



Title

CoMIAR v1.0: a hybrid system for multi-species adjoint inversion of atmospheric emissions with online adaptive regularization

Authors

Zelin Mai^{1,2}, Huizhong Shen^{1,2}, Qiming Liu^{1,2}, Ruixin Zhang^{1,2}, Peng Guo^{1,2}, Han Shen^{1,2},
Tiancheng Luan^{1,2}, Junpeng Xu^{1,2}, Jinling He^{1,2}, Zhiyu Zheng^{1,2}, Yilin Chen³, Shunliu Zhao⁴,
Amir Hakami⁴

Affiliations

¹State Key Laboratory of Soil Pollution Control and Safety, Shenzhen Key Laboratory of Precision Measurement and Early Warning Technology for Urban Environmental Health Risks, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China.

²Guangdong Provincial Observation and Research Station for Coastal Atmosphere and Climate of the Greater Bay Area, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China.

³School of Urban Planning and Design, Peking University, Shenzhen Graduate School, Shenzhen 100871, China.

⁴Department of Civil and Environmental Engineering, Carleton University, Ottawa, Ontario K1S5B6, Canada.

Correspondence: Huizhong Shen (shenhz@sustech.edu.cn)



ABSTRACT

Accurate estimates of anthropogenic emissions are fundamental to air quality and climate assessment. Satellite-based top-down inversion offers an independent constraint on bottom-up emission inventories, but two key limitations remain. First, dynamically constrained 4D-Var optimization often relies on fixed regularization parameters, which can destabilize convergence during iteration. Second, most inversions treat chemical species separately, even though atmospheric species are chemically coupled. Here, we develop CoMIAR v1.0 (Coupled Multi-species Inversion with Adaptive Regularization), a hybrid system implemented with CMAQ-Adjoint to address these limitations. CoMIAR combines mass-balance initialization with 4D-Var optimization, extends the adjoint formulation to a fully coupled multi-species control system, and incorporates an online Iterative-Adaptive Regularization (IAR) strategy that updates the species-specific regularization factor at each iteration using the L-curve criterion, thereby avoiding fixed empirical tuning. In observing system simulation experiments over China, IAR suppressed the convergence oscillations observed under fixed-parameter schemes. In the single-species SO₂ regularization experiment, IAR reduced emission NRMSE by 50.5% from the 4D-Var starting point and achieved the lowest final error among the tested schemes. Joint multi-species inversion further exploited sulfate-nitrate-ammonium thermodynamic coupling to separate emission biases that single-species inversions could not resolve. CoMIAR remained robust under spatially heterogeneous prior errors. Together, CoMIAR provides a stable and chemically consistent approach for emission inversion in complex atmospheric systems.

Keywords: CoMIAR; top-down emission estimation; CMAQ-Adjoint; four-dimensional variational inversion; satellite constraints; sulfate-nitrate-ammonium thermodynamics; aerosol optical depth; observing system simulation experiments; emission inventory uncertainty

1 INTRODUCTION

Emission inventories are key to source attribution, air-quality forecasting, and the design of emission-control policies. Bottom-up inventories depend on lagged activity statistics and



generalized emission factors, which introduce persistent biases and often lag behind actual emission changes by several years (Li et al., 2017). Satellite observations therefore provide a critical top-down complement for diagnosing inventory biases and updating emissions on regional to continental scales (Martin et al., 2003; Arellano et al., 2004; Vijayaraghavan et al., 2008; Lamsal et al., 2011; Koukouli et al., 2018; Zhang et al., 2025).

Over the past two decades, several families of inversion methods have been developed. Mass-balance (MB) and scaling-factor approaches relate observed column enhancements to emissions under simplified transport assumptions. They are computationally efficient but limited by steady-state and linearity assumptions that reduce transport and chemical fidelity (Martin et al., 2003; Zhang et al., 2025). Ensemble Kalman filter methods propagate flow-dependent error covariances without an adjoint model, yet require careful localization and inflation tuning and can struggle with strong nonlinearities in reactive chemistry (Miyazaki et al., 2012; Miyazaki et al., 2015; Miyazaki et al., 2017; Miyazaki et al., 2020). Adjoint-based four-dimensional variational (4D-Var) inversion minimizes a cost function using gradients from the adjoint of the chemical transport model, enabling optimization over high-dimensional emission fields and nonlinear chemical processes simultaneously (Müller and Stavrakou, 2005; Brasseur and Jacob, 2017; Qu et al., 2017; Chen et al., 2021). Of the three, the adjoint-based inversion best resolves nonlinear chemical feedbacks, though at higher computational cost.

Building on this capability, adjoint-based 4D-Var has been used to constrain emissions of single species and, increasingly, of jointly optimized species. Single-species studies cover both reactive gases and aerosol precursors, including NO_x over China from OMI NO_2 (Qu et al., 2017), U.S. NH_3 from IASI (Chen et al., 2021) and from TES with the GEOS-Chem adjoint (Zhu et al., 2013), and European NH_3 from CrIS (Cao et al., 2022). Where gas–aerosol thermodynamic partitioning is relevant, current adjoint models propagate the partitioning through ISORROPIA or equivalent modules in both the forward and backward calculation (Hakami et al., 2007; Zhu et al., 2013; Zhao et al., 2020a), so that ammonium-nitrate formation, for instance, contributes to the NH_3 adjoint sensitivity even when only NH_3 emissions are optimized. Joint multi-species 4D-Var inversions, by contrast, have so far been attempted only



for small sets of chemically interacting species (Müller and Stavrakou, 2005; Fortems-Cheiney et al., 2012; Xu et al., 2013; Qu et al., 2019; Opacka et al., 2025; Sfindla et al., 2026). For example, Qu et al. (2019) extended the hybrid framework to jointly invert NO_x and SO_2 over East Asia, showing that two-species inversion reduces cross-species error compensation. Müller and Stavrakou (2005) jointly optimized CO and NO_x emissions using the IMAGES adjoint, and Zheng et al. (2019) simultaneously assimilated CO, HCHO, and CH_4 satellite observations within a variational framework to close the global CO budget. Together, these studies demonstrate that multi-species constraints improve emission estimates.

In the sulfate–nitrate–ammonium (SNA) system, atmospheric NH_3 is governed not only by NH_3 emissions but also by the sulfur and nitric acid environment that controls gas–particle partitioning (Lachatre et al., 2019; Liu et al., 2022). Reductions in SO_2 over China, for example, have altered aerosol acidity and shifted the thermodynamic equilibrium, increasing ambient NH_3 and modifying the partitioning among sulfate, nitrate, and ammonium (Luo et al., 2022). A single-species NH_3 inversion operating under biased SO_2 or sulfate fields may compensate for errors in sulfur-related chemistry rather than isolate true NH_3 source biases, resulting in chemical aliasing and incorrect source attribution (Fu et al., 2017; Luo et al., 2022; Wen et al., 2023). Multi-species inversion is therefore needed not merely to add observational information but to ensure that posterior emissions are chemically self-consistent.

