC data were divided into PFTs based on canopy height and climate rules, and these were subsequently used as inputs for GEE-MEGAN.

Dynamic EFMAP. The dynamic EFMAP module adjusts BVOC emission factors in real-time to account for environmental fluctuations. This dynamic refinement enhances the accuracy of BVOC emission estimates, as detailed in Note S7 in the supplementary materials.

LAI. The LAI estimation was improved by combining MODIS-derived LAI data with Landsat-based LAI interpolated on matching dates (Note S9). This hybrid approach reduces temporal gaps and enhances spatial resolution, achieving 9.1% reduction in RMSE at 500-m resolution and 3.6% reduction in RMSE at 3-km resolution. The refined LAI method was validated against the global DIRECT ground-based LAI dataset (Fig. S3A), ensuring reliability and accuracy for both local and global applications.

Multiscale simulation framework

To balance near-real-time applicability across spatial scales, GEE-MEGAN uses consistent meteorological inputs from ERA5-Land hourly while tailoring land-use and vegetation inputs across 100 m, 500 m, and 5 km scales, with varying data sources for LUCC and LAI to reflect spatial heterogeneity in emissions. For scales smaller than 500 m, high-resolution Landsat data were utilized to calculate LUCC and PFTs, while the LAI was derived using an enhanced MCD15 method with Landsat observations. At the 500 m scale, LUCC data were sourced from MCD12Q1 dataset, and LAI data were obtained using the MODIS MCD15A3H product. At the coarser 5 km scale, LUCC inputs were extracted from the MODIS MCD12C1 dataset, and LAI was modeled using the LAI/FPAR Climate Data Record (CDR)⁷⁶. At coarser spatial resolutions (e.g., 500 m), urban green space may be underrepresented or misclassified as impervious surfaces, resulting in negligible urban BVOC emissions. Accurate estimation of emissions in urban areas therefore requires high-resolution LUCC and LAI inputs, which GEE-MEGAN supports. The spatial resolution of GEE-MEGAN outputs aligns with the resolution of the input datasets and can be flexibly selected depending on the application. For urban or rapidly changing landscapes, we recommend using high resolution inputs (10–30 m); for regional or global assessments, coarser inputs (500 m to 5 km) are more practical and computationally efficient.

Validation

The accuracy of GEE-MEGAN was evaluated by comparing emission estimates with those of MEGAN2.1 across multiple sites and scales. Validation metrics included R^2 , root mean squared error (RMSE), MAE, and Pearson correlation, for a comprehensive assessment of the model's performance. These metrics were also applied to assess improvements in LAI estimates.

The GEE-MEGAN framework (Fig. S1) significantly improves the computational efficiency and scalability of BVOC emission modeling by leveraging the cloud-based capabilities of GEE^{21,22}. This innovation enables near-real-time, high-resolution simulations across large spatial domains while minimizing both computational demands and data transfer requirements. For instance, using a consistent input dataset for South America (60° S–13° N, 35–90° W) at a 500-m resolution and operating on a local Intel Xeon Gold 6230R CPU (2.1 GHz), MEGAN2.1 required 736.9 s per species to complete the simulation while GEE-MEGAN achieved comparable results in just 7.2 to 9.1 s per species, delivering an 81- to 102-fold increase in processing speed; moreover, GEE-MEGAN reduced data transfer by 97–99% compared to traditional models, facilitating efficient processing even for large-scale or high-resolution applications.

GEE-MEGAN accurately replicates the core dynamics of MEGAN2.1 while significantly improving efficiency and scalability⁷. Simulations using input datasets from North America (<https://bai.ess.uci.edu/megan/data-and-code/megan21>, last retrieved September 14, 2024) revealed high consistency between MEGAN2.1 (Fig. S8a, e, i, m) and GEE-MEGAN (Fig. S8b, f, g, n), with a correlation coefficient of 1.0 and a total error of less than 0.1 across 16,576 data points. Minor discrepancies (Fig. S8d, h, l, p) were observed due to differences in computational environments but remained within acceptable margins (RMSE < 3.0×10^{-2} ; Fig. S8 and Table S12).

