

Anonymous Referee #1

Received and published: 23 April 2018

The authors have presented an evaluation of the CAMS prediction system, focusing on CO and CO₂, during the KORUS-AQ campaign. They evaluated three different CO and CO₂ forecast and analysis products: 16-km CO and CO₂ forecasts, 9-km CO and CO₂ forecasts, and analyses of CO and CO₂ at 80 km and 40 km, respectively. The CAMS products were compared to the KORUS-AQ aircraft data as well as to ground-based and satellite measurements of CO and CO₂. They found that CAM overestimated CO₂, suggesting a positive bias in background CO₂, whereas it underestimated CO, with the underestimate confined mainly to the lower troposphere. The authors also found that CAMS underestimates the outflow of pollution from China, possibly due to an underestimate of Chinese emissions. The study is a nice evaluation of CAMS CO and CO₂ under unique conditions. I have no major concerns about the analysis.

Response: Thank you!

My main concern is about the appropriateness of the manuscript for ACP. As a model evaluation study, I think it is better suited for GMD than ACP. My comments below are relatively minor, but must be address before the manuscript can be accepted for publication, if the Editor decides it is suitable for ACP.

Response: Thank you. We understand the reviewer's concern. But we think that our manuscript is still within the scope of ACP for the following reasons. We double-checked the main subject areas of ACP, which includes atmospheric modelling, field measurements, remote sensing, and laboratory studies of gases, aerosols, clouds and precipitation, radiation, and so on. In this study, we assess the performance of the Copernicus Atmosphere Monitoring Service (CAMS) global prediction system using field measurements from aircraft, ground sites, and ships, and remote sensing data during the KORUS-AQ field campaign. In addition to model evaluation, this study also addresses a few scientific topics on atmospheric chemistry and physics, including (1) anthropogenic combustion characteristics in Korea and China (as well as how well CAMS captures it), (2) impacts of different model configurations and environmental conditions on CO simulations, and (3) implications for CO emissions in CAMS. Thus, this manuscript is in line with the research focuses of ACP. Furthermore, we believe that the findings of this manuscript will be of interest not only to CAMS developers and users, but also to the general atmospheric chemistry community. Therefore, ACP is a perfect platform for us to share these results with the community. We sincerely hope that the editor will consider publishing this manuscript in ACP.

Comments

1. There is no mention of the **CAMS OH field**, which is critical for the simulation of CO. What is the global mean OH from the analyses and forecasts? On page 5, lines 28, it is mentioned that the 16-km CO forecasts use a linear chemistry scheme. A brief description of the scheme, either in the manuscript or in the supplement, would be helpful.

Response: The two high-resolution forecast products (FC16s and FC9s) employ a linear chemistry scheme, without the direct use of model OH.

The OH fields are only used in the CAMS ANs for CO which has full chemistry. In the ANs, the global and Northern Hemisphere mean are 0.98×10^{-6} molecules/cm³ and 1.20×10^{-6} molecules/cm³ during May 2016, respectively. We have extended description of the linear chemistry scheme at the end of Section 2.1. Specifically, we extended

“A linear chemistry scheme is used in FC16s for CO (C-IFS-LINCO) for computationally expediency (Claeyman et al., 2010; Flemming et al., 2012; Massart et al. 2015; Eskes et al., 2017). Key aspects of the three CAMS configurations evaluated in this study are listed in Table 1.”

to:

“ANs for CO use the on-line implemented chemical mechanism (C-IFS-CB05, Flemming et al., 2015) that is an extended version of the Carbon Bond mechanism 5 (CB05, Yarwood et al., 2005). Because hydroxyl radical (OH) is an important sink for CO, modeled OH is critical for the simulation of CO (Gaubert et al., 2016, 2017). In the ANs for CO, the global and NH mean of air mass-weighted OH are 0.98×10^{-6} molecules/cm³ and 1.20×10^{-6} molecules/cm³ during May 2016, respectively (calculated following recommendations from Lawrence et al. (2001)). The mean OH from the ANs for CO is consistent with previous studies (e.g., Lawrence et al., 2001; Lelieveld et al., 2016; Gaubert et al., 2016, 2017). A linear chemistry scheme is (C-IFS-LINCO) used in FC16s and FC9s for CO for computationally expediency (Claeyman et al., 2010; Flemming et al., 2012; Massart et al. 2015; Eskes et al., 2017). C-IFS-LINCO computes CO sources and sinks using the approach developed by Cariolle and Déqué (1986) and updated by Cariolle and Teyssède (2007), without direct use of modeled OH. C-IFS-LINCO is less computationally demanding than the full chemistry, permitting simulations at higher resolutions (Massart et al. 2015). Key aspects of the three CAMS configurations evaluated in this study are listed in Table 1.”

2. It is stated that the overestimate in CO₂ is associated with the bias correction in the biogenic source of CO₂, but there is no discussion of this “bias correction”. Furthermore, in Agusti-Panareda et al. (2016) CAMS CO₂ was underestimating CO₂ observations from the surface in situ network and from TCCON, which the “bias correction” (the biospheric flux adjustment) reduced. Why is CAMS overestimating CO₂ here? A discussion is needed about the **treatment of the biospheric fluxes in CAMS and its possible impact on the modeled CO₂ over Korea**.

Response: First of all, we appreciate the reviewer for noticing this. According to Agusti-Panareda et al (2016), in the Northern Hemisphere there is a growing overestimation of the atmospheric CO₂ at the end of winter and throughout spring (from March to May); while at the end of the growing season in both the Northern Hemisphere and the Southern Hemisphere (August and March, respectively) there is a growing negative bias, i.e. an overestimation of the sink based on observations from NOAA/ESRL and TCCON (Section 5.1 of Agusti-Panareda et al (2016)). This is consistent with our finding. Agusti-Panareda et al (2016) also implies that the CO₂ overestimation by CAMS is enhanced in the BFAS simulation (Section 5.1 of Agusti-Panareda et al (2016)).

However, we note that the statement “As found by Agusti-Panareda et al (2016), the overall overestimation of CO₂ is associated with the biogenic bias correction” is inappropriate. We have changed it to “Agusti-Panareda et al. [2016] also suggests CO₂ is overestimated by CAMS in the Northern Hemisphere at the end of winter and throughout spring”.

We have also included more discussions on the bias correction section 2.1 (where we introduce biogenic flux adjustment scheme (BFAS)), including treatment of the biospheric fluxes in CAMS, and its possible impact on the modeled CO₂ over Korea.

3. On page 11 it was shown that the model produced steeper vertical gradients in CO and CO₂ than observed over Seoul, which the authors suggested may be due to weak boundary layer mixing. Since CAMS seems to perform better over Taehwa, it would be interesting to compare the vertical

gradients over Seoul and Taehwa in CAMS and in the observations to see if the issue is mainly an inability of CAMS to capture the PBL heights over the Seoul urban environment.

Response: We compared the observed vertical gradient of CO₂ over Seoul and Taehwa. We also analyzed the PBL height derived from observations and CAMS.

We have added the following analysis in the revised manuscript (Section 3.3.1):

“We further find that compared with the Seoul metropolitan, the observed vertical gradient of CO₂ over Taehwa (~0.03 ppmv/hPa) below 925 hPa is smaller, which is relatively better captured by CAMS (0.02–0.12 ppmv/hPa). This again implies the possible inefficient boundary layer mixing in CAMS over the Seoul urban environment.”

“CO over Taehwa is more likely to be due to regional transport, as Taehwa is not a strong CO source region. Thus, the vertical gradient of CO over Taehwa does not necessarily reflect the impact of boundary layer mixing over Taehwa.”

4. I am surprised that the analyses are not much better than the forecasts. Indeed, it seems as though the 9-km forecast is better than the analyses in some cases. I think it would be helpful for the reader if the authors expanded the description of the analyses to give the reader **more information about the configure and quality of the analyses**. Figure S1 and the brief text on page 5 are not enough.

Response: Analyses of both CO and CO₂ (ANs) do show improvement from the free running simulation (i.e., FC16s). However, the new 9-km forecast product (i.e., FC9s) is expected to have a better performance of FC9s because:

- (1) FC9s is initialized with the analysis product every 24 hours (i.e., it incorporates information of analyses every 24 hours);
- (2) FC9s has a much higher horizontal resolution (9 km) than the analyses (80 km for CO and 40 km for CO₂).

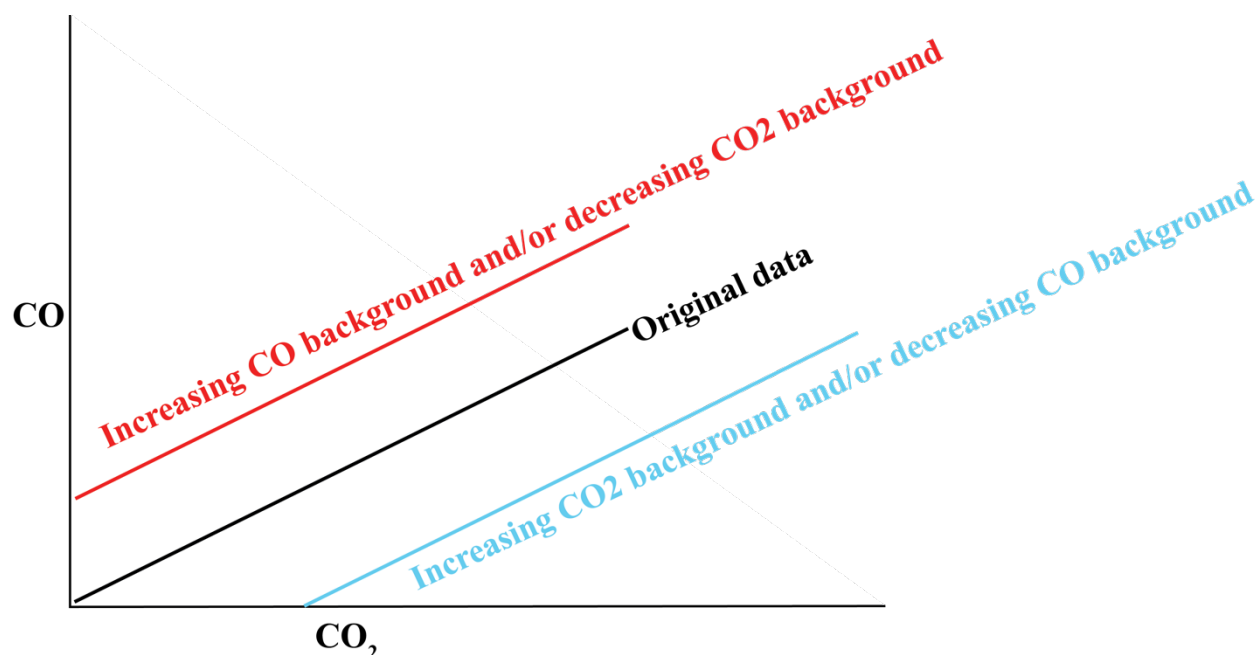
In addition to Figure S1 (i.e., the new Figure S2), Table 1 also summarized configurations of CAMS global atmospheric composition products including the analysis product of CO (AN_CHEM) and CO₂ (AN_GHG). We have also added more information about the configure and quality of the analyses:

“Observations of both CO and CO₂ are assimilated in 12-hour assimilation windows. Inness et al. (2015) found that CO total column field, vertical distribution, and concentrations in the lower troposphere are improved by assimilating the CO total column from MOPITT. Assimilation of the GOSAT XCO₂ lead to improvements in mean absolute error and bias variability in XCO₂ fields during the year 2013 (Massart et al., 2016).”

5. The discussion of **enhancement ratios** is confusing. It is unclear if the authors are using the slope of the CO/CO₂ relationship or the slope of delta CO/delta CO₂ relationship. The two approaches are different. The description in the text suggests that they are using the RMA regression of CO/CO₂ to assess the combustion sources, but throughout the text there is use of the delta CO/delta CO₂ notation. If they are indeed calculating an enhancement ratio (delta CO/delta CO₂) above the background, how is the background being calculated? How sensitive is the analysis to the definition of the background?

Response: Thank you for pointing this out. We used slopes of the CO to CO₂ regressions as our enhancement ratios.

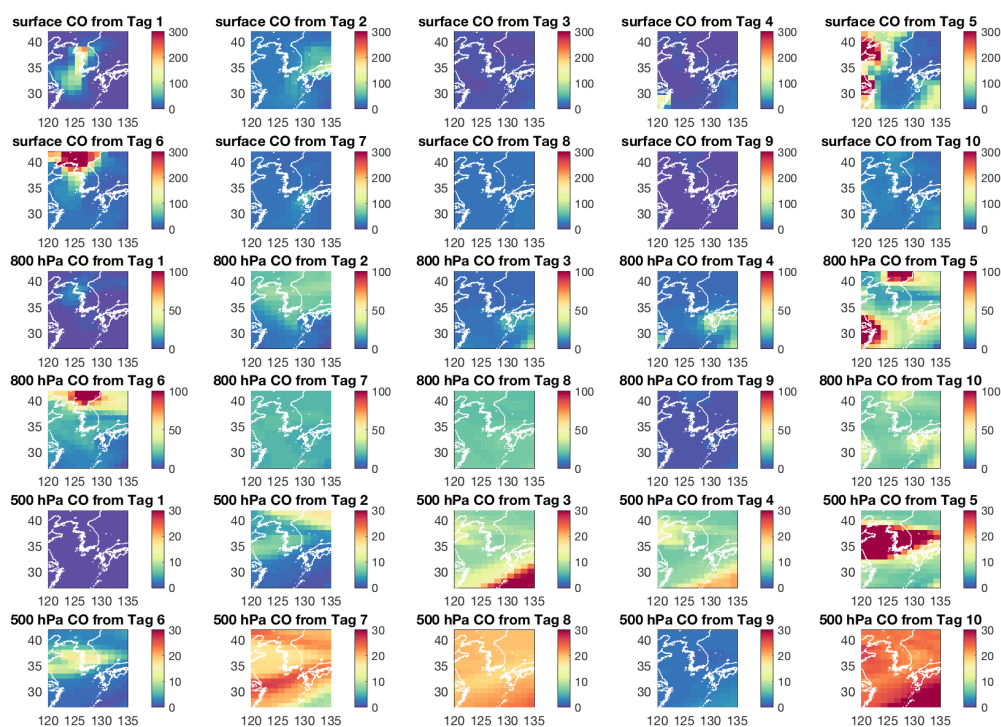
The estimated regression slope in the RMA corresponds to enhancement ratio of CO and CO₂ (dCO/dCO₂). No background values are used here. In fact, the definition of the background does not change the regression slope. Please see the following figure for a demonstration.



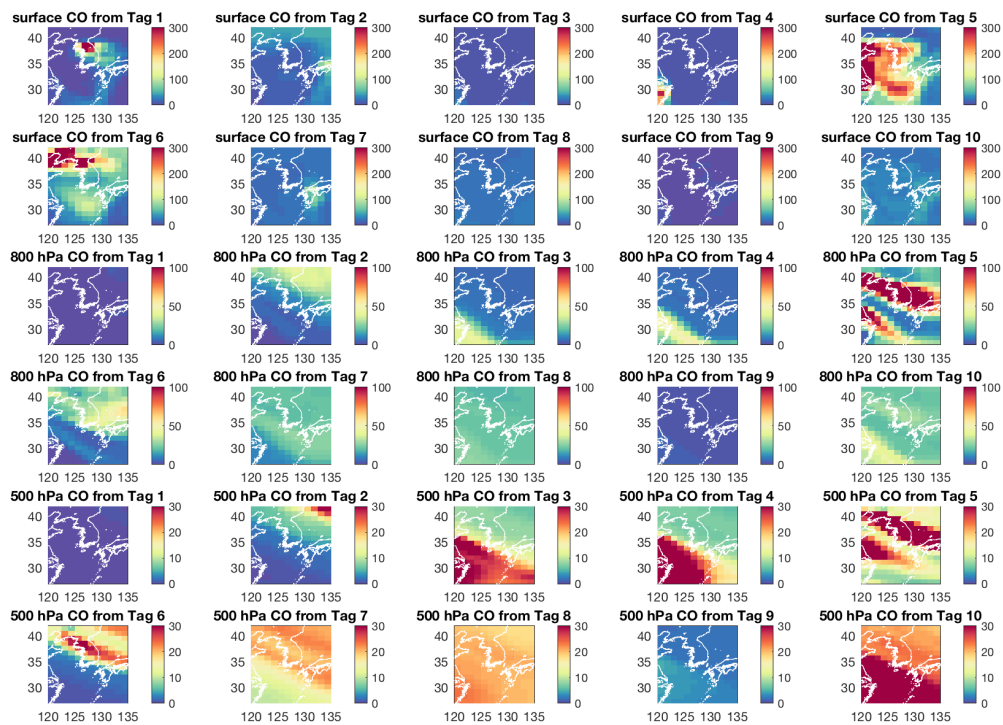
Nevertheless, we agree with the reviewer that the usage of $\Delta\text{CO}/\Delta\text{CO}_2$ in this manuscript could be misleading, and have changed $\Delta\text{CO}/\Delta\text{CO}_2$ to $d\text{CO}/d\text{CO}_2$.

6. The authors found that CAMS underestimated CO during China outflow events, but overestimated it under normal conditions. What are the **different source regions** for air reach the West Sea during “outflow” and “normal” conditions? To what degree is the model bias due to CAMS not capturing this difference in transport as compared to it not have the correct balance of emissions in China?

Our recent work with Community Atmosphere Model with chemistry (CAM-chem) tagged CO tracers studied the different source regions. Taking condition on June 5th (corresponding to the June 4th flight) as an example of normal conditions, both China and Korea contribute to the CO over the West Sea at surface. At 800 hPa, Japan, Russia, China, the rest of the world, and ocean all contribute to CO over the West Sea. However, CO concentrations over the West Sea is relatively low at these conditions. The following figure (Fig. S6 of Tang et al., 2018) shows spatial distributions of the tagged CO (ppbv) on June 5th, 2016 at model surface, 800 hPa, and 500 hPa (Tag 1: Korea; Tag 2: Japan+Russia; Tag 3: Indonesia+India; Tag 4: EA-S; Tag 5: EA-M; Tag 6: EA-N; Tag 7: the rest of the world+ocean; Tag 8: CH₄ oxidation; Tag 9: biogenic; Tag 10: chemical production besides CH₄):



During outflow events (e.g., conditions during the May 30th flight), contribution from China are largely enhanced and becomes dominant at surface, 800 hPa, and 500 hPa. CO concentrations over the West Sea is relatively high during China outflow events. The following figure (Fig. S7 of Tang et al., 2018) shows spatial distributions of the tagged CO (ppbv) on May 31st, 2016 (corresponding to the May 30th flight) at model surface, 800 hPa, and 500 hPa (the Tags are the same as in the Fig. S6).



In Tang et al. (2018), we found that estimates of CO emissions in China rather than transport is potentially the main source of model bias over the West Sea. We have included the discussions and reference in the revised manuscript:

“More elaborate analysis of source contributions during KORUS-AQ is beyond the scope of this study and can be found in Tang et al. (2018), which suggested that during China outflow events, the contribution from Chinese direct emissions to CO over the West Sea is largely enhanced and dominant.”

Technical Comments

1. Page 4, line 26: add a comma between “forecast” and “CO2”.

Response: We have added the comma.

2. Page 4, line 28: add “the” between “on” and “free-running”.

Response: Thank you. We have edited accordingly.

3. Page 5, line 1: Is Figure 1 really necessary? I don’t think it adds much to the manuscript. Since there are already 11 figures, I would suggest removing Figure 1.

Response: We have moved Figure 1 to supplement (the New Figure S2).

4. Page 6, lines 1 and 2; add “the” before “South Korean peninsula”.

Response: Correction made.

5. Page 6, line 7: remove “including” before the list of the three questions.

Response: Thank you. We have deleted the “including”.