Extending adjoint-based inversion beyond a limited number of species to a fully coupled multi-species system is not a routine expansion of the control vector. A recognized challenge arises in the regularization of the inverse problem itself. The regularization parameter (γ) controls the weight of prior information relative to observational constraints in the cost function and directly affects whether the posterior emissions remain physically interpretable (Brasseur and Jacob, 2017; Chen et al., 2021). For single-species inversions, a fixed γ chosen from empirical experience or offline diagnostics is often adequate. When multiple chemically coupled species are optimized simultaneously, however, the cross-species error structures evolve during iteration, and a static γ can cause convergence stagnation, oscillatory behavior, or misallocation of emission corrections across species and regions (Hansen, 2001). Adaptive



regularization schemes that update γ in response to the evolving solution for individual species are therefore needed in multi-species adjoint inversion frameworks.

Here, we develop CoMIAR v1.0 (Coupled Multi-species Inversion with Adaptive Regularization), a hybrid system implemented with CMAQ-Adjoint, and evaluate it through observing system simulation experiments (OSSEs) over China. CoMIAR couples mass-balance initialization with 4D-Var optimization and incorporates an online Iterative-Adaptive Regularization (IAR) strategy, in which the species-specific γ is updated at each iteration using the L-curve criterion (Hansen, 2001). We use these experiments to assess whether adaptive regularization stabilizes convergence under spatially heterogeneous prior errors and whether joint constraints improve the recovery of chemically coupled emissions.

2 METHODS

2.1 Model configuration and CoMIAR workflow

We implemented CoMIAR v1.0 with CMAQ v5.0 and its adjoint system (Hakami et al., 2007; Zhao et al., 2020a). The simulation domain covers China at a horizontal resolution of 36 km. OSSEs were conducted for two days from July 1, 2017 to July 2, 2017. Meteorological fields were generated by the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008). Meteorological initial and boundary conditions were taken from the NCEP FNL Operational Global Analysis data on a $1^\circ \times 1^\circ$ grid, and chemical initial and lateral boundary conditions were provided by a hemispheric CMAQ simulation at 108 km resolution with 44 vertical layers (Shen et al., 2021; Chen et al., 2024; Zheng et al., 2024; Zhang et al., 2025).

For emissions, anthropogenic sources outside China were taken from HTAP_v3 ($0.1^\circ \times 0.1^\circ$), whereas anthropogenic emissions inside China were based on the MEIC integrated inventory ($0.1^\circ \times 0.1^\circ$) (Zhang et al., 2025). The China inventory combines the MEIC v1.3 baseline with several specialized high-resolution inventories for key economic zones and source categories, including ammonia, open burning, and shipping. The detailed composition of the integrated inventory is summarized in Table S3.



The WRF-CMAQ simulations used a physics and chemistry configuration previously applied and evaluated for regional air-quality modeling over China (Sarwar et al., 2008). WRF employed RRTMG for longwave and shortwave radiation, MYJ for the surface layer and planetary boundary layer, Noah for land surface processes, Lin microphysics, and the Grell 3D ensemble cumulus scheme. CMAQ used the CB05_AE5_Aq chemical mechanism so that gas-phase chemistry, aerosol processes, and aqueous reactions relevant to inorganic aerosol partitioning were represented consistently.

The inversion proceeded in two sequential phases (Figure 1). Phase I was a mass-balance update of the prior emissions. Under the assumption of a linear local relationship between emissions and column concentrations, this step efficiently removes large-scale first-order bias and provides an improved initial state for the subsequent variational optimization (Qu et al., 2017; Li et al., 2019; Chen et al., 2021). The update is written as:

$$E_t = E_a \times \frac{\Omega_o}{\Omega_a} \quad \text{Eq. 1}$$

where E represents emissions, Ω_o and Ω_a denote observed and simulated column concentrations, respectively.

Phase II used the mass-balance output as the a priori state in the CMAQ-Adjoint 4D-Var system. In this phase, gradients of the cost function were calculated with respect to multiple species simultaneously. This step resolves nonlinear chemical feedbacks, including aerosol thermodynamics and oxidation chemistry, that cannot be represented by the simplified mass-balance step.

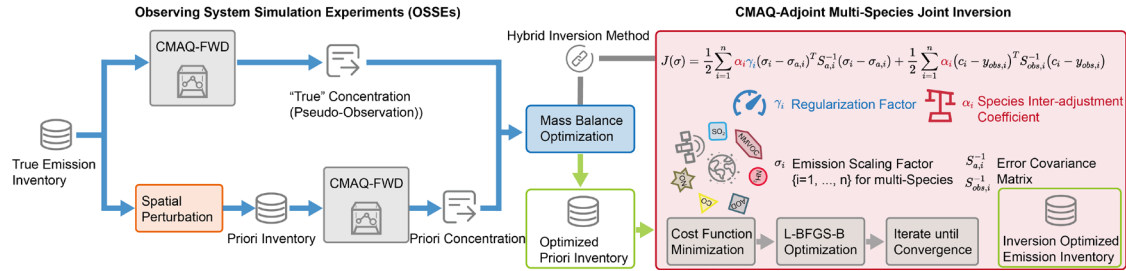


Figure 1. Workflow of CoMIAR v1.0 in the observing system simulation experiments (OSSEs). The upper branch generates pseudo-observations by driving the forward CMAQ model with the true emission inventory, whereas the lower branch perturbs the true inventory to construct the a priori inventory and prior concentrations. A mass-balance step first updates the a priori emissions, and the resulting optimized prior is then passed to the CMAQ-Adjoint 4D-Var system, where the multi-species cost function is minimized iteratively to obtain the final posterior emissions.

2.2 Cost function, species adjustment coefficient, and adaptive regularization

The 4D-Var inversion minimizes a scalar cost function that balances the deviation from the a priori emissions against the mismatch between simulations and observations. For the coupled multi-species system considered here, the cost function is written as:

$$J(\sigma) = \frac{1}{2} \sum_{i=1}^n \gamma_i \alpha_i (\sigma_i - \sigma_{a,i})^T S_{a,i}^{-1} (\sigma_i - \sigma_{a,i}) + \frac{1}{2} \sum_{i=1}^n \alpha_i [H_i(\sigma) - y_{obs,i}]^T S_{obs,i}^{-1} [H_i(\sigma) - y_{obs,i}] \quad \text{Eq. 2}$$

where σ_i represents the emission scaling factors for all optimized species i . The first term is the background constraint, in which the regularization factor γ_i controls the weight of the a priori information relative to the observational term. Because the relative importance of the background and observation terms changes during chemically coupled optimization, we did not prescribe a fixed γ_i . Instead, we used the IAR strategy, in which γ_i is updated during the 4D-Var iterations using an L-curve criterion. At each update, candidate γ_i values are evaluated by



comparing the background penalty with the observation mismatch, and the value near the corner of the L-curve is selected to balance emission adjustment against observational fitting. A detailed description of the L-curve implementation is provided in Sect. S1.

The second term aggregates constraints from multiple satellite datasets through the forward chemical transport operator and the observation error covariance matrix. The coefficient $\alpha_{a,i}$ is a species-level weighting factor applied to both the background and observation terms of species i , so that species with larger concentration magnitudes do not dominate the total multi-species cost function. Taking NH_3 as the reference, $\alpha_{a,i}$ is defined as the ratio of the regional concentration standard deviation of species i to that of NH_3 ; therefore, $\alpha_{\text{NH}_3} = 1$. This definition places the multi-species constraints on a more comparable scale. For computational efficiency, the background error covariance matrix $S_{a,i}$ and the observation error covariance matrix $S_{\text{obs},i}$ were both represented as diagonal matrices, whereas inter-species coupling was represented explicitly by the forward chemical transport operator.