Data availability

All data supporting the findings of this study are openly available. The datasets used in GEE-MEGAN are detailed in Table 3 of the Supplementary Information. The accuracy of the LAI estimation method was validated against the global ground-based observation dataset DIRECT (<https://calvalportal.ceos.org/lpv-direct-v2.1>, last access: 21 December 2023). The access links for ERA5-Land hourly, MERRA-2, CFSv2, and GLDAS-2.1 data are provided in Table 7 of the Supplementary

Information. The Landsat surface reflectance, Sentinel-2, GLC_FCS3D⁷¹, Global Sentinel-2 10 m Canopy Height⁷², EAS WorldCover⁷³, FROM-GLC10⁷⁴ and sensor-independent LAI/FPAR CDR⁷⁶ are all accessible via the Google Earth Engine platform (<https://earthengine.google.com/>). Data related to MEGAN 2.1 MOHYCAN-OMI and IMAGESv2-GOME2 are available from <https://emissions.aeronomie.be/>, and the MEGAN 2.1 test-case inputs from <https://bai.ess.uci.edu/megan/data-and-code/megan21>. All the raw data supporting the figures and tables presented in this study are stored in: <https://doi.org/10.6084/m9.figshare.29377247>⁷⁷. Source data are provided with this paper.

Code availability

The MEGAN2.1 model source code is available from <https://bai.ess.uci.edu/megan/data-and-code/megan21>. The WRF v4.3.1 code can be cloned from <https://github.com/wrf-model/WRF.git>, and the BOXMOX model framework is available for download from <https://mbees.med.uni-augsburg.de/boxmodeling/download.html>. The code for GEE-MEGAN framework, including the scripts and documentation in this study, is publicly accessible at the following Zenodo repository: <https://doi.org/10.5281/zenodo.15714886>⁷⁸.

References

1. Sanaei, A. et al. Changes in biodiversity impact atmospheric chemistry and climate through plant volatiles and particles. *Commun. Earth Environ.* **4**, 445 (2023).
2. Hewitt, C. N. et al. Ground-level ozone influenced by circadian control of isoprene emissions. *Nat. Geosci.* **4**, 671–674 (2011).
3. Park, J. H. et al. Active atmosphere-ecosystem exchange of the vast majority of detected volatile organic compounds. *Science* **341**, 643–647 (2013).
4. IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Masson-Delmotte, V. et al.) (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021).
5. Danabasoglu, G. et al. The community earth system model version 2 (CESM2). *J. Adv. Model. Earth Syst.* **12**, e2019MS001916 (2020).
6. Arneth, A. et al. Terrestrial biogeochemical feedbacks in the climate system. *Nat. Geosci.* **3**, 525–532 (2010).
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. Ma, M. et al. Development and assessment of a high-resolution biogenic emission inventory from urban green spaces in China. *Environ. Sci. Technol.* **56**, 175–184 (2021).
9. Maison, A. et al. Significant impact of urban tree biogenic emissions on air quality estimated by a bottom-up inventory and chemistry transport modeling. *Atmos. Chem. Phys.* **24**, 6011–6046 (2024).
10. Ciccioli, P. et al. The potential impact of biogenic volatile organic compounds (BVOCs) from terrestrial vegetation on a Mediterranean area using two different emission models. *Agric. Meteorol.* **328**, 109255 (2023).
11. Bao, X. et al. Tree species classification improves the estimation of BVOCs from urban greenspace. *Sci. Total Environ.* **914**, 169762 (2024).
12. Wei, D. et al. High-resolution modeling of summertime biogenic isoprene emissions in New York City. *Environ. Sci. Technol.* **58**, 13783–13794 (2024).
13. Kuang, W. et al. Global observation of urban expansion and land-cover dynamics using satellite big-data. *Sci. Bull.* **66**, 297–300 (2021).