6. Page 6, line 8: The wording for question (3) needs improving. The English is not quite correct.

Response: Thank you for pointing this out. We have changed the sentence to “(3) *how well do models perform and what improvements are needed to better represent atmospheric composition over Korea and its connection to the larger global atmosphere (Kim and Park, 2014, KORUS-AQ White Paper).*”

7. Page 8, lines 22 and 23: The revisit and overpass times seem to be used interchangeably here. The revisit time of OCO-2, for example, is 16 days since it is in the A-Train orbit. However, the local overpass time is around 1:30 pm. For GOSAT the revisit time is 3 days.

Response: Thank you. We have changed “revisit” to “overpass”.

8. Page 8, line 26: Change 0.09e18 from e-notation to standard SI notation.

Response: We have changed the notation accordingly.

9. Page 9, line 5: Please add “is” between “CO₂” and “associated”.

Response: We have deleted this sentence.

10. Page 10, line 1: The variance in CO in the May 3rd data does not seem larger than average to me. In fact, it seems to be smaller than average.

Response: We have changed the sentences “*For example, the flights in May 3rd, May 17th, May 24th, May 29th, and May 30th were specifically designed to capture Chinese pollution outflow. In these days, the variances in CAMS biases for CO (but not CO₂) are larger than the average*” to

“*For example, parts of flight tracks on May 3rd, May 17th, May 24th, May 29th, and May 30th were specifically designed to capture Chinese pollution outflow. In these days, the variances in CAMS biases for CO (but not CO₂) are generally larger than the average except for the flight tracks on May 3rd when Chinese influences were expected to be weak*”

11. Page 10, lines 12 and 14: Please change “tale” to “tail”.

Response: Thank you. We have changed it.

12. Page 12, line 4: It is unclear what is meant by the statement that “the wind speeds dominate the transport flux variations in CO₂.” It the argument here that the meteorological uncertainty is the dominant contribution to the uncertainty in the forecast and analysis fields? If so, how does one come to that conclusion from Figure 7?

Response: Because the variations in CO₂ density are very low relative to CO₂ background, the pattern of the CO₂ fluxes in the Figure (CO₂ fluxes = wind speed × CO₂ density) mostly display pattern of wind speed instead of pattern of CO₂ density. However, we find the implication of “*the wind speeds dominate the transport flux variations in CO₂*” to be redundant with previous sentences, and this sentence itself is confusing. We have deleted it.

13. Page 15, line 23: Section 4.2 only discusses the comparison to ship data of CO, not CO and CO₂.

Response: Thank you. We have deleted “*and CO₂*”.

14. Page 15, line 23: change “shop tracks” to “ship tracks”.

Response: Correction made.

15. Page 17, line 7: Change “size of CO data” to “amount of CO data”.

Thank you. We have changed “*size*” to “*amount*”.

16. Page 40, Table 3: The title for the table is wrong. This is the same title as for Table 4.

Response: Thank you for pointing this out. We have changed the title to “**Table 3. Enhancement ratios of CO to CO₂ (ppbv/ppmv), CO and CO₂ correlations, and bias of CO to bias of CO₂ correlations from airborne measurements, CAMS FC16s, ANs, and FC9s.**”

17. Figure S1: Should the labels “FX9s” and “FX16s” in the figure be “FC9s” and “FC16s”?

Response: Thank you for noticing this. The labels should be “*FC9s*” and “*FC16s*”, and we have corrected them.

Anonymous Referee #2

Received and published: 22 May 2018

This manuscript presents a comparison between three modelling products of the Copernicus Atmosphere Monitoring Service (CAMS) and in situ measurements from the KORUS-AQ (and KORUS-OC) campaign in the vicinity of the Korean peninsula during May/June 2016. Airborne, surface, ship-based, and satellite measurements of CO and CO₂ are compared to the CAMS analysis and two forecast products at different spatial resolutions. The statistical analysis is relatively straightforward and clearly laid out, and some patterns of over- and underestimation are found for the two tracers under different conditions. The importance of vertical transport in the understanding of these differences could be further explored, as outlined below.

While I understand the other reviewer’s comment that this manuscript might be a better fit for GMD, as it is assessing the capabilities of a specific modelling system, the general conclusions about the potential underestimation of CO emissions from China make it relevant for a broader audience as well. This is ultimately an editorial decision. However the quality of the manuscript, datasets and analysis is good, and appropriate for publication.

Response: Thank you!

Below are some suggestions, some major, some minor, on how the analysis might be slightly extended in order to better understand the processes driving the model-data mismatch.

Section 3.3: In the discussion about the **relative agreement in the profile for CO₂ vs. the disagreement between the lower atmosphere values between the observations and the model**, a discussion of the relevance of the mixing height and/or **planetary boundary layer** was

somewhat lacking. A difference in profile shape can be attributed to **incorrect fluxes, incorrect mixing**, or a combination of the two. By having two tracers with differing results, it should be possible to deepen this analysis a bit.

There is further discussion about the vertical gradients of the tracers, but no attempt is made to diagnose the **PBL height**. Given the model data and the meteorological information from the aircraft profiles, this should be possible. Could you at least comment on this, and why such an approach was not undertaken? It is even suggested that there might be a "**possible weaker boundary layer mixing in CAMS**". Here **diagnosing the PBL height (as a function of time) from both the model fields and the profiles might be enlightening**.

Response: We have added the following discussion of analyzing differences in profile shapes by having two tracers in Section 3.3: "*CO over Taehwa is more likely to be due to regional transport, as Taehwa is not a strong CO source region. Thus, the vertical gradient of CO over Taehwa does not necessarily reflect the impact of BL mixing over Taehwa.*"

We have also added analyses and discussions of the relevance of the mixing height and/or planetary boundary layer, by adding the new Figure S7 (time series of the mixing layer heights from both the model fields and the measurements), and the following discussion:

"We further compared the mixing layer (ML) height derived from the KORUS-AQ airborne DIAL-HSRL measurements of aerosol backscatter following the technique from Brooks et al. (2003), and the BL heights from CAMS. We note that ML height is only approximately equal to BL height. We find that CAMS generally underestimates BL heights during KORUS-AQ (Fig. S6). The model underestimation of BL over the Seoul metropolitan (-761.3 ± 39.7 m) is stronger than that over Taehwa (721.7 ± 38.6 m) which is covered by forests instead of urban. This is consistent with the CAMS's relatively better capability of capturing vertical gradient of CO₂ over Taehwa compared to that over Seoul, supporting our previous implication of the possible inefficient BL mixing in CAMS over the Seoul urban environment."

Another interesting point might be the representation of urban effects for Seoul in particular. Here it would be interesting to compare the **PBL height** as modelled vs. measured in the vicinity of Seoul compared to other less rural sites. However, this may be beyond the scope of this study.

Response: Thank you. We compared modeled profiles of CO and CO₂ over Seoul and Taehwa to imply possible inefficient modeled boundary layer mixing over the Seoul metropolitan. Please see the response to the previous comment and the response to Reviewer #1, Comment 3 for details.

We also added Figure S8, Model bias of boundary layer heights over Seoul and Taehwa (a less rural site).

For the special case of Seoul, the low altitude measurements were taken during missed approaches at the airport. Given all the air traffic in the region, might it be that the CO in this area is locally very much enhanced, and as such not representative of even the relatively small spatial footprint of the CAMS model? Here perhaps a referral to a relevant paper by Boschetti et al. (Tellus B, 2015) looking at **enhancements of CO in the boundary layer** from commercial airline measurements might be relevant.

Response: Thank you. We have added the reference (Boschetti et al., 2015) as well as the following discussion in the Section 3.3.1:

"In addition, given the air traffic over the Seoul Air Base (where the DC-8 aircraft frequently conducted missed approaches), emissions from airplanes may also contribute to the model biases."

Regarding the assessment of the outflow over the West Sea, I was confused by the phrase: "Hence, the wind speeds dominate the transport flux variations in CO₂." I'm not sure what is meant here. Is this because the outflow pattern wasn't as strong as for CO? But aren't both flux variations (more or less) linearly dependent on wind speed anyhow? Please clarify.

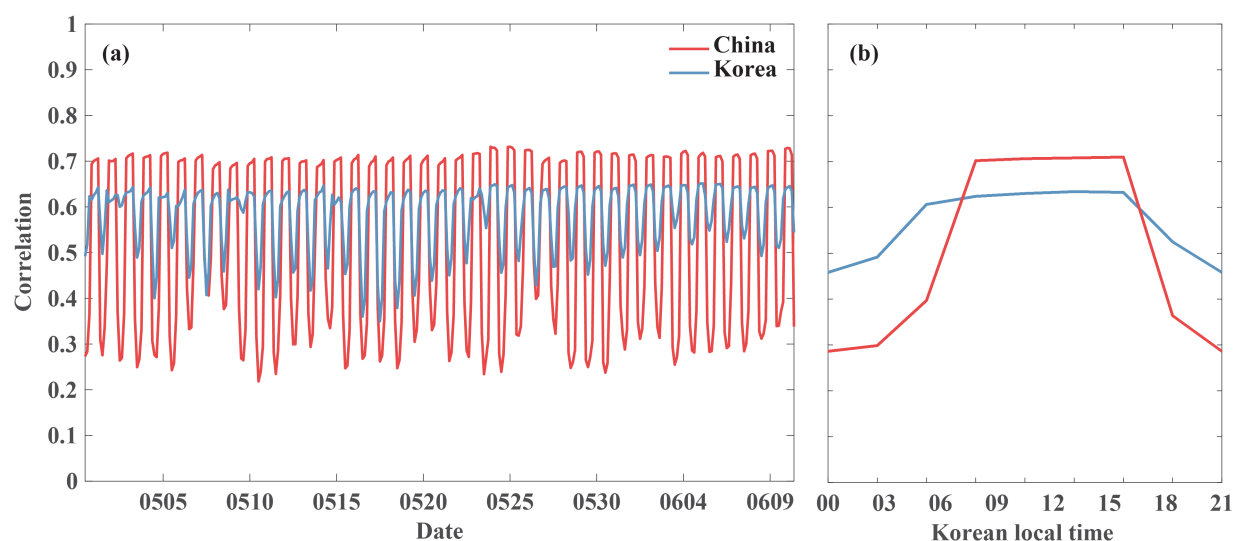
Response: Thank you. We have deleted this sentence. Please see the response to Reviewer #1, Comment 12 for details.

The discussion about the correlation of CO and CO₂ over the West Sea is quite interesting, and invites further inquiry. The suggestion that the difference in time factors for anthropogenic CO and CO₂ (with the former having constant monthly values and the latter having diurnal variability) should effect the correlation over Korea as well. Could it be explained by the differences in transport times, e.g. diurnal CO₂ emissions peak in daytime while measurements are being made over Korea, whereas daytime measurements over the West Sea represent nighttime emissions from China, where the difference in time factors is at a maximum? In terms of just the correlation in the fluxes, it should be easy to test if EDGAR has a higher spatial correlation between CO₂ and CO in Korea vs. China.

Response: Thank you. The high observed CO and CO₂ correlations over the West Sea and Seoul suggest that CO and CO₂ are likely from common sources. We found that the difference in the time factors of CO emissions and CO₂ fluxes may contribute to the model's inability of capturing such high correlation over the West Sea. However, such difference in the time factors is unlikely to impact the modeled correlation over Seoul as much as it impacts the West Sea. In fact, the diurnal variability of CO₂ fluxes comes from the CO₂ net ecosystem exchange rather than the anthropogenic part. Since the flights over Seoul are close to the strong common anthropogenic sources of CO and CO₂ (i.e., the Seoul metropolitan area), the correlation over Seoul is dominated by anthropogenic emissions and unlikely impacted by diurnal variability of CO₂ fluxes that comes from the CO₂ net ecosystem exchange. This is supported by the consistency between observed and modeled correlations.

As for the other 3 groups over Korea (i.e., Taehwa, Seoul-Jeju jetway, Seoul-Busan jetway), their observed correlations are not high (i.e., 0.68, 0.62, 0.60, respectively) at the first place compared to the observed high correlations over the West Sea (0.89) and Seoul (0.78). This implies that CO₂ and CO observed over these three flight groups may not come from common sources and/or have been mixed with the environment.

We agree with the reviewer that correlation in the fluxes may provide valuable insights to explain the correlations in the modeled abundance correlations. The following figure shows that time series of spatial correlations between CO emissions and CO₂ fluxes in CAMS over East China (which dominates Chinese contribution to the West Sea (Tang et al., 2018)) and Korea. There is a strong diurnal cycle in the correlations caused by the difference in time factors.



The diurnal cycle of spatial correlations between CO emissions and CO₂ fluxes over Korea in CAMS peaks (~0.7) in daytime while measurements over Korea were made. On the other hand, during the nighttime, the correlations between CO emissions and CO₂ fluxes in CAMS are relatively low over East China (<0.4). This implies that the relatively low correlations between the CO and CO₂ abundances over the West Sea in CAMS may reflect the effect of nighttime emissions from East China in CAMS.

We thank the reviewer for pointing this out and have included this in the manuscript (text in Section 3.4 and the new Fig. S8).

The analysis of the satellite data is not particularly illuminating, with the exception of the separation of MOPITT data into those influence by outflow. Regarding the use of the OCO-2 data, most of the data references are pre-launch, and should be updated. Wunch et al. 2017 would be a better up to date reference than those from 2011, and an updated estimate of the OCO-2 precision, even if it is coming from grey literature (such as the ACOS OCO-2 User's Guide) would be preferable to a largely theoretical assessment from Boesch et al., 2011.

Response: Thank you. We have included the following two references:

Osterman, G. B., Eldering, A., Avis, C., Chafin, B., O'Dell, C., Frankenberg, C., ... & Crisp, D. (2015). Orbiting Carbon Observatory-2 (OCO-2) Data Product User's Guide, Operational L1 and L2 Data Versions 7 and 7R. NASA Jet Propulsion Laboratory, California Institute of Technology.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., ... & Griffith, D. W. (2017). Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) X CO₂ measurements with TCCON. Atmospheric Measurement Techniques, 10(6), 2209.

It is unclear what is meant by the "recommended quality control" in section 2.2.4. Does this mean the standard quality flag? Was the bias correction applied? Was a certain warn-level threshold used? Please elaborate.

Response: "recommended quality control" means the standard quality flag, and we have changed the term in the manuscript. The standard quality flag we used is from Table 1 and Table 2 of Mandrake et al. (2015)

(https://co2.jpl.nasa.gov/static/docs/OCO2_XCO2_Lite_Files_and_Bias_Correction_0716.docx):

Table 1: Quality Filters Applied to Land Soundings

All Land Soundings		
Field	Lower Limit (> or =)	Upper Limit (< or =)
Warn level	N/A	15
Outcome flag (not in lite file)	N/A	2
Preprocessors/h2o_ratio	0.700	1.030
Preprocessors/co2_ratio	0.995	1.025
Preprocessors/dp_apb	-15.00	5.00
Retrieval/dp	-5.00	10.0
Retrieval/aod_ice	N/A	0.050
Retrieval/Aod_sulfate	N/A	0.400
Retrieval/Aod_dust*	0.001	0.30
Retrieval/Co2_grad_del	-70.0	70.0
Retrieval/albedo_2	0.10	N/A
Blended albedo (2.4*albedo_3 – 1.13*albedo_1) (both in retrieval group)	N/A	0.8
dof_co2 (not in lite product)	1.8	N/A
Sounding/airmass	N/A	3.6
* or AOD dust = 0.0		

Table 2: Quality Filters Applied to Ocean Glint Soundings

Ocean Glint Soundings		
Field	Lower Limit (> or =)	Upper Limit (< or =)
Warn level	N/A	15
Outcome flag (not in lite file)	N/A	2
Preprocessors/co2_ratio	0.994	1.020
Preprocessors/dp_apb	N/A	0.00
Retrieval/dp	-3.00	9.0
Retrieval/Co2_grad_del	-30.0	5.0
Retrieval/albedo_slope $3 \cdot 10^5$	1.0	10.0
Retrieval/windspeed	2.0	N/a
Sounding/snr_weak_co2	380	N/A
Sounding/airmass	N/A	3.5

We used the suggested warn-level threshold in the Table (i.e., ≤ 15). We have added the warn_level information in the manuscript. The bias correction is not applied to the data used in this study, as we used Standard Data files (L2Std) instead of Lite files.

If the Taylor skill score is being used for the assessment of the forecasting skill as in section 3.1, the equation should be in the main paper, and not just in the supplement. Please include it here as well.

Response: We have moved the equation from supplement to the section 3.1 of the paper.

P4, L18-20: The text here states that the CO analysis runs at "approximately 40 km horizontal resolution", but in Figure 1 it is shown to be 80 km horizontal resolution. Later on page 5 80 km is given again, and the text on P4 refers to that fact that the CO₂ analysis is at a higher spatial resolution (in both the horizontal and vertical). Please ensure that the information is consistent and correct.

Response: Thank you for pointing it out. The CO analysis is at approximately 80 km horizontal resolution while the CO₂ analysis is at the approximately 40 km horizontal resolution. We have corrected the P4, L18-20.

Minor/typographical comments:

P2, L11: show -> shows

Response: We have corrected it.

P2, L12: "over Seoul metropolitan" -> either "over the Seoul metropolitan area" or "over Seoul"

Response: Thank you. We have changed that to "*over the Seoul metropolitan area*".

P3, L6: near-real time -> near-real-time

Response: Correction made.

P3, L16: field -> field campaign

Response: Thank you. We have added "*campaign*" there.

P4, L20: Perhaps this should be one sentence?

Response: Thank you for noticing this. This should be one sentence and we have corrected it.

P4, L25: 4-days shouldn't be hyphenated (four days)

Response: Correction made.

P4, L26: 16km -> 16 km

Response: Thank you for pointing this out. We noticed that in the text we used "*16km forecast*" or "*16 km forecast*" a few times. To be consistent, we changed them into "*16-km forecast*". Same for the "*9-km forecast*".

P5, L26: The -> the

Response: Correction made.

P6, L1, L2, and often afterwards: South Korean peninsula -> the South Korean peninsula

Response: Thank you. We have edited the manuscript accordingly.

P6, L8-10: The third scientific question needs to be restated. It doesn't make sense as it is written here.

Response: Thank you. We have rephrased the question (3):

“(3) how well do models perform and what improvements are needed to better represent atmospheric composition over Korea and its connection to the larger global atmosphere (Kim and Park, 2014, KORUS-AQ White Paper).”

P6, L27: data is -> data are

Response: Correction made.

P6, L30-31: Wouldn't UTC be one day behind local time?

Response: Thank you for pointing this out. Korea time = UTC time +9. We have changed “*UTC time is one day ahead of Korea local time*” to “*UTC time is one day behind Korea local time*”.

P7, L7: combustion signatureS (plural needed to match grammar)

Response: Correction made.

P7, L10-12: Not sure which preposition should be used to describe the jetway flights, I would suggest "in", but consistency is more important. Also check the grammar: "Flights in the Seoul-Busan jetway were designed to capture... Flight in the Seoul-Jeju jetway, on the other hand, sampled air over..."

Response: We have changed “*Flights over Seoul-Busan jetway*” to “*Flights in Seoul-Busan jetway*”; and “*The flights in Seoul-Jeju jetway, on the other hand, samples air over local power*” to “*The flights in Seoul-Jeju jetway, on the other hand, sample air over local power*”.

P7, L17: Baengnyeong site is located in less populated Baengnyeong Island, Incheon which is northwest of Seoul. -> The Baengnyeong site is located on the sparsely populated Baengnyeong Island, Incheon, northwest of Seoul.

Response: Thank you. We have changed accordingly.

P7, L19: "on remote" -> "on the remote"

Response: Correction made.

P8, L21: resolutions -> the resolutions

Response: Correction made.

P9, L2: Here is the first of many instances of referring to the in situ measurements collected from the DC-8 aircraft as simply "DC-8". As a reader I found this jarring. Perhaps instead refer to the dataset as the "DC-8 in situ data" or the "aircraft data" or "the airborne measurements"?

Response: Thank you for pointing it out. We have changed "DC-8" to “*the DC-8 aircraft data*” or “*the airborne measurements*” in the manuscript to refer the airborne measurements from the DC-8 aircraft.

