RC1: 'Comment on egusphere-2025-4938', Anonymous Referee #1, 01 Dec 2025

This manuscript combines WRF-Chem with the DART system to perform a top-down inversion of CO2 and CH4 using atmospheric observations, constraining concentrations and fluxes at a 9-km resolution. Given the growing need to improve national-scale greenhouse gas estimates within the MMRV framework, the topic is timely and of scientific relevance. However, for the reasons outlined below, I believe the manuscript in its current form is not yet suitable for publication in Geoscientific Model Development.

>> We sincerely thank you for your insightful evaluation and for acknowledging the importance of high-resolution greenhouse gas (GHG) monitoring within the MMRV framework. We have carefully addressed your concerns regarding technical novelty, the necessity of an OSSE, and the justification for our modeling configurations

1. Sustainability for GMD

The manuscript does not clearly articulate the novelty of the WRF-Chem/DART modeling framework or how it differs from existing inversion systems. Details necessary for reproducibility—an essential requirement for GMD—are limited. The numerical settings and methodological choices require clearer justification to meet the journal's standards for model development papers.

>> We thank the reviewer for constructive comments and we revised our manuscript as the reviewer suggested. Our revision is based on the following:

1. As the reviewer mentioned, reproducibility is an important requirement of GMD. Our paper corresponds to the type of "MODEL DESCRIPTION PAPERS" at GMD (https://www.geoscientific-model-development.net/about/manuscript_types.html#item1) and please consider that reproducibility has been checked before the review process. At the submission and discussion state, the full model and DA codes and input data configurations have been published on a persistent public archive (zenodo with DOI number) with a unique identifier as described in the manuscript. We check if we can make better readability for the reproducibility during the revision and we further substantially expanded technical details throughout the manuscript (key DA settings including the augmented state vector updates, localization/inflation and perturbations regarding IC/BCs and emissions). As a type of "Model Description Papers in GMD", we put more focus on presenting a general framework development and introducing its usefulness and reproducibility of the system, while reserving broader case-driven evaluations and pure scientific conclusions for follow-on work. We also document the coupled WRF-Chem/DART cycling workflow step-by-step and summarize key configuration settings used in this study for reproducibility.

2. As we described in the manuscript, our system is a regional CO_2/CH_4 top-down framework that couples online Eulerian transport (WRF-Chem) model with ensemble data assimilation

(DART/EAKF), incorporating model-physics extensions and configurations tailored to the complex urban–mountain–coastal heterogeneity of East Asia. We also revised the manuscript to better articulate novelty of our system to better show the following characteristics:

- Complementary benefits of Eulerian DA systems relative to Lagrangian inversions:

For regional inversions, an online Eulerian system offers (1) spatially continuous 4D fields of both meteorology and tracer concentrations over the full domain (rather than only receptor-based footprints), (2) dynamical–physical consistency because transport and key sub-grid physical processes (e.g., PBL mixing, mesoscale circulations) are evolved interactively within the coupled model rather than treated offline, and (3) a natural pathway to multivariate cycling in which transport-state errors (e.g., initial conditions, boundary-layer mixing, mesoscale flow errors) can be corrected within the same assimilation framework, reducing the risk that residual transport errors are aliased into flux adjustments.

- Novelty within the Eulerian paradigm: Korea-focused transport model (WRF-Chem) physics extensions and dual-species multivariate cycling. We also emphasize that this physics extension is generally applicable to other complex landscapes. Building on the Eulerian advantages above, our framework adds two system-level innovations tailored to the Korean Peninsula:

(1) Terrain-aware boundary-layer dispersion physics: we incorporate surface heterogeneity parameterizations relevant to Korea's urban–mountain–coastal heterogeneity, including roughness sublayer effects and canopy-height adjustments. These options are implemented in our WRF-Chem configuration to better represent near-surface mixing and dispersion over complex terrain, building on previous works (Lee and Hong, 2016; Lee et al., 2020; Kim et al., 2024).

(2) Dual-species, dual-state joint updates with meteorology: Unlike standard flux-only inversions, our top-down system jointly updates a comprehensive state vector comprising CO_2 and CH_4 concentrations, their respective emissions, and meteorological fields. This multivariate framework ensures physical consistency between the atmospheric states and surface fluxes of both gases.

These two elements are explicitly documented in the revised manuscript and Supplement to enable reproducible reuse of the configuration for Korea-focused regional applications.

Kim, J., Shin, H.-J., Lee, K., and Hong, J.: Enhancement of ANN-based wind power forecasting by modification of surface roughness parameterization over complex terrain, Journal of Environmental Management, **362**, 121246, https://doi.org/10.1016/j.jenvman.2024.121246, 2024.

