Supplement of

Global aerosol composition constraints from simultaneous data assimilation of satellite AOD and trace gas observations

Takashi Sekiya et al.

Correspondence: Takashi Sekiya (tsekiya@jamstec.go.jp)

S1 Data assimilation parameters

The optimal data assimilation setting can vary among different species and different observations, depending on typical lifetime in the atmosphere and spatial coverage of observations. Table S1 summarizes the evaluation results of the sensitivity data assimilation runs after a spin-up against AOD derived from MODIS. The RMS innovation was more than 23% larger in the assimilation with a localization length of 500 km than the other localization length (150 km and 1500 km), whereas correlation coefficient with a localization length of 500 km was larger than those with other lengths. Using the same localization length (500 km), increasing and decreasing emission inflation from the baseline setting degraded the correlation coefficient and RMS innovation. Thus, we selected the data assimilation parameter settings for assimilating AOD as described in Section **??**.

S2 Consistency with the assimilated AOD observations

The performance of AOD assimilation scheme newly implemented in this study was evaluated using the χ^2 test (Ménard and Chang, 2000; Zupanski and Zupanski, 2006). The χ^2 value was used to diagnose the balance between the actual and estimated errors and was calculated as the ratio of the observation minus the forecast (OmF; i.e., $\boldsymbol{y}^o - H(\boldsymbol{x}^f)$) to the estimated error covariance in the observation space ($\boldsymbol{H}\boldsymbol{P}^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}$):

$$\boldsymbol{Y} = \frac{1}{\sqrt{\mathrm{N}}}\left(\boldsymbol{H}\boldsymbol{P}^{\mathrm{f}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}\right)^{-1/2}\left(\boldsymbol{y}^o - H(\boldsymbol{x}^f)\right), \tag{1}$$

$$\chi^2 = \mathrm{trace}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}. \tag{2}$$

The mean χ^2 value with the standard deviation in GasAero DA was 0.99 ± 0.08, which suggests that the background and observational error covariances were well estimated.

As summarized in Table S2, we confirmed the self-consistency of GasAero DA with assimilated AOD observations over the globe. Compared to the control model simulation, the GasAero DA reduced the mean bias by 60%. Spatial correlation coefficient was increased from 0.7 in the control simulation to 0.88 due to GasAero DA, while RMSE was reduced by 36%.

S3 Emission estimation scheme

To determine the emission optimization approach, the Observing System Simulation Experiments (OSSEs) with standard state augmentation and simplified version of a two-stage bias estimation approaches were performed. As summarized in Table S3, the two-stage bias estimation was characterized by better agreements with global dust emissions predicted in Nature run (true value) compared to the state augmentation (by 86%). For the global total emission of carbonaceous aerosols, these two approaches showed comparable performance. The errors in regional emission estimates over East Asia and Russia were reduced by 31–80% in the two-stage bias estimation, compared to the state augmentation, while the errors over South and Central Africa were increased by a factor of 2.

S4 Potential impacts of optimized ammonia emissions on aerosol analyses

Simultaneous assimilation of NH_3 as well as NO_2, SO_2, and AOD would be the efficient way to optimize sulfate-nitrate-ammonium aerosols comprehensively. Thus, we estimate NH_3 emissions using Suomi-NPP/CrIS NH_3 retrievals (REF) with a simple mass balance approach as did in Marais et al. (2021), and investigate the sensitivity of nitrate and ammonium aerosols to the optimized NH_3 emissions. We calculate the optimized monthly mean NH_3 emissions (E_{CrIS}) at each model grid based on the CrIS NH_3 retrievals, as follows:

$$E_{\mathrm{CrIS}} = C_{\mathrm{CrIS}} \times \left(\frac{E}{C}\right)_{\mathrm{DA}}, \tag{3}$$

where C_{CrIS} indicates surface NH_3 concentrations derived from CrIS, $(E/C)_{\mathrm{DA}}$ is the ratio of monthly mean NH_3 emissions to monthly mean surface NH_3 concentrations sampled at 13:30LT obtained from the data assimilation calculation. The performance of the mass balance approach is affected by a failure of spatial attribution of emission sources owing to horizontal displacement of NH_3. The length of displacement was estimated to vary within 10–105 km, depending on meteorological and photochemical conditions Marais et al. (2021). The longer limit of length is comparable to the model grid size.