Anthropogenic volatile organic compounds (VOCs) emissions were optimized as a lumped control variable, with the corresponding CMAQ species listed in Table S2. The aerosol optical depth (AOD) constraint was linked to aerosol-related emission controls in CMAQ, including primary elemental carbon (PEC), primary organic carbon (POC), and coarse particulate matter (PMC). AOD pseudo-observations were diagnostically reconstructed from simulated aerosol concentrations using a simplified extinction formulation:

$$AOD_{\text{model}} = \sum_{i=1}^N (\sigma_{\text{sp}} + \sigma_{\text{sp}})_i \Delta Z_i \quad \text{Eq.3}$$

where AOD_{model} is calculated from the vertically integrated sum of scattering and absorption contributions across model layers. The scattering term was parameterized as a function of hygroscopic inorganic aerosol, organic matter (OM), fine soil (FS), and coarse mass (CM), whereas the absorption term was linked to light-absorbing carbon (LAC). Here, CM denotes the coarse-mass term used in the AOD diagnostic and should be distinguished from the CMAQ



emission control variable PMC. The full expressions and variable definitions are provided in Sect. S2.

2.3 OSSE experimental design

We evaluated CoMIAR using a series of simulation experiments, in which the truth is known by construction. The original MEIC inventory was treated as the true emissions, and a CMAQ forward simulation driven by those emissions provided the reference concentration fields. All OSSE simulations covered July 1-2, 2017.

A priori emissions were created by perturbing the true inventory, and pseudo-observations were generated from the forward simulation driven by the true emissions. This design created a controlled difference between the prior and the truth, so that inversion performance could be evaluated directly against known emissions and concentration fields.

Three groups of OSSEs were designed to evaluate CoMIAR stepwise (Table S1). The first group used spatially heterogeneous prior perturbations in a single-species SO₂ inversion to compare fixed and adaptive regularization strategies. The fixed- γ experiments included four static regularization levels, denoted Fixed-init, Fixed-mean, Fixed-max, and Fixed-min, whereas the adaptive experiment used IAR to update γ during the 4D-Var iterations. The second group used uniform 200% perturbations to isolate the effect of SO₂-NH₃ chemical coupling in an idealized setting, comparing single-species SO₂ and NH₃ inversions with the corresponding joint SO₂-NH₃ inversion. The third group used spatially heterogeneous perturbations to represent inventory-like prior errors and compared the multi-species joint inversion with the corresponding single-species baselines for SO₂, NH₃, NO_x/NO₂, CO, lumped VOCs/formaldehyde (FORM), and aerosol/AOD-related controls. No strict convergence criterion was applied. Instead, all inversions were run beyond nominal convergence to allow fuller analysis of convergence behavior, oscillation, and cross-species responses.



3 RESULTS AND DISCUSSION

3.1 Single-species inversion with adaptive regularization

To evaluate adaptive regularization, we compared IAR with four fixed- γ strategies (Fixed-init, Fixed-mean, Fixed-max, and Fixed-min), each holding γ constant throughout its iterations at the initial, mean, maximum, or minimum value taken from the IAR γ trajectory (Table S1). All experiments shared the same mass-balance initialization, during which the normalized root mean square error (NRMSE) of SO₂ emissions declined from 8.04 to 6.54 over iterations 0–7, indicating that the preprocessing step removed a substantial portion of the domain-wide bias prior to the start of 4D-Var optimization. During the subsequent 4D-Var stage, the SO₂ emission NRMSE under IAR decreased from 6.5 at iteration 7 to 3.2 by iteration 50, corresponding to a relative reduction of 50.5%.

Compared to IAR, Fixed-init and Fixed-max converged slowly and sustained higher residual errors throughout 4D-Var (Figure 2a), leaving local hotspots undercorrected. Fixed-mean provided an intermediate response but did not reach the posterior error level of IAR. Fixed-min decreased rapidly at first and then oscillated, indicating overfitting. The spatial patterns confirm this diagnosis (Figure S1). Strong constraints (higher γ) produced broad residual structures, whereas weak constraints (lower γ) generated noisy artifacts around major source regions (Figures S1–S3).

Overall, IAR met the requirement for dynamic regularization. γ dropped rapidly at the start of 4D-Var when prior errors were large, and then adjusted iteratively as the inversion approached convergence (Figure 2b). This adaptive behavior preserved strong corrections where observational information was sufficient while suppressing high-frequency fluctuations once the major bias was removed, a differentiated response unattainable with static regularization.

It should be noted that the γ values for Fixed-init, Fixed-mean, Fixed-max, and Fixed-min could be assigned only because the γ range was already known from IAR. In practice, these values are unknown and must be set manually based on expert judgment. This task becomes increasingly difficult for multi-species inversion since each species requires its own γ



specification. By eliminating this manual tuning, IAR not only improves inversion efficiency
(Figures 2, S1–S3) but also makes multi-species joint inversion practically feasible.