Lee, J. and Hong, J.: Implementation of spaceborne lidar-retrieved canopy height in the WRF model, Journal of Geophysical Research: Atmospheres, **121**, 6863–6876, https://doi.org/10.1002/2015JD024299, 2016.

Lee, J., Hong, J., Noh, Y., and Jiménez, P. A.: Implementation of a roughness sublayer parameterization in the Weather Research and Forecasting model (WRF version 3.7.1) and its evaluation for regional climate simulations, Geoscientific Model Development, **13**, 521–536, https://doi.org/10.5194/gmd-13-521-2020, 2020.

2. Insufficient OSSE-based accuracy evaluation

Although the study focuses on developing an inversion framework, it lacks prior validation of its performance. An OSSE experiment using synthetic observations would allow the authors to quantitatively assess the accuracy and robustness of the system before applying it to real observations.

>> We agree that an OSSE (twin experiment with a known truth) provides a quantitative check of whether the inversion machinery behaves as intended. To address this point, we added a targeted OSSE to the revised manuscript, using an expanded 50-site synthetic surface network to test whether the framework can recover the intentionally imposed “missing power-industry” emission signal (a highly heterogeneous, point-source-dominated target) and to diagnose filter stability/consistency (including uncertainty reduction and innovation behavior). All OSSE configuration details and results are documented in Section 6 of the revised manuscript and Supplement for reproducibility.

We emphasize that this OSSE is included to demonstrate algorithmic consistency and stability under controlled conditions; real-world performance remains influenced by transport errors, boundary-condition biases, and representativeness errors, which we discuss explicitly. The main focus of the manuscript remains the real-data demonstration.

3. Limited description of DART and EAKF methodology

The manuscript provides only a brief description of the DART system and the Ensemble Adjustment Kalman Filter. Further explanation of the theoretical basis, implementation steps, and internal processing is needed to help readers understand how the assimilation framework functions.

>> As suggested by the reviewer, we substantially expanded the description of the DART system and the Ensemble Adjustment Kalman Filter (EAKF). In the revised manuscript (section 2.2), we explicitly define the augmented state vector used in this study and provide a concise mathematical summary of the EAKF analysis update. We also document the coupled WRF-Chem/DART cycling workflow step-by-step and summarize key configuration settings used in this study for reproducibility. Finally, we include more standard DART/EAKF references and documentation for readers who want additional algorithmic details.

4. Unclear scope of assimilated variables

Lines 133–134 suggest that both greenhouse gases and meteorological fields are assimilated, but the specific meteorological variables (e.g., temperature, wind components, humidity, surface variables) are not identified. Clarifying the full set of assimilated variables is essential.

>> As suggested by the reviewer, we revised the manuscript to explicitly list the model state variables directly updated by the EAKF analysis and the observation data used to constrain each variable. This revision clarifies that our configuration is a joint state and parameter estimation system: meteorology and tracer concentrations are updated directly, while fluxes are updated via ensemble-estimated cross-covariances with the observed tracers.

Specifically, the system assimilates conventional meteorological observations provided via NCEP PREPBUFR (e.g., surface and upper-air observation reports), which update the 3-D meteorological state variables (zonal and meridional winds, air temperature, and specific humidity) and the 2-D surface pressure in the state vector. In addition, the analysis updates the CO_2 and CH_4 tracer mixing ratios using high precision in situ observations from WMO/GAW stations. A complete list of assimilated state variables and corresponding observation types is now provided in Section 2.2.

5. Insufficient justification for uncertainty settings

The rationale for applying 5% uncertainty to initial/lateral boundary conditions and 30% to anthropogenic emissions is not provided. Given that CO2 and CH4 exhibit different source characteristics and variability, using identical uncertainty ratios is difficult to justify. Supporting evidence from previous studies or sensitivity analyses is needed. Moreover, observational and model uncertainties are not adequately discussed.

>> As the reviewer suggested, we revised the manuscript to clarify the prior uncertainty magnitudes using region-specific information and to document how observational and model uncertainties are represented in the assimilation.

1. Ensemble perturbations: The initial and lateral boundary conditions for CO_2/CH_4 concentrations are prescribed from the CAMS EGG4 reanalysis, which is observation-constrained rather than a free-running forecast. Previous evaluation studies indicate that EGG4 concentration errors are typically on the order of ~1% on large scales, although regional performance can vary (e.g., over East Asia) depending on period and diagnostics (Agustí-Panareda et al., 2023). In our DART/EAKF cycling, we apply a small multiplicative boundary perturbation (5%) primarily as a spread-maintenance choice to ensure non-zero ensemble variance at the domain inflow and to avoid filter collapse, while preventing unrealistically large boundary-driven dispersion that could dominate the regional increments driven by in-domain observations. We clarify in the revised text that this setting represents a

pragmatic balance between acknowledging residual boundary uncertainty and avoiding over-spreading.