Figure S1 exhibited global distribution of surface NH_3 concentrations during May–August 2016 derived from the CrIS retrievals and data assimilation. The CrIS retrievals show high NH_3 concentrations over North and Central Africa, northern India, Malay Peninsula, Central Eastern China, the midwest United States, and South America. Data assimilation with the CEDS emission inventory (v_2021_04_21) overestimated the observed surface NH_3 concentrations over Central Africa, northern India, mainland China, Malay Peninsula, and the southeastern United States typically by 3–10 ppbv, whereas the data assimilation underestimated the observed NH_3 over North Africa, the Great Plains, and South America typically by 3–5 ppbv. In the model simulation using the optimized NH_3 emissions, positive and negative biases were substantially reduced compared to the data assimilation, except for North Africa.

The global total NH_3 emission during May–August 2016 was increased to 45.8 Tg N yr^{-1} from 27.5 Tg N yr^{-1} in the a priori emissions. The optimized global total emission was smaller but closer to those (55–64 Tg N yr^{-1}) derived from previous studies (Zhu et al., 2013; Paulot et al., 2014) than the a priori emissions. The optimized regional total emission in the United States was increased from 1.6 Tg N yr^{-1} in the a priori emissions to 3.3 Tg N yr^{-1}, which ranged within the range of 2.8–7.9 Tg N yr^{-1} obtained from previous studies (Zhu et al., 2013; Paulot et al., 2014; Cao et al., 2020; Sitwell et al., 2022). Furthermore, nitrate and ammonium aerosol concentrations were increased typically by 0.5–1.0 μg m^{-3} over North and South Africa, West Asia, South America, and were decreased typically by 0.5 ppbv over Europe, China, and eastern United States, and Middle East by the NH_3 emission optimization. These modifications partially reduced remaining biases relative to surface in-situ observations of nitrate and ammonium aerosols.

Further investigation would be required to elucidate whether these model biases are attributed to NH_3 emissions or model processes. The AeroCom model intercomparison highlighted that negative biases in the simulations can result from significant uncertainties in the wet deposition process, where the effective Henry's law constant for ammonia varies by an order of six among different global chemical transport models (Bian et al., 2017). In addition, the lack of NH_3 recycling processes between the atmosphere and soil further contributes to these biases. Zhu et al. (2015) and Cao et al. (2022) demonstrated that the imple-

mentation of a bi-directional flux scheme mitigated these To further improve the representation of sulfate-nitrate-ammonium aerosols in the data assimilation, the effective way would be to incorporate improved model schemes and to simultaneously assimilate SO_2, NO_2, along with NH_3 observations.

Table S1. Correlation Coefficients and Root Mean Square (RMS) Innovations of Observation-Minus-Forecast (OmF) for AOD during July 2016. L and INF indicate the horizontal localization length for AOD and inflation factor for aerosol emissions.

Experiment description	correlation coefficient	RMS innovation
Standard setting (L: 500 km, INF: baseline)	0.852	0.080
Shorter localization length (L: 150 km, INF: baseline)	0.788	0.104
Longer localization length (L: 1500 km, INF: baseline)	0.748	0.120
Smaller emission inflation (L: 500 km, INF: baseline$\times$0.33)	0.723	0.125
Larger emission inflation (L: 500 km, INF: baseline$\times$3)	0.099	87.8

Table S2. Statistical scores of self-consistency check with assimilated observations for the GasAero DA during May–August 2016. MB, S-Corr, and RMSE indicate mean bias, spatial correlation coefficient, and root mean square error, respectively.