P9, L8: inconsistent description of correlation range (to vs. -)

Response: We have changed the “*to*” to “*-*”.

P9, L12: CAMS have -> CAMS has

Response: Correction made.

P9, L15: those for -> that of

Response: Correction made.

P10, L14: tale -> tail (Please change later instances as well.)

Response: Correction made.

P10, L25: West Sea -> the West Sea

Response: Thank you. We have changed all the “*West Sea*” to “*the West Sea*” in the revised manuscript.

P11, L28: than in Korea -> as in Korea

Response: Correction made.

P12, L1: West Sea -> the West Sea

Response: Correction made.

P13, L27: West Sea -> the West Sea

Response: Correction made.

P14, L11: Baengnyeong -> the Baengnyeong

Response: Correction made.

P15, L10 (and other locations): Olympic Park should always be capitalized (both words)

Response: Thank you. Correction made.

P15, L13: exhibit -> exhibits

Response: Correction made.

P15, L23: shop tracks -> ship tracks

Response: Correction made.

P17, L7-9: There are a few disjointed short sentences here. (e.g. "Because the size of CO data () is much larger than that of CO₂ ().") Perhaps they could be joined together to make more sense?

Response: We have joined this sentence and the two followed sentences into one. I.e., we have changed:

“Because the amount of CO data (13612 retrievals for MOPITT and 25509 for IASI over our study domain during KORUS-AQ) is much larger than that of CO₂ (42 for GOSAT over our domain during KORUS-AQ). This is illustrated in Fig. 10 and listed in Table 4. There are more observational constraints for CO in CAMS resulting to better performance of ANs CO.”

to:

“Because the amount of CO data (13612 retrievals for MOPITT and 25509 for IASI over our study domain during KORUS-AQ) is much larger than that of CO₂ (42 for GOSAT over our domain during KORUS-AQ), there are more observational constraints for CO in CAMS resulting to better performance of ANs CO (Fig. 10 and Table 4).”

P17, L24-25: near Korean coast -> near the Korean coast

Response: Correction made.

P18, L15: "(by -2 to -4 ppmv for CO₂ and -86 to -88 ppbv)" -> "(by -2 to -4 ppmv for CO₂ and -86 to -88 ppbv for CO)"

Response: Correction made.

P27, caption label: I would suggest using "bright" instead of "luminous" to describe the colours. Also add some articles when describing the sites, i.e. "The Olympic Park and Yonsei sites are located in an urban region (Seoul) while the Baengnyeong and Fukue sites are located in remote regions. The Taehwa site is located in a forest near Seoul."

Response: Thank you. We have substituted "*luminous*" with "*bright*". We have also added references to the caption when describing the sites.

P34, Figure 9: The figure label includes DC-8 still, but I believe this is in fact surfacebased in situ data. If so, please remove these confusing labels.

Response: Thank you for pointing this out. We have changed "*DC-8*" to "*Observations*" in the label.

P40, Table 3: The label refers to satellite measurements, but it should be in situ measurements.

Response: Thank you. Correction made.

Evaluating High-Resolution Forecasts of Atmospheric CO and CO₂ from a Global Prediction System during KORUS-AQ Field Campaign

Wenfu Tang¹, Avelino F. Arellano¹, Joshua P. DiGangi², Yonghoon Choi^{2,3}, Glenn S. Diskin², Anna Agustí-Panareda⁴, Mark Parrington⁴, Sebastien Massart⁴, Benjamin Gaubert⁵, Youngjae Lee⁶, Danbi Kim⁶, Jinsang Jung⁷, Jinkyu Hong⁸, Je-Woo Hong⁸, Yugo Kanaya⁹, Mindo Lee⁶, Ryan M. Stauffer^{10,11}, Anne M. Thompson¹¹, James H. Flynn¹², and Jung-Hun Woo¹³

¹Dept. of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

²NASA Langley Research Center, Hampton, VA, USA

³Science Systems and Applications, Inc., Hampton, VA, USA

⁴European Centre for Medium-Range Weather Forecasts, Reading, UK

⁵Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

⁶National Institute of Environmental Research, Korea

⁷Korea Research Institute of Standards and Science, Korea

⁸Department of Atmospheric Sciences, Yonsei University, Korea

⁹Japan Agency for Marine-Earth Science and Technology, Japan

¹⁰Universities Space Research Association, Columbia, MD, USA

¹¹Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA

¹²Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, USA

¹³Dept. of Advanced Technology Fusion, Konkuk University, Korea

Correspondence to: Wenfu Tang (wenfutang@email.arizona.edu)

Abstract. Accurate and consistent monitoring of anthropogenic combustion is imperative because of its significant health and environmental impacts, especially at city-to-regional scale. Here, we assess the performance of the Copernicus Atmosphere Monitoring Service (CAMS) global prediction system using measurements from aircraft, ground sites, and ships during the Korea United States Air Quality (KORUS-AQ) field study in May to June 2016. Our evaluation focuses on CAMS CO and CO₂ analyses plus two higher resolution forecasts (16-km and 9-km horizontal resolution), to assess their capability in predicting combustion signatures over East Asia. Our results show a slight overestimation of CAMS CO₂ with a mean bias against airborne CO₂ measurements of 2.2, 0.7, and 0.3 ppmv for 16-km and 9-km CO₂ forecasts, and analyses, respectively. The positive CO₂ mean bias in the 16-km forecast appears to be consistent across the vertical profile of the measurements. In contrast, we find a moderate underestimation of CAMS CO with an overall bias against airborne CO measurements of -19.2 (16-km), -16.7 (9-km), and -20.7 ppbv (analysis). This negative CO mean bias is mostly seen below 750 hPa for all three forecast/analysis configurations. Despite these biases, CAMS shows a remarkable agreement with observed enhancement ratios of CO with CO₂ over the Seoul metropolitan area and over the West Sea, where East Asian outflows were sampled during the study period. More efficient combustion is observed over Seoul ($\frac{dCO}{dCO_2} = 9$ ppbv/ppmv) compared to the West Sea ($\frac{dCO}{dCO_2} = 28$ ppbv/ppmv). This ‘combustion signature contrast’ is consistent with previous studies in these two regions. CAMS captured this difference in enhancement ratios (Seoul: 8-12 ppbv/ppmv, the West Sea: ~30 ppbv/ppmv) regardless of forecast/analysis configurations. The correlation of CAMS CO bias with CO₂ bias is relatively high over these two regions (Seoul: 0.64-0.90, the West Sea: ~0.80) suggesting that the contrast captured by CAMS may be dominated by anthropogenic emission ratios used in CAMS. However, CAMS shows poorer performance in terms of capturing local-to-urban CO and CO₂ variability. Along with measurements at ground sites over the Korean peninsula, CAMS produces too high CO and CO₂ concentrations at the surface with steeper vertical gradients (~0.4 ppmv/hPa for CO₂ and 3.5 ppbv/hPa for CO) in the morning samples than observed (~0.25 ppmv/hPa for CO₂ and 1.7 ppbv/hPa for CO), suggesting weaker boundary layer mixing in the model. Lastly, we find that the combination of CO analyses (i.e., improved initial condition) and use of finer resolution (9-km vs 16-km) generally produce better forecasts.

1. Introduction

Anthropogenic combustion significantly impacts air quality, climate, ecosystem, agriculture, and public health at local to global scales (Charlson et al., 1992; Doney et al., 2007; Feely et al., 2004; Heald et al., 2006; Maher et al., 2016). This is especially the case in megacities where human activities are most intense, accompanied by immense energy consumption, mainly in the form of fossil-fuel combustion, which directly leading to enhanced emissions of air pollutants, greenhouse gases, and waste energy. In particular, cities in the Asian region that are rapidly developing in recent decades are subject to more frequent severe pollution conditions (Yang et al., 2013; Guo et al., 2014; Ohara et al., 2007; Shindell et al., 2008, 2011). It is imperative therefore that we enhance our current capability to monitor, verify, and assess anthropogenic

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

combustion and its impacts as the number of megacities across the globe is expected to rapidly grow in the following decades (United Nations, 2016). The Copernicus Atmosphere Monitoring Service (CAMS) has a state-of-art global and integrated prediction systems that is currently being implemented to meet this need. The Service is funded by the European Union and it builds upon a legacy of projects such as the Monitoring Atmospheric Composition and Climate (MACC) and GEMS (Hollingsworth et al, 2008).

For nearly a decade, CAMS has been operationally producing daily global near-real-time forecasts and analyses of reactive trace gases, greenhouse gases, and aerosols including global reanalyses and estimation of emissions of these atmospheric constituents (Morcrette et al., 2009; Benedetti et al., 2009; Kaiser et al., 2012; Flemming et al., 2015; Flemming et al., 2017; Massart et al., 2016; Agusti-Panareda et al. 2014, Agusti-Panareda et al. 2017). CAMS global forecasts and analyses are based on the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF), which is also used for Numerical Weather Prediction (NWP). CAMS recently developed 2 forecasts at higher resolution, which have potential advantages compared to lower resolution analysis and/or forecast, in terms of local-to-regional air quality (Table 1).

The Korea United States Air Quality (KORUS-AQ) field measurement campaign offers a unique opportunity to assess the accuracy and consistency of the high resolution forecast and analysis system of CAMS and its skill in simulating atmospheric CO₂ from anthropogenic combustion. During May to June 2016, the KORUS-AQ field campaign collected comprehensive measurements of air quality (including CO₂ and tracers of fossil-fuel combustion) over the South Korean peninsula and its surrounding waters. KORUS-AQ is an international collaboration between U.S. and South Korea to better understand the factors controlling air quality in the region across urban, rural, and coastal interfaces (Kim and Park, 2014, KORUS-AQ White Paper). This field campaign follows several NASA-led sub-orbital missions in the past focusing on air quality in the United States (e.g., DISCOVER-AQ, SEAC⁴RS), and pollution outflows from Asia (e.g., TRACE-P, INTEX-B, ARCTAS) and integrating the measurements from these campaigns to satellite retrievals and air quality models (Crawford et al., 2014; Toon et al., 2016; Jacob et al., 2003; Singh et al., 2009; Jacob et al., 2010). Local measurements over the West Sea, often representative of Chinese pollution outflow, and over the Seoul metropolitan area provide a rich dataset that is very useful in evaluating global prediction and analysis systems like CAMS at city-to-regional scale.

In this study, we evaluate CAMS forecast and analysis of fossil-fuel combustion signatures over the KORUS-AQ spatial and temporal domain. In particular, we use measurements of the main products of combustion (i.e., CO and CO₂) from the NASA DC-8 aircraft, along with observations from five ground sites, two research ships, and four satellites to assess the capability of CAMS to monitor anthropogenic combustion. Although CAMS CO and CO₂ forecasts and analyses have been evaluated previously (Agusti-Panareda et al., 2014; Agusti-Panareda et al., 2016; Agusti-Panareda et al., 2017; Claeysman et al., 2010; Massart et al., 2016; Flemming et al., 2009; Flemming et al., 2015; Flemming et al., 2017), this study is unique for the following reasons: (1) This study is a joint evaluation of CO and CO₂ species, including their associated enhancement ratios which provide insights on CAMS representation of anthropogenic combustion processes; (2) A focus on megacities provides an important baseline investigation. This is especially the case in East Asia where there is still lack of

Deleted:

detailed information and measurements to constrain emission inventories; (3) KORUS-AQ provides a unique opportunity to evaluate the new high resolution global CAMS forecasts of CO and CO₂ at local-to-regional scale. This paper begins with a brief description of CAMS and KORUS-AQ (Section 2), followed by an evaluation of CAMS with airborne measurements (Section 3) and with ground sites, ships, and satellites (Section 4). We provide a summary of our findings in Section 5.

2. Descriptions of CAMS and KORUS-AQ CO and CO₂

2.1 CAMS CO and CO₂ forecasts and analysis

The Copernicus Atmosphere Monitoring Service (CAMS) has been providing global forecasts and analysis of atmospheric composition on a daily basis at ECMWF for nearly a decade with applications on air quality and monitoring of long-lived greenhouse gases. CAMS uses the Integrated Forecasting System (IFS) for Numerical Weather Prediction (NWP) to assimilate a wealth of meteorological observations plus satellite products of atmospheric composition to produce atmospheric analysis of reactive gases (e.g. CO, O₃, NO₂, SO₂), aerosols and long-lived greenhouse gases (e.g. CO₂, CH₄) on the NWP model grid which are then used as initial conditions to forecast the atmospheric composition with a 5-day lead time. The IFS simulates transport of the chemical species (Flemming et al. 2009, Agusti-Panareda et al. 2017) and includes the on-line integration of modules for atmospheric chemistry (Flemming et al. 2015, 2017) and biogenic CO₂ fluxes from terrestrial vegetation (Boussetta et al., 2013) to model atmospheric composition in conjunction with an assimilation system based on four-dimensional variational (4D-VAR) data assimilation (Rabier et al., 2000, Inness et al., 2015). The CAMS global atmospheric analysis and prediction system runs at different resolutions and at a different lag times for the various atmospheric species depending on the use of chemistry in the model and the timeliness of the satellite retrievals used in the analysis. The system providing reactive trace gases and aerosols runs at approximately 80 km horizontal resolution with 60 vertical levels and its analysis is available less than 1-day behind real time. While higher horizontal and vertical resolution is used for the analysis and forecasts of greenhouse gases, the analysis of CO₂ and CH₄ is available at around 40 km in the horizontal and 137 vertical levels. Currently the forecasts of CO₂ and CH₄ have the same resolution as the operational weather forecast at ECMWF (137 levels with 9 km horizontal resolution) but previously their resolution was 16 km (from 2015 to 2016). A CO tracer with simplified chemistry based on a linear CO scheme (Massart et al., 2015) is also available in the high resolution forecasts. However, the CO₂ and CH₄ analysis is only available four days behind real time as the satellite retrievals are not available closer to real time. Because of this, in the 16-km resolution forecast, CO₂, CH₄ and linear CO are free running and only the meteorology is initialised with the meteorological operational analysis (see Agusti-Panareda et al. (2014) for further details on the free-running forecast configuration). Following a recent improvement in the timeliness of the satellite retrievals, the linear CO is initialised with CO analysis, while CO₂ and CH₄ are initialised with a 4-day forecast from the CO₂ and CH₄ 40 km analysis in the 9-km forecasts. In order not to lose the small-scale features in the initialization process, a spectral filter is applied to only adjust the large scales in the initial conditions of the forecast (Massart, 2016, personal communication). Table 1 (as well as Fig. S1) provides a summary of the three CAMS configurations and five

Deleted: 4

Deleted: .

Deleted: T

Deleted: 4-days

Deleted:

Deleted: 1

resulting CAMS products evaluated in this paper and Fig. S2 depicts the different vertical and horizontal resolutions used in the different CAMS configurations.

For this study, we focus on evaluating the three CO and CO₂ forecasts and analysis products listed above, namely, CO₂ and CO 16-km forecast (FC16s), analyses (ANs) of CO₂ (at 40 km) and CO (at 80 km), and a relatively recent CAMS 9-km CO₂ and CO forecast product (FC9s) which are initialized from its respective analysis. The FC9s are different from FC16s in terms of both resolution and initialization as described above (e.g. the FC16s are produced from a free-running simulation of CO₂ and CO). The near-real time ANs of CO and CO₂ are also different from FC16s and FC9s as these ANs continuously assimilate satellite retrievals of CO total column from Measurements Of Pollution In The Troposphere (MOPITT V5-TIR) and the Infrared Atmospheric Sounding Interferometer (IASI) (Inness et al., 2015), and column averaged dry-air mole fractions of CO₂ (XCO₂) from the Greenhouse gases Observing Satellite (GOSAT) (Massart et al., 2016), in addition to the available meteorological data. Observations of both CO and CO₂ are assimilated in 12-hour assimilation windows. Inness et al. (2015) found that CO total column field, vertical distribution, and concentrations in the lower troposphere are improved by assimilating the CO total column from MOPITT. Assimilation of the GOSAT XCO₂ lead to improvements in mean absolute error and bias variability in XCO₂ fields during the year 2013 (Massart et al., 2016). FC9s CO are initialized from MOPITT and IASI CO analysis at a previous time, which are then downscaled from 80 km to 9 km by a spectral filtering scheme. Due to observational and computing constraints, FC9s of CO₂ are initialized and downscaled from a 96-hour forecast of CO₂ initialized by GOSAT analysis 4 days earlier.

The IFS contains several components, including an atmospheric general circulation model, a land surface model, an ocean wave model, an ocean general circulation model, and perturbation models for the data assimilation and forecast (Persson, 2001). Model dynamics and numerical procedures, and physical processes are documented in IFS documentation-Cy43r3 (ECMWF, 2017, https://www.ecmwf.int/search/elibrary/part?title=part&year=2017&secondary_title=IFS). Detailed cloud and precipitation physics of the IFS benefits the calculation of wet deposition (Flemming et al., 2017). As for emissions and surface fluxes, CAMS uses the Global Fire Assimilation System (GFAS) for biomass burning fluxes of CO₂ (Kaiser et al., 2012). CAMS uses the anthropogenic CO₂ fluxes that are based on the annual mean of the Emission Database for Global Atmospheric Research version 4.2 (EDGARv4.2). As the most recent year available for EDGARv4.2 is 2008, estimated and climatological trends are used to extrapolate to the years after 2008. The land vegetation fluxes for CO₂ are calculated online by the carbon module of the land surface model in IFS CTESSEL (Boussetta et al., 2013). A biogenic flux adjustment scheme (BFAS) is employed in CAMS to improve the continental budget of CO₂ fluxes (Agusti-Panareda et al., 2014; Agusti-Panareda et al., 2015; Agusti-Panareda et al., 2016). Specifically, (1) BFAS computes the scaling factors for the model net ecosystem exchange (NEE) based on reference (NEE climatology from the optimized fluxes); (2) the scaling factors are used to adjust biogenic CO₂ fluxes from the land surface model (i.e., flux bias correction); (3) the bias-corrected fluxes are then used to simulate the atmospheric CO₂. According to Agusti-Panareda et al. (2016), in Northern Asia, the employment of BFAS slightly decreases NEE in May and has negligible impacts on NEE in June. CO₂ overestimation by CAMS over the Northern Hemisphere (NH) in winter and spring is enhanced by BFAS. For CO, CAMS uses anthropogenic

Deleted: 1

Deleted: ;

Deleted:)

Formatted: Subscript

Formatted: Subscript

Formatted: Subscript

Formatted: Subscript

and biogenic emissions that are based on the MACC/CityZEN EU projects (MACCity) (Granier et al., 2011), and a climatology of the Model of Emissions of Gases and Aerosols from Nature developed under the MACC (MEGAN-MACC) emission inventories (Sindelarova et al., 2014). GFAS is also used for fire emissions. ANs for CO use the on-line implemented chemical mechanism (C-IFS-CB05, Flemming et al., 2015) that is an extended version of the Carbon Bond mechanism 5 (CB05, Yarwood et al., 2005). Because hydroxyl radical (OH) is an important sink for CO, modeled OH is critical for the simulation of CO (Gaubert et al., 2016, 2017). In the ANs for CO, the global and NH means of air mass-weighted OH are 0.98×10^{-6} molecules/cm³ and 1.20×10^{-6} molecules/cm³ during May 2016, respectively (calculated following recommendations from Lawrence et al. (2001)). The mean OH from the ANs for CO is consistent with previous studies (e.g., Lawrence et al., 2001; Lelieveld et al., 2016; Gaubert et al., 2016, 2017). A linear chemistry scheme is (C-IFS-LINCO) used in FC16s and FC9s for CO for computationally expediency (Claeyman et al., 2010; Flemming et al., 2012; Massart et al. 2015; Eskes et al., 2017). C-IFS-LINCO computes CO sources and sinks using the approach developed by Cariolle and Déqué (1986) and updated by Cariolle and Teyssède (2007), without direct use of modeled OH. C-IFS-LINCO is less computationally demanding than the full chemistry, permitting simulations at higher resolutions (Massart et al. 2015). Key aspects of the three CAMS configurations evaluated in this study are listed in Table 1.