2. Prior emissions uncertainty for CO_2 and CH_4: Previous studies indicate that anthropogenic emission at national scale has relatively smaller uncertainties (< 5%) but increased to 30 - 250% at city scales (e.g., Solazzo et al., 2021). In this case study, the 30% prior spread was selected to encompass the empirically observed inter-inventory spread for South Korea (2020), as summarized in Fig. 11. We therefore use 30% as a conservative, inventory-informed prior flexibility that is large enough to admit meaningful correction while remaining within the envelope of plausible bottom-up discrepancies for both gases in the study region. We clarify that this value is tied to the observed inter-inventory spread (rather than an arbitrary tuning choice) and note that gas- and sector-specific uncertainty characterization would be preferable when robust spatially resolved uncertainty information becomes available.
3. Observational and model uncertainties: We now explicitly define the observation error treatment for GHG data as $\sigma_{\text{obs}}^2 = \sigma_{\text{meas}}^2 + \sigma_{\text{repr}}^2$, where σ_{meas} is consistent with the WMO/GAW station measurement reports and σ_{repr} represents the dominant representativeness and model–observation mismatch when mapping point observations to 9-km grid-cell means. In the revision, we clarify how σ_{repr} is parameterized and the adopted values (Section 2.2). Uncertainties in model and prior are represented through the ensemble perturbations and the standard DART/EAKF controls (localization and adaptive inflation) too.

As a model description paper, we note that the uncertainty specifications in the DART-based system are configurable and can be adapted to the application scale (e.g., national versus city). Nevertheless, in this study we document and justify the specific settings used (5% boundary perturbations and 30% prior emission uncertainty) based on published evaluations and region-specific inventory spread (e.g. Zhang et al., 2021a, b).

Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., Ades, M., Baier, B. C., Balsamo, G., Borsdorff, T., Bousserez, N., Boussetta, S., Buchwitz, M., Cantarello, L., Crevoisier, C., Engelen, R., Eskes, H., Flemming, J., Garrigues, S., Hasekamp, O., Huijnen, V., Jones, L., Kipling, Z., Langerock, B., McNorton, J., Meilhac, N., Noël, S., Parrington, M., Peuch, V.-H., Ramonet, M., Razinger, M., Reuter, M., Ribas, R., Suttie, M., Sweeney, C., Tarniewicz, J., and Wu, L.: Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020, Atmospheric Chemistry and Physics, 23, 3829–3859, https://doi.org/10.5194/acp-23-3829-2023, 2023.

Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M., and Janssens-Maenhout, G.: Uncertainties in the Emissions Database for Global Atmospheric Research (EDGAR) emission inventory of greenhouse gases, Atmospheric Chemistry and Physics, 21, 5655–5683, https://doi.org/10.5194/acp-21-5655-2021, 2021.

Zhang, Q., Li, M., Wei, C., Mizzi, A. P., Huang, Y., and Gu, Q.: Assimilation of OCO-2 retrievals with WRF-Chem/DART: A case study for the Midwestern United States, Atmospheric Environment, 246, 118106, https://doi.org/10.1016/j.atmosenv.2020.118106, 2021a.

Zhang, Q., Li, M., Wang, M., Mizzi, A. P., Huang, Y., Wei, C., Jin, J., and Gu, Q.: CO2 Flux over the Contiguous United States in 2016 Inverted by WRF-Chem/DART from OCO-2 XCO2 Retrievals, Remote Sensing, 13, 2996, https://doi.org/10.3390/rs13152996, 2021b.

6. Lack of meteorological field evaluation

The manuscript does not evaluate the WRF-Chem meteorological fields (e.g., wind speed and direction), which directly influence transport and mixing. A comparison of modeled meteorology with observational data is necessary to quantify biases and assess their effects on the inversion.

>> As suggested by the reviewer, we agree that evaluating meteorological fields is necessary because transport and mixing directly affect CO_2/CH_4 simulations and the inferred fluxes. Accordingly, our coupled system assimilates meteorological observations together with GHG concentrations so that meteorological analyses are continuously constrained during the inversion period. We also emphasize that the new surface parameterizations adopted in this study improve wind simulations near the surface (e.g., Lee et al., 2015; Lee and Hong, 2016; Lim et al., 2019; Lee et al., 2020; Kim et al., 2024; Lee et al., 2024; Jo et al., 2025; WMO, 2025). In the revised manuscript, we added a concise meteorological verification section alongside the inversion framework. Specifically, we evaluate near-surface winds (speed and direction) and report bias in Supplement (Fig. S12, Fig. S13 and Table S1). The results show that the cycling assimilation of NCEP PREPBUFR observations in the WRF-Chem/DART EAKF configuration substantially reduces transport errors during the inversion period, thereby increasing confidence that the inferred flux adjustments are not driven by gross meteorological biases. Relevant prior WRF-Chem/DART studies using similar configurations are now cited for context (Mizzi et al., 2016; Mizzi et al., 2018; Pouyaei et al., 2023).

