We thank both reviewers for their careful reading of our paper and their very valuable comments. We repeat the reviewer comments below in italics, then give our responses.

Reviewer 2, Comment 1(i)

Reviewer comment

i) the methodology for box model is not clear.
It is not clear how to set up the Box-model emissions. As shown in L218, the emissions are tuned to satisfy observed precursor constraints. What kind of observed precursors? I didn't find the measurements for NO_x and VOC that should be important for your emission constraint.

Figure 2 shows very low AVOC emissions relative to BVOC emissions in Shanghai. Is it reasonable for megacity?

Tuning Box-model ozone to match observations is also not clear. As said in L256-258, it is hard to understand how to tune local O_3 emissions and depositions.

Response

We thank the reviewer for identifying that the previous description of the box-model setup was insufficiently clear. We have now rewritten Sect. 2.2.1 to provide a complete account of the model initialisation, inventory inputs, observational constraints, emission and deposition adjustments, and the relationship between the baseline construction and the subsequent temperature experiments (lines 232-308). Table 2 now presents the principal observational constraints directly in the manuscript, Fig. 2 defines precisely which anthropogenic VOC species are displayed, and the Supplement provides the complete tuning framework, scaling factors, reference concentrations, and achieved baseline concentrations in Tables S1-S4.

The baseline construction began from established model inputs rather than arbitrary values. Anthropogenic emissions were prescribed from CEDS (based on selected CEDS cells shown in Fig. 2), biogenic emissions from the MEGAN-based China inventory of Wu et al. (2020) and dry deposition from the UKCA implementation of the Wesely (1989) resistance scheme. All tuning runs used the same environmental conditions: 30 °C, specific humidity of 0.0184 kg kg^{-1}, RH of 70 %, surface pressure of 1000 hPa, and a PBL height of 590 m. Each configuration was integrated for 1 year with a 10 min chemical time step and a repeating 1 July solar cycle.

Three types of effective adjustment were used, for different reasons.

First, emission scaling was used where the selected inventory cells did not fully represent the intended chemical environment or where the available emitted tracer set represented only part of the actual precursor mixture. These factors define effective box-model source terms and should not be interpreted as revised estimates of the true regional emissions.

Second, emissions and depositions are adjusted to represent the unsolved source/inflow or loss/outflow affected by processes that are absent from the closed 0-D model.

Third, where a species continued to accumulate because no suitable direct loss term was available, its effective source was reduced. This was required for species such as NO, because

increasing NO_2 deposition could not directly remove accumulated NO, and there is no NO_2 emission and NO deposition set in the model.

The procedure used explicit stopping criteria. A species was considered stable when the change in its daily mean concentration was less than 0.1 % relative to the preceding day. Where suitable observations were available, the target was agreement within ±20 %. Where only observations from chemically comparable but non-colocated environments were available, the model value was required to remain within the same order of magnitude; these values were used as contextual references rather than strict ±20 % tuning targets. These criteria, together with the expected VOC-limited or NO_x-limited chemical regime, determined the final settings.

The Supplement now organises the procedure into four functional types. Type 1 covers CO, C_2H_6, and C_3H_8, which are less active species in building the O_3 precursor environment. CO was constrained within ±20 % of observations, while C_2H_6 and C_3H_8 were adjusted only to prevent persistent accumulation or excessive removal where suitable observations were unavailable. Type 2 covers NO emissions and NO_2 deposition and establishes the contrasting NO_x environments. Type 3 covers C_5H_8 and monoterpene emissions and represents the different biogenic influences in Shanghai and Lishui. Type 4 covers HCHO emissions and O_x deposition other than NO_2. These inputs were retained at their original values and used to evaluate the resulting chemical environment rather than to force agreement with ozone.

The adjustment was conducted through successive steady-state sensitivity runs. Each candidate configuration was integrated until the change in the daily mean concentration of the monitored species was less than 0.1 % relative to the preceding day. Inputs were then adjusted by functional group: first CO and the less reactive alkanes, then NO emissions and NO_2 deposition, followed by C_5H_8 and monoterpene emissions. After each adjustment, the complete chemical system was reintegrated, and the previously constrained species were checked again. The procedure ended when the quantitative concentration criteria, stability requirements and intended chemical regime were satisfied simultaneously. Type 4 inputs were not adjusted.