14. Lewis, A. C. The changing face of urban air pollution. *Science* **359**, 744–745 (2018).
15. Barwise, Y. & Kumar, P. Designing vegetation barriers for urban air pollution abatement: a practical review for appropriate plant species selection. *Npj Clim. Atmos. Sci.* **3**, 12 (2020).
16. Yin, Y., He, L., Wennberg, P. O. & Frankenberg, C. Unequal exposure to heatwaves in Los Angeles: Impact of uneven green spaces. *Sci. Adv.* **9**, eade8501 (2023).
17. Frolking, S. et al. Global urban structural growth shows a profound shift from spreading out to building up. *Nat. Cities* **1**, 555–566 (2024).
18. Li, D. et al. The effect of urbanization on plant phenology depends on regional temperature. *Nat. Ecol. Evol.* **3**, 1661–1667 (2019).
19. Pfannerstill, E. Y. et al. Comparison between spatially resolved airborne flux measurements and emission inventories of volatile organic compounds in Los Angeles. *Environ. Sci. Technol.* **57**, 15533–15545 (2023).
20. Pfannerstill, E. Y. et al. Temperature-dependent emissions dominate aerosol and ozone formation in Los Angeles. *Science* **384**, 1324–1329 (2024).
21. Gorelick, N. et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
22. Amani, M. et al. Google earth engine cloud computing platform for remote sensing big data applications: a comprehensive review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13**, 5326–5350 (2020).
23. Eyring, V. et al. AI-empowered next-generation multiscale climate modelling for mitigation and adaptation. *Nat. Geosci.* **17**, 963–971 (2024).
24. Vance, T. C. et al. Big data in Earth science: emerging practice and promise. *Science* **383**, eadh9607 (2024).
25. Bauwens, M. et al. Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations. *Atmos. Chem. Phys.* **16**, 10133–10158 (2016).
26. Opacka, B. et al. Global and regional impacts of land cover changes on isoprene emissions derived from spaceborne data and the MEGAN model. *Atmos. Chem. Phys.* **21**, 8413–8436 (2021).
27. Stavrou, T. et al. How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI? *Atmos. Chem. Phys.* **15**, 11861–11884 (2015).
28. Müller, J.-F. et al. Bias correction of OMI HCHO columns based on FTIR and aircraft measurements and impact on top-down emission estimates. *Atmos. Chem. Phys.* **24**, 2207–2237 (2024).
29. Vettikatt, L. et al. High emission rates and strong temperature response make boreal wetlands a large source of isoprene and terpenes. *Atmos. Chem. Phys.* **23**, 2683–2698 (2023).
30. 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).
31. Acton, W. J. F. et al. Surface-atmosphere fluxes of volatile organic compounds in Beijing. *Atmos. Chem. Phys.* **20**, 15101–15125 (2020).
32. Bai, J. et al. Seasonal and interannual variations in whole-ecosystem BVOC emissions from a subtropical plantation in China. *Atmos. Environ.* **161**, 176–190 (2017).
33. Sarkar, C. et al. A portable, low-cost relaxed eddy accumulation (REA) system for quantifying ecosystem-level fluxes of volatile organics. *Atmos. Environ.* **242**, 117764 (2020).
34. Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008).
35. Csillik, O. et al. A large net carbon loss attributed to anthropogenic and natural disturbances in the Amazon Arc of Deforestation. *Proc. Natl. Acad. Sci. USA* **121**, e2310157121 (2024).
36. Yáñez-Serrano, A. M. et al. Amazonian biogenic volatile organic compounds under global change. *Glob. Chang. Biol.* **26**, 4722–4751 (2020).
37. Reydon, B. P. et al. Land governance as a precondition for decreasing deforestation in the Brazilian Amazon. *Land use policy* **94**, 104313 (2020).