Jo, Y., Kim, S., Lee, Y., Kim, C., Hong, J., Lee, J., and Jang, K.: Evaluation of topographic effect parameterizations in Weather Research and Forecasting model over complex mountainous terrain in wildfire-prone regions, Fire, 8, 196, https://doi.org/10.3390/fire8050196, 2025.

Kim, J. W., Shin, H., Lee, K., and Hong, J.: Enhancement of ANN-based wind power forecasting by modification of surface roughness parameterizations over complex terrain, J. Environ. Manage., 362, 121246, https://doi.org/10.1016/j.jenvman.2024.121246, 2024.

Lee, J. and Hong, J.: Implementation of space borne LiDAR-retrieved forest canopy height in the WRF model, J. Geophys. Res. Atmos., 121, https://doi.org/10.1002/2015JD024299, 2016.

Lee, J., Shin, H. H., Hong, S.-Y., Jimenez, P. A., Dudhia, J., and Hong, J.: Impacts of sub-grid scale orography parameterization on simulated surface-layer wind and monsoonal precipitation over Korea, J. Geophys. Res. Atmos., 120, 644–653, https://doi.org/10.1002/2014JD022747, 2015.

Lee, J., Hong, J., Noh, Y., and Jiménez, P. A.: Implementation of a roughness sublayer parameterization in the Weather Research and Forecasting model (WRF version 3.7.1) and its validation for regional climate simulations, Geosci. Model Dev., 13, 521–536, https://doi.org/10.5194/gmd-13-521-2020, 2020.

Lee, K. M., Park, B., Kim, J. W., and Hong, J.: Day-ahead wind power forecasting based on feature extraction integrating vertical layer wind characteristics in complex terrain, Energy, 288, 129713, 2024.

Lim, K., Lim, J., Shin, H., Hong, J., and Ji, Y.: Impacts of subgrid-scale orography parameterization on simulated atmospheric fields in the high-resolution atmospheric forecast model, Meteorol. Atmos. Phys., https://doi.org/10.1007/s00703-018-0615-4, 2019.

Mizzi, A. P., Arellano Jr., A. F., Edwards, D. P., Anderson, J. L., and Pfister, G. G.: Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: a regional chemical transport/ensemble Kalman filter data assimilation system, Geoscientific Model Development, 9, 965–978, https://doi.org/10.5194/gmd-9-965-2016, 2016.

Mizzi, A. P., Edwards, D. P., and Anderson, J. L.: Assimilating compact phase space retrievals (CPSRs): comparison with independent observations (MOZAIC in situ and IASI retrievals) and extension to assimilation of truncated retrieval profiles, Geoscientific Model Development, **11**, 3727–3745, https://doi.org/10.5194/gmd-11-3727-2018, 2018.

Pouyaei, A., Mizzi, A. P., Choi, Y., Mousavinezhad, S., and Khorshidian, N.: Downwind Ozone Changes of the 2019 Williams Flats Wildfire: Insights From WRF-Chem/DART Assimilation of OMI NO2, HCHO, and MODIS AOD Retrievals, Journal of Geophysical Research: Atmospheres, **128**, e2022JD038019, https://doi.org/10.1029/2022JD038019, 2023.

World Meteorological Organization (WMO): Integrated Global Greenhouse Gas Information System Good Practice Guidance for Estimating National-scale Greenhouse Gas Emissions using Atmospheric Observations, 2025 Edition, GAW Report No. 319, World Meteorological Organization, Geneva, Switzerland, 2025.

7. Unclear treatment of initial and boundary conditions & need for sensitivity analysis

It is not specified whether WRF-Chem/DART updates initial and boundary conditions in the single-domain configuration. And additional justification is needed to demonstrate that a single-domain setup is sufficient for obtaining reliable posterior estimates. Given that regional inversions are highly sensitive to boundary conditions—and considering the known negative CH_4 bias in EGG4—

additional sensitivity tests using alternative boundary conditions would help evaluate the robustness of the posterior estimates.

>> We acknowledge that boundary-condition biases (including CH_4 biases documented for EGG4) remain an important structural uncertainty for regional inversions. We have revised Section 2.2 to remove ambiguity about how initial and lateral boundary conditions are treated in our single-domain WRF-Chem/DART configuration. Initial conditions are updated every cycle through standard cycling (the next-cycle background is initialized from the previous analyzed ensemble). Time-varying lateral boundary conditions are prescribed from the external product and perturbed across the ensemble to maintain boundary spread. After each analysis, the DART "update_wrf_bc" procedure is used to apply the analysis increments consistently near the lateral boundaries to avoid spurious discontinuities, rather than to "optimize" the external boundary forcing itself.