2.2 CO and CO₂ measurements during KORUS-AQ

KORUS-AQ is a comprehensive field campaign based on international collaboration between U.S. and South Korea (<https://espo.nasa.gov/korus-aq>). The goal is to better understand the factors controlling air quality (AQ) in the region across urban, rural, and coastal interfaces. The field campaign was conducted over the South Korean peninsula and surrounding waters from May to June 2016. the South Korean peninsula and its surrounding waters is a desirable region to conduct the campaign because: (1) Korea's urban/rural sectors are distinct, which is advantageous for distinguishing anthropogenic and natural emissions; (2) Korea is embedded in a rapidly changing region; (3) the region allows studies of local versus trans-boundary pollution; and (4) air quality monitoring and ground-based measurements are provided by Korea. AQ measurements (including CO₂) from aircrafts, ships, and ground sites were obtained during this period. The campaign was designed to answer three scientific questions: (1) what are the challenges and opportunities for satellite observations of air quality; (2) what are the factors governing ozone photochemistry and aerosol evolution; (3) how well do models perform and what improvements are needed to better represent atmospheric composition over Korea and its connection to the larger global atmosphere (Kim and Park, 2014, KORUS-AQ White Paper). Fig. 1 shows the study domain (30°N – 39°N, 123°E – 133°E) along with the tracks from DC-8 aircraft flights and research ship deployments. The locations of ground sites are also added in Fig. 1. Satellite retrievals from MOPITT CO and Orbiting Carbon Observatory-2 (OCO-2) CO₂ are shown in Fig. 1 to provide spatial context and coverage of remote sensing measurements during the campaign. All the observational data used in this study are summarized in Table 2.

Deleted: T

Formatted: Superscript

Deleted: (C-IFS-LINCO)

Formatted: Font:(Asian) SimSun, Not Bold, Font color: Auto, English (US)

Formatted: English (US)

Deleted: ,

Deleted: including

Deleted: ance

Deleted: needed

Deleted: 2

Deleted: 2

Deleted: 2

2.2.1 Airborne CO and CO₂ measurements

We use measurements of CO₂ and CO from the DC-8 aircraft. CO₂ was measured by Atmospheric Vertical Observations of CO₂ in the Earth's Troposphere (AVOCET) using a modified LI-COR model 6252 non-dispersive infrared spectrometer (NDIR). This instrument provides CO₂ concentrations with high precision by sensing the difference in light absorption between the continuously flowing sample and reference gases (Vay et al., 2003, 2011; <https://airbornescience.nasa.gov/instrument/AVOCET>). CO₂ 1 Hz 1 σ precision and accuracy are ± 0.1 ppm and ± 0.25 ppm, respectively. CO was measured by the Differential Absorption CO Measurement (DACOM) instrument via infrared wavelength modulation spectroscopy. The system uses three tunable diode lasers providing 4.7, 4.5, and 3.3 μ m radiation for accessing absorption lines of CO, N₂O, and CH₄. The time response for CO measurements is 1 s; the precision is < 1% or 0.1 ppbv; the accuracy is 2% (Warner et al., 2010; <https://airbornescience.nasa.gov/instrument/DACOM>). Calibrations for both instruments were performed during flight at regular intervals using gas standards traceable to the WMO scale (CO₂: x2012; CO: x2008) and certified by NOAA ESRL. Details about the two instruments are listed in Table 2. Note that we use the 1 min (60 s) merged DC-8 data in this study. The data ~~are~~ available at NASA Langley Research Center archive (www-air.larc.nasa.gov/missions/korus-aq/).

There were 20 formal DC-8 science flights. Note that for time reference, the ‘Date’ in this paper refers to the day on which the flight started in UTC time instead of Korean local time, unless the term ‘Local time’ is explicitly used. This ‘date’ in UTC time is one day ~~behind~~ Korea local time as all flights typically start at 8am local time. We also divide the flight measurements into five groups based on the land cover below the flight tracks and types of pollution sources with which they can be broadly associated with. These groups are classified as: Seoul metropolitan, Taehwa, ~~the~~ West (Yellow) Sea, Seoul-Jeju jetway and Seoul-Busan jetway (Please refer to Fig. 1 for an illustration of these flight groups). The Seoul metropolitan represents air samples over the large city of Seoul which can have a dominant signature from anthropogenic combustion processes. On the other hand, Taehwa represents air samples over a forest area near Seoul, which can be influenced by both surface carbon fluxes from the local forest as well as anthropogenic emissions from Seoul. Measurements over the West Sea were designed to capture China pollution outflows. The flight tracks over the West Sea were typically zonal tracks forming a ‘wall’ between China and Korea (see Fig. 1). These flights are conducted only when a China outflow is expected to be present based on weather and AQ forecasts during the campaign. These measurements enable us to investigate combustion signatures from China and differentiate them from Seoul. The Seoul-Jeju jetway and Seoul-Busan jetway groups are two jetway flights on which the DC-8 aircraft frequently obtain measurements. The two jetways are both above the Korean peninsula, connecting Seoul to Jeju and Busan, respectively. Flights ~~in~~ Seoul-Busan jetway is designed to capture activities in forest, rural, and Busan urban regions. The flights in Seoul-Jeju jetway, on the other hand, sample air over local power plants, transported air from the West Sea, and over nearby croplands. We will discuss our evaluation CAMS for each of these five groups in Section 3.

Deleted: is

Deleted: ahead

Deleted: of

Deleted: 2

Deleted: 2

Deleted: over

Deleted: s

2.2.2 Ground-based CO and CO₂ measurements

Observations from the following ground sites are used for comparison with CAMS CO and CO₂: Baengnyeong, Fukue, Olympic Park, Taehwa, and Yonsei University (see Fig. 1₄ for the site locations). The sites in Baengnyeong and Taehwa are managed by the National Institute of Environmental Research (NIER). The Baengnyeong site is located on the sparsely populated Baengnyeong Island, Incheon, northwest of Seoul. The Fukue site belongs to the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) and is located on the remote island of Fukue, Japan (Kanaya et al., 2016). The Olympic Park and Yonsei University sites belong to Korea Research Institute of Standards and Science and Yonsei University, respectively. Both sites are located within the Seoul Metropolitan area. These five ground sites cover different environments, which allows us to differentiate between urban (Olympic Park and Yonsei University) and remote (Baengnyeong and Fukue) air quality conditions during the campaign. The sites in Baengnyeong, Fukue, and Olympic Park provide measurements of CO (in ppbv), while the site in Yonsei University provides measurements of CO₂ (in ppmv). Only the site in Taehwa provides measurements of both CO (in ppbv) and CO₂ (in mg/m³) (Kim et al., 2013). Locations of the five sites, and corresponding instruments and data intervals are provided in the Table 2. Note that we use data from these sites taken during the KORUS-AQ campaign period to provide the ground context of our evaluation.

Deleted: 2

Deleted: Baengnyeong site is located in less populated Baengnyeong Island, Incheon which is northwest of Seoul

2.2.3 Ship observations

We use ship measurements of CO from Jangmok and Onnuri. Both of them are research vessels owned by Korea Institute of Ocean Science and Technology. The ship deployments are part of the Korea-United States Ocean Color (KORUS-OC) field study coinciding with KORUS-AQ. KORUS-OC was led by NASA and the Korean Institute of Ocean Science and Technology, focusing on the ocean color, biology and biogeochemistry as well as atmospheric composition in coastal waters adjacent to Korea (https://www.asp.ucar.edu/sites/default/files/4_Emmons_07_27_2016.pdf). The two ships sailed along the Korean coast from May 20th to June 5th. Tracks of the two ships are shown in Fig. 1₄ by dark grey (Jangmok) and light grey (Onnuri). CO measurements in Jangmok and Onnuri were taken from the Thermo 48i-TLE CO analyzer and Thermo 48C CO analyser, respectively (<http://www.kiost.ac.kr/kor.do>), and are provided every minute.

Deleted: 2

2.2.4 Satellite-derived CO and CO₂ retrievals

We use four sets of satellite-derived measurements for comparison with CAMS CO and CO₂. We use retrievals of CO₂ column-averaged dry air mole fraction (XCO₂) from NASA OCO-2, version 7, Level 2 (L2) full product with the standard quality flag and warn level <15 (Crisp et al. 2004; Boesch et al., 2011; Wunch et al. 2011a, 2011b, 2017; Osterman et al., 2015; Mandrake et al., 2015; <https://oco.jpl.nasa.gov/>), and from Japan Aerospace Exploration Agency (JAXA) GOSAT, Level 2 (L2), version 2 (Yokota et al., 2004, 2009; Morino et al., 2011; Crisp et al. 2012; <http://global.jaxa.jp/projects/sat/gosat/>). Short-wavelength Infrared observations measured by the Thermal And Near-infrared Sensor for carbon Observation (TANSO) onboard the GOSAT satellite are used to retrieve XCO₂. OCO-2 also has three

Deleted: recommended quality control

specific Near Infrared (NIR) wavelength bands to retrieve XCO₂ (<https://oco.jpl.nasa.gov/>). For CO, we use the NASA Terra MOPITT version 6, Level 2, multispectral (Thermal Infrared/Near Infrared; TIR/NIR) total column retrievals (MOP02J, L2, V6) with [the standard quality flag](#). Compared to thermal infrared only retrievals (TIR), these retrievals have an enhanced sensitivity to the lower tropospheric CO (Deeter et al., 2014; <https://www2.acom.ucar.edu/mopitt>). In addition, we also use total column mole fractions of CO from IASI, Level 2 data with [the standard quality flag](#) (George et al., 2009; Clerbaux et al., 2009). IASI is on board MetOp-A and B satellites and uses Fast Optimal Retrievals on Layers for IASI (FORLI) to retrieve CO distributions from the thermal infrared (TIR) spectra. We applied the associated averaging kernels from MOPITT and IASI to CAMS CO before comparison as these retrievals exhibit large sensitivities in the free troposphere. We also note that both IASI and MOPITT have significantly more observations than OCO-2 and GOSAT. As summarized in Table 2, [the](#) resolutions of OCO-2, GOSAT, MOPITT, and IASI are 2.25×1.29 km, 10.5×10.5 km, 22×22 km, and 12×12 km, respectively. The [overpass](#) times for the four satellites are also different. OCO-2 overpasses at 1:18 - 1:33 pm, GOSAT overpasses at around 1 pm. [Overpass](#) time is 10:30 am for MOPITT, and 9:30 am for IASI. Uncertainties have also been reported for these satellite products. OCO-2 XCO₂ has uncertainties of 1-2 ppm (Boesch et al., 2011) while GOSAT XCO₂ has retrieval errors of 2 ppm (Griffith et al. 2011; Crisp et al. 2012). Deeter et al. (2014) reported 0.09×10¹⁸ molecules/cm² for total column retrieval for MOPITT. Wachter et al. (2012) reported uncertainties to be <13% for IASI FORLI.

3. Comparison with airborne measurements

Here, we evaluate CAMS forecasts and analysis of CO and CO₂ with NASA DC-8 aircraft observations. We interpolate the 4-D fields of CAMS CO and CO₂ model output to collocate with flight measurements in both space and time. The equivalent model data for all flights and for the three configurations (FC16s, FC9s, ANs) are made available in the same file format as the 1-min merged DC-8 dataset to facilitate model to observation comparison. We also estimate enhancement ratios of CO and CO₂ from both airborne and model data and analyse its spatial and temporal variations across different flights. We present in the following subsections the summary statistics of our comparison of CAMS data with [the DC-8 aircraft data](#).

3.1 Performance across all flights

Across all flight data, CAMS overestimates CO₂, with mean biases of 2.2, 0.7, and 0.3 ppmv for FC16s, FC9s, and ANs, respectively. [Agusti-Panareda et al. \(2016\) also suggested CO₂ is overestimated by CAMS in the NH at the end of winter and throughout spring.](#) In contrast, CAMS underestimates CO with mean biases for FC16s, FC9s, and ANs against [the DC-8 aircraft data](#) of -19.2, -16.7, and -20.7 ppbv, respectively. The mean bias is calculated as the average across all data of CAMS minus [the DC-8 aircraft data](#). We also find that the overall pairwise correlation between [the DC-8 aircraft data](#) and CAMS is moderately high (CO₂: 0.52–0.57, CO: 0.65–0.73) while the root-mean-square-errors (RMSEs) in CAMS relative to [the DC-8 aircraft data](#) are about 7 ppmv for CO₂ and 80 ppbv for CO. These statistics can be summarized using a Taylor

Deleted: recommended quality control

Deleted: recommended quality control

Deleted: revisit

Deleted: Revisit

Deleted: °

Formatted: Superscript

Deleted: As found by Agusti-Panareda et al (2016), the overall overestimation of CO₂ associated with the biogenic bias correction.

Deleted: DC-8

Deleted: DC-8

Deleted: DC-8

Deleted: to

Deleted: DC-8

diagram as shown in Fig. S3 and Fig. S4 of the supplementary material. We also calculated the associated Taylor scores to summarize the skill of CAMS in capturing the observed CO₂ or CO variations. Taylor score (Taylor, 2001) is defined by

$$S = \frac{4(1+R)}{(\hat{\sigma}_r+1/\hat{\sigma}_r)^2(1+R_0)} \quad (1)$$

where $\hat{\sigma}_r$ is the ratio of σ_r (standard deviation of the model) and σ_r (standard deviation of observations), R is correlation between model and observations, and R_0 is the maximum potentially realizable correlation (= 0.9 in this study).

We find that CAMS has relatively good skill regardless of configuration: for CO₂, the skill scores are 0.82 (FC16s), 0.82 (FC9s), and 0.75 (ANs); while for CO, the skill scores are 0.85 (FC16s), 0.86 (FC9s), and 0.83 (ANs). However, it is important to note that these statistics can vary from flight to flight and the skill for CO₂ is not necessarily related to that of CO. For instance, for the May 10th flight, where a southern peninsula outflow was expected, CAMS ANs show higher skill than those from FC9s in terms of both CO₂ and CO, while the scores of FC16s are higher than those of FC9s in terms of CO (Fig. S5). Yet, for the May 3rd flight, where a weak Chinese influence was expected, the scores of FC16s and FC9s are higher for CO₂ than CO, while we find the opposite for the June 2nd flight, where the DC-8 aircraft sampled local influences. Lastly, we note that the skill of CAMS during the June 4th flight is not high for either species. This flight was designed to measure local point sources with large variations at much finer scales.

3.2 Performance across individual flights

We present in Fig. 2 the summary statistics of CAMS against the DC-8 aircraft data for all 20 individual flights. This is shown in the second to fourth rows of Fig. 2 as boxplots of the bias for FC16s, ANs and FC9s, respectively. We also show the boxplot of the airborne measurements of CO₂ (first row left column) and CO (first row right column) for each flight as points of comparison. The overall mean, median, interquantile range (IQR), and standard deviation (sigma) of the airborne measurements of CO₂ mixing ratios (in ppmv) are 410.37, 408.25, 5.97, and 7.73 respectively. The overall mixing ratio, which varies within 1 to 2 percent, are slightly higher than the month median observed in Mauna Loa (NOAA <https://www.esrl.noaa.gov/gmd/ccgg/>) for May 2016 (408±1 ppmv). For the airborne measurements of CO mixing ratios (in ppbv), the corresponding statistics (mean: 204.59, median:183.90, IQR:127.97, sigma: 101.74) show enhanced CO (and larger variance) than the background value observed in Mauna Loa (100±24 ppbv). In general, CAMS overestimates CO₂ and underestimates CO for most flights. Differences also exist among the 20 flights in terms of both measured mixing ratios and model biases from the DC-8 aircraft. For flights with higher observed variances, CAMS biases and the corresponding variance of the biases tend to be also larger. This is related to variations in weather conditions during the campaign along with variations in sampling goals of the science flights. For example, parts of flight tracks on May 3rd, May 17th, May 24th, May 29th, and May 30th were specifically designed to capture Chinese pollution outflow. In these days, the variances in CAMS biases for CO (but not CO₂) are generally larger than the average except for the flight tracks on May 3rd when Chinese influences were expected to be weak. The colored shades in Fig. 2 indicate flights for 'special conditions'. The grey and yellow shades indicate two special cases that we study in detail in later sections. In particular, DC-8 flew a 'wall' over

Deleted: 2
Deleted: 3
Deleted: (please refer to Equation S1 in the supplement)
Formatted: Centered
Formatted: Indent: First line: 0"
Formatted: Font color: Auto, (Asian) Chinese (PRC), (Other) English (US)
Deleted: have

Deleted: those for

Deleted: 4

Deleted: 3
Deleted: DC-8 measurements
Deleted: 3
Deleted: DC-8

Deleted: DC-8

Deleted: DC-8

Deleted: DC-8

Deleted: the flights in

Deleted: 3

the West Sea on May 24th to investigate the transport of Chinese pollution. On June 4th, DC-8 flew near Seoul to measure pollution from local point sources (e.g., power plants). The other shades indicate that the flights were conducted during a frontal passage (purple) and that the flights may possibly be affected by fires in Siberia (orange). These flights were not further analyzed in this study since for example the May 26th flight (with frontal passage influence) and the May 17th and May 19th flights (with possible fire influence) do not clearly stand out from the other flights (see Fig. 2).

3.3 Performance across flight groups

Here, we evaluate CAMS per flight group as described in section 2.2.1. We show in Fig. 3 the probability density functions (pdfs) of CO and CO₂ for the DC-8 aircraft data and CAMS per flight group. The pdf of CAMS CO₂ (which exhibits a longer tail to higher values) show a general offset to higher values relative to the DC-8 aircraft data (except for the West Sea). There is a systematic overestimation of CAMS CO₂ against the DC-8 aircraft data. Accordingly, the ‘apparent local background’ of CO₂ (lower tails of the pdfs) is relatively high in CAMS than the DC-8 aircraft data. In contrast, CO is underestimated in CAMS across all of the five groups. The pdfs of CO in CAMS show a bi-modal distribution (except in Taehwa and the West Sea) indicative of two dominant AQ conditions sampled by DC-8 over this region. The shapes of the CO pdfs of CAMS largely differ from those of the DC-8 aircraft data (except in Taehwa). We see a higher frequency of occurrence of the two to three modes in the West Sea in CAMS that is not apparent in the DC-8 aircraft data while the opposite is the case in Seoul-Busan. This suggests that the underestimation of CO in CAMS may not be systematic or may be caused by biases in CO background values. The pdf over the West Sea also show that CAMS underestimates (or even misses) the more elevated CO observed by the DC-8 aircraft.

We further investigate the differences between CAMS and the DC-8 aircraft data by looking at the bias in the mean profiles. We show in Fig. 4 the mean profiles for all data and each individual group. We find that the overall bias in CAMS CO₂ is systematic and close to uniform across all layers (FC16s: ~2.2 ppmv, FC9s: ~1 ppmv, and ANs: ~0.8 ppmv). This overestimation is true for all flight groups except over the West Sea. On the other hand, for CO, the overall bias in CAMS is mostly evident in the lower troposphere (about -20 to 25 ppbv below 700 hPa). This underestimation is especially the case over the West Sea and is consistent with the pdfs in Fig. 3.