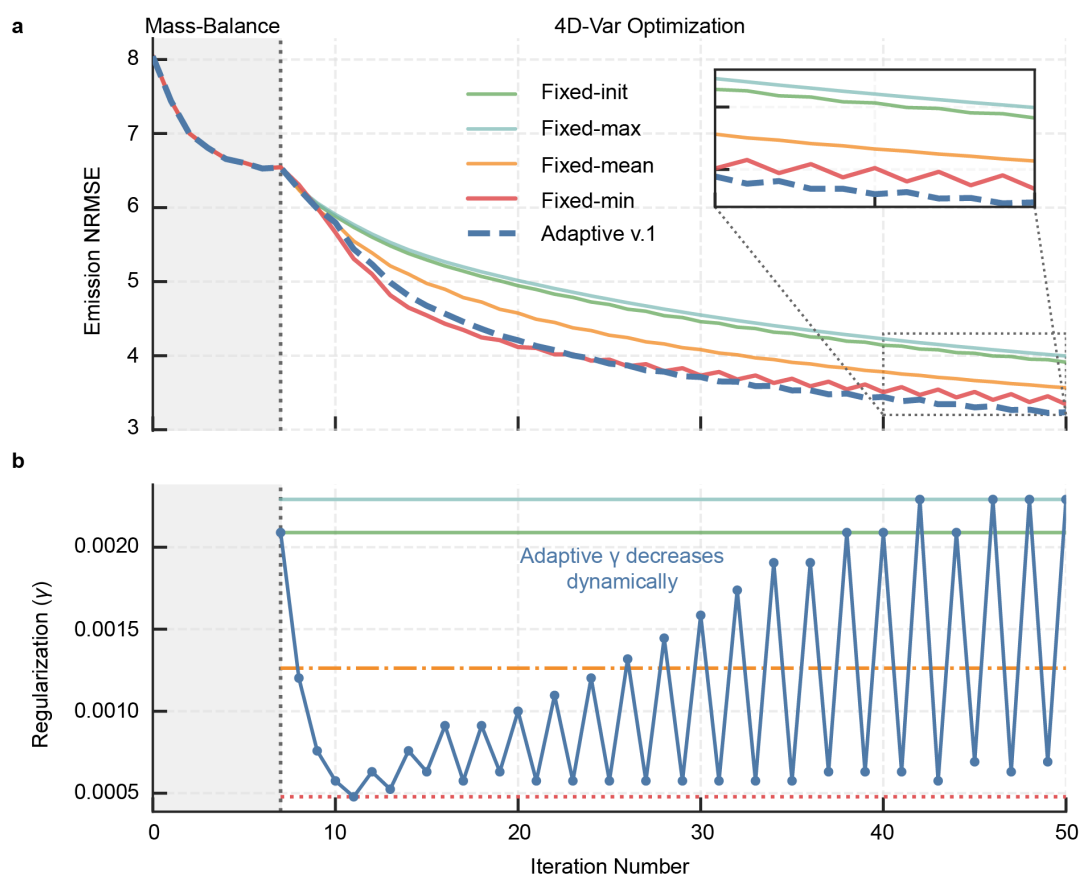


Figure 2. Comparison of inversion performance across regularization strategies. (a) Evolution of emission normalized root mean square error (NRMSE) during the shared mass-balance stage and the subsequent 4D-Var stage. The adaptive strategy achieves the lowest posterior error, whereas strong fixed constraints (Fixed-max and Fixed-init) slowed convergence, and the weak fixed constraint (Fixed-min) led to oscillatory convergence. **(b)** The



evolution of the regularization factor γ . The adaptive strategy (IAR) relaxes γ rapidly at the onset of 4D-Var and subsequently adjusts it dynamically, whereas the fixed strategies maintain constant values throughout.

3.2 Pairwise inversion under sulfate-ammonium coupling

Using IAR, we further conducted pairwise inversions to isolate and evaluate the effect of chemical coupling on inversion performance. We applied a uniform 200% perturbation to SO_2 and NH_3 emissions and compared single-species SO_2 , single-species NH_3 , and joint SO_2 – NH_3 inversions (Figure 3). The results showed distinct convergence behavior for SO_2 and NH_3 . SO_2 was constrained primarily by its own column abundance. Single-species and joint inversions produced nearly identical convergence trajectories, with emission NRMSE decreasing from ~ 4.0 to ~ 1.1 within 15–20 iterations and remaining stable thereafter (Figure 3a). Thus, SO_2 emission errors were largely identifiable from SO_2 observations alone, and additional NH_3 constraints provided little benefit.

In contrast, NH_3 inversion was strongly affected by cross-species chemical coupling. The single-species inversion oscillated and stalled at an emission NRMSE of approximately 1.7–2.0, whereas the joint inversion continued to decline to approximately 1.0 (Figure 3c). This divergence reflected chemical inconsistency rather than numerical noise. When SO_2 remained positively biased, as in the single-species NH_3 inversion, excess sulfate strengthened the thermodynamic sink for NH_3 , causing the inversion to compensate for sulfate errors through spurious NH_3 adjustments (Figures S4–S6). This chemical aliasing degraded not only NH_3 emission recovery but also secondary inorganic aerosol columns and the reconstructed ammonium partitioning state (Figures S4–S6).

Joint inversion substantially reduced this issue by correcting SO_2 and NH_3 simultaneously. Once the sulfate field was corrected, gas-particle partitioning was recovered more consistently, as shown by substantially lower secondary inorganic aerosol errors under the joint inversion (Figures S5 and S6). The NRMSE values for ANH_4 , ASO_4 , and ANO_3 were 60–90% lower under pairwise inversion than those under single-species inversion (Figure S5). The



reconstructed ammonium partitioning state also improved, with the domain-mean absolute ammonium partitioning coefficient error 19.5% lower under pairwise inversion than under the SO_2 -only inversion, although it was 4.4% higher than that under the NH_3 -only inversion; nevertheless, pairwise inversion produced lower absolute errors than the average of the two single-species inversions at 87.1% of valid grid cells (Figure S6). The largest improvements occurred in high-emission regions, where the additional error reduction under pairwise inversion relative to the NH_3 -only inversion reached 12.8%, compared with a domain-mean improvement of 8.1% (Figure 3d). These results demonstrate that for chemically or thermodynamically coupled species, joint inversion can be necessary to obtain physically consistent emissions and secondary aerosol fields.

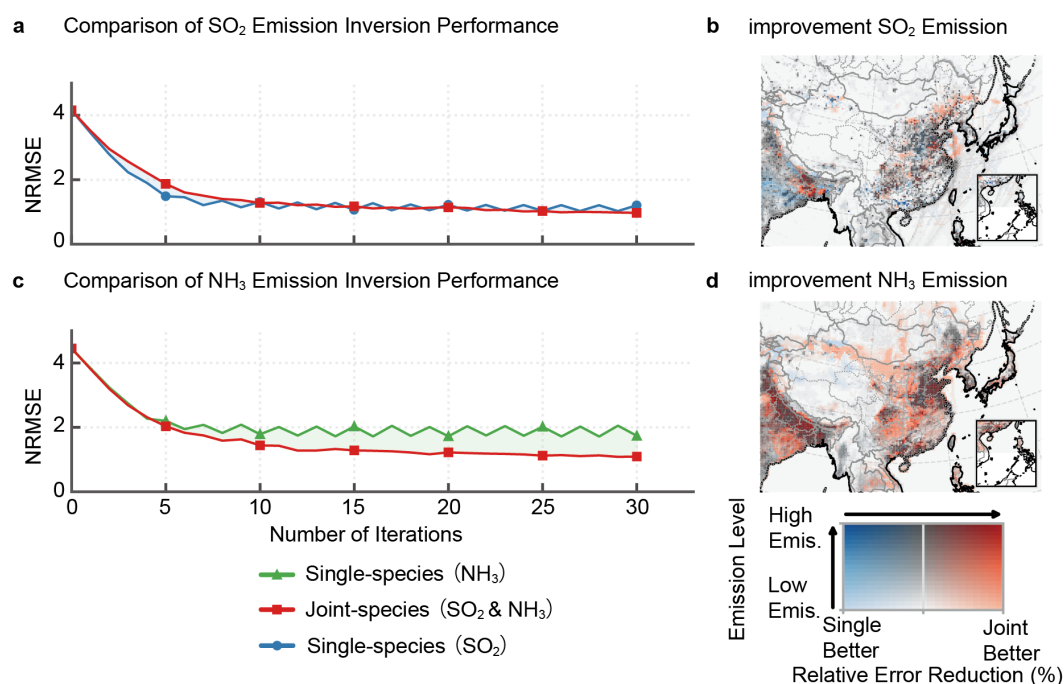


Figure 3. Joint inversion for NH_3 and SO_2 in an idealized experiment. Only SO_2 and NH_3 emissions were perturbed and optimized under controlled prior-error conditions, allowing direct comparison between the joint inversion and the corresponding single-species inversions. (a, c)



Evolution of emission NRMSE for SO₂ (a) and NH₃ (c). **(b, d)** Spatial distribution of the relative emission-error reduction achieved by joint inversion compared with single-species inversion for SO₂ (b) and NH₃ (d), overlaid with emission intensity. Red indicates regions where the joint inversion performs better, blue indicates regions where the single-species inversion performs better, and darker shading denotes higher-emission areas.

3.3 Multi-species inversion under heterogeneous prior errors

A practical advantage of adjoint-based inversion is that its computational cost scales primarily with the number of optimization iterations rather than with the number of constrained species. Single-species and multi-species adjoint-based inversions therefore require comparable resources at the same iteration count. However, combining multiple observational constraints into a single cost function can degrade convergence efficiency for individual species. To assess this trade-off, we compared the joint inversion of multiple species, including SO₂, NO_x, NH₃, CO, VOCs, and PM compounds, with the corresponding single-species inversions from a common 4D-Var starting point, using heterogeneous prior errors that mimic practical inversion conditions.