38. Xu, H. et al. Global patterns and drivers of post-fire vegetation productivity recovery. *Nat. Geosci.* **17**, 874–881 (2024).
39. Williams, B. A. et al. Global potential for natural regeneration in deforested tropical regions. *Nature* **636**, 131–137 (2024).
40. Brandt, M. et al. Severe decline in large farmland trees in India over the past decade. *Nat. Sustain.* **7**, 860–868 (2024).
41. Reiner, F. et al. More than one quarter of Africa's tree cover is found outside areas previously classified as forest. *Nat. Commun.* **14**, 2258 (2023).
42. Coggon, M. M. et al. Volatile chemical product emissions enhance ozone and modulate urban chemistry. *Proc. Natl. Acad. Sci. USA* **118**, e2026653118 (2021).
43. Guenther, A. Biological and chemical diversity of biogenic volatile organic emissions into the atmosphere. *Int. Sch. Res. Not.* **2013**, e786290 (2013).
44. Blichner, S. M. et al. Process-evaluation of forest aerosol-cloud-climate feedback shows clear evidence from observations and large uncertainty in models. *Nat. Commun.* **15**, 969 (2024).
45. Pöschl, U. et al. Rainforest aerosols as biogenic nuclei of clouds and precipitation in the Amazon. *Science* **329**, 1513–1516 (2010).
46. Kulmala, M. et al. *Climate feedbacks linking the increasing atmospheric CO₂ concentration, BVOC emissions, aerosols and clouds in forest ecosystems: Biology, Controls and Models of Tree Volatile Organic Compound Emissions.* (eds Niinemets, Ü. & Monson, R.) 489–508 (Springer Press, 2013).
47. Nölscher, A. et al. Unexpected seasonality in quantity and composition of Amazon rainforest air reactivity. *Nat. Commun.* **7**, 10383 (2016).
48. Weber, J. et al. Chemistry-driven changes strongly influence climate forcing from vegetation emissions. *Nat. Commun.* **13**, 7202 (2022).
49. Armstrong McKay, D. I. et al. Exceeding 1.5 °C global warming could trigger multiple climate tipping points. *Science* **377**, eabn7950 (2022).
50. Shrivastava, M. et al. Urban pollution greatly enhances formation of natural aerosols over the Amazon rainforest. *Nat. Commun.* **10**, 1046 (2019).
51. Curtius, J. et al. Isoprene nitrates drive new particle formation in Amazon's upper troposphere. *Nature* **636**, 124–130 (2024).
52. Shen, J. et al. New particle formation from isoprene under upper-tropospheric conditions. *Nature* **636**, 115–123 (2024).
53. Engert, J. E. et al. Ghost roads and the destruction of Asia-Pacific tropical forests. *Nature* **629**, 370–375 (2024).
54. Forzieri, G. et al. Emerging signals of declining forest resilience under climate change. *Nature* **608**, 534–539 (2022).
55. Rap, A. et al. Enhanced global primary production by biogenic aerosol via diffuse radiation fertilization. *Nat. Geosci.* **11**, 640–644 (2018).
56. Bastin, J.-F. et al. The global tree restoration potential. *Science* **365**, 76–79 (2019).
57. Tölgyesi, C. et al. Urgent need for updating the slogan of global climate actions from “tree planting” to “restore native vegetation”. *Restor. Ecol.* **30**, e13594 (2022).
58. Venter, Z. S. et al. Reassessing the role of urban green space in air pollution control. *Proc. Natl. Acad. Sci. USA* **121**, e2306200121 (2024).
59. Xu, L. et al. Effects of anthropogenic emissions on aerosol formation from isoprene and monoterpenes in the southeastern United States. *Proc. Natl. Acad. Sci. USA* **112**, 37–42 (2015).
60. Li, M. et al. Coping with the concurrent heatwaves and ozone extremes in China under a warming climate. *Sci. Bull.* **69**, 2938–2947 (2024).