3.3.1 the Seoul metropolitan and Taehwa

The airborne measurements over the Seoul metropolitan area were mostly during frequent aborted landing maneuvers (i.e. missed approaches) over the Seoul Air Base. More than 90% of the measurements in this group are taken below 850 hPa. Fig. 3 shows that the performance of FC16s, FC9s, and ANs are alike over Seoul for both CO and CO₂, in contrast to the other four flight groups. Given that the measurements over Seoul are dominated by boundary layer (BL) and anthropogenic emissions in Seoul, the model performance over Seoul are most likely to be driven by local emissions. We show in Fig. 5 the mean vertical profiles over Seoul below 800hPa. For CO₂, FC9s profiles agree well with the observations. This is not the case for CO, where FC16s, FC9s, and ANs do not agree well with the DC-8 aircraft data, but with the bias in

Deleted: 3

Deleted: 4

Deleted: DC-8

Deleted: tale

Deleted: DC-8

Deleted: DC-8

Deleted: tales

Deleted: DC-8

Deleted: in

Deleted: s

Deleted: DC-8

Deleted: DC-8

Deleted: in

Deleted:

Deleted: DC-8

Deleted: 5

Deleted: 4

Deleted: a

Deleted: b

Deleted: 4

Deleted: 6

Deleted: DC

Deleted: -8

ANs being relatively smaller. However, the near surface temporal variations (changes in the profile from morning to afternoon) observed by [the DC-8 aircraft](#) are captured by FC16s, FC9s, and ANs. It is worth noting that over Seoul, there is an abrupt change in the profile at around 925 hPa for both CO and CO₂ of the morning samples. Accordingly, CO is overestimated below 925 hPa and underestimated above 925 hPa. This vertical gradient [below 925 hPa](#) (i.e., change in mixing ratios divided by change in pressure) in the [averaged profiles of the DC-8 aircraft data](#) CO₂ and CO are about 0.25 ppmv/hPa and 1.7 ppbv/hPa, respectively. In contrast, the gradients of CO₂ in CAMS are 0.50 ppmv/hPa for FC16s, 0.34 ppmv/hPa for FC9s, and 0.45 ppmv/hPa for ANs while the gradients of CO in CAMS are 4.2 ppbv/hPa for FC16s, 3.4 ppbv/hPa for FC9s, and 3.3 for ANs. It is evident that these gradients (CO and CO₂) regardless of CAMS configuration are significantly steeper than observed. While in part this may be attributed to overestimation of emissions during rush hours (and night-time) in Seoul along with model representativeness errors in the [BL](#), we attribute this steep gradient to a possible weaker [BL](#) mixing in CAMS since there is an important contrast between near surface CO (overestimation) and CO aloft (underestimation) which cannot be explained by emissions alone. This is not very apparent in CO₂ since there is an overestimation of background CO₂ superimposed on this difference. [In addition, given the air traffic over the Seoul Air Base \(where the DC-8 aircraft frequently conducted missed approaches\), emissions from airplanes may also contribute to the model biases \(Boschetti et al., 2015\).](#)

In Taehwa, the differences between morning and afternoon samples are not as large compared to [the Seoul metropolitan](#). The CO₂ profiles from ANs and FC9s are apparently closer to [the DC-8 aircraft data](#) than from FC16s. However, this difference is not obvious for the CO profiles. Note that in the afternoon (2-4pm), measured CO₂ mixing ratio near surface (at 975 hPa) becomes lower than the layer above, indicating a possible drawdown of CO₂ by underlying vegetation in Taehwa. This change is captured by CAMS, especially in FC9s. [We further find that compared with the Seoul metropolitan, the observed vertical gradient of CO₂ over Taehwa \(-0.03 ppmv/hPa\) below 925 hPa is smaller, which is relatively better captured by CAMS \(0.02-0.12 ppmv/hPa\). This again implies the possible inefficient BL mixing in CAMS over the Seoul urban environment. CO over Taehwa is more likely to be due to regional transport, as Taehwa is not a strong CO source region. Thus, the vertical gradient of CO over Taehwa does not necessarily reflect the impact of BL mixing over Taehwa. We further compared the mixing layer \(ML\) height derived from the KORUS-AQ airborne DIAL-HSRL measurements of aerosol backscatter following the technique from Brooks et al. \(2003\), and the BL heights from CAMS. We note that ML height is only approximately equal to BL height. We find that CAMS generally underestimates BL heights during KORUS-AQ \(Fig. S6\). The model underestimation of BL over the Seoul metropolitan \(-761.3±39.7 m\) is stronger than that over Taehwa \(721.7±38.6 m\) which is covered by forests instead of urban. This is consistent with the CAMS's relatively better capability of capturing vertical gradient of CO₂ over Taehwa compared to that over Seoul, supporting our previous implication of the possible inefficient BL mixing in CAMS over the Seoul urban environment.](#)

3.3.2 West (Yellow) Sea

As previously mentioned, the flights over the West (Yellow) Sea are focused on capturing pollution outflow from

Deleted: morning samples

Deleted: boundary layer

Deleted: boundary layer

Formatted: Pattern: Clear

Formatted: Indent: First line: 0"

Deleted: DC-8

Formatted: Subscript

Deleted: -

Formatted: Pattern: Clear (White)

China. Both CO and CO₂ in this flight group are underestimated by CAMS below 900 hPa (Fig. 4). It is the only group in which near surface CO₂ is underestimated by all the three CAMS configuration. In addition, the underestimation of CAMS CO over the West Sea is more significant than that over the other groups. We list two possible reasons for this unique model performance over the West Sea considering that the Chinese outflows constitute the dominant influence of CO and CO₂ samples in this group. First, the transport of surface pollution from China to the West Sea is not well represented in CAMS. Second, emissions in China may not be as well quantified as in Korea. During the May 24th flight, a strong outflow from China was expected, so DC-8 aircraft flew an extended sampling “wall” over the West Sea to sample transport from China. We show in Fig. 6 some of the details of this flight. In particular, we show the vertical cross sections of meridional (panel a) and zonal (panel b) fluxes of CO and CO₂ in CAMS FC9s. These fluxes are calculated as the product of meridional (from west to east) or zonal (from south to north) wind speed with simulated species density (i.e. in terms of units, $\frac{m}{s} \times \frac{kg}{m^3} = \frac{kg}{m^2 \cdot s}$). The China outflow moving towards the West Sea and Seoul is well demonstrated in the fluxes of CO in panel (a) and (b) especially in the region marked by the black rectangles. This outflow is not apparent in the fluxes of CO₂. This is because the variations in CO₂ density are very low relative to CO₂ background in contrast to CO variations. We also show in Fig. 6, panel (c) the measurements from the DC-8 aircraft and the bias of FC9s over the West Sea on that day. As can be seen in Fig. 6, CAMS CO₂ and CO are largely underestimated (CO₂: 2-4 ppmv, CO: 86-88 ppbv) for this flight. This underestimation in both species is consistent with Fig. 4. Note that the underestimation of CO₂ over the West Sea is not consistent with other flights and the overall results. This underestimation could be associated with an underestimation of anthropogenic emissions in China, and/or transport from China to the West Sea. This is discussed in Section 3.4 in more details. In summary, the transport pattern of China outflow (CO and CO₂) to the West Sea is captured but the abundances of both CO and CO₂ are underestimated by CAMS especially near the surface.

Deleted: 5

Deleted: than

Deleted: 7

Deleted: Hence, the wind speeds dominate the transport flux variations in CO₂.

Deleted: 7

Deleted: of

Deleted: 7

Deleted: 5

3.3.3 Seoul-Jeju and Seoul-Busan jetways

Measurements in the Seoul-Jeju and Seoul-Busan jetways are both above the South Korean peninsula, connecting Seoul to Jeju and Busan, respectively. While both flight groups share some common features, they are treated here as two distinct groups for the following reasons: (1) Seoul-Jeju jetway is close to the west coast of South Korea, whereas Seoul-Busan jetway sampled air southeast of Seoul and more inland; (2) There are more croplands, urban, and build-up areas along Seoul-Jeju jetway while there are more forested areas along Seoul-Busan jetway; (3) There are some important point sources along Seoul-Jeju jetway such as power plants and the Daesan chemical facility. In fact, the June 4th flight was designed to survey point sources west of Seoul and focused more to the Seoul-Jeju jetway. Details of the June 4th flight are summarized in Fig. 7. In contrast to the overall statistics across all flight groups, FC16s, FC9s, and ANs for this flight clearly overestimate CO near point sources. We also note that measurements for this flight are mostly taken below 900 hPa. As such, the spatial variations are larger near point sources than in other conditions. Nevertheless, these variations are not well captured by CAMS, especially by ANs. This may be due to its coarser grid representation (i.e., 40 km for CO₂ and 80 km for

Deleted: 8

CO). In addition, we find a difference in terms of mean bias in CO₂ between CAMS FC9s and FC16s. This difference is not apparent in CO. This implies there might be large spatiotemporal errors existing in CO emission inventories in the region, since higher emission resolution does not result in an improvement. In this case, increasing the spatiotemporal resolution might even weaken the simulation results, whereas lower resolution usually agrees better with observations as it “diffuses” the error of the emissions.

3.4 Enhancement ratios of CO to CO₂

We also evaluate the three CAMS configuration against the DC-8 aircraft data in terms of enhancement ratios of CO to CO₂ (dCO/dCO_2) for all flights and for each flight group. We conduct a reduced major axis (RMA) regression to estimate the sensitivity of CO to CO₂ (i.e., dCO/dCO_2) with the 1 minute merges. We use RMA instead of ordinary least squares (OLS) regression as the two variables (CO and CO₂) are both subject to error (Smith, 2009). The estimated regression slope in the RMA corresponds to enhancement ratio of CO and CO₂. This ratio can reflect the emission ratios of a particular area especially when using near field data (Parrish et al. 2002). Despite its limitations (Yokelson et al., 2013), such analysis has been used in previous studies for surface CO and NO_x (Parrish et al. 2002), emission factors for biomass burning (Wofsy et al., 1992; Lefer et al., 1994), flask samples of CO and CO₂ in East Asia (Turnbull et al., 2011), airborne measurements of CO and CO₂ during TRACE-P (Suntharalingam et al. 2004), surface CO and CO₂ in rural Beijing (Wang et al. 2010) and more recently with satellite retrievals of CO (MOPITT) and CO₂ (GOSAT) (Silva et al., 2013). We present our estimates of dCO/dCO_2 (with units of ppbv/ppmv) from the DC-8 aircraft data and CAMS FC16s, FC9s and ANs in Table 3. Overall, the observed dCO/dCO_2 during the KORUS-AQ campaign is ~13 ppbv/ppmv (or ~1.3%). This is a relatively low value compared to reported ratios in more polluted megacities such as Beijing. The lowest dCO/dCO_2 among the five flight groups is observed over Seoul (~9 ppbv/ppmv). The observed dCO/dCO_2 for other groups within Korea ranges from ~10 ppbv/ppmv (Seoul-Jeju) to ~16 ppbv/ppmv (Seoul-Busan and Taehwa). Taehwa is close to and sometimes downwind of Seoul, but has higher observed dCO/dCO_2 than Seoul. We attribute this difference to biogenic CO sources and biospheric influence on CO₂ over Taehwa. The highest dCO/dCO_2 (~28 ppbv/ppmv) is observed over the West Sea. This ratio is a sharp contrast to Seoul and other flight groups over Korea. This indicates that the bulk combustion efficiency over Seoul is higher in Seoul than in the China pollution outflows over the West Sea. The ratio over the West Sea is very consistent with dCO/dCO_2 observed over China (upwind of the West Sea) during KORUS-AQ by ARIAs (20-100 ppbv/ppmv (REF). Such ‘combustion signature contrast’ is consistent with previous studies in the region. During TRACE-P in 2001, the observed ratio over Japan is ~12-17 ppbv/ppmv and ~50-100 ppbv/ppmv over northern China (Suntharalingam et al. 2004). Over Shangdianzi, China and Tae-Ahn Peninsula (TAP), Korea, Turnbull et al. (2011) reported CO:CO₂ff ratios (which are derived from measurements of CO and $\Delta^{14}CO_2$ in flask samples taken during winter 2009/2010), of ~47 and ~44 ppbv/ppmv, respectively. They also reported that the South Korea samples from TAP have CO:CO₂ff of ~13 ppbv/ppmv. Wang et al. (2010) reported a change in observed dCO/dCO_2 near Beijing from 34-42 ppbv/ppmv in 2005-2007 to 22

Deleted: DC-8

Deleted: ($\Delta CO/\Delta CO_2$)

Deleted: estimate

Deleted: S

Deleted: Δ

Deleted: Δ

Deleted: DC-8

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

ppbv/ppmv in 2008. Finally, $\Delta\text{CO}/\Delta\text{CO}_2$ derived from satellite retrievals in 2010 indicate a similar contrast between Beijing/Tianjin (~25-50 ppbv/ppmv) and Seoul (~7-9 ppbv/ppmv). Despite the differences in the data sources (satellites, airborne measurements, flask samples) and time period, these $\Delta\text{CO}/\Delta\text{CO}_2$ values are consistent and all point to a 'combustion signature contrast' between Korea and China. We expect that this contrast may be decreasing over time as Chinese combustion activities become more efficient.

These observed ratios are remarkably consistent with $\Delta\text{CO}/\Delta\text{CO}_2$ from CAMS (see Table 3). The three CAMS configurations have $\Delta\text{CO}/\Delta\text{CO}_2$ over Seoul metropolitan of ~8 to 12 ppbv/ppmv and over the West Sea of ~31-32 ppbv/ppmv. Our rough estimates of CO to CO₂ emission ratios in CAMS over Seoul and China during KORUS-AQ also show marked similarity with CAMS enhancement ratios. The CO to CO₂ emission ratios over China is about 28 (1000 mole/mole) and about 10 (1000 mole/mole) over Korea. Our results suggest that CAMS emission ratios reflect this contrast and that the modeled $\Delta\text{CO}/\Delta\text{CO}_2$ is indicative of emissions of Seoul and China. To further understand the skill of CAMS in capturing this contrast, we compare the observed correlation between CO and CO₂ and the correlation from CAMS FC16s, FC9s, and ANs. This corr(CO₂,CO) is presented in the second row of Table 3. Over Seoul, the observed corr(CO₂,CO) is moderately high (~0.8), which is likely driven by common CO and CO₂ sources (mostly local anthropogenic emissions from Seoul). This correlation is well captured by ANs and FC9s but not FC16s. We attribute this difference to a better initialization in ANs and FC9s due to assimilation. The observed corr(CO₂,CO) over the West Sea is even higher (0.89), indicating that CO and CO₂ comes from common sources in China. However, this corr(CO₂,CO) is not captured by any of the three configurations (0.25-0.42). A few factors may contribute to this low corr(CO₂,CO) over the West Sea. First, the flight on May 12th is a noteworthy source of low corr(CO₂,CO) in CAMS. We have shown in Fig. 2 that the major goal of this flight is to study AQ conditions during a frontal passage instead of sampling China outflows. Even though part of the track during May 12th is located in the West Sea, the AQ features of that day are evidently different from China outflow events. After excluding measurements during May 12th, the corr(CO₂,CO) in CAMS (FC16s-0.51, FC9s-0.43, and ANs-0.29) are now higher albeit still lower than observed (0.9). Uncertainties in model transport can be a likely cause as the corr(CO₂,CO) can be subject to transport errors even though $\Delta\text{CO}/\Delta\text{CO}_2$ may not necessarily be affected. Performance of CAMS over the Baengnyeong site (discussed in Section 4.1) also implies possible issues with transport of China pollution towards the West Sea. Furthermore, the difference in temporal representation of China emissions in CAMS may contribute to this mismatch in timing and hence resulting to low correlation. As mentioned in Section 2, CAMS uses prescribed monthly emission for CO while the diurnal cycle of CO₂ fluxes is calculated online in CAMS. In fact, there is a strong diurnal cycle in the spatial correlations between CO emissions and CO₂ fluxes in CAMS caused by diurnal cycles of the CO₂ NEE (Fig. S8). The diurnal cycle of spatial correlations between CO emissions and CO₂ fluxes over Korea in CAMS peaks (~0.7) in daytime when measurements over Korea were made. On the other hand, during the nighttime, the correlations between CO emissions and CO₂ fluxes in CAMS are relatively low over East China (<0.4). This implies that the relatively low correlations between the CO and CO₂ abundances over the West Sea in CAMS may reflect the effect of nighttime emissions from East China in CAMS. Lastly, the

- Deleted: Δ
- Deleted: Δ
- Deleted: Δ
- Deleted: Δ
- Formatted: Indent: First line: 0"
- Deleted: Δ
- Deleted: Δ
- Deleted: Δ
- Deleted: Δ
- Deleted: Δ
- Deleted: 3
- Deleted: Δ
- Deleted: Δ
- Formatted: Subscript

corr(CO₂,CO) in FC16s and FC9s are closer to observed corr(CO₂,CO) than in ANs suggesting that resolution may also play a role. For the other three flight groups, the observed corr(CO₂,CO) are not as high as those over Seoul and the West Sea. This implies that CO₂ and CO observed over these three flight groups may not come from common sources and/or have been mixed with the environment. CAMS corr(CO₂,CO) do not always agree with observed corr(CO₂,CO). Overall, corr(CO₂,CO) from FC16s is higher than observed while corr(CO₂,CO) from FC9s and ANs agree well with observed corr(CO₂,CO). Again, this may be related to the fact that FC16s ~~is generated from a free running simulation (i.e. not initialized with analyses).~~

Finally, we present the correlation between the biases of CAMS for the two species (corr(Bias_{CO},Bias_{CO₂})) (please see the third row of Table 3). This correlation provides another piece of information on whether the performance of CAMS in CO₂ and CO are related. We find that corr(Bias_{CO},Bias_{CO₂}) are high over Seoul and the West Sea, indicating that the performance of CAMS in CO and CO₂ are related for the two groups. Over the West Sea, FC16s, FC9s, and ANs perform similarly. However, the corr(Bias_{CO},Bias_{CO₂}) are lower in the other three groups relative to Seoul and the West Sea. In addition, our results show that ANs and FC9s usually have lower corr(Bias_{CO},Bias_{CO₂})) than FC16s, especially over Seoul. This implies that FC16s performance in CO₂ and CO are more strongly related than in FC9s and ANs performance, which could be associated again with the fact that FC16s comes from a free running simulation while FC9s and ANs are both initialized from analyses. The assimilation of CO and CO₂ satellite retrievals may reduce the interdependence of CAMS CO₂ and CO performance.

4 Comparison with other measurements

In this section, we evaluate CAMS FC16s and FC9s, and ANs against CO and/or CO₂ measurements from five ground sites, two ships, and four satellites. Unlike ~~the data from the DC-8 aircraft,~~ data on CO₂ or CO in these cases may not be jointly available. In particular, each ground site (except Taehwa) only measures one of the two species. The ships also provide measurements for CO only while the four sets of satellite retrievals of CO₂ and CO are from four different instruments on board four different satellites. Therefore, in this section, CO₂ and CO are evaluated separately, and relationships between CO₂ and CO inferred from some of these sites are only indicative of a larger pattern that we see in ~~the~~ DC-8 ~~aircraft data.~~

4.1 Comparison with ground observations

Here, we focus our evaluation on CAMS performance in capturing surface conditions and diurnal cycle of CO₂ and/or CO. Data from the following five ground sites are used in this study: Baengnyeong, Fukue, Olympic Park, Taehwa, and Yonsei University (Fig. 1 and Table 2). It can be seen in Fig. 8 that CO from Olympic Park and CO₂ from Yonsei and Taehwa clearly show a diurnal cycle during KORUS-AQ. This feature is well captured by CAMS. CO at Taehwa on the other hand, exhibits a very weak diurnal cycle that is not captured by CAMS. At this site, CO in CAMS (especially ANs) shows a strong diurnal cycle. Variations of CO in the remote sites of Baengnyeong and Fukue also appears to be irregular

Deleted: comes

Formatted: Pattern: Clear

Deleted: DC-8

Deleted: 2

Deleted: 9

Deleted: p

and episodic. Signatures of elevated CO can also be seen at these sites, some of which coinciding with pollution transport from China sampled by the DC-8 aircraft. The mean diurnal cycle for these five ground sites can be found in Fig. S9.