Observation data	GasAero DA			Aero-only DA			CTL		
	MB	S-Corr	RMSE	MB	S-Corr	RMSE	MB	S-Corr	RMSE
Terra+Aqua/MODIS AOD	-0.02	0.88	0.07				-0.05	0.7	0.11

Table S3. Estimates of carbonaceous and dust aerosols during August 2016 using the Observing System Simulation Experiments (OSSEs) framework. The unit is Tg month^{-1}. Nature run is treated as true value and synthetic observations were generated from Nature run with the typical MODIS observation errors. Control run is the same as Nature run, but the global emissions were reduced by 68% with random spatial variations. Conventional and New scheme indicate the OSSE DA calculation with emission inversion schemes using state augmentation methods and simplified version of a two-stage bias estimation algorithm, respectively.

	Carbonaceous aerosol				Dust			
Experiment ID	Global	East Asia	South/Central Africa	Russia	Global	East Asia	North Africa	Middle East
Nature run	10.4	0.62	3.7	0.88	460	186	70	57
Control run	4.4	0.47	0.78	0.14	197	58	79	19
Conventional	10.3	0.42	4.0	1.2	571	227	101	69
New scheme	11.1	0.59	4.3	1.1	476	189	82	59

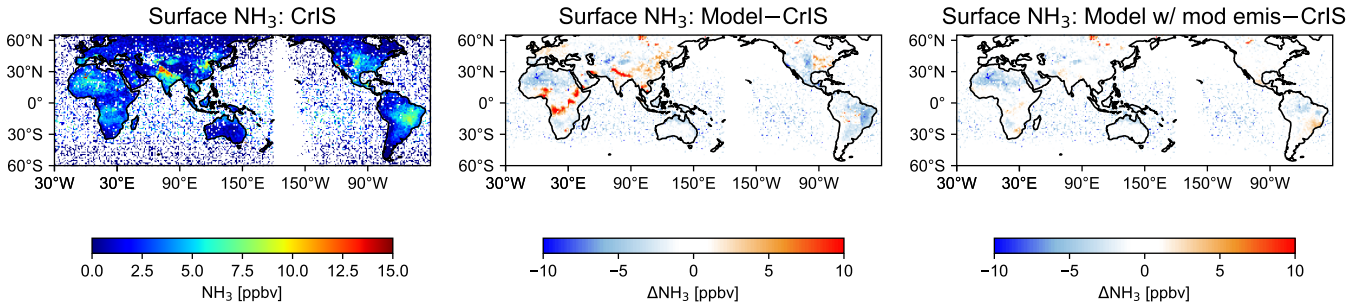

Figure S1. Surface NH_3 concentrations derived from Suomi-NPP/CrIS (left), and the differences between CrIS and the model simulations using the optimized aerosol and its precursor emissions (center and right). A model simulation used a priori NH_3 emissions (center), whereas another model simulation used the optimized NH_3 emissions based on the comparison to the CrIS retrievals. The unit is ppbv.