We also clarify why a single 9-km domain is adequate for the framework demonstration in this manuscript. First, in our WRF configuration the lateral boundary forcing is applied through a boundary blending ("sponge") region (specified zone + relaxation zone); we use 1 grid point for the specified zone and 4 grid points for the relaxation zone, and therefore restrict posterior diagnostics and station-based evaluations to the domain interior, excluding the outer 5 grid points where boundary relaxation is active. Second, the observing network and region of interest are well separated from the lateral boundaries, reducing the likelihood that boundary errors directly dominate the interior increments. This is consistent with recent theoretical work showing that BC-induced posterior flux errors depend on transport and information content and can be mitigated either by directly optimizing BC terms (boundary method) or by excluding an outer buffer region from the final analysis (Nesser et al., 2025).

Finally, our cycling strategy follows standard WRF/DART practice with 6-hour synoptic analysis times, with observations collected within approximately ±3 hours of each analysis time. While a comprehensive exploration of alternative boundary products and domain/nesting configurations is beyond the scope of this model-description paper, we now explicitly outline a robustness protocol (alternative boundary prescriptions and bias-correction sensitivity tests) and identify multi-BC and/or nesting sensitivity experiments as an important follow-on step for application studies. The single-domain design was chosen to balance computational cost and resolution for high-resolution inversions, and the framework can be readily extended to alternative domain configurations.

Nesser, H., K. Bowman, M. Thill, D. Varon, C. Randles, A. Tewari, F. Cardoso-Saldaña, E. Reidy, J. Maasakkers, and D. Jacob (2025) Predicting and correcting the influence of boundary conditions in regional inverse analyses, Geoscientific Model Development, 18, 9279-9291, https://doi.org/10.5194/gmd-18-9279-2025.

8. Limited ground observation network & questionable ability to constrain emissions

The number and spatial distribution of ground sites appear insufficient to constrain national-scale emissions. The sites used are located near the national borders and represent background conditions, making it uncertain whether they provide meaningful constraints for South Korea's total emissions. Figures 6 and 7 also indicate limited uncertainty reduction in several high-uncertainty regions. Inclusion of satellite observations would likely enhance spatial representativeness.

>> We appreciate the reviewer's focus on observational representativeness and network adequacy. To avoid any confusion, we clarify that the inversion system presented here is not a Lagrangian (trajectory/footprint) inversion. Instead, it is an Eulerian regional inversion based on WRF-Chem forward transport coupled with DART's ensemble Kalman filtering, in which meteorological fields and tracer concentrations are evolved and updated on the model grid through cycling data assimilation. We have revised Section 2.2 to state this more explicitly and to distinguish our approach from commonly used Lagrangian frameworks. That said, the reviewer's underlying point, "how well the available observations constrain national-scale emissions" applies to both Eulerian and Lagrangian inversions. We therefore address it directly using our system's diagnostics (e.g., flow-dependent sensitivity/footprint implied by the ensemble, localization, and OSSE-style tests) and by discussing the implications for spatially non-uniform uncertainty reduction.

We agree that the current surface network limits fine-scale spatial attribution, and we do not claim that the present station set can uniquely resolve grid-scale emissions everywhere. However, limited station numbers do not necessarily preclude useful national-scale constraints when background handling, data selection, and transport sensitivity are carefully designed (e.g., Kenea et al., 2024; Steiner et al., 2024; Vardag and Maiwald, 2024). For example, national inversions have demonstrated robust national-scale CO_2 flux estimates using only a small number of well-characterized sites by focusing on well-mixed periods and carefully treating background conditions (e.g., the New Zealand national inversion using Baring Head and Lauder; Bukosa et al., 2025 and Supplement).

In our configuration, the effective constraint is expected to be spatially non-uniform because it depends on (1) the flow-dependent footprint of each site under the prevailing synoptic conditions, (2) boundary-layer mixing, and (3) covariance localization. Therefore, regions that are weakly coupled to the observing sites during the study period can show limited uncertainty reduction even when prior uncertainty is large. Transport and dispersion uncertainties further limit the recoverable spatial detail and can dominate the model–data mismatch, particularly at fine scales.

To separate observation limitations from methodological limitations, we add an OSSE-style capability test with a substantially denser surface-network scenario, showing that the same framework

yields stronger and more localized uncertainty reduction when observational coverage is increased, consistent with previous OSSE-based network-design studies.
We agree that satellite observations can enhance spatial representativeness and are a natural extension of the framework. We therefore state this explicitly as a planned follow-on development, while noting that satellite integration requires careful treatment of retrieval biases and vertical sensitivity and is beyond the scope of the present model-description manuscript.