While CAMS is able to get the observed timing of CO₂, the modelled magnitudes of CO₂ (and CO) at these sites from CAMS are too high (especially for the sites in and nearby Seoul). We took the average value across a few layers near the model surface in CAMS to provide a reasonable comparison at these sites. We use model vertical layers below 95% of the model surface pressure (i.e., if surface pressure is 1000hPa, we average the layers below 950 hPa) to account for potential weak BL mixing (especially near source regions). This feature in CAMS has been discussed in section 3.3.1. Since this averaging may introduce errors in our comparison, we only evaluate CAMS in terms of relative patterns (diurnal cycle and spatial variability across sites). Note that CAMS CO along the ship tracks (to be discussed in the succeeding section) are also averaged across a few layers in the same way for consistency. We show in Fig. 8 the summary statistics of the bias in CAMS relative to ground observations. The boxplots show that the variability of model bias in CO is in general smaller for remote sites and larger for the two sites in Seoul metropolitan. The bias in CAMS is also smaller in Fukue than in Baengnyeong, where a larger influence of pollution transport from China is observed but not well captured in CAMS. It is also worth mentioning that relative to other sites, CAMS significantly overestimates both CO and CO₂ at Taehwa. This may be due to the proximity of Taehwa to Seoul. The model grid spacing may not be able to resolve well the subgrid-scale processes (emissions) and variations between Seoul and Taehwa. This overestimation is most apparent in CAMS ANs which has a coarser grid spacing (40 km for CO₂ and 80 km for CO) than FC16s and FC9s. In the case of CO₂ at Yonsei, we find lower bias in CAMS FC9s and ANs than FC16s suggesting improvements of CAMS due to better initialization.

We take advantage of the location of the sites in Olympic Park (CO) and Yonsei University (CO₂) which are within Seoul metropolitan and the collocated measurements of CO and CO₂ in Taehwa to investigate patterns of ground-based $\Delta\text{CO}/\Delta\text{CO}_2$ in Seoul and Taehwa. Here, we only discuss observed $\Delta\text{CO}/\Delta\text{CO}_2$ since the modeled $\Delta\text{CO}/\Delta\text{CO}_2$ at these ground sites may not be accurate given CAMS issues with vertical mixing near the surface and representativeness errors. Following similar analysis with the $\Delta\text{CO}/\Delta\text{CO}_2$ of the DC-8 aircraft data, regressions of CO to CO₂ at these sites can represent emission ratios of CO to CO₂ in Seoul metropolitan. Our estimate of $\Delta\text{CO}/\Delta\text{CO}_2$ from Olympic Park and Yonsei sites is 11.32 ppbv/ppmv. This is consistent with $\Delta\text{CO}/\Delta\text{CO}_2$ calculated from the DC-8 aircraft data, which sampled air closely above these sites (~9 ppbv/ppmv). Our estimate of $\Delta\text{CO}/\Delta\text{CO}_2$ from the Taehwa site is 6.57 ppbv/ppmv. This is different from our estimate of 15.3 ppbv/ppmv based on the DC-8 aircraft data. Unlike Seoul, 70% of the airborne measurements over Taehwa are taken above 800 hPa. Over Taehwa, airborne $\Delta\text{CO}/\Delta\text{CO}_2$ varies with altitude from 8.92 ppbv/ppmv below 950 hPa, 10.28 ppbv/ppmv below 900 hPa, and 14.74 ppbv/ppmv above 400 hPa.

4.2 Comparison with ship observations

Two research vessels (Jangmok and Onnuri) were deployed during KORUS-OC. The two ships travelled along the Korean coast and measured CO from May 20th to June 5th (as marked in Fig. 1). Measurements of CO from ships, and biases of CAMS FC16s, ANs, and FC9s are shown in Fig. 9. Note that CAMS values along ship tracks are also averaged across a

Deleted: 5

Deleted: boundary layer

Deleted: and CO₂

Deleted: o

Deleted: 9

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: DC-8 Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: Δ

Deleted: DC-8

Deleted: Δ

Deleted: Δ

Deleted: DC-8

Deleted: DC-8

Deleted: Δ

Deleted: Δ

Deleted: 2

Deleted: 10

few layers near surface in the same way CAMS at ground sites were processed. CAMS at three (out of four) ground sites tend to underestimate CO, while CAMS overestimates CO relative to ship measurements. This seems to be inconsistent with our findings with airborne measurements (i.e., CO is underestimated by CAMS at lowermost troposphere (Fig. 4 and Fig. 6). This is likely due to the differences in sampling between the airborne and ship measurements. Over sea, the DC-8 aircraft often sampled air from China outflow while the two ships continuously sampled air over the waters regardless of the presence of China outflows. The ship measurements reflect surface conditions over waters which may also be different from what is observed by the DC-8 aircraft along the vertical profile. This inconsistency is further discussed in the next section with satellite data.

4.3 Comparison with satellite retrievals

The total column dry air mole fractions of CO₂ and CO (XCO₂ and XCO) derived from CAMS are compared here to XCO₂ from OCO-2 and GOSAT, and XCO from MOPITT and IASI. It is worth noting that satellite retrievals may have associated bias and uncertainties, which are generally larger than those of ground and airborne measurements. Slight inconsistencies also exist between MOPITT XCO and IASI XCO (George et al., 2009; 2015). We show in Fig. 10 the spatial distribution of CAMS biases against these retrievals. We also summarize the statistics in Table 4. Overall, ANs tend to agree better with satellite observations than the forecasts. For CO, CAMS XCO tends to be higher than MOPITT but lower than IASI. In addition, CAMS XCO agrees better with MOPITT than IASI. For CO₂, CAMS XCO₂ tend to be higher than GOSAT but lower than OCO-2. FC16s, FC9s, and ANs differ from each other in terms of bias when compared to any of the four satellite retrievals although there is no clear difference in terms of RMSE. For XCO, when compared to MOPITT, ANs are better than the two forecasts in terms of bias, RMSE, and correlation. When compared to IASI, ANs are better in terms of RMSE and correlation, but not its bias. For XCO₂, ANs do not show improvements from the two forecasts when compared to both OCO-2 and GOSAT retrievals. For both XCO and XCO₂, FC9s is not necessarily better than FC16s. In summary, ANs XCO show better agreement with satellite retrievals but this is not the case for XCO₂. Differences in the resolution and amount of satellite data of XCO and XCO₂ could be two possible causes. The spatial and temporal resolutions of FC16s and FC9s are higher than those of ANs while ANs assimilate observational data from these satellite retrievals (except OCO-2). These two factors compete against each other. Because the amount of CO data (13612 retrievals for MOPITT and 25509 for IASI over our study domain during KORUS-AQ) is much larger than that of CO₂ (42 for GOSAT over our domain during KORUS-AQ), there are more observational constraints for CO in CAMS resulting to better performance of ANs CO (Fig. 9 and Table 4). The opposite is the case for CO₂. The model resolution dominates for CAMS CO₂ performance especially with regards to capturing spatiotemporal variability. Scatter plots of CAMS XCO and XCO₂ against satellite observations are also presented in Fig. S10 of the supplementary material.

We note that CAMS overestimates XCO when compared with MOPITT XCO over the West Sea (Fig. 10). This appears to be contradictory to our conclusions in section 3 and the similar inconsistency also exists when we compare CAMS CO with ship measurements (as mentioned in Section 4.2). To further explain this inconsistency, we compare CAMS

Deleted: 5

Deleted: 7

Deleted: 1

Deleted: size

Moved (insertion) [1]

Deleted: .

Moved down [1]: This is illustrated in Fig. 10 and listed in Table 4.

Deleted: T

Deleted: This is illustrated in

Deleted: 10

Deleted: listed in

Deleted: .

Deleted: 6

Deleted: 1

FC9s with ship measurements and satellite XCO. Because the West Sea flight group in the DC-8 aircraft data forms a zonal ‘wall’ and such measurements over the West Sea are only conducted when a China outflow is expected, we separate the days when China outflows are present. The following are the days during the campaign when China outflows were expected to occur and DC-8 flights measured walls over the West Sea: May 3rd, May 17th, May 24th, May 29th, and May 30th. On May 3rd, May 17th, May 24th, and May 29th, there are no MOPITT observations over the West Sea (Fig. S11). Therefore, the overall differences between CAMS FC9s and MOPITT observations are driven by the non-outflow days. On May 30th, however, there are MOPITT observations over the West Sea. Unlike the overall picture (Fig. 10), we find that CAMS actually underestimates the outflows over the West Sea on that day, which is consistent with our findings in Section 3. On June 1st (a non-China outflow day), comparison with ship measurements indicates that CAMS FC9s overestimates CO near the Korean coast. It is also consistent with MOPITT XCO in June 1st (Fig. S11). This overestimation in CAMS FC9s is also captured in our comparison with Baengnyeong (highlighted by a black box in Fig. 9). We find similar overestimation using CAMS FC16s and ANs. Hence, during ‘normal’ conditions, CAMS tend to overestimate CO over the West Sea, whereas during China outflow events, CAMS tend to underestimate CO. More elaborate analysis of source contributions during KORUS-AQ is beyond the scope of this study and can be found in Tang et al. (2018), which suggested that during China outflow events, the contribution from Chinese direct emissions to CO over the West Sea is largely enhanced and dominant.

5 Discussions and Conclusions

We use measurements from the NASA DC-8 aircraft, five ground sites (Baengnyeong, Fukue, Olympic Park, Taehwa, and Yonsei University), and two ships (Jangmok and Onnuri) during the KORUS-AQ field campaign, along with four sets of satellite retrievals (MOPITT XCO, IASI XCO, OCO-2 XCO₂, and GOSAT XCO₂) to evaluate the capability of a high-resolution global modeling system (CAMS) in simulating anthropogenic combustion. Specifically, we evaluate the performance of CAMS FC16s, FC9s, and ANs of CO₂, CO, and their relationships. Our assessment of the overall performance of CAMS against the DC-8 aircraft data show that: (1) The nominal background CO₂ in CAMS is slightly overestimated (bias is 2.2 ppmv for FC16s, 0.7 ppmv for FC9s, and 0.3 ppmv for ANs), which is further improved by CO₂ analysis. On the other hand, CO is generally underestimated by CAMS (bias is -19.2 ppbv for FC16s, -16.7 ppbv for FC9s, and -20.7 ppbv for ANs); and (2) Among the three forecasts/analysis configurations, FC9s are more accurate and consistent overall than FC16s and ANs because of the finer model resolution and improved initialization. While ANs are coarser in resolution, they generally perform better than FC16s as the impact of initialization surpasses the impact of resolution (Fig. S3). We also classify the airborne measurements into five groups based on land cover below the flight tracks and associated pollution sources. While CO₂, CO, and their relationships vary across these five groups, CAMS perform well in terms of simulating regional pattern of anthropogenic combustion. This is because: 1) CAMS simulations of both species have relatively low bias; and 2) CAMS reproduces dCO/dCO_2 observed by the DC-8 aircraft. Both CAMS and the DC-8 aircraft data show more efficient combustion (low dCO/dCO_2) over Seoul than over the West Sea which is representative of Chinese outflows. Our case study on the May 24th flight over the West Sea indicates that the Chinese outflow is captured by CAMS.

Deleted: DC-8

Deleted: measurements

Deleted: S7

Deleted: 1

Deleted: 7

Deleted: 10

Formatted: English (US)

Deleted: DC-8 measurements

Deleted: . The overall overestimation of CO₂ might be associated with the biogenic bias correction

Deleted: 2

Deleted: Δ

Deleted: Δ

Deleted: DC-8

Deleted: DC-8

Deleted: Δ

Deleted: Δ

However, the modeled CO and CO₂ concentrations are significantly underestimated (by -2 to -4 ppmv for CO₂ and -86 to -88 ppbv for CO) especially within the lowermost troposphere. This suggests that, although CAMS emission ratios are relatively consistent with $\delta\text{CO}/\delta\text{CO}_2$, the absolute magnitude of China emissions are still underestimated. CAMS also show poorer performance at local-to-urban scales as exemplified by our case study in the June 4th flight where larger variations near point sources were not represented in CAMS. Our comparisons with measurements from ground sites and two ships indicate that: (1) the diurnal cycle of CO and CO₂ are stronger over urban environments and such periodic features are reasonably captured by CAMS; (2) vertical mixing near sources (such as Seoul) is too weak in CAMS and needs to be improved; and (3) in some cases, FC9s do not show improvements from FC16s (such as over Seoul and the point sources during the June 4th flight), implying large spatiotemporal errors in emission inventories. In these cases, increasing the spatiotemporal resolution might even weaken the simulation results, whereas lower resolution usually agrees better with observations as it “diffuses” the error of the emissions. We also compared XCO and XCO₂ derived from CAMS to satellite retrievals from four instruments (MOPITT CO, IASI CO, OCO-2 CO₂, and GOSAT CO₂). We find that ANs XCO show better agreement with satellite retrievals compared to the forecasts, while ANs CO₂ is no better than the forecasts. We attribute this contrast to significant differences in the number of XCO and XCO₂ satellite data potentially available for assimilation.

We recognize the following limitations of this work. (1) The temporal distribution of airborne measurements are not completely independent from their spatial distributions. For example, most of the measurements in the West Sea group are conducted before noon, whereas measurements in Seoul-Busan jetway are concentrated in the afternoon. (2) CAMS is only evaluated over the South Korean peninsula and surrounding waters during the campaign (May 1st to June 10th). More work is needed to determine if our findings are valid over other regions. For example, Agusti-Panareda et al. (2014) reported the overall overestimation of CO₂ in spring over the whole NH and it is enhanced by biogenic flux correction. (3) Inconsistencies exist even among different satellite products (George et al., 2009; 2015), thus limiting our comparisons with CAMS to relative differences; and 4) Our comparison of CAMS with ground and ship measurements are only qualitative and indicative as CAMS surface concentrations are significantly higher than surface observations and not comparable.

Finally, this study has important implications on the design and implementation of current and future prediction system for atmospheric composition and air quality. Although CAMS captured the regional combustion signatures, it still has difficulty representing the variability at local-to-urban scales even at finer resolution. This suggests both improvements in observational constraints and model representation of relevant processes (e.g., emissions and BL mixing).

Data availability. CAMS 16-km forecasts, and analyses are available online (<http://apps.ecmwf.int/datasets/data/cams-nrealtime/levtype=sfc/>). CAMS 9-km forecasts are available upon request. Observational data from KORUS-AQ will be open to public soon (<https://www-air.larc.nasa.gov/cgi-bin/ArcView/korusaq>). All the satellite data used in this study are available online. MOPITT CO and OCO-2 CO₂ can be downloaded at <https://reverb.echo.nasa.gov/reverb/>. IASI CO can be

Deleted: Δ

Deleted: Δ

Deleted: over

Deleted: Northern Hemisphere

Deleted: associated with

Deleted: boundary layer

found at http://ether.ipsl.jussieu.fr/ether/pubipsl/iasi_CO_uk.jsp. GOSAT CO₂ data after 2014 is available at <http://www.gosat.nies.go.jp/en/>.

Acknowledgements. This work is supported by NASA KORUSAQ NNX16AE16G and NNX16AD96G. We thank the KORUS-AQ team for observational data, the CAMS global production team for the model products of CO and CO₂, MOPITT, IASI, OCO-2, and GOSAT data teams for satellite data. IASI CO is provided by LATMOS/CNRS and ULB. We acknowledge NASA and the OCO-2 project for OCO-2 CO₂ data. We thank the DIAL/HSRL team for the Mixed Layer Heights product. The authors thank Dr. Cenlin He and Dr. Kazuyuki Miyazaki for helpful comments on improving the paper. NCAR is sponsored by the National Science Foundation. Yugo Kanaya was supported by the Environment Research and Technology Development Fund (2-1505 and 2-1803) of the Ministry of the Environment, Japan. The authors thank the anonymous reviewers for their comments and suggestions. The CAMS data was generated using Copernicus Atmosphere Monitoring Service Information [2016].

15 References

- Agusti-Panareda, A., Massart, S., Chevallier, F., Bousetta, S., Balsamo, G., Beljaars, A., Ciais, P., Deutscher, N. M., Engelen, R., Jones, L. and Kivi, R.: Forecasting global atmospheric CO₂. *Atmospheric Chemistry and Physics*, 14(21), 11959-11983, 2014.
- Agusti-Panareda, A., Monitoring upgrades of analysis/forecast system, MACC-III Deliverable D44.04, June 2015.
- Agusti-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Bousetta, S., Dutra, E., and Beljaars, A.: A biogenic CO₂ flux adjustment scheme for the mitigation of large-scale biases in global atmospheric CO₂ analyses and forecasts. *Atmospheric Chemistry and Physics*, 16(16), 10399-10418, 2016.
- Agusti-Panareda, A., Diamantakis, M., Bayona, V., Klappenbach, F., and Butz, A.: Improving the inter-hemispheric gradient of total column atmospheric CO₂ and CH₄ in simulations with the ECMWF semi-Lagrangian atmospheric global model. *Geoscientific Model Development*, 10(1), 1, 2017.
- Andreae, M. O., and Merlet, P.: Emission of trace gases and aerosols from biomass burning. *Global biogeochemical cycles*, 15(4), 955-966, 2001.
- Benedetti, A., Morcrette, J.J., Boucher, O., Dethof, A., Engelen, R.J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J.W. and Kinne, S.: Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation. *Journal of Geophysical Research: Atmospheres*, 114(D13), 2009.

- Boesch, H., Baker, D., Connor, B., Crisp, D., and Miller, C.: Global characterization of CO₂ column retrievals from shortwave-infrared satellite observations of the Orbiting Carbon Observatory-2 mission. *Remote Sensing*, 3(2), 270-304, 2011.
- 5 [Boschetti, F., Chen, H., Thouret, V., Nedelec, P., Janssens-Maenhout, G., & Gerbig, C. \(2015\). On the representation of IAGOS/MOZAIC vertical profiles in chemical transport models: contribution of different error sources in the example of carbon monoxide. *Tellus B: Chemical and Physical Meteorology*, 67\(1\), 28292.](#)
- Boussetta, S., Balsamo, G., Beljaars, A., Panareda, A. A., Calvet, J. C., Jacobs, C., Hurk, B., Viterbo, P., Lafont, S., Dutra, E. and Jarlan, L. (2013). Natural land carbon dioxide exchanges in the ECMWF Integrated Forecasting System: Implementation and offline validation. *Journal of Geophysical Research: Atmospheres*, 118(12), 5923-5946.
- 10 [Brooks, I. M.: Finding boundary layer top: Application of wavelet covariance transform to Lidar backscatter profiles, *J. Atmos. Ocean. Tech.*, 20, 1092–1105, 2003.](#)
- [Cariolle, D., & Déqué, M. \(1986\). Southern Hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. *Journal of Geophysical Research: Atmospheres*, 91\(D10\), 10825-10846.](#)
- [Cariolle, D., & Teyssedre, H. \(2007\). A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations. *Atmospheric chemistry and physics*, 7\(9\), 2183-2196.](#)
- 15 [Charlson, R. J., and Schwartz, S. E: Climate forcing by anthropogenic aerosols. *Science*, 255\(5043\), 423, 1992.](#)
- Clayman, M., Attié, J. L., El Amraoui, L., Cariolle, D., Peuch, V. H., Teyssède, H., Josse, B., Ricaud, P., Massart, S., Piacentini, A. and Cammas, J. P.: A linear CO chemistry parameterization in a chemistry-transport model: evaluation and application to data assimilation. *Atmospheric Chemistry and Physics*, 10(13), 6097, 2010.
- 20 Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S. and Wespes, C. (2009). Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmospheric Chemistry and Physics*, 9(16), 6041-6054.
- Crawford, J.H. and Pickering, K.E.: DISCOVER-AQ: Advancing strategies for air quality observations in the next decade. *Environ. Manage*, pp.4-7, 2014.
- 25 [Crisp, D., Atlas, R.M., Breon, F.M., Brown, L.R., Burrows, J.P., Ciais, P., Connor, B.J., Doney, S.C., Fung, I.Y., Jacob, D.J. and Miller, C.E.: The orbiting carbon observatory \(OCO\) mission. *Advances in Space Research*, 34\(4\), pp.700-709, 2004.](#)
- [Crisp, D., Fisher, B., O'Dell, C., Frankenberg, C., Basilio, R., Bosch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M. and Eldering, A.: The ACOS CO₂ retrieval algorithm-Part II: Global XCO₂ data characterization, 2012.](#)
- [Deeter, M. N., Martínez-Alonso, S., Edwards, D. P., Emmons, L. K., Gille, J. C., Worden, H. M., Sweeney, C., Pittman, J.V.,](#)
- 30 [Daube, B.C. and Wofsy, S. C.: The MOPITT Version 6 product: algorithm enhancements and validation. *Atmospheric Measurement Techniques*, 7\(11\), 3623-3632, 2014.](#)
- [Doney, S. C., Mahowald, N., Lima, I., Feely, R. A., Mackenzie, F. T., Lamarque, J. F., and Rasch, P. J.: Impact of anthropogenic atmospheric nitrogen and sulfur deposition on ocean acidification and the inorganic carbon system. *Proceedings of the National Academy of Sciences*, 104\(37\), 14580-14585, 2007.](#)