61. Wang, H. et al. Heat waves may trigger unexpected surge in aerosol and ozone precursor emissions from sedges in urban landscapes. *Proc. Natl. Acad. Sci. USA* **121**, e2412817121 (2024).
 62. Qin, M. et al. Increased urban ozone in heatwaves due to temperature-induced emissions of anthropogenic volatile organic compounds. *Nat. Geosci.* **18**, 50–56 (2025).
 63. Rentschler, J. et al. Global evidence of rapid urban growth in flood zones since 1985. *Nature* **622**, 87–92 (2023).
 64. Li, J. et al. Future increase in compound soil drought-heat extremes exacerbated by vegetation greening. *Nat. Commun.* **15**, 10875 (2024).
 65. Ren, X. et al. Deep learning-based weather prediction: a survey. *Big Data Res* **23**, 100178 (2021).
 66. Bi, K. et al. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **619**, 533–538 (2023).
 67. Knote, C. et al. Influence of the choice of gas-phase mechanism on predictions of key gaseous pollutants during the AQMEII phase-2 intercomparison. *Atmos. Environ.* **115**, 553–568 (2015).
 68. Beck, H. E. et al. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 180214 (2018).
 69. Friedl, M., Sulla-Menashe, D. MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*. (2019).
 70. Lawrence, D. M. et al. The community land model version 5: description of new features, benchmarking, and impact of forcing uncertainty. *J. Adv. Model. Earth Syst.* **11**, 4245–4287 (2019).
 71. Zhang, X. et al. GLC_FCS30D: The first global 30-m land-cover dynamic monitoring product with a fine classification system from 1985 to 2022 using dense time-series Landsat imagery and continuous change-detection method. *Earth Syst. Sci. Data Discuss* **1**, 32 (2023).
 72. Lang, N. et al. A high-resolution canopy height model of the Earth. *Nat. Ecol. Evol.* **7**, 1778–1789 (2023).
 73. Zanaga, D. et al. ESA WorldCover 10 m 2020 v100. Zenodo <https://doi.org/10.5281/ZENODO.5571936> (2021).
 74. Gong, P. et al. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Sci. Bull.* **64**, 370–373 (2019).
 75. Zeng, Y. et al. Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nat. Rev. Earth Environ.* **3**, 477–493 (2022).
 76. Pu, J. et al. Sensor-independent LAI/FPAR CDR: reconstructing a global sensor-independent climate data record of MODIS and VIIRS LAI/FPAR from 2000 to 2022. *Earth Syst. Sci. Data* **16**, 15–34 (2024).
 77. Zhang, Y. et al. Improved modelling of biogenic emissions in human-disturbed forest edges and urban areas. <https://doi.org/10.6084/m9.figshare.29377247> (2025).
 78. Zhang, Y. et al. Improved modelling of biogenic emissions in human-disturbed forest edges and urban areas. Zenodo. <https://doi.org/10.5281/zenodo.15714886> (2025).
- Development Program (2022YFC3701103 to X.W. and 2024YFC3714300 to Y.Z.), and the Department of Science and Technology of Guangdong (2023B0303000007 to X.W.).

Author contributions

Y.Z., A.G., Q.Z. and X.W. contributed to the design of the study. H.R. carried out the modeling. Y.Z., H.R. and X.W. analyzed and interpreted the results. Y.Z., H.R. and X.W. prepared the original manuscript. A.G., Q.Z., W.M., C.G., G.S. and P.P. contributed to the discussion and interpretation of the results.

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/s41467-025-63437-8>.

Correspondence and requests for materials should be addressed to Alex Guenther, Qiang Zhang or Xinming Wang.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions 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.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Acknowledgements

This study was supported by the National Natural Science Foundation of China (42321003 to X.W.), the National Key Research and