The two environments were adjusted separately. In Shanghai, the objective was to construct a stable high-NO_x, VOC-limited environment. NO_2 deposition was increased to represent unresolved NO_x removal by ventilation and dilution, while NO emissions were reduced because the inventory-derived NO source accumulated in the closed box. This produced NO_2 within ±20 % of the observed Shanghai concentration while maintaining stable NO. C_5H_8 and monoterpene emissions were increased to represent the documented regional biogenic influence from forests south and southwest of Shanghai.

This transported contribution should not be expected to respond as strongly to the instantaneous temperature prescribed in the Shanghai box as a local emission. The C_5H_8 reaching Shanghai has been emitted over a wider upwind region and over an earlier period, and its contribution is weakened during transport by atmospheric dilution and chemical loss. Applying the full local temperature response to the entire additional source would therefore treat transported C_5H_8 as though it were emitted locally and instantaneously, exaggerating its contribution to the simulated high-temperature response.

We consequently applied a conservative 50 % reduction to the additional regional source before applying the temperature-dependent emission factor. The source used in the simulations was therefore $E_{base} = E_{MEGAN} + 0.5E_{transported}$.

,In Lishui, the objective was to construct a low-NO_x, NO_x-limited, vegetation-influenced regional environment. The available Lishui monitoring stations are situated in populated urban or suburban locations and therefore do not directly represent the broader forest-dominated environment used in the box model. Forest-canopy and mountain-background NO_2 measurements were consequently used as order-of-magnitude references. The original NO_2 deposition was retained, and NO emissions were reduced only because NO accumulated. C_5H_8 emissions were reduced to represent unresolved export and dilution from the vegetation-rich source region, producing a daytime C_5H_8 concentration within ±20 % of the available Lishui measurement. Monoterpene emissions were retained because no suitable local observational constraint was available.

We agree that the observational basis for the Lishui configuration involves an unavoidable representativeness limitation. We could not identify a single site that simultaneously provided both rural or background NO_2 representative of the intended regional chemical environment and suitable C_5H_8 measurements for constraining the biogenic influence. Lishui was therefore selected because it combines a relatively low-NO_x regional setting with an available C_5H_8 measurement from its forested environment. We now state this limitation explicitly in the manuscript.

The final scaling factors are now reported explicitly. For Shanghai, the emission factors relative to the inventory values are 0.045 for CO, 0.50 for NO, 0.50 for C_2H_6, 0.40 for C_3H_8, 2.00 for C_5H_8, and 4.00 for monoterpenes; the CO and NO_2 deposition factors are 4 and 5, respectively. For Lishui, the emission factors are 0.05 for CO, 0.50 for NO, and 0.50 for C_5H_8, while the CO deposition factor is 2. Inputs not listed in Tables S1 and S3 were retained at their original values. Tables S2 and S4 then compare the resulting baseline concentrations with the observational constraints and independent evaluation data.

The observational constraints are also now stated explicitly. For Shanghai, the principal references include NO_2 and CO observations from CNEMC, published urban C_5H_8 measurements from Gu et al. (2022), and a published monoterpene reference from Cheng et al. (2018). For Lishui, CO is constrained by the CNEMC observation, C_5H_8 by the daytime forest measurements of Geng et al. (2011), and NO_2 by forest-canopy and mountain-background observations from Min et al. (2014) and Yin et al. (2022). HCHO provides an integrated indicator of the carbonyl and VOC-oxidation environment. Its agreement within ±20 % of the Shanghai observation shows that the model reproduces the overall reactive VOC environment well and supports the adopted effective AVOC representation.