Bukosa, B., Mikaloff-Fletcher, S., Brailsford, G., Smale, D., Keller, E. D., Baisden, W. T., Kirschbaum, M. U. F., Giltrap, D. L., Liáng, L., Moore, S., Moss, R., Nichol, S., Turnbull, J., Geddes, A., Kennett, D., Hidy, D., Barcza, Z., Schipper, L. A., Wall, A. M., Nakaoka, S.-I., Mukai, H., and Brandon, A.: Inverse modelling of New Zealand's carbon dioxide balance estimates a larger than expected carbon sink, Atmospheric Chemistry and Physics, 25, 6445–6473, https://doi.org/10.5194/acp-25-6445-2025, 2025.

Kenea, S., Shin, D., Li, S., Joo, S., Kim, S., and Labzovskii, L.: Designing additional CO$_2$ in-situ surface observation networks over south Korea using Bayesian inversion coupled with Lagrangian modelling, Atmos. Environ., 326, 120471, https://doi.org/10.1016/j.atmosenv.2024.120471, 2024.

Steiner, M., Cantarello, L., Henne, S., and Brunner, D.: Flow-dependent observation errors for greenhouse gas inversions in an ensemble Kalman smoother, Atmos. Chem. Phys., 24, 12447–12463, https://doi.org/10.5194/acp-24-12447-2024, 2024.

Vardag, S. N. and Maiwald, R.: Optimising urban measurement networks for CO$_2$ flux estimation: a high-resolution OSSE using GRAMM/GRAL, Geosci. Model Dev., 17, 1885–1902, https://doi.org/10.5194/gmd-17-1885-2024, 2024.

9. Filtering of WMO/GAW background-site observations
WMO/GAW stations typically apply data filtering to remove local source influences. It is unclear whether filtered or unfiltered measurements were used in this study, and this point should be clarified. If filtered data were used, the reduced local variability may limit the ability of the assimilation system to effectively constrain regional emissions.
>> We clarify that we assimilated the Level-2 quality-assured WMO/GAW in situ mole-fraction products. As documented in the station measurement reports (Lee et al., 2019; Lee et al., 2023), Level-2 processing focuses on QA/QC to remove instrument-related artifacts (e.g., malfunction periods, calibration intervals, and sampling instabilities) and is distinct from Level-3 "background-selected" products that apply additional selection to isolate baseline conditions. No additional Level-3 baseline selection/background filtering was applied in this study, and therefore short-term enhancements potentially attributable to regional emissions were retained for assimilation. We revised Section 3 to remove any ambiguity about the data product level, the applied filtering, and the rationale for retaining enhancements.

Lee, H., Han, S.-O., Ryoo, S.-B., Lee, J.-S., and Lee, G.-W.: The measurement of atmospheric CO2 at KMA GAW regional stations, its characteristics, and comparisons with other East Asian sites, Atmospheric Chemistry and Physics, **19**, 2149–2163, https://doi.org/10.5194/acp-19-2149-2019, 2019.

Lee, H., Seo, W., Li, S., Lee, S., Kenea, S. T., and Joo, S.: Measurement report: Atmospheric CH 4 at regional stations of the Korea Meteorological Administration–Global Atmosphere Watch Programme: measurement, characteristics, and long-term changes of its drivers, Atmos. Chem. Phys., **23**, 7141–7159, https://doi.org/10.5194/acp-23-7141-2023, 2023.

10. Missing wetland emissions for CH4 and lack of chemistry description

Wetland CH4 emissions—one of the dominant natural CH4 sources—are not included, and the potential impact of this omission is not discussed. In addition, CH4 oxidation by OH radicals should be addressed, but the manuscript does not describe how WRF-Chem handles this chemistry.

>> We thank the reviewer for highlighting the missing treatment of natural wetland CH_4 emissions and the need to clarify CH_4 chemistry. The inventory data used in this study, EDGAR does not include natural wetland methane emission and it is reported that annual CH_4 emission from natural wetlands in Korea is about 0.001 – 0.01 Tg CH_4 (Zhang et al., 2025). This is less than 0.5% of national total anthropogenic CH_4 emissions. OH chemistry is still challenging to simulate in the atmospheric chemistry model and many methane-focused 3-D inversion systems and even CAMS EGG4 uses prescribed OH fields (or climatological OH-driven loss rates) rather than simulating OH interactively, therefore treat the CH_4 sink in an empirical/parameterized manner; this can contribute to systematic biases in CH_4, particularly when the assumed loss rate is imperfect (e.g., Zhao et al., 2020; Agustí-Panareda et al., 2023). CO_2 and CH_4 can be treated as a passive tracer for our regional application because the chemical lifetime of CH_4 against OH is on the order of years (about 9 years) (See WRF-GHG official document and Saunois et al. (2020) for more information), whereas typical air-mass residence times across our regional domain are <1 day. The implied fractional loss over 1 day is therefore $< 1 / (9 \times 365) \approx 0.03\%$, which is negligible compared to transport and emission uncertainties. Finally, large-scale chemical aging and seasonality are already reflected in the CAMS EGG4 concentration fields used to prescribe the initial and lateral boundary conditions; thus our regional system focuses on resolving transport and source-driven enhancements within the domain rather than re-simulating a negligible local chemical sink.