The joint inversion did not slow convergence relative to single-species inversion at matched iteration counts (Figure 4a–c; whole-domain summary in Figure S7). By iteration 20, the mean emission NRMSE across all evaluated constraints was comparable between the joint and single-species inversions (7.01 versus 6.63), indicating no substantial convergence penalty from the joint inversion. By iteration 50, the joint inversion yielded a lower mean concentration NRMSE (0.086 versus 0.095) while the mean emission NRMSE remained comparable (4.79 versus 5.15). The corresponding posterior spatial error fields show that the broad prior column biases were largely reduced for the gas-phase and integrated AOD constraints under the joint inversion (Figures S8 and S9). Performance across individual species and spatial scales was broadly consistent between the two configurations, with most NRMSE values over the last five iterations differing by less than 10% (Figure 4c and Figure S10). Both configurations also exhibited similar error structures across emission magnitudes (Figure 4d). Relative errors were



largest in the lowest-emission bins and decreased rapidly toward higher-emission bins (Figure 4d), indicating that low emissions remain difficult to constrain regardless of inversion configuration, likely because background concentrations sustained by transport from external sources mask the local emission signal. The joint inversion thus preserved its computational advantage without compromising convergence.

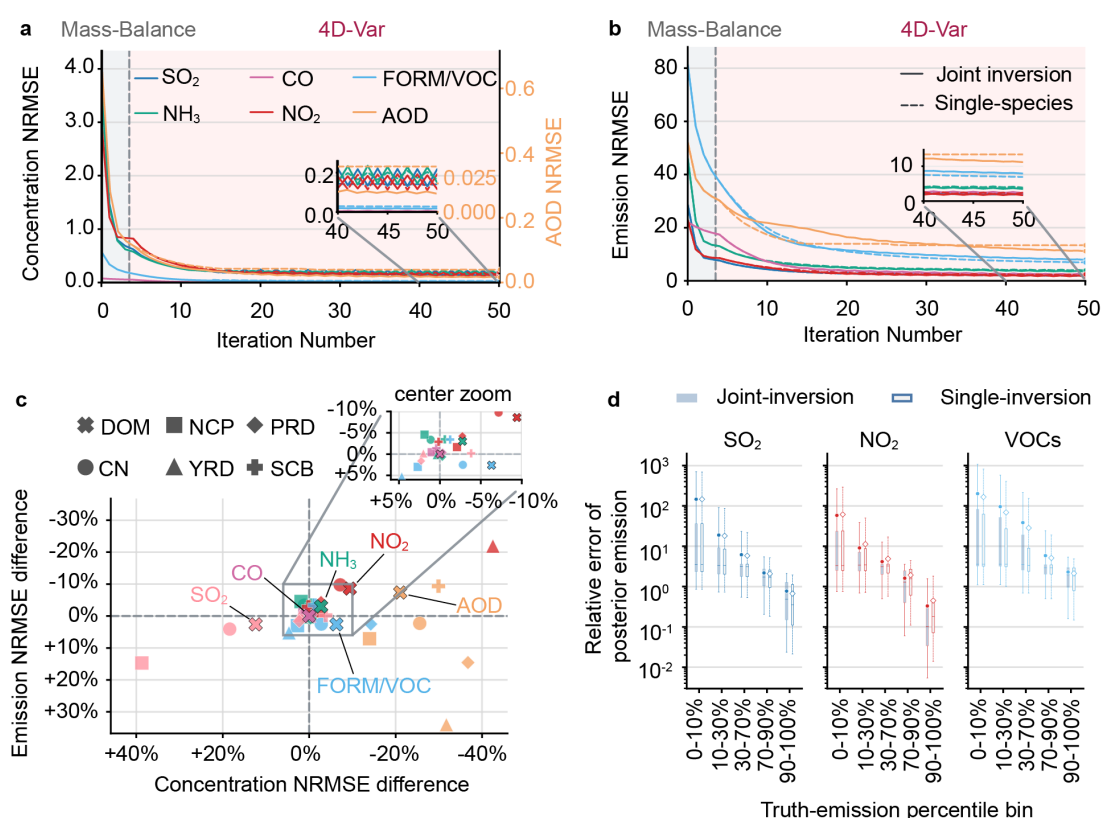


Figure 4. Performance of the joint and single-species inversions under heterogeneous prior errors. (a, b) Evolution of concentration NRMSE (a) and emission NRMSE (b) for the joint inversion and the corresponding single-species inversions. Gray and pink shading denote the Mass-Balance and 4D-Var phases, respectively. The inset magnifies iterations 40–50. (c) Relative differences in NRMSE between the joint and single-species inversions for concentration (x-axis) and emission (y-axis), averaged over the last five iterations, for the full



model domain, China, and key regions (NCP, YRD, PRD/GBA, and the Sichuan Basin). Negative values indicate that the joint inversion outperforms the single-species inversion. Colors denote the evaluated constraints or species. For formaldehyde (FORM), the concentration metric is based on FORM columns, while the emission metric reflects the recovery of the lumped VOCs control constrained by FORM. The inset enlarges the central cluster. **(d)** Distributions of posterior emission relative errors against “true” emissions, binned by “true” emission magnitude, for SO₂, NO₂, and VOCs. Relative error is defined as $|\text{posterior} - \text{true}|/\text{true}$ and is therefore dimensionless. A value of 10 corresponds to an error ten times the true emission. Boxes show the interquartile range, whiskers the 5th–95th percentile range, and the circle and diamond mark the means of the joint and single-species inversions, respectively.

4 Implications

In this study, we developed CoMIAR v1.0, an adjoint-based hybrid inversion system that integrates multiple observational constraints to enable simultaneous multi-species emission optimization. Single-species inversion experiments showed that the automatic adaptive regularization strategy, IAR, improved convergence efficiency relative to static regularization. This capability is critical for multi-species joint inversion, in which manual tuning of regularization parameters becomes increasingly impractical as the number of constrained species increases. Pairwise inversion experiments further demonstrated that joint inversion can produce emission estimates closer to truth than single-species inversion for chemically coupled species. Finally, joint inversion experiments for six groups of species showed that, with IAR and appropriate inter-species weighting, joint inversion achieved error reduction comparable to that of single-species inversions. Because adjoint-based inversion has a similar computational cost per iteration regardless of the number of constrained species, joint inversion substantially improves overall inversion efficiency, by approximately a factor of six in our case.

The development of CoMIAR carries important environmental implications for emission characterization, air quality management, and integrated impact assessment. Bottom-up inventories carry substantial uncertainties for reactive species such as NH₃ and VOCs, and top-



down constraints derived from single-species inversions are often biased by errors propagated from co-emitted or chemically coupled pollutants (Henze et al., 2009; Qu et al., 2019; Wang et al., 2020; Chen et al., 2021; Opacka et al., 2025). By constraining chemically interacting species simultaneously, multi-species inversion produces emission estimates that are internally consistent across species and provide a stronger basis for designing coordinated control strategies for PM_{2.5}, O₃, and their precursors, rather than optimizing each precursor in isolation. The resulting inventories also support more reliable source attribution and exposure assessment, both of which are essential for evaluating the health and climate co-benefits of emission mitigation. As ground-based and satellite observations continue to expand in coverage, resolution, and species diversity, CoMIAR offers a principled means of exploiting these complementary constraints to refine bottom-up inventories.