Regarding the apparently low AVOC emissions in Fig. 2, the plotted quantity is not total anthropogenic VOC emissions in Shanghai. It is specifically the sum of the three CEDS VOC tracers emitted explicitly in the present StratTrop box-model configuration: C_2H_6, C_3H_8, and HCHO. It excludes substantial components of the actual urban AVOC mixture, including many alkanes, alkenes, aromatics, solvent-related compounds, and evaporative emissions. The ratio shown in Fig. 2 must therefore not be interpreted as the real-world ratio of total AVOC to

BVOC emissions in Shanghai. This definition is now stated directly in Sect. 2.2.1 and the Fig. 2 caption (lines 250-269), and the implications of the restricted AVOC representation are quantified and discussed further in Sect. 4.1.

Finally, ozone was not tuned. No direct ozone emission is prescribed in the UKCA Box model. Ozone is produced and destroyed through the StratTrop chemical mechanism. Ozone deposition was retained at its original UKCA value and was not modified to reproduce observed ozone. O_3, PAN, and other Type 4 species were used as independent checks on the chemical environment. The final emission and deposition settings were established once at the 30 °C baseline and then fixed before all temperature, humidity, diurnal-cycle, and isoprene-temperature experiments. The model was not retuned at different temperatures. The simulated ozone-temperature response therefore arises from the controlled experimental perturbations applied to fixed baseline environments, not from fitting ozone to observations.

The revised manuscript now presents the essential methodology in Sect. 2.2.1, while the Supplement provides a transparent and reproducible record of the tuning criteria, species-specific decisions, final scaling factors, observational references, and achieved baseline concentrations.

Reviewer 2, Comment 1(ii)

Reviewer comment

1. *ii) the O_3 chemistry in Box-model may be inconsistent with measurements.*

In Fig.6, the simulated response of O_3 to NO_x/VOC emission perturbations is almost unchanged across the temperature range. However, previous studies have shown that O_3 formation regime is changing under different temperature conditions.

In addition, the comparisons in Table 3 are not helpful. All of the species (PAN, OH, HO_2, and C_5H_8) are reactive and of short lifetime. The observations were conducted in different seasons and different years. I think such an observation-model comparison can be used to evaluate the UKCA Box Model.

The conclusions show that In Shanghai, the isoprene temperature response explained ~70 % of the observed O_3 increase from 30 to 40 °C. I don't agree with this argument. This is misleading if the model used very high BVOC emissions, as shown in Fig.2. The role of AVOC emissions should be strongly underestimated in current model.

Response

We thank the reviewer for this important comment. We have strengthened the chemical-regime analysis by adding O_3-NO_x-isoprene response surfaces to Fig. 6 and revising the accompanying interpretation in Sect. 3.2 (lines 463-500). We have also restructured Table 4 and clarified its purpose and the comparability of the reported values (lines 501-510). The isoprene-temperature parameterisation, BVOC concentrations and temperature-dependent AVOC emissions are now examined in detail in Sects. 4.1-4.2 (lines 770-825). The Conclusions now state explicitly that the approximately 70 % value is the fraction diagnosed within the adopted factorial UKCA Box configuration, rather than a complete real-world attribution. This numerical fraction depends on the adopted isoprene-temperature response, the tuned Shanghai baseline environment and the restricted AVOC representation in the StratTrop configuration, in which only C_2H_6, C_3H_8 and HCHO are emitted explicitly, and AVOC emissions are held temperature-independent (lines 846-856).