Callewaert, S.: User guide for the greenhouse gas option of WRF-Chem (WRF-GHG), Royal Belgian Institute for Space Aeronomy (BIRA-IASB), https://doi.org/10.18758/q6rapneu, 2024.

Zhang, Z., Poulter, B., Melton, J. R., Riley, W. J., Allen, G. H., Beerling, D. J., Bousquet, P., Canadell, J. G., Fluet-Chouinard, E., Ciais, P., et al.: Ensemble estimates of global wetland methane emissions over 2000–2020, Biogeosciences, 22, 305–321, https://doi.org/10.5194/bg-22-305-2025, 2025.

Zhao, Y., Saunois, M., Bousquet, P., Lin, X., Berchet, A., Hegglin, M. I., Canadell, J. G., Jackson, R. B., Dlugokencky, E. J., Langenfelds, R. L., Ramonet, M., Worthy, D., and Zheng, B.: Influences of hydroxyl radicals (OH) on top-down estimates of the global and regional methane budgets, Atmos. Chem. Phys., 20, 9525–9546, https://doi.org/10.5194/acp-20-9525-2020, 2020.

11. Insufficient discussion of results

The evaluation of top-down estimates lacks depth. For example, the reasons behind the strong improvements at AMY and GSN, the pronounced summer behavior, and the underestimation of CH4 in vertical profiles are not explained. In particular, the manuscript mentions "underlying mechanisms for the maritime CH4" but does not clarify what those mechanisms are or how they influence the vertical structure. Similar gaps appear in the CO2 increment patterns across different regions; for instance, negative increments in SMA and MWI and positive increments in SCI and SEI are reported without any discussion of potential drivers. Without addressing these points, it is difficult to assess the reliability of the results.

>> We agree that additional interpretation is needed to help readers assess the reliability of the top-down results. We therefore expanded Section 7 (formerly Section 6 in the preprint) and added mechanism-oriented explanations for each point raised by the reviewer, with explicit cross-references to the relevant diagnostics.

(1) The strong improvements at AMY and GSN and the associated seasonal (summer) behavior are now discussed in Section 7.1 using the monthly distributions and performance statistics (Fig. 5 and Table 1). We also explicitly relate the pronounced summertime CH_4 features to large-scale/background behavior in the boundary forcing (Fig. S9).

(2) The systematic CH_4 underestimation in vertical profiles is now documented and discussed in Section 7.2 using independent aircraft vertical profiles (Fig. 6; flight information in Fig. S7). In the same subsection, we clarify the maritime CH_4 issue as a possible discrepancy between EGG4 and EDGAR over ocean grids that can affect coastal/marine baseline concentrations and thus the vertical structure (Section 7.2; Fig. 6).

(3) The regional CO_2 increment sign patterns (negative in SMA/MWI and positive in SCI/SEI) are now explicitly described in Section 7.4 with the corresponding increment map (Fig. 10b). The associated CH_4 increment behavior is discussed alongside (Fig. 10d), including a brief cautionary interpretation regarding potential boundary-related effects (Section 7.4).

These additions provide the essential physical context needed to assess consistency and limitations of the presented results, while keeping the overall scope aligned with a model-description manuscript.

12. Limitations of bottom-up inventory comparisons

When comparing total anthropogenic emissions with bottom-up inventories, sectoral coverage should be consistent across datasets. If they are not harmonized, the manuscript must explain the sectoral differences for proper interpretation. The discussion of why posterior emissions align with ROK-BTR but differ substantially from ODIAC is also insufficient. Lines 399–400 are unclear and require revision.

>> As suggested by the reviewer, we revised the manuscript to make explicit what is compared: national-total anthropogenic CO_2 excluding LULUCF, aggregated consistently over the South Korea domain and the analysis period. We also clarify that full sector-by-sector harmonization is not always possible across products with fundamentally different constructions; however, we now document the remaining scope differences so that the comparison is interpreted correctly.