Several directions remain open. First, CoMIAR is less sensitive in low-emission regions, a limitation likely amplified under operational conditions because satellite retrievals over low-concentration areas typically carry larger relative biases. Introducing spatially differentiated regularization, such as smaller regularization weights in low-emission regions, once the optimization has reached intermediate convergence, may relax prior constraints and improve posterior adjustments, although the optimal scheme requires further testing. Second, extending CoMIAR to the O₃–NO_x–VOCs system is a natural next step. O₃ production responds nonlinearly to precursor ratios, and current top-down source apportionment struggles to separate NO_x-limited and VOCs-limited regimes. Incorporating O₃ observations into the cost function would constrain NO_x and VOCs emissions jointly and help disentangle regime-dependent sensitivities. Third, because this study is based on OSSEs, future applications should evaluate CoMIAR with real satellite retrievals, where correlated observation errors, cloud-induced sampling biases, and model representation errors in meteorology, mixing, and chemistry may all affect inversion performance. These uncertainties should be addressed through more realistic error characterization, sampling-bias correction, and observation-space diagnostics to ensure that the inversion remains physically meaningful under operational conditions. Finally, the computational cost of adjoint inversion remains a barrier to routine use.



397 Machine-learning surrogates could accelerate gradient evaluation either by replacing the full
398 forward–adjoint system with an end-to-end differentiable emulator or by selectively
399 substituting expensive modules, such as gas-phase chemistry, aerosol thermodynamics, and
400 photolysis, while retaining the physically based dynamical core (Zheng et al., 2026). Together,
401 these advances would extend CoMIAR from controlled OSSE evaluation toward a robust and
402 scalable system for chemically consistent emission estimation.
403



5 Code and data availability

The exact version of CoMIAR used in this study (v1.0) is archived on Zenodo at <https://doi.org/10.5281/zenodo.21389865> (Mai et al., 2026) under the MIT License. The archive contains the custom Python inversion drivers and helper modules, forward and adjoint runtime wrappers, configuration arrays, dependency specifications, and README documentation mapping the simulations underlying Figs. 2-4 to the corresponding code. It also contains the processed source data and plotting scripts required to reproduce these figures, which are available under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. The plotting workflow can be run directly from the archive, whereas reproducing the inversions additionally requires the CMAQ 5.0 Adjoint and an MPI-enabled high-performance computing environment. The forward and adjoint source code of CMAQ 5.0 Adjoint is archived on Zenodo at <https://doi.org/10.5281/zenodo.3780216> (Zhao et al., 2020b) under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. The baseline MEIC product used in the integrated Chinese emission inventory was MEIC v1.3. It is available to registered users for download from the MEIC website (<http://meicmodel.org.cn/>, last access: 26 July 2026) under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. The full model outputs are not archived because of their volume, but the processed outputs used in Figs. 2-4 are included in the CoMIAR archive.

6 Supplement

The supplement contains two supplementary text sections, 10 figures, and 3 tables with additional information on the L-curve-based adaptive regularization strategy, AOD reconstruction equations and notation definitions, posterior emission fidelity under different regularization constraints, simulated-versus-observed SO₂ column comparisons, posterior SO₂ emission recovery, error compensation in single-species and joint-species inversions, posterior secondary inorganic aerosol columns, aerosol ammonium partitioning ratio diagnostics, joint-versus-single inversion performance under heterogeneous prior errors, prior and posterior concentration error distributions, regional contrasts in joint-inversion benefits, experimental configurations for the OSSEs, VOC species included in the lumped VOC control variable, and the composition of the integrated MEIC emission inventory.



7 Author contribution

ZM and HZS designed the study. ZM developed CoMIAR, performed the simulations and formal analysis, prepared the figures, and drafted the manuscript. QL, RZ, PG, HS, TL, JX, JH and ZZ contributed to the model experiments, data processing, validation, and interpretation of the results. YC, SZ, and AH contributed to the inversion methodology, model interpretation, and manuscript revision. HZS supervised the study and acquired funding. All authors discussed the results and contributed to the final manuscript.

8 Competing interests

The authors declare that they have no conflict of interest.

9 Acknowledgements

The authors acknowledge the Center for Computational Science and Engineering at Southern University of Science and Technology for providing computational resources.

10 Financial support

This research has been supported by the National Key R&D Program of China (2023YFE0112901), the National Natural Science Foundation of China (42475108, 42192512, 42330709, and 42571087), the Shenzhen Science and Technology Program (KQTD20240729102048052 and JCYJ20241202152804007), the European Union's Horizon Europe research and innovation programme (101137905), the Guangdong Provincial Field Observation and Research Station for Coastal Atmosphere and Climate of the Greater Bay Area (2021B1212050024), and High-level Special Funds (G03034K006).



11 References

- Arellano Jr., A. F., Kasibhatla, P. S., Giglio, L., van der Werf, G. R., and Randerson, J. T.: Top-down Estimates of Global CO Sources Using MOPITT Measurements, *Geophys. Res. Lett.*, 31, L01104, <https://doi.org/10.1029/2003GL018609>, 2004.
- Brasseur, G. P. and Jacob, D. J.: *Modeling of Atmospheric Chemistry*, Cambridge University Press, Cambridge, <https://doi.org/10.1017/9781316544754>, 2017.
- Cao, H., Henze, D. K., Zhu, L., Shephard, M. W., Cady-Pereira, K., Dammers, E., Sitwell, M., Heath, N., Lonsdale, C., Bash, J. O., Miyazaki, K., Flechard, C., Fauvel, Y., Kruit, R. W., Feigenspan, S., Brümmner, C., Schrader, F., Twigg, M. M., Leeson, S., Tang, Y. S., Stephens, A. C. M., Braban, C., Vincent, K., Meier, M., Seidler, E., Geels, C., Ellermann, T., Sanocka, A., and Capps, S. L.: 4D-Var Inversion of European NH₃ Emissions Using CrIS NH₃ Measurements and GEOS-Chem Adjoint With Bi-Directional and Uni-Directional Flux Schemes, *J. Geophys. Res.-Atmos.*, 127, e2021JD035687, <https://doi.org/10.1029/2021JD035687>, 2022.
- Chen, Y., Shen, H., Kaiser, J., Hu, Y., Capps, S. L., Zhao, S., Hakami, A., Shih, J.-S., Pavur, G. K., Turner, M. D., Henze, D. K., Resler, J., Nenes, A., Napelenok, S. L., Bash, J. O., Fahey, K. M., Carmichael, G. R., Chai, T., Clarisse, L., Coheur, P.-F., Van Damme, M., and Russell, A. G.: High-Resolution Hybrid Inversion of IASI Ammonia Columns to Constrain US Ammonia Emissions Using the CMAQ Adjoint Model, *Atmos. Chem. Phys.*, 21, 2067–2082, <https://doi.org/10.5194/acp-21-2067-2021>, 2021.
- Chen, Y., Shen, H., Shen, G., Ma, J., Cheng, Y., Russell, A. G., Zhao, S., Hakami, A., and Tao, S.: Substantial Differences in Source Contributions to Carbon Emissions and Health Damage Necessitate Balanced Synergistic Control Plans in China, *Nat. Commun.*, 15, 5880, <https://doi.org/10.1038/s41467-024-50327-8>, 2024.
- Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunio, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The Formaldehyde Budget as Seen by a Global-Scale Multi-Constraint and Multi-Species Inversion System, *Atmos. Chem. Phys.*, 12, 6699–6721, <https://doi.org/10.5194/acp-12-6699-2012>, 2012.
- Fu, X., Wang, S., Xing, J., Zhang, X., Wang, T., and Hao, J.: Increasing Ammonia Concentrations Reduce the Effectiveness of Particle Pollution Control Achieved via SO₂ and NO_x Emissions Reduction in East China, *Environ. Sci. Technol. Lett.*, 4, 221–227, <https://doi.org/10.1021/acs.estlett.7b00143>, 2017.
- Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D., and Li, Q.: The Adjoint of CMAQ, *Environ. Sci. Technol.*, 41, 7807–7817, <https://doi.org/10.1021/es070944p>, 2007.
- Hansen, P. C.: The L-Curve and Its Use in the Numerical Treatment of Inverse Problems, in: *Computational Inverse Problems in Electrocardiology*, *Advances in Computational Bioengineering*, Vol. 5, edited by: Johnston, P. R., WIT Press, Southampton, UK, 119–142, 2001.
- Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse Modeling and Mapping US Air Quality Influences of Inorganic PM_{2.5} Precursor Emissions Using the Adjoint of GEOS-