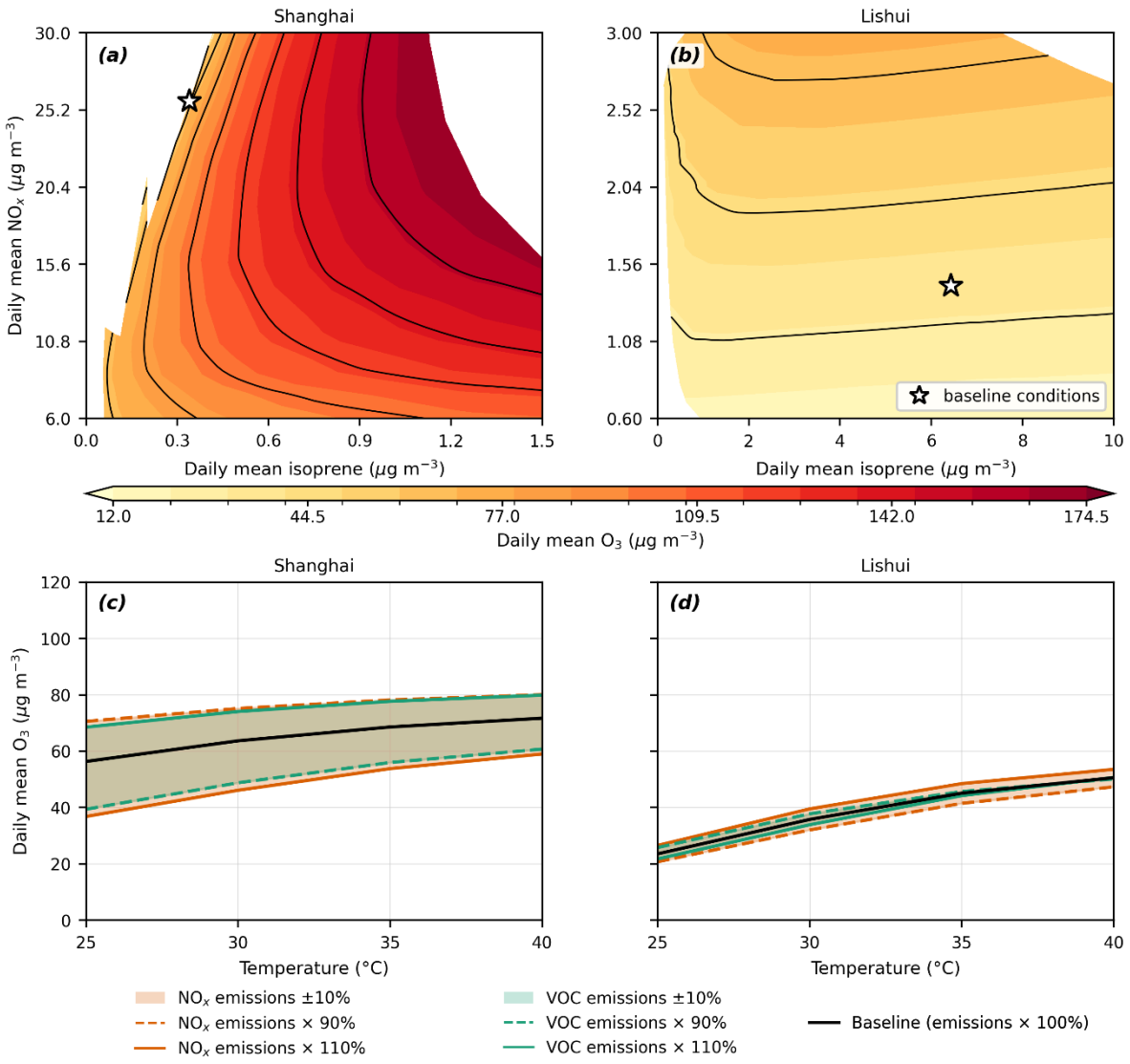

Figure 6. Simulated precursor sensitivity of daily mean surface O_3 in Shanghai and Lishui using the UKCA Box Model under the VRH_Ori baseline configuration with emission changes. Panels (a) and (b) show box model simulations of the evolution of ozone as a function of NO_x (NO + NO_2) and isoprene at 30 °C. Colours indicate daily mean O_3, black lines denote O_3 contours, and white stars mark the baseline conditions. Panels (c) and (d) show the temperature dependence of O_3 under ±10 % perturbations in NO_x and VOC emissions at 25, 30, 35, and 40 °C.

The original ±10 % NO_x and VOC perturbations were designed to diagnose the local precursor sensitivity around each baseline chemical environment, rather than to map the complete chemical-regime space. We agree that these small perturbations alone did not show sufficiently clearly whether the baseline environments could move between chemical regimes. We have therefore added two-dimensional O_3 response surfaces over substantially wider NO_x and isoprene ranges in the revised Fig. 6a-b.