Specifically, ROK-BTR is an official national inventory compiled from detailed activity data and reported following IPCC sector definitions, whereas ODIAC is a global fossil-fuel CO_2 product that estimates emissions and downscales them spatially using proxy information (e.g., point sources and nightlights). Therefore, differences can arise from (1) scope/coverage (e.g., treatment of industrial process CO_2, international bunkers, and other non-combustion categories), (2) methodological choices in spatial disaggregation and temporal allocation, and (3) uncertainty in proxy assumptions that can affect both the national total and its distribution. We frame the inventory comparison as a consistency check for policy-relevant totals rather than as a validation claim. This is consistent with prior inter-comparisons showing that high-resolution fossil-fuel CO_2 inventories can differ substantially due to differing methodological choices and proxy assumptions (Hutchins et al., 2017). In this context, Lines 399–400 have been rewritten to explicitly state these scope and methodological differences and to avoid implying that ODIAC is directly comparable to a full IPCC inventory without caveats (Lines 592, "*We note that these products differ in scope and construction …*").

Hutchins, M. G., Colby, J. D., Marland, G., and Marland, E.: A comparison of five high-resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United States, Mitig Adapt Strateg Glob Change, **22**, 947–972, https://doi.org/10.1007/s11027-016-9709-9, 2017.

13. Additional issues

Lack of contextual clarity for EGG4 CO2 and CH4 errors: The discussion of EGG4 biases in Lines 185–190 does not clearly connect with the surrounding text. The purpose and relevance of this information within the flow of the manuscript are unclear.

>> As suggested by the reviewer, we revised Section 4 our manuscript to clarify why EGG4 CO_2/CH_4 bias information is introduced at this point in the manuscript. Specifically, we now state explicitly that the EGG4 bias information is provided as diagnostic context for boundary-condition uncertainty, which is a known driver of structural error in regional inversions, and that it is used to interpret posterior behaviors and limitations discussed later in Section 7. To improve the flow, we also added a short transition sentence linking this diagnostic context to the subsequent configuration description (boundary/initial condition specification and perturbation strategy). The revised text makes clear that the EGG4 bias discussion motivates our boundary-uncertainty treatment and frames the later sensitivity/limitation discussion.

Input data preprocessing: The manuscript does not describe how datasets with differing spatial and temporal resolutions were regridded and harmonized.

>> As suggested by the reviewer, we revised the manuscript to describe the harmonization procedures at the points where each input is introduced (Sections 4 and 5.1–5.4). In brief, all heterogeneous inputs are mapped to the 9-km WRF grid and hourly cycling timeline using the following steps:

(1) Spatial mapping: gridded flux inventories are reprojected to the WRF Lambert Conformal Conic (LCC) grid and regridded using area-weighted conservative remapping to preserve domain-integrated emissions; fields that represent state variables (e.g., mixing ratios) are interpolated using bilinear methods.

(2) Temporal harmonization: sub-daily cycling inputs are obtained by applying documented temporal profiles (e.g., sector-dependent diurnal/weekly scaling for EDGAR; hourly injection for FINN; monthly-to-hourly prescription for SeaFlux; VPRM computed at the model time step), ensuring that the time-integrated totals are conserved over the parent interval (daily/monthly).

(3) EGG4 GHG IC/BCs: EGG4 3-D mole fractions are horizontally interpolated to the WRF boundary/initial fields and linearly interpolated in time between analysis times; vertical interpolation is performed from the native EGG4 vertical levels to WRF's hybrid vertical coordinate used for chemical/tracer initialization.

These additions clarify how differing spatial/temporal resolutions are handled without changing the underlying methodology.

VPRM parameters: The text states that parameters were adopted but does not specify which ones or why

>> We used the VPRM parameter set calibrated for East Asia reported by Li et al. (2020). As suggested by the reviewer, we revised Section 5.4 to explicitly list the adopted VPRM parameters (e.g., $\lambda, \alpha, \beta, PAR_0$) for each land-cover/vegetation class used in this study (now provided in Table Sx), and to explain the rationale for this choice. Specifically, we adopted Li et al. (2020) because it provides a regionally robust parameterizations evaluated with East Asian ecosystems.

Figure 7: Prior flux exists over the ocean, yet its uncertainty is zero; the rationale should be explained. The reduction of uncertainty over North Korea and China also requires discussion.

>> We acknowledge that Fig. 7 (now Fig. 8 after revision) can be misread without clarification and we revised the texts and figure. The prior uncertainty over the ocean is not set to zero; rather, the ensemble spread is prescribed as a fractional perturbation of the flux (e.g., 30%), so oceanic flux components with much smaller magnitudes yield correspondingly small absolute uncertainties. In this figure, these values can appear near-zero because the color scale is dominated by large terrestrial/urban sources; however, they are treated as non-zero variances in the filter under the same multiplicative uncertainty formulation. We have updated the figure caption to clarify this interpretation.

Regarding the small uncertainty reductions over North Korea and some parts of China, these arise from weak but non-zero concentration–flux cross-covariances within the localization radius under occasional transport connectivity across the domain boundaries. We interpret these as marginal numerical effects rather than robust observational constraints. We therefore add a brief note stating that transboundary updates are not quantitatively interpreted in this study and that our discussion focuses on the South Korean domain where observational constraints are strongest.