- ECMWF: IFS Documentation CY43R3. European Centre for Medium-Range Weather Forecasts: Reading, UK, 2017.
- Eskes, H.J., Wagner, A., Schulz, M., Christophe, Y., Ramonet, M., Basart, S., Benedictow, A., Blechschmidt, A.-M., Chabrilat, S., Clark, H., Cuevas, E., Flentje, H., Hansen, K.M., Im, U., Kapsomenakis, J., Langerock, B., Petersen, K., Richter, A., Sudarchikova, N., Thouret, V., Warneke, T., Zerefos, C.: Validation report of the CAMS near-real-time global atmospheric composition service: Period June - August 2017. Copernicus Atmosphere Monitoring Service (Cams) Report, Cams84_2015sc2_D84.1.1.7_2017djf_V1.Pdf, May 2017.
- Feely, R. A., Sabine, C. L., Lee, K., Berelson, W., Kleypas, J., Fabry, V. J., and Millero, F. J.: Impact of anthropogenic CO₂ on the CaCO₃ system in the oceans. *Science*, 305(5682), 362-366, 2004.
- Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system. *Geoscientific Model Development*, 2, 253-265, 2009.
- Flemming, J., Peuch, V. H., Engelen, R., and Kaiser, J. W.: A European global-to-regional air pollution forecasting system that combines modeling with satellite observations. *EM Magazine*, November, 610, 2013.
- Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A. M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A. and Jones, L.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF. *Geoscientific model development*, 8(4), 975-1003, 2015.
- Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A. and Peuch, V. H.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003–2015. *Atmospheric Chemistry and Physics*, 17(3), 1945-1983, 2017.
- Gamnitzer, U., Karstens, U., Kromer, B., Neubert, R. E., Meijer, H. A., Schroeder, H., and Levin, I.: Carbon monoxide: A quantitative tracer for fossil fuel CO₂? *Journal of Geophysical Research: Atmospheres*, 111(D22), 2006.
- Gaubert, B., Arellano, A. F., Barré, J., Worden, H. M., Emmons, L. K., Tilmes, S., ... Jones, N. (2016). *Toward a chemical reanalysis in a coupled chemistry climate model: An evaluation of MOPITT CO assimilation and its impact on tropospheric composition*. *Journal of Geophysical Research: Atmospheres*, 121, 7310–7343. <https://doi.org/10.1002/2016JD024863>.
- Gaubert, B., Worden, H. M., Arellano, A. F. J., Emmons, L. K., Tilmes, S., Barré, J., ... & Houweling, S. (2017). *Chemical Feedback From Decreasing Carbon Monoxide Emissions*. *Geophysical Research Letters*, 44(19), 9985-9995.
- George, M., Clerbaux, C., Hurtmans, D., Turquety, S., Coheur, P. F., Pommier, M., Hadji-Lazaro, J., Edwards, D. P., Worden, H., Luo, M. and Rinsland, C.: Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors. *Atmospheric chemistry and physics*, 9(21), 8317-8330, 2009.
- George, M., Clerbaux, C., Bouarar, I., Coheur, P.F., Deeter, M.N., Edwards, D.P., Francis, G., Gille, J.C., Hadji-Lazaro, J., Hurtmans, D. and Inness, A.: An examination of the long-term CO records from MOPITT and IASI: comparison of retrieval methodology. *Atmospheric Measurement Techniques*, 8, pp.4313-4328, 2015.
- Granier, C. et al.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period. *Climatic Change* (109), 2011.

- Griffith, D. W., Toon, G. C., Connor, B., Sussmann, R., Warneke, T., Deutscher, N. M., Wennberg, P.O., Notholt, J., Sherlock, V., Robinson, J. and Uchino, O.: Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra, 2011.
- Guan, D., Liu, Z., Geng, Y., Lindner, S. and Hubacek, K.: The gigatonne gap in China's carbon dioxide inventories. *Nature Climate Change*, 2(9), pp.672-675, 2012.
- Guo, S., Hu, M., Zamora, M. L., Peng, J., Shang, D., Zheng, J., Du, Z., Wu, Z., Shao, M., Zeng, L. and Molina, M. J. (2014). Elucidating severe urban haze formation in China. *Proceedings of the National Academy of Sciences*, 111(49), 17373-17378.
- Heald, C. L., Jacob, D. J., Park, R. J., Alexander, B., Fairlie, T. D., Yantosca, R. M., and Chu, D. A.: Transpacific transport of Asian anthropogenic aerosols and its impact on surface air quality in the United States. *Journal of Geophysical Research: Atmospheres*, 111(D14), 2006.
- Hollingsworth, A., Engelen, R. J., Benedetti, A., Dethof, A., Flemming, J., Kaiser, J. W., Morcrette, J.J., Simmons, A.J., Textor, C., Boucher, O. and Chevallier, F.: Toward a monitoring and forecasting system for atmospheric composition: The GEMS project. *Bulletin of the American Meteorological Society*, 89(8), 1147-1164, 2008.
- [Inness, A., Blechschmidt, A. M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., ... & Huijnen, V. \(2015\). Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS. *Atmospheric chemistry and physics*, 15\(9\), 5275-5303.](#)
- Jacob, D. J., Crawford, J. H., Kleb, M. M., Connors, V. S., Bendura, R. J., Raper, J. L., Sachse, G.W., Gille, J.C., Emmons, L. and Heald, C. L.: Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft mission: Design, execution, and first results. *Journal of Geophysical Research: Atmospheres*, 108(D20), 2003.
- Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B. and Thompson, A. M.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results. *Atmospheric Chemistry and Physics*, 10(11), 5191-5212, 2010.
- Kaiser, J.W., Heil, A., Andreae, M.O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.J., Razinger, M., Schultz, M.G., Suttie, M. and Van Der Werf, G.R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9(1), p.527, 2012.
- [Kanaya, Y., Pan, X., Miyakawa, T., Komazaki, Y., Taketani, F., Uno, I., & Kondo, Y. \(2016\). Long-term observations of black carbon mass concentrations at Fukue Island, western Japan, during 2009–2015: constraining wet removal rates and emission strengths from East Asia. *Atmospheric Chemistry and Physics*, 16\(16\), 10689-10705.](#)
- [Kim, S. Y., Jiang, X., Lee, M., Turnipseed, A., Guenther, A., Kim, J. C., ... & Kim, S. \(2013\). Impact of biogenic volatile organic compounds on ozone production at the Taehwa Research Forest near Seoul, South Korea. *Atmospheric environment*, 70, 447-453.](#)
- Kim, S., and Park, R. KORUS-AQ: An International Cooperative Air Quality Field Study in Korea, the KORUS-AQ white paper, 2014.

- Lawrence M. G., Jöckel P., and von Kuhlmann R., What does the global mean OH concentration tell us, *Atmos. Chem. Phys.*, 1, 37–49, 2001.
- Lefer, B. L., Talbot, R. W., Harriss, R. H., Bradshaw, J. D., Sandholm, S. T., Olson, J. O., ... & Klemm, K. I. (1994). Enhancement of acidic gases in biomass burning impacted air masses over Canada. *Journal of Geophysical Research: Atmospheres*, 99(D1), 1721-1737.
- 5 Lelieveld, J., Gromov, S., Pozzer, A., & Taraborrelli, D. (2016). Global tropospheric hydroxyl distribution, budget and reactivity. *Atmospheric Chemistry and Physics*, 16, 12,477–12,493. <https://doi.org/10.5194/acp-16-12477-2016>.
- Maher, B. A., Ahmed, I. A., Karloukovski, V., MacLaren, D. A., Foulds, P. G., Allsop, D., Mann, D. M., Torres-Jardón, R. and Calderon-Garciduenas, L.: Magnetite pollution nanoparticles in the human brain. *Proceedings of the National Academy of Sciences*, 113(39), 10797-10801, 2016.
- 10 Mandrake, L., O'Dell, C. W., Wunch, D., Wennberg, P. O., Fisher, B., Osterman, G. B., & Eldering, A. (2015). Orbiting Carbon Observatory-2 (OCO-2) Warn Level, Bias Correction, and Lite File Product Description. *Jet Propulsion Laboratory: Pasadena, CA, USA*.
- Massart, Flemming J., Cariolle, D., Jones, L.: High resolution CO tracer forecasts, MACC-III Deliverable D022.04, 2015.
- 15 Massart, S., Agustí-Panareda, A., Heymann, J., Buchwitz, M., Chevallier, F., Reuter, M., Hilker, M., Burrows, J. P., Deutscher, N. M., Feist, D. G. and Hase, F.: Ability of the 4-D-Var analysis of the GOSAT BESD XCO₂ retrievals to characterize atmospheric CO₂ at large and synoptic scales. *Atmospheric Chemistry and Physics*, 16(3), 1653-1671, 2016.
- Morcrette, J.J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J.W., Razinger, M. and Schulz, M.: Aerosol analysis and forecast in the European Centre for medium-range weather forecasts integrated forecast system: Forward modeling. *Journal of Geophysical Research: Atmospheres*, 114(D6), 2009.
- 20 Morino, I., Uchino, O., Inoue, M., Yoshida, Y., Yokota, T., Wennberg, P., Toon, G.C., Wunch, D., Roehl, C.M., Notholt, J. and Warneke, T.: Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra, 2011.
- OCO-2 Science Team/Michael Gunson, Annmarie Eldering (2015), OCO-2 Level 2 bias-corrected XCO₂ and other select fields from the full-physics retrieval aggregated as daily files, Retrospective processing V7r, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed 07/2016
- 25 https://disc.gsfc.nasa.gov/datacollection/OCO2_L2_Lite_FP_7r.html.
- Ohara, T. A. H. K., Akimoto, H., Kurokawa, J. I., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020. *Atmospheric Chemistry and Physics*, 7(16), 4419-4444, 2007.
- 30 Parrish, D. D., Trainer, M., Hereid, D., Williams, E. J., Olszyna, K. J., Harley, R. A., Meagher, J. F. and Fehsenfeld, F. C.: Decadal change in carbon monoxide to nitrogen oxide ratio in US vehicular emissions. *Journal of Geophysical Research: Atmospheres*, 107(D12), 2002.
- Persson, A. User Guide to ECMWF forecast products, 2001.

- Rabier F, Jarvinen H, Klinker E, Mahfouf J-F, Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. Q. J. R. Meteorol. Soc. 126, 1143 – 1170, 2000.
- Shindell, D., Faluvegi, G., Walsh, M., Anenberg, S. C., Van Dingenen, R., Muller, N. Z., Austin, J., Koch, D. and Milly, G.:
5 Climate, health, agricultural and economic impacts of tighter vehicle-emission standards. Nature Climate Change, 1(1), 59-66, 2011.
- Shindell, D., Lamarque, J. F., Unger, N., Koch, D., Faluvegi, G., Bauer, S., Ammann, M., Cofala, J. and Teich, H.: Climate forcing and air quality change due to regional emissions reductions by economic sector. Atmospheric Chemistry and Physics, 8(23), 7101-7113, 2008.
- 10 Silva, S. J., Arellano, A. F., and Worden, H.: Toward anthropogenic combustion emission constraints from space-based analysis of urban CO₂/CO sensitivity, Geophys. Res. Lett., 40, 4971–4976, doi:10.1002/grl.50954, 2013.
- Singh, H. B., Brune, W. H., Crawford, J. H., Flocke, F., and Jacob, D. J.: Chemistry and transport of pollution over the Gulf of Mexico and the Pacific: spring 2006 INTEX-B campaign overview and first results. Atmospheric Chemistry and Physics, 9(7), 2301-2318, 2009.
- 15 Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317-9341, doi:10.5194/acp-14-9317-2014, 2014.
- Smith, R.J.: Use and misuse of the reduced major axis for line-fitting. American journal of physical anthropology, 140(3), pp.476-486, 2009.
- 20 Streets, D.G., Bond, T.C., Carmichael, G.R., Fernandes, S.D., Fu, Q., He, D., Klimont, Z., Nelson, S.M., Tsai, N.Y., Wang, M.Q. and Woo, J.H.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000. Journal of Geophysical Research: Atmospheres, 108(D21), 2003.
- Streets, D.G., Zhang, Q., Wang, L., He, K., Hao, J., Wu, Y., Tang, Y. and Carmichael, G.R.: Revisiting China's CO emissions after the transport and chemical evolution over the Pacific (TRACE-P) mission: synthesis of inventories,
25 atmospheric modeling, and observations. Journal of Geophysical Research: Atmospheres, 111(D14), 2006.
- [Tang, W. et al., KORUS-AQ Carbon Monoxide Source Attribution using CAM-chem Tagged Tracers, JGR, 2018, in revision.](#)
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research: Atmospheres, 106(D7), 7183-7192, 2001.
- 30 Toon, O.B., Maring, H., Dibb, J., Ferrare, R., Jacob, D.J., Jensen, E.J., Luo, Z.J., Mace, G.G., Pan, L.L., Pfister, L. and Rosenlof, K.H.: Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) field mission. Journal of Geophysical Research: Atmospheres, 121(9), pp.4967-5009, 2016.

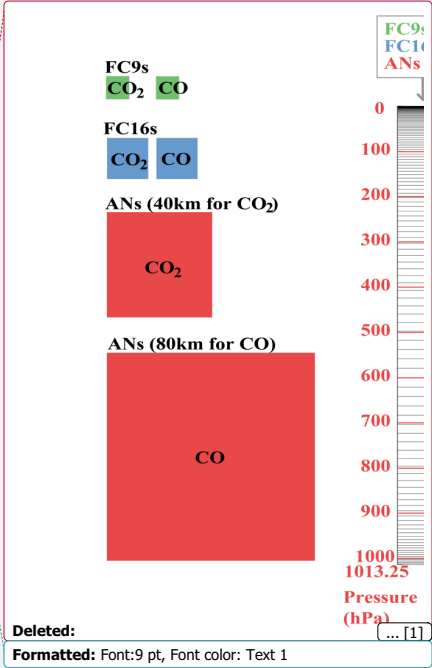
- Turnbull, J. C., Tans, P. P., Lehman, S. J., Baker, D., Conway, T. J., Chung, Y. S., Gregg, J., Miller, J. B., Southon, J. R., and Zhou, L. X.: Atmospheric observations of carbon monoxide and fossil fuel CO₂ emissions from East Asia. *Journal of Geophysical Research: Atmospheres*, 116(D24), 2011.
- United Nations, Department of Economic and Social Affairs, Population Division: The World's Cities in 2016 – Data Booklet (ST/ESA/SER.A/392), 2016.
- 5 Van Leeuwen, T. T., and Van Der Werf, G. R.: Spatial and temporal variability in the ratio of trace gases emitted from biomass burning. *Atmospheric Chemistry and Physics*, 11(8), 3611-3629, 2011.
- Vay, S.A., Woo, J.H., Anderson, B.E., Thornhill, K.L., Blake, D.R., Westberg, D.J., Kiley, C.M., Avery, M.A., Sachse, G.W., Streets, D.G. and Tsutsumi, Y.: Influence of regional-scale anthropogenic emissions on CO₂ distributions over the western North Pacific. *Journal of Geophysical Research: Atmospheres* (1984–2012), 108(D20), 2003.
- 10 Vay, S.A., Choi, Y., Vadrevu, K.P., Blake, D.R., Tyler, S.C., Wisthaler, A., Hecobian, A., Kondo, Y., Diskin, G.S., Sachse, G.W. and Woo, J.H.: Patterns of CO₂ and radiocarbon across high northern latitudes during International Polar Year 2008. *Journal of Geophysical Research: Atmospheres*, 116(D14), 2011.
- Wachter, E. D., Barret, B., Flochmoën, E. L., Pavelin, E., Matricardi, M., Clerbaux, C., Hadji-Lazaro, J., George, M., Hurtmans, D., Coheur, P. F. and Nedelec, P.: Retrieval of MetOp-A/IASI CO profiles and validation with MOZAIC data. *Atmospheric Measurement Techniques*, 5(11), 2843-2857, 2012.
- 15 Wang, Y., Munger, J. W., Xu, S., McElroy, M. B., Hao, J., Nielsen, C. P., and Ma, H.: CO₂ and its correlation with CO at a rural site near Beijing: implications for combustion efficiency in China. *Atmospheric Chemistry and Physics*, 10(18), 8881-8897, 2010.
- 20 Ward, D. E., and Hardy, C. C.: Smoke emissions from wildland fires. *Environment International*, 17(2-3), 117-134, 1991.
- Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P. and Biraud, S. C. A method for evaluating bias in global measurements of CO₂ total columns from space. *Atmospheric Chemistry and Physics*, 11(23), 12317-12337, 2011.
- Warner, J.X., Wei, Z., Strow, L.L., Barnett, C.D., Sparling, L.C., Diskin, G. and Sachse, G.: Improved agreement of AIRS tropospheric carbon monoxide products with other EOS sensors using optimal estimation retrievals. *Atmospheric Chemistry and Physics*, 10(19), pp.9521-9533, 2010.
- 25 [Wofsy, S. C., Sachse, G. W., Gregory, G. L., Blake, D. R., Bradshaw, J. D., Sandholm, S. T., ... & Shipham, M. A. \(1992\). Atmospheric chemistry in the Arctic and Subarctic: Influence of natural fires, industrial emissions, and stratospheric inputs. *Journal of Geophysical Research: Atmospheres*, 97\(D15\), 16731-16746.](#)
- 30 Wunch, D., Toon, G. C., Blavier, J. F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D.W., Sherlock, V. and Wennberg, P. O.: The total carbon column observing network. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369(1943), 2087-2112, 2011.
- Yang, X. J.: China's rapid urbanization. *Science*, 342(6156), 310-310, 2013.

Yarwood, G., Rao, S., Yocke, M., & Whitten, G. (2005). Updates to the carbon bond chemical mechanism: CB05. Final report to the US EPA, EPA Report Number: RT-0400675.

Yokelson, R. J., Andreae, M. O., & Akagi, S. K. (2013). Pitfalls with the use of enhancement ratios or normalized excess mixing ratios measured in plumes to characterize pollution sources and aging. Atmospheric Measurement Techniques, 6, 2155.

Yokota, T., Oguma, H., Morino, I., and Inoue, G.: A nadir looking SWIR FTS to monitor CO2 column density for Japanese GOSAT project, Proc. Twenty-fourth Int. Sympo. on Space Technol. And Sci. (Selected Papers), 887–889, 2004.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H. and Maksyutov, S. Global concentrations of CO2 and CH4 retrieved from GOSAT: First preliminary results. Sola, 5, pp.160-163, 2009.