We additionally tested the chemical regime at the highest-VOC condition reached by the temperature experiments. In the Shanghai Dft simulation at 40 °C, NO_x is 21.1 μg m^{-3} and

C_5H_8 is 0.5 μg m^{-3}. Placing this state on the wider Fig. 6a-b precursor space shows that it remains within the VOC-limited region. The temperature-dependent increase in isoprene therefore strengthens O_3 formation without moving the Shanghai case across the regime boundary.

The new response surfaces show that the Shanghai baseline lies within a VOC-limited regime, where O_3 increases with isoprene and decreases with increasing NO_x. In contrast, the Lishui baseline lies within a NO_x-limited regime, where O_3 increases with NO_x and responds only weakly to additional isoprene (lines 484-488). These results provide a direct diagnosis of the contrasting chemical environments rather than relying only on the original ±10 % experiments.

Figures 6c-d retain the local perturbation experiments at 25, 30, 35 and 40 °C. In Shanghai, enhanced NO_x consistently decreases O_3, while enhanced VOC emissions increase O_3. Lishui shows the opposite NO_x response and remains only weakly sensitive to additional VOC emissions. The signs of these responses remain unchanged from 25 to 40 °C, demonstrating that no regime reversal occurs around the adopted Shanghai and Lishui baseline states within the tested temperature range (lines 489-500). Temperature changes the magnitude of O_3 production, but it does not necessarily cause a regime transition where the precursor environment remains on the same side of the broader NO_x-VOC response surface.

We agree that PAN, OH, HO_2 and C_5H_8 are short-lived species whose concentrations depend on location, season and averaging period. Table 4 is used as an observation-based check on the magnitude and character of the simulated chemical environments. Its purpose is to provide observational context for the order of magnitude of the simulated chemical environments. This purpose is stated in Sect. 3.2 (lines 463-468).

Table 4 now reports the environment, location, year, season and value type for every comparison, and the simulated values use matched averaging periods wherever the published data allow. The source studies also account explicitly for the local nature of these reactive species. The Shanghai PAN campaign used a mixed urban site without an immediate major roadside or industrial source and interpreted PAN together with simultaneous O_3, NO_x, CO, particles, meteorology and air-mass direction. The Hangzhou and Nanjing measurements provide suburban and regional-background YRD references, respectively, so the simulated Shanghai and Lishui PAN values can be placed within a documented regional range.

The radical comparisons are treated similarly. The direct Shanghai OH and HO_2 measurements were made at an urban site with simultaneous NO_x, VOC, OVOC, HONO, photolysis and meteorological constraints, while the Taizhou campaign provides a warm-season suburban YRD reference influenced by both regional and biogenic chemistry. To match the reported metric, the simulated Shanghai OH and HO_2 values were recalculated as 11:00-13:00 means rather than 24 h means.

For C_5H_8, both source studies were designed around its short lifetime and spatial heterogeneity. The Shanghai measurement used a 75 m long-path DOAS configuration at approximately 25 m height in a representative urban environment, reducing sensitivity to an individual tree or street source, and analysed temperature, radiation, humidity, wind and chemical loss. The Lishui study sampled multiple forest locations and interpreted the observations with MEGAN-WRF-Chem while recognising local vegetation effects. These campaign designs make the published values appropriate order-of-magnitude references for the urban and vegetation-

influenced environments represented here. The revised values and their sampling definitions are reported in Table 4 and discussed at lines 501-510.

The concerns in the final paragraph regarding the isoprene-temperature response and temperature-dependent AVOCs overlap directly with Reviewer 1 Comment 4, where we provide our full response and supporting calculations. As explained there, the revised manuscript compares the JULES isoprene-temperature formulation with MEGAN2.1 and shows that JULES produces a smaller high-temperature response over 30-40 °C. The available C_5H_8 observations also do not indicate excessive simulated isoprene: the Shanghai simulation is lower than the corresponding urban observation, while the Lishui simulation is close to the available regional measurement. These results are presented in Table 4 and discussed fully in our response to Reviewer 1 Comment 4.

As also clarified in our response to Reviewer 2 Comment 1(i), the AVOC category in Fig. 2 contains only the explicitly emitted C_2H_6, C_3H_8 and HCHO species represented in the selected StratTrop configuration. It does not represent the complete anthropogenic VOC mixture in Shanghai and therefore cannot be used to infer the actual ratio of total AVOC to BVOC emissions in the megacity.

Temperature-dependent AVOC emissions provide an additional urban O_3 pathway. This issue is addressed fully in Sect. 4.1 and in our response to Reviewer 1 Comment 4. The revised manuscript now quantifies the potential temperature response of selected evaporative AVOC sources and states that fixed AVOC emissions can underestimate anthropogenic VOC availability under hot conditions (lines 770-797).

We also assessed whether the estimated temperature-dependent AVOC enhancement could be introduced directly into the present UKCA Box configuration. To do this, we compared the relevant CEDS species and source sectors with the CEDS-to-StratTrop lumping approach (Archer-Nicholls et al., 2021; Qin et al., 2025). The overlap is extremely limited: only propane (VOC03) and propene (VOC08), both represented by the C_3H_8 tracer, are retained. Together, these species account for only 3.56 % of the selected gasoline-evaporation emissions, while none of the selected VCP-related solvent emissions are represented. The dominant temperature-dependent butanes, pentanes, higher alkanes, other alkenes and aromatic compounds are absent from the current emitted-tracer mapping. We therefore did not conduct a box-model sensitivity experiment using this small subset, because its O_3 response would not represent the full temperature-dependent AVOC source. Instead, we retained the Qin-based emission calculation using JJA 2022 Shanghai conditions as a quantitative diagnostic of the potential AVOC enhancement (Qin et al., 2025). This analysis recognises temperature-dependent AVOCs as an additional urban O_3 pathway, while demonstrating that their chemical contribution requires a more comprehensive VOC mechanism and emission mapping, which we will develop in our following work.

The Conclusions state the approximately 70 % result conditionally: under the adopted isoprene-temperature parameterisation and Shanghai baseline emission setting, temperature-dependent isoprene accounts for approximately 70 % of the observed 30-40 °C ΔO_{3max} represented within the factorial box-model framework (lines 846-856). Temperature-dependent AVOCs can alter this quantitative fraction in a more comprehensive VOC mechanism, but they do not remove the strong isoprene-driven O_3 response diagnosed in the VOC-limited Shanghai environment.

Reviewer 2, Other Comments

Other Comment 1

Reviewer comment

L42: add NO_x for "(i) emissions of O_3 precursors (VOCs)"

Response

We thank the reviewer for identifying this omission. We have revised the text to read "emissions of O_3 precursors (NO_x and VOCs)" in the Introduction (lines 43-45).

Other Comment 2

Reviewer comment

L178: it is better to specify the temperature values of the 90th percentile threshold

Response

We thank the reviewer for this comment. There is no single unified 90th-percentile temperature threshold for either Shanghai or Lishui. The threshold is calculated separately for each calendar day from the 2000-2019 reference-period daily maximum temperatures, using a centred 15-day window. It therefore varies by both date and location.

Reporting one temperature value for each location would not represent the heatwave definition used in the analysis. We have clarified the calendar-day-dependent calculation in Sect. 2.1 (lines 185-192).

Other Comment 3

Reviewer comment

The reason for focusing on the summer of 2022 should be well justified. In addition, anthropogenic emissions are changing rapidly in China and if the current results are supportive for future ozone controls is questionable. More model scenario simulations are needed.

Response

We thank the reviewer for this comment. This issue overlaps with Reviewer 1 Comment 2, where we provide a detailed response. We have revised Sect. 2.1 to explain that summer 2022 was selected because it represents a severe recent YRD heatwave with reliable routine surface O_3 observations (lines 167-170).

The box-model experiments are controlled process-sensitivity tests based on the chemical environments defined for 2022, rather than simulations of a climatological summer or future emission scenario (lines 305-308). We have also revised the Conclusions to state that the

quantitative results are specific to the adopted 2022 baseline and should not be interpreted as direct projections of future ozone-control effectiveness (lines 893-908).

Future emission scenarios would address a different research question and are therefore outside the scope of the present process-attribution study. The revised manuscript presents the results as mechanistic evidence for temperature responses under contrasting chemical regimes rather than as quantitative predictions of future ozone-control outcomes.

References

Archer-Nicholls, S., Abraham, N. L., Shin, Y. M., Weber, J., Russo, M. R., Lowe, D., Utembe, S. R., O'Connor, F. M., Kerridge, B., Latter, B., Siddans, R., Jenkin, M., Wild, O., and Archibald, A. T.: The Common Representative Intermediates Mechanism Version 2 in the United Kingdom Chemistry and Aerosols Model, Journal of Advances in Modeling Earth Systems, 13, e2020MS002420, https://doi.org/10.1029/2020MS002420, 2021.
Cheng, X., Li, H., Zhang, Y., Li, Y., Zhang, W., Wang, X., Bi, F., Zhang, H., Gao, J., Chai, F., Lun, X., Chen, Y., Gao, J., and Lv, J.: Atmospheric isoprene and monoterpenes in a typical urban area of Beijing: Pollution characterization, chemical reactivity and source identification, Journal of Environmental Sciences, 71, 150-167, https://doi.org/10.1016/j.jes.2017.12.017, 2018.
Geng, F., Tie, X., Guenther, A., Li, G., Cao, J., and Harley, P.: Effect of isoprene emissions from major forests on ozone formation in the city of Shanghai, China, Atmos. Chem. Phys., 11, 10449-10459, 10.5194/acp-11-10449-2011, 2011.
Gu, C., Wang, S., Zhu, J., Wu, S., Duan, Y., Gao, S., and Zhou, B.: Investigation on the urban ambient isoprene and its oxidation processes, Atmospheric Environment, 270, 118870, https://doi.org/10.1016/j.atmosenv.2021.118870, 2022.
Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: 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, 10.5194/gmd-5-1471-2012, 2012.
Min, K. E., Pusede, S. E., Browne, E. C., LaFranchi, B. W., and Cohen, R. C.: Eddy covariance fluxes and vertical concentration gradient measurements of NO and NO₂ over a ponderosa pine ecosystem: observational evidence for within-canopy chemical removal of NO_x, Atmos. Chem. Phys., 14, 5495-5512, 10.5194/acp-14-5495-2014, 2014.
Pacifico, F., Harrison, S. P., Jones, C. D., Arneth, A., Sitch, S., Weedon, G. P., Barkley, M. P., Palmer, P. I., Serça, D., Potosnak, M., Fu, T. M., Goldstein, A., Bai, J., and Schurgers, G.: Evaluation of a photosynthesis-based biogenic isoprene emission scheme in JULES and simulation of isoprene emissions under present-day climate conditions, Atmos. Chem. Phys., 11, 4371-4389, 10.5194/acp-11-4371-2011, 2011.
Qin, M., She, Y., Wang, M., Wang, H., Chang, Y., Tan, Z., An, J., Huang, J., Yuan, Z., Lu, J., Wang, Q., Liu, C., Liu, Z., Xie, X., Li, J., Liao, H., Pye, H. O. T., Huang, C., Guo, S., Hu, M., Zhang, Y., Jacob, D. J., and Hu, J.: Increased urban ozone in heatwaves due to temperature-induced emissions of anthropogenic volatile organic compounds, Nature Geoscience, 18, 50-56, 10.1038/s41561-024-01608-w, 2025.
Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmospheric Environment (1967), 23, 1293-1304, https://doi.org/10.1016/0004-6981(89)90153-4, 1989.
Wu, K., Yang, X., Chen, D., Gu, S., Lu, Y., Jiang, Q., Wang, K., Ou, Y., Qian, Y., Shao, P., and Lu, S.: Estimation of biogenic VOC emissions and their corresponding impact on ozone and secondary organic aerosol formation in China, Atmospheric Research, 231, 104656, https://doi.org/10.1016/j.atmosres.2019.104656, 2020.
Yin, Q., Ma, Q., Lin, W., Xu, X., and Yao, J.: Measurement report: Long-term variations in surface NOx and SO2 mixing ratios from 2006 to 2016 at a background site in the Yangtze River Delta region, China, Atmos. Chem. Phys., 22, 1015-1033, 10.5194/acp-22-1015-2022, 2022.