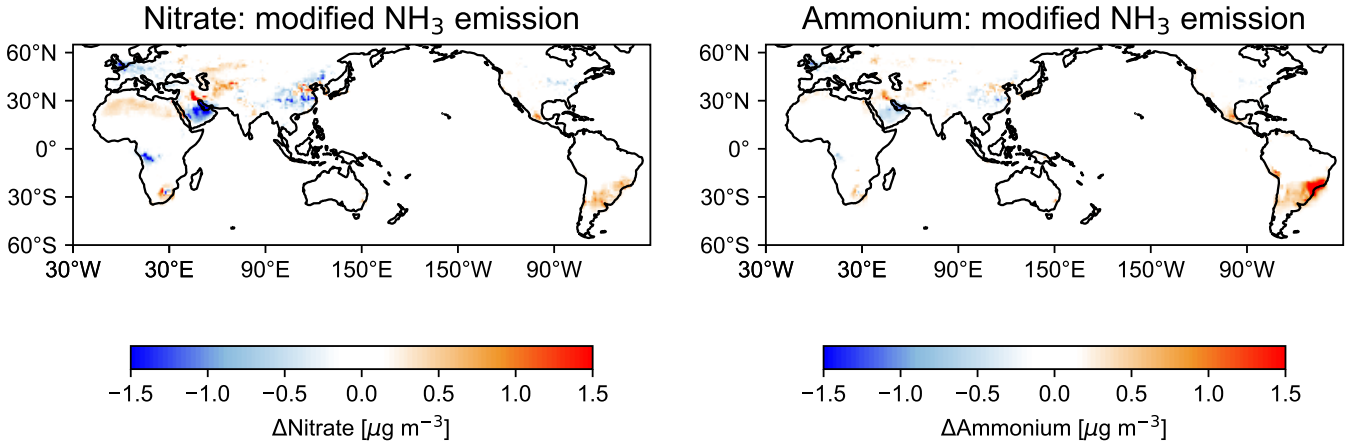

Figure S2. Changes in surface nitrate (left) and ammonium concentrations (right) due to NH_3 emission optimization based on the comparison to the CrIS NH_3 retrievals. The unit is μg m^{-3}.

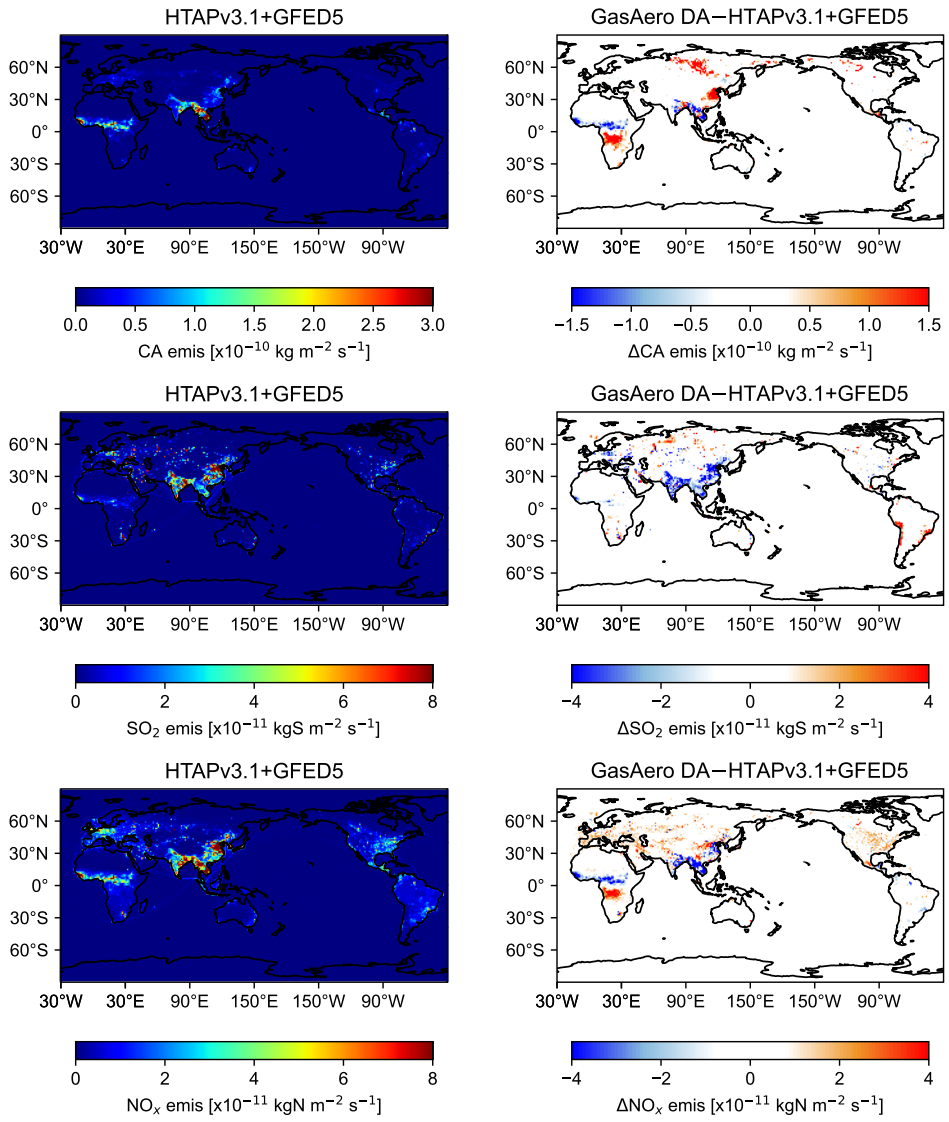

Figure S3. Bottom-up emission estimates obtained from the HTAPv3.1 and GFED5 inventories (left) and their differences from top-down emission estimates derived from GasAero DA (right). The top, middle, and bottom rows are carbonaceous aerosol, SO_2, and NO_x emissions, respectively. The units of carbonaceous aerosol, SO_2, and NO_x emissions are kg m^{-2} s^{-1}, kg S m^{-2} s^{-1}, and kg N m^{-2} s^{-1}, respectively.

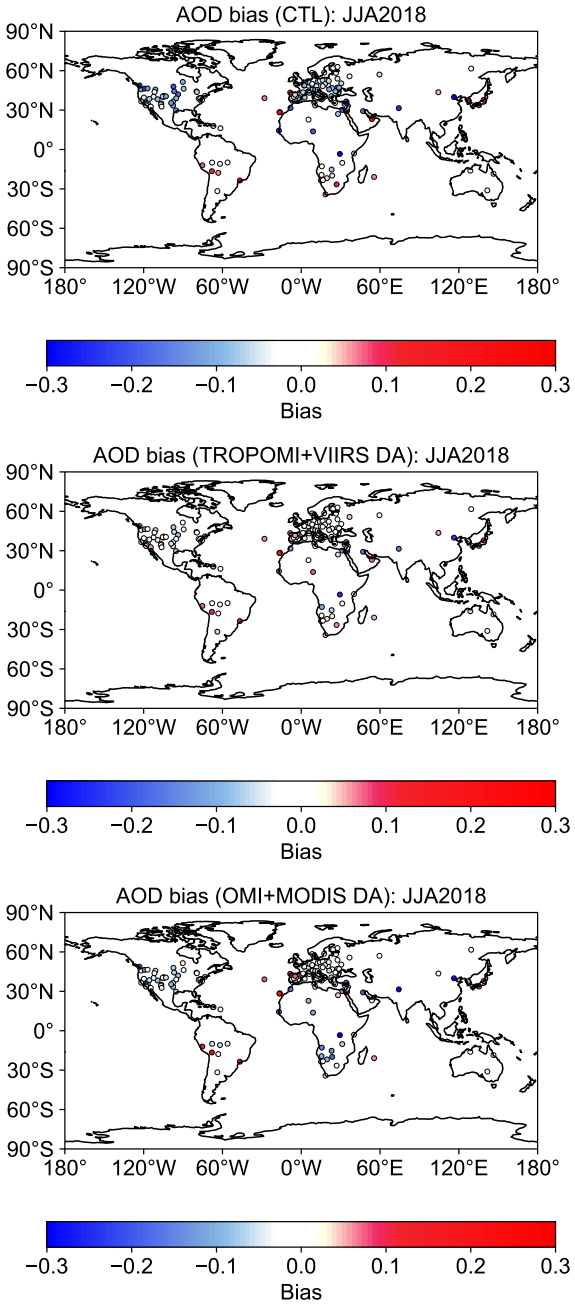

Figure S4. Mean AOD biases against AERONET at individual sites obtained from the CTL simulation (top), TROPOMI+VIIRS DA (middle), and OMI+MODIS DA (bottom) for June–August, 2018.

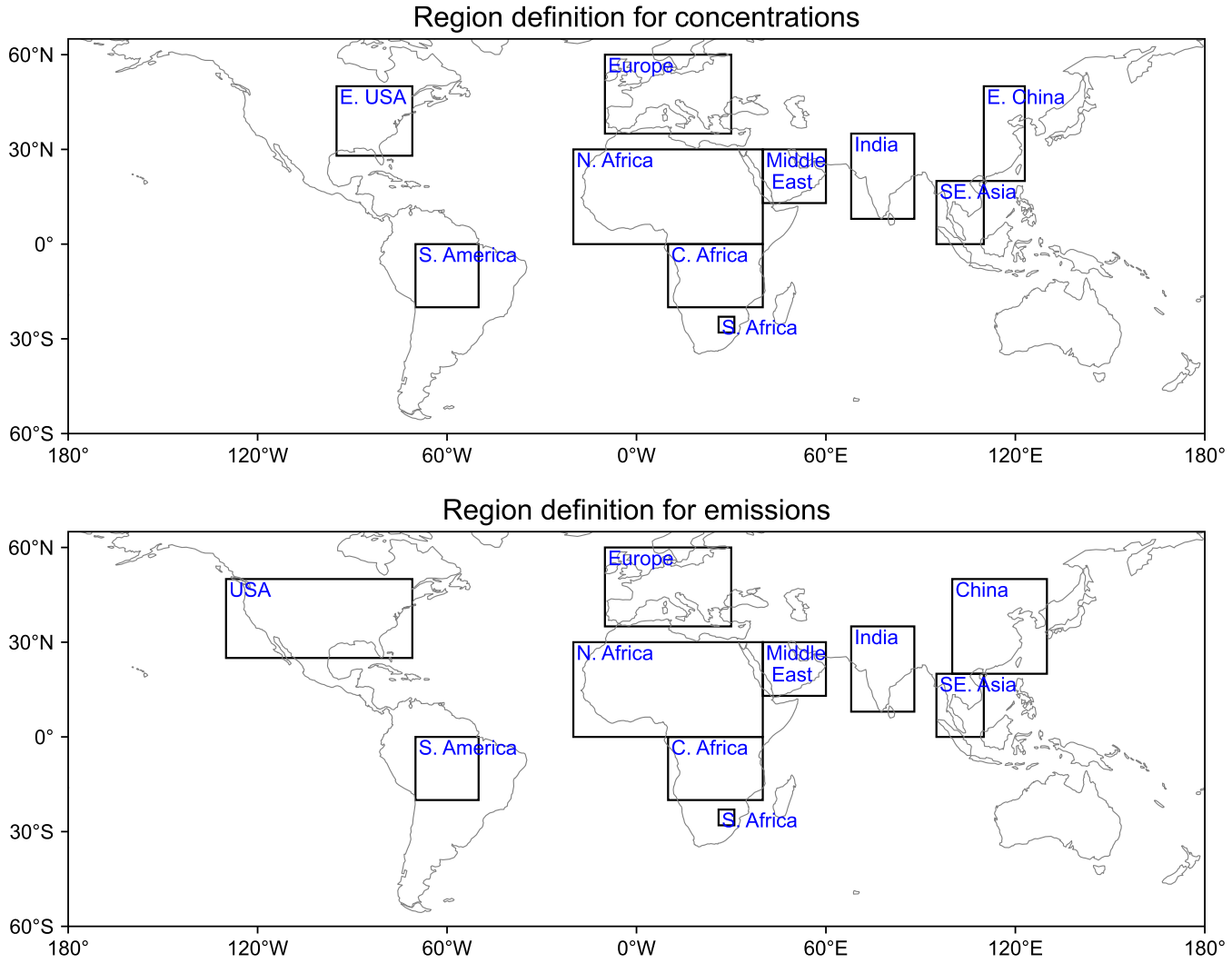

Figure S5. Region definitions used for $PM_{2.5}$ concentrations and their composition (top) and emission estimates (bottom).

References

Bian, H., Chin, M., Hauglustaine, D. A., Schulz, M., Myhre, G., Bauer, S. E., Lund, M. T., Karydis, V. A., Kucsera, T. L., Pan, X., Pozzer, A., Skeie, R. B., Steenrod, S. D., Sudo, K., Tsigaridis, K., Tsimpidi, A. P., and Tsyro, S. G.: Investigation of global particulate nitrate from the AeroCom phase III experiment, Atmos. Chem. Phys., 17, 12 911–12 940, https://doi.org/10.5194/acp-17-12911-2017, 2017.

Cao, H., Henze, D. K., Shephard, M. W., Dammers, E., Cady-Pereira, K., Alvarado, M., Lonsdale, C., Luo, G., Yu, F., Zhu, L., Danielson, C. G., and Edgerton, E. S.: Inverse modeling of NH3 sources using CrIS remote sensing measurements, Environ. Res. Lett., 15, 104 082, https://doi.org/10.1088/1748-9326/abb5cc, 2020.

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., Brümmer, C., Schrader, F., Twigg, M. M., Leeson, S., Tang, Y. S., Stephens, A. C. M., Braban, C., Vincent, K., Meier, M., Seitler, 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., 127, e2021JD035 687, https://doi.org/10.1029/2021JD035687, 2022.

Marais, E. A., Pandey, A. K., Van Damme, M., Clarisse, L., Coheur, P.-F., Shephard, M. W., Cady-Pereira, K. E., Misselbrook, T., Zhu, L., Luo, G., and Yu, F.: UK Ammonia Emissions Estimated With Satellite Observations and GEOS-Chem, J. Geophys. Res., 126, e2021JD035 237, https://doi.org/10.1029/2021JD035237, 2021.

Ménard, R. and Chang, L.-P.: Assimilation of Stratospheric Chemical Tracer Observations Using a Kalman Filter. Part II: χ2-Validated Results and Analysis of Variance and Correlation Dynamics, Mon. Weather Rev., 128, 2672–2686, https://doi.org/10.1175/1520-0493(2000)128<2672:AOSCTO>2.0.CO;2, 2000.

Paulot, F., Jacob, D. J., Pinder, R. W., Bash, J. O., Travis, K., and Henze, D. K.: Ammonia emissions in the United States, European Union, and China derived by high-resolution inversion of ammonium wet deposition data: Interpretation with a new agricultural emissions inventory (MASAGE_NH3), J. Geophys. Res., 119, 4343–4364, https://doi.org/https://doi.org/10.1002/2013JD021130, 2014.

Sitwell, M., Shephard, M. W., Rochon, Y., Cady-Pereira, K., and Dammers, E.: An ensemble-variational inversion system for the estimation of ammonia emissions using CrIS satellite ammonia retrievals, Atmos. Chem. Phys., 22, 6595–6624, https://doi.org/10.5194/acp-22-6595-2022, 2022.

Zhu, L., Henze, D. K., Cady-Pereira, K. E., Shephard, M. W., Luo, M., Pinder, R. W., Bash, J. O., and Jeong, G.-R.: Constraining U.S. ammonia emissions using TES remote sensing observations and the GEOS-Chem adjoint model, J. Geophys. Res., 118, 3355–3368, https://doi.org/https://doi.org/10.1002/jgrd.50166, 2013.

Zhu, L., Henze, D., Bash, J., Jeong, G.-R., Cady-Pereira, K., Shephard, M., Luo, M., Paulot, F., and Capps, S.: Global evaluation of ammonia bidirectional exchange and livestock diurnal variation schemes, Atmos. Chem. Phys., 15, 12 823–12 843, https://doi.org/10.5194/acp-15-12823-2015, 2015.

Zupanski, D. and Zupanski, M.: Model Error Estimation Employing an Ensemble Data Assimilation Approach, Mon. Weather Rev., 134, 1337–1354, https://doi.org/10.1175/MWR3125.1, 2006.