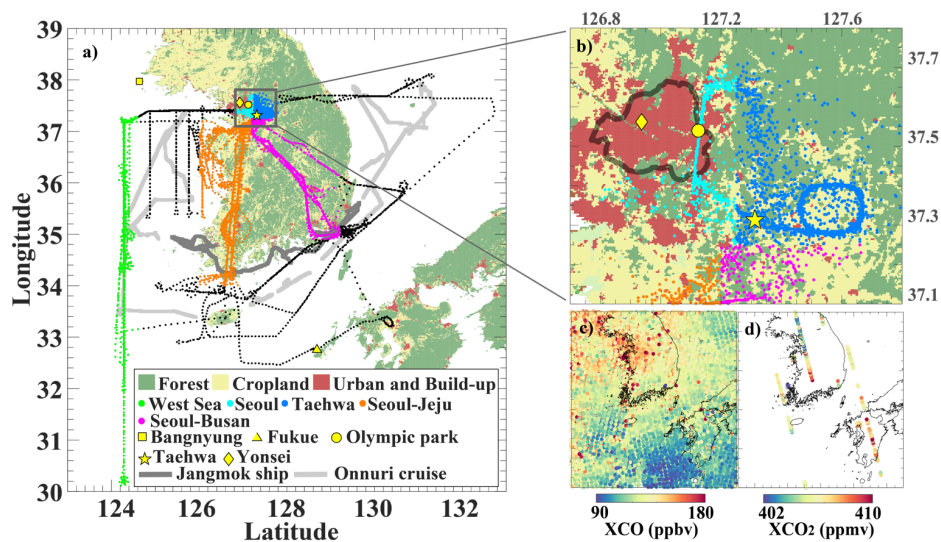


Figure 1. Domain of the study and KORUS-AQ measurements used in this study. Panel (a) shows land cover of the domain (Broxton et al., 2014), DC-8 aircraft tracks, ship tracks, and location of ground sites. The airborne measurements are classified into 5 groups (the West Sea, Seoul, Taehwa, Seoul-Jeju jetway, and Seoul-Busan jetway), as marked in bright green, bright blue, mazarine blue, orange, and magenta. The ground sites are labelled with bright yellow markers. Olympic Park and Yonsei sites are located in urban regions (Seoul) while Baengnyeong and Fukue (Kanaya et al., 2016) site are located in remote regions. Taehwa (Kim et al., 2013) site is located in a forest nearby Seoul. Tracks of the two ships are marked in dark grey (Jangmok ship) and light grey (Onnuri ship). Also shown in (b) is the zoomed-in version of the grey box in panel (a). Panel (c) shows a composite MOPITT XCO retrievals during KORUS-AQ campaign while panel (d) shows OCO-2 XCO₂ retrievals in the same time period.

Deleted: 2

Deleted: luminous

Deleted: luminous

Formatted: Font:9 pt

Deleted: luminous

Deleted: p

Formatted: Font:9 pt

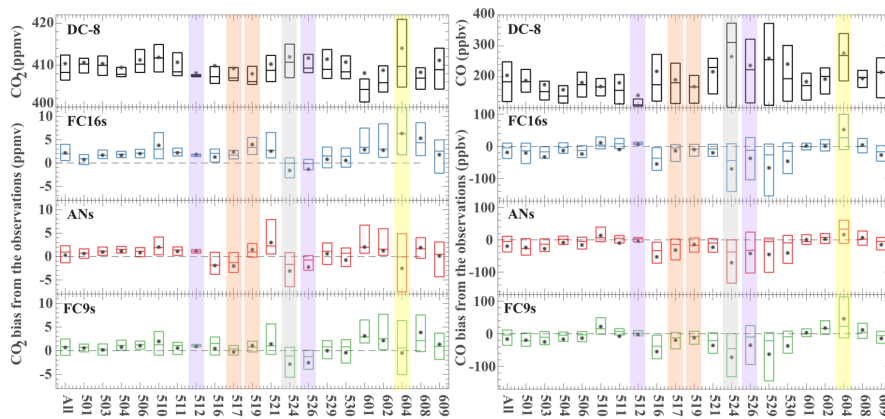


Figure 2. Boxplot for each individual flight. The flight date (MDD) for each boxplot is indicated in the bottom x-axis. Note that the dates here are in UTC time instead of Korea time. The left panel is for CO₂ and the right panel is for CO. The first row corresponds to the boxplot of the abundances measured by DC-8 aircraft. The second, third, and fourth rows correspond to the boxplot of the bias of FC16s, ANs, and FC9s relative to the DC-8 aircraft data, respectively. The purple shade marks the flights with frontal passage, and orange shade marks the flights that may possibly be affected by biomass burning. The grey shade marks the flight measuring China outflow while yellow shade marks the flight surveying point emission sources.

Deleted: 3

Deleted: DC-8

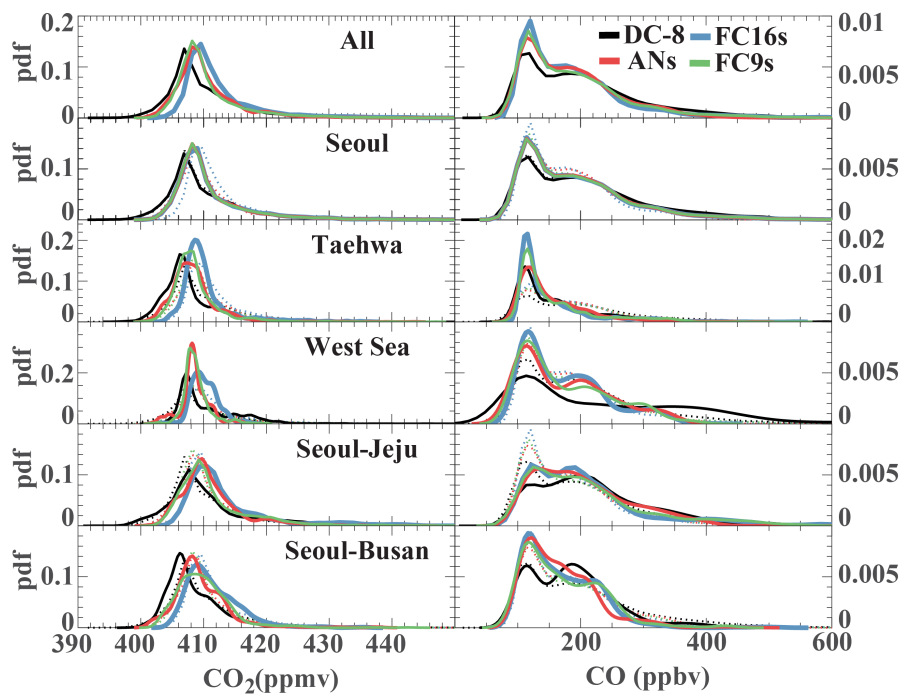


Figure 3. Probability density functions (pdfs) of CO_2 and CO for each flight group. Solid lines are pdfs for each group while the dashed lines are pdfs for all groups.

Deleted: 4

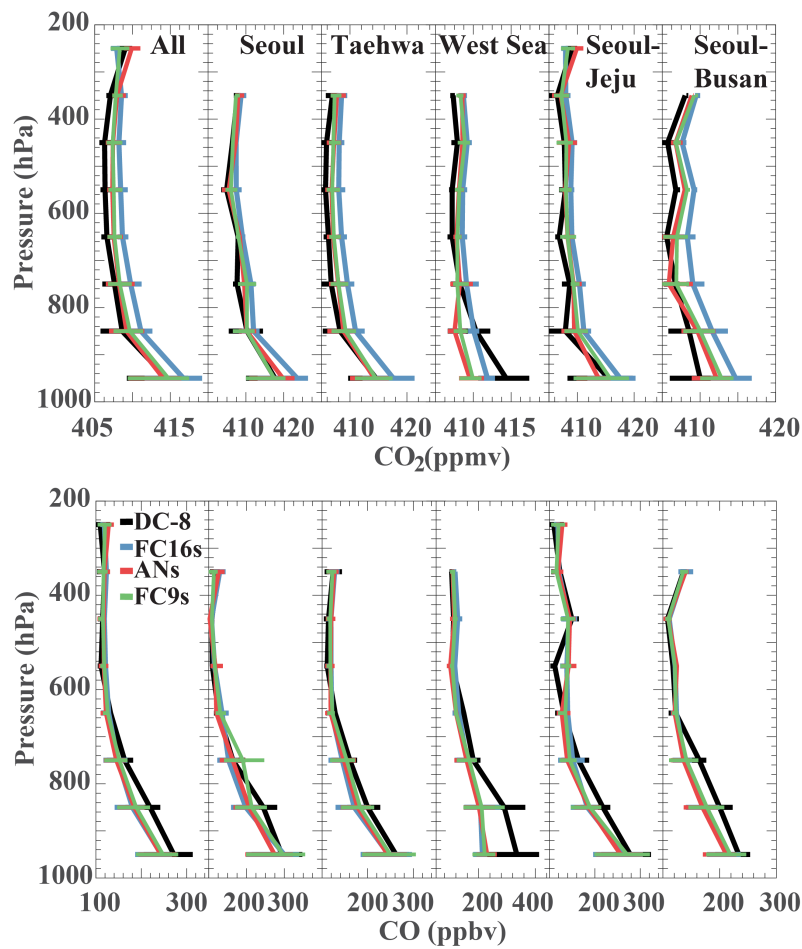


Figure 4. Averaged vertical profiles of CO₂ and CO mixing ratios from the DC-8 aircraft data and CAMS for each flight group. Horizontal bars correspond to the interquartile ranges (between 25th and 75th percentiles) of the layer bin.

Deleted: 5

Deleted: DC-8

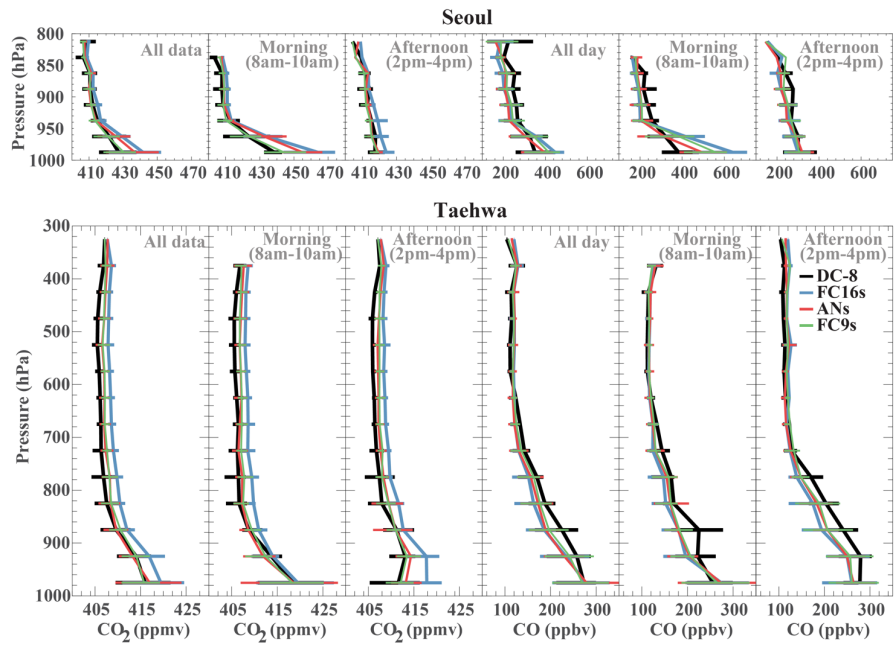


Figure 5. Temporal variation of averaged vertical profiles of CO₂ and CO mixing ratios from the DC-8 aircraft data and CAMS over Seoul and Taehwa flight groups. The first, second, and third columns are averaged CO₂ profiles for all day, morning (8-10am), and afternoon (2-4pm), respectively. Horizontal bars correspond to interquartile ranges (between 25th and 75th percentiles) of the profiles. The fourth, fifth, and sixth column are the same as the first three columns but for CO.

Deleted: 6

Deleted: DC-8

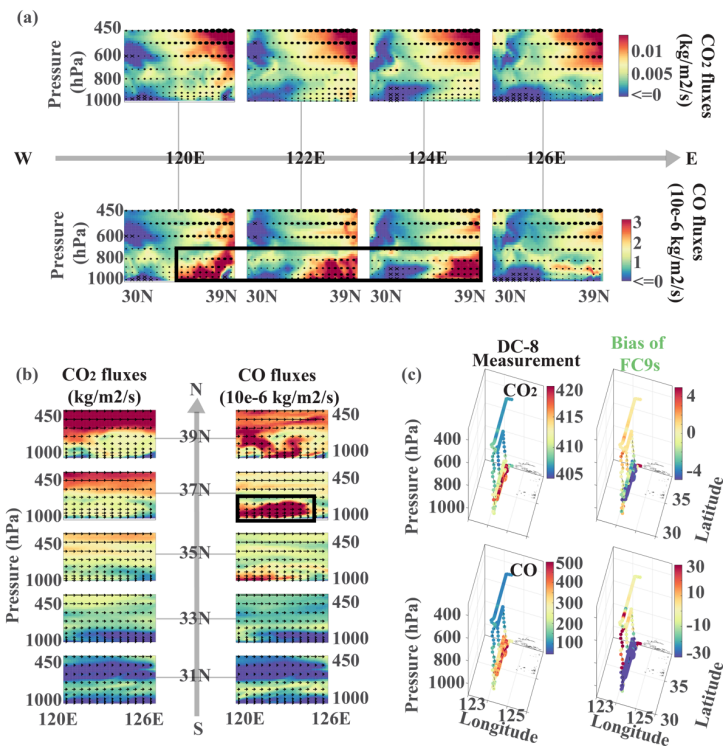


Figure 6. Case study for the flight on May 24th (UTC time). (a) Vertical distributions (hereafter denoted as ‘sections’) of fluxes (kg/m²/s) at 9:00 am on May 25th (Korea time) in meridional direction. Dots represent meridional winds going from west to east (i.e., from China to Korea) and crosses represent meridional winds with the opposite direction. Sizes of the dots and crosses are proportional to the wind speed. ‘Sections’ on the top are for CO₂ fluxes and the bottom are for CO fluxes. (b) ‘Sections’ of fluxes (kg/m²/s) at 9:00 am on May 25th (Korea time) in zonal direction. Arrows represent meridional winds. ‘Sections’ in panel (b) share the same colorbar as panel (a). (c) the DC-8 aircraft measurements (left column) and bias of CAMS along the flight track over the West Sea (right column). The top row is for CO₂ and bottom row is for CO.

Deleted: 7

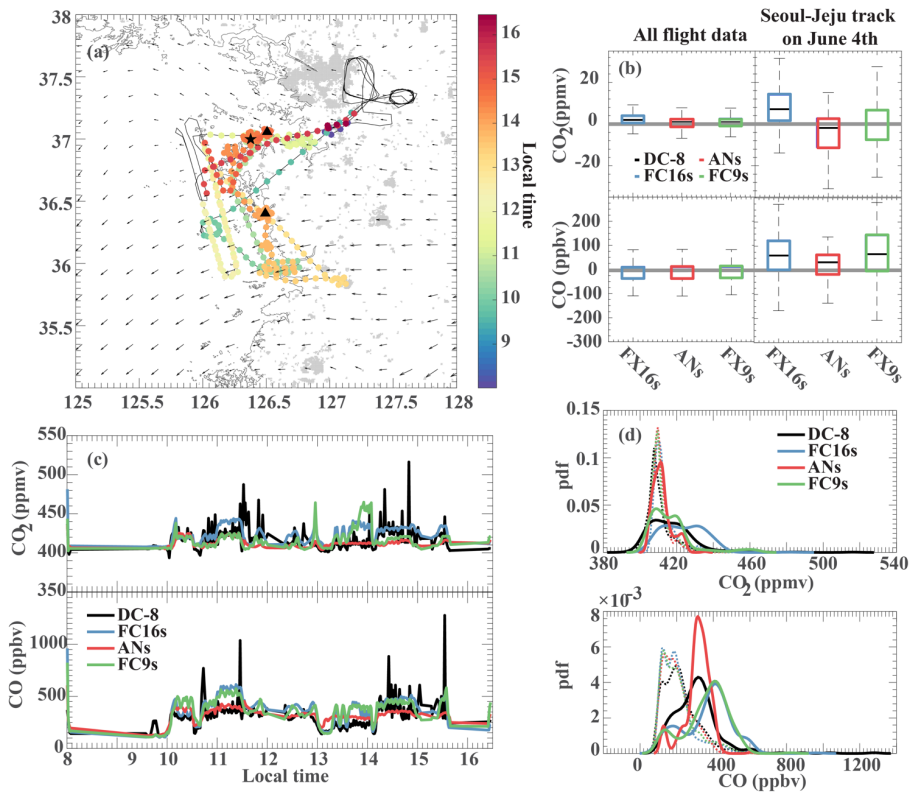


Figure 7. Case study for the flight on June 4th (UTC time). (a) Flight track of DC-8 aircraft in the Seoul-Jeju jetway group for this day. The Daesan chemical facility is marked as black pentagram and two power plants are marked as black triangles. Arrows correspond to 950 hPa wind field at 12:00pm local time. (b) Boxplot of CAMS bias from all the DC-8 aircraft measurements during the campaign (left), and from measurements on June 4th in the Seoul-Jeju jetway group (right). Top row is for CO₂ and bottom row is for CO. (c) Time series of the DC-8 aircraft measurements and CAMS during the flight. (d) pdfs of CO and CO₂ for measurements on June 4th of the Seoul-Jeju jetway group (solid) and for all groups (dashed).

Deleted: 8

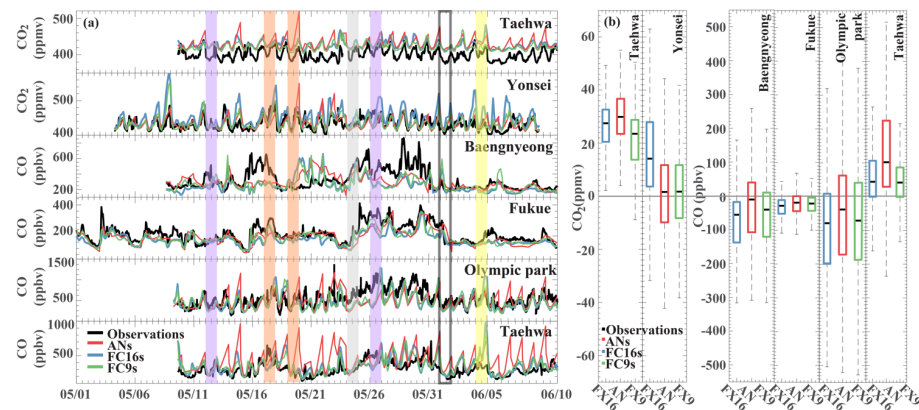
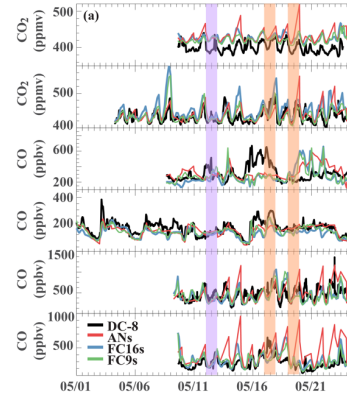


Figure 8. Comparisons of CAMS against ground site measurements. Values of CAMS are averages across layers with pressure higher than 95% of the surface pressure. (a) Time series of measured and CAMS CO₂ from the Taehwa and Yonsei sites, and CO from the Baengnyeong, Fukue, Olympic Park, and Taehwa sites. Shades denote same events as they do in Fig. 2. (b) Boxplot of CAMS bias for CO₂ at the Taehwa and Yonsei site measurements, and for CO at the Baengnyeong, Fukue, Olympic Park, and Taehwa sites.



Deleted:

Formatted: Font:9 pt, Bold, Font color: Text 1

Deleted: 9

Deleted: p

Deleted: 3

Deleted: p

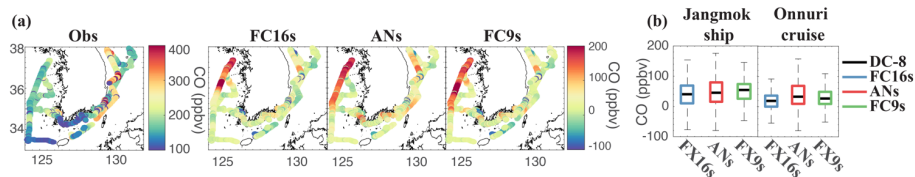


Figure 2. Comparisons of CAMS CO against ship measurements. Values of CAMS are averages across layers with pressure higher than 95% of the surface pressure. (a) Bias of CAMS CO against ship measurements along the ship track. (b) Boxplot of CAMS bias for CO compared with ship measurements.

Deleted: 10