- 494 Chem, Atmos. Chem. Phys., 9, 5877–5903, <https://doi.org/10.5194/acp-9-5877-2009>,
495 2009.
- 496 Koukouli, M. E., Theys, N., Ding, J., Zyrichidou, I., Mijling, B., Balis, D., and van der A, R.
497 J.: Updated SO₂ Emission Estimates over China Using OMI/Aura Observations, Atmos.
498 Meas. Tech., 11, 1817–1832, <https://doi.org/10.5194/amt-11-1817-2018>, 2018.
- 499 Lachatre, M., Fortems-Cheiney, A., Foret, G., Siour, G., Dufour, G., Clarisse, L., Clerbaux, C.,
500 Coheur, P.-F., Van Damme, M., and Beekmann, M.: The Unintended Consequence of
501 SO₂ and NO₂ Regulations over China: Increase of Ammonia Levels and Impact on
502 PM_{2.5} Concentrations, Atmos. Chem. Phys., 19, 6701–6716, [https://doi.org/10.5194/acp-](https://doi.org/10.5194/acp-19-6701-2019)
503 19-6701-2019, 2019.
- 504 Lamsal, L. N., Martin, R. V., Padmanabhan, A., van Donkelaar, A., Zhang, Q., Sioris, C. E.,
505 Chance, K., Kurosu, T. P., and Newchurch, M. J.: Application of Satellite Observations
506 for Timely Updates to Global Anthropogenic NO_x Emission Inventories, Geophys. Res.
507 Lett., 38, L05810, <https://doi.org/10.1029/2010GL046476>, 2011.
- 508 Li, C., Martin, R. V., Shephard, M. W., Cady-Pereira, K., Cooper, M. J., Kaiser, J., Lee, C. J.,
509 Zhang, L., and Henze, D. K.: Assessing the Iterative Finite Difference Mass Balance and
510 4D-Var Methods to Derive Ammonia Emissions Over North America Using Synthetic
511 Observations, J. Geophys. Res.-Atmos., 124, 4222–4236,
512 <https://doi.org/10.1029/2018JD030183>, 2019.
- 513 Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H.,
514 Zhang, Q., and He, K.: Anthropogenic Emission Inventories in China: A Review, Natl.
515 Sci. Rev., 4, 834–866, <https://doi.org/10.1093/nsr/nwx150>, 2017.
- 516 Liu, Y., Zhan, J., Zheng, F., Song, B., Zhang, Y., Ma, W., Hua, C., Xie, J., Bao, X., Yan, C.,
517 Bianchi, F., Petäjä, T., Ding, A., Song, Y., He, H., and Kulmala, M.: Dust Emission
518 Reduction Enhanced Gas-to-Particle Conversion of Ammonia in the North China Plain,
519 Nat. Commun., 13, 6887, <https://doi.org/10.1038/s41467-022-34733-4>, 2022.
- 520 Luo, Z., Zhang, Y., Chen, W., Van Damme, M., Coheur, P.-F., and Clarisse, L.: Estimating
521 Global Ammonia (NH₃) Emissions Based on IASI Observations from 2008 to 2018,
522 Atmos. Chem. Phys., 22, 10375–10388, <https://doi.org/10.5194/acp-22-10375-2022>,
523 2022.
- 524 Mai, Z., Shen, H., Liu, Q., Zhang, R., Guo, P., Shen, H., Luan, T., Xu, J., He, J., Zheng, Z.,
525 Chen, Y., Zhao, S., and Hakami, A.: CoMIAR v1.0: source code and figure source data
526 for multi-species atmospheric emission inversion, Zenodo [code],
527 <https://doi.org/10.5281/zenodo.21389865>, 2026.
- 528 Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: Global
529 Inventory of Nitrogen Oxide Emissions Constrained by Space-Based Observations of
530 NO₂ Columns, J. Geophys. Res.-Atmos., 108, 4537,
531 <https://doi.org/10.1029/2003JD003453>, 2003.
- 532 Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.:
533 Simultaneous Assimilation of Satellite NO₂, O₃, CO, and HNO₃ Data for the Analysis
534 of Tropospheric Chemical Composition and Emissions, Atmos. Chem. Phys., 12, 9545–
535 9579, <https://doi.org/10.5194/acp-12-9545-2012>, 2012.



- Miyazaki, K., Eskes, H. J., and Sudo, K.: A Tropospheric Chemistry Reanalysis for the Years 2005–2012 Based on an Assimilation of OMI, MLS, TES, and MOPITT Satellite Data, *Atmos. Chem. Phys.*, 15, 8315–8348, <https://doi.org/10.5194/acp-15-8315-2015>, 2015.
- Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., and Kanaya, Y.: Decadal Changes in Global Surface NO_x Emissions from Multi-Constituent Satellite Data Assimilation, *Atmos. Chem. Phys.*, 17, 807–837, <https://doi.org/10.5194/acp-17-807-2017>, 2017.
- Miyazaki, K., Bowman, K. W., Yumimoto, K., Walker, T., and Sudo, K.: Evaluation of a Multi-Model, Multi-Constituent Assimilation Framework for Tropospheric Chemical Reanalysis, *Atmos. Chem. Phys.*, 20, 931–967, <https://doi.org/10.5194/acp-20-931-2020>, 2020.
- Müller, J.-F. and Stavrou, T.: Inversion of CO and NO_x Emissions Using the Adjoint of the IMAGES Model, *Atmos. Chem. Phys.*, 5, 1157–1186, <https://doi.org/10.5194/acp-5-1157-2005>, 2005.
- Opacka, B., Stavrou, T., Müller, J.-F., De Smedt, I., van Geffen, J., Marais, E. A., Horner, R. P., Millet, D. B., Wells, K. C., and Guenther, A. B.: Natural Emissions of VOC and NO_x over Africa Constrained by TROPOMI HCHO and NO₂ Data Using the MAGRITTEv1.1 Model, *Atmos. Chem. Phys.*, 25, 2863–2894, <https://doi.org/10.5194/acp-25-2863-2025>, 2025.
- Qu, Z., Henze, D. K., Capps, S. L., Wang, Y., Xu, X., Wang, J., and Keller, M.: Monthly Top-down NO_x Emissions for China (2005–2012): A Hybrid Inversion Method and Trend Analysis, *J. Geophys. Res.-Atmos.*, 122, 4600–4625, <https://doi.org/10.1002/2016JD025852>, 2017.
- Qu, Z., Henze, D. K., Theys, N., Wang, J., and Wang, W.: Hybrid Mass Balance/4D-Var Joint Inversion of NO_x and SO₂ Emissions in East Asia, *J. Geophys. Res.-Atmos.*, 124, 8203–8224, <https://doi.org/10.1029/2018JD030240>, 2019.
- Sarwar, G., Luecken, D., Yarwood, G., Whitten, G. Z., and Carter, W. P. L.: Impact of an Updated Carbon Bond Mechanism on Predictions from the CMAQ Modeling System: Preliminary Assessment, *J. Appl. Meteorol. Clim.*, 47, 3–14, <https://doi.org/10.1175/2007JAMC1393.1>, 2008.
- Sfendla, Y., Stavrou, T., Müller, J.-F., Oomen, G.-M., Opacka, B., Danckaert, T., De Smedt, I., and Lerot, C.: Global VOC Emissions Quantified from Inversion of TROPOMI Spaceborne Formaldehyde and Glyoxal Data, *Atmos. Chem. Phys.*, 26, 733–767, <https://doi.org/10.5194/acp-26-733-2026>, 2026.
- Shen, H., Shen, G., Chen, Y., Russell, A. G., Hu, Y., Duan, X., Meng, W., Xu, Y., Yun, X., Lyu, B., Zhao, S., Hakami, A., Guo, J., Tao, S., and Smith, K. R.: Increased Air Pollution Exposure among the Chinese Population during the National Quarantine in 2020, *Nat. Hum. Behav.*, 5, 239–246, <https://doi.org/10.1038/s41562-020-01018-z>, 2021.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR Technical Note NCAR/TN-475+STR, National Center for Atmospheric Research, Boulder, CO, <https://doi.org/10.5065/D68S4MVH>, 2008.



- Vijayaraghavan, K., Snell, H. E., and Seigneur, C.: Practical Aspects of Using Satellite Data in Air Quality Modeling, *Environ. Sci. Technol.*, 42, 8187–8192, <https://doi.org/10.1021/es7031339>, 2008.
- Wang, Y., Wang, J., Xu, X., Henze, D. K., Qu, Z., and Yang, K.: Inverse Modeling of SO₂ and NO_x Emissions over China Using Multisensor Satellite Data – Part 1: Formulation and Sensitivity Analysis, *Atmos. Chem. Phys.*, 20, 6631–6650, <https://doi.org/10.5194/acp-20-6631-2020>, 2020.
- Wen, L., Xue, L., Dong, C., Wang, X., Chen, T., Jiang, Y., Gu, R., Zheng, P., Li, H., Shan, Y., Zhu, Y., Zhao, Y., Yin, X., Liu, H., Gao, J., Wu, Z., Wang, T., Herrmann, H., and Wang, W.: Reduced Atmospheric Sulfate Enhances Fine Particulate Nitrate Formation in Eastern China, *Sci. Total Environ.*, 898, 165303, <https://doi.org/10.1016/j.scitotenv.2023.165303>, 2023.
- Xu, X., Wang, J., Henze, D. K., Qu, W., and Kopacz, M.: Constraints on Aerosol Sources Using GEOS-Chem Adjoint and MODIS Radiances, and Evaluation with Multisensor (OMI, MISR) Data, *J. Geophys. Res.-Atmos.*, 118, 6396–6413, <https://doi.org/10.1002/jgrd.50515>, 2013.
- Zhang, R., Chen, Z., Wu, X., Liu, Q., Mai, Z., Zheng, Z., Chen, Y., Tao, S., Hu, Y., Zhao, S., Hakami, A., Russell, A. G., and Shen, H.: Adjoint Analysis of PM_{2.5} and O₃ Episodes in Priority Control Zones in China, *Environ. Sci. Ecotechnol.*, 27, 100612, <https://doi.org/10.1016/j.ese.2025.100612>, 2025.
- Zhao, S., Russell, M. G., Hakami, A., Capps, S. L., Turner, M. D., Henze, D. K., Percell, P. B., Resler, J., Shen, H., Russell, A. G., Nenes, A., Pappin, A. J., Napelenok, S. L., Bash, J. O., Fahey, K. M., Carmichael, G. R., Stanier, C. O., and Chai, T.: A Multiphase CMAQ Version 5.0 Adjoint, *Geosci. Model Dev.*, 13, 2925–2944, <https://doi.org/10.5194/gmd-13-2925-2020>, 2020a.
- Zhao, S., Russell, M. G., Hakami, A., Capps, S. L., Turner, M. D., Henze, D. K., Percell, P. B., Resler, J., Shen, H., Russell, A. G., Nenes, A., Pappin, A. J., Napelenok, S. L., Bash, J. O., Fahey, K. M., Carmichael, G. R., Stanier, C. O., and Chai, T.: CMAQ 5.0 Adjoint, Zenodo [code], <https://doi.org/10.5281/zenodo.3780216>, 2020b.
- Zheng, B., Chevallier, F., Yin, Y., Ciais, P., Fortems-Cheiney, A., Deeter, M. N., Parker, R. J., Wang, Y., Worden, H. M., and Zhao, Y.: Global Atmospheric Carbon Monoxide Budget 2000–2017 Inferred from Multi-Species Atmospheric Inversions, *Earth Syst. Sci. Data*, 11, 1411–1436, <https://doi.org/10.5194/essd-11-1411-2019>, 2019.
- Zheng, L., Adalibieke, W., Zhou, F., He, P., Chen, Y., Guo, P., He, J., Zhang, Y., Xu, P., Wang, C., Ye, J., Zhu, L., Shen, G., Fu, T.-M., Yang, X., Zhao, S., Hakami, A., Russell, A. G., Tao, S., Meng, J., and Shen, H.: Health Burden from Food Systems Is Highly Unequal across Income Groups, *Nat. Food*, 5, 251–261, <https://doi.org/10.1038/s43016-024-00946-7>, 2024.
- Zheng, Z., Shen, H., Chen, Y., Guo, P., Wang, Z., Zheng, L., Liu, Q., Luan, T., Sun, H. Z., and Tao, S.: Advancing Finely Resolved Adjoint Analysis of Global Premature Mortality From Ambient Exposure to Fine Particulate Matter, *J. Geophys. Res.-Atmos.*, 131, e2025JD043629, <https://doi.org/10.1029/2025JD043629>, 2026.



620 Zhu, L., Henze, D. K., Cady-Pereira, K. E., Shephard, M. W., Luo, M., Pinder, R. W., Bash, J.
621 O., and Jeong, G.-R.: Constraining U.S. Ammonia Emissions Using TES Remote Sensing
622 Observations and the GEOS-Chem Adjoint Model, *J. Geophys. Res.-Atmos.*, 118, 3355–
623 3368, <https://doi.org/10.1002/jgrd.50166>, 2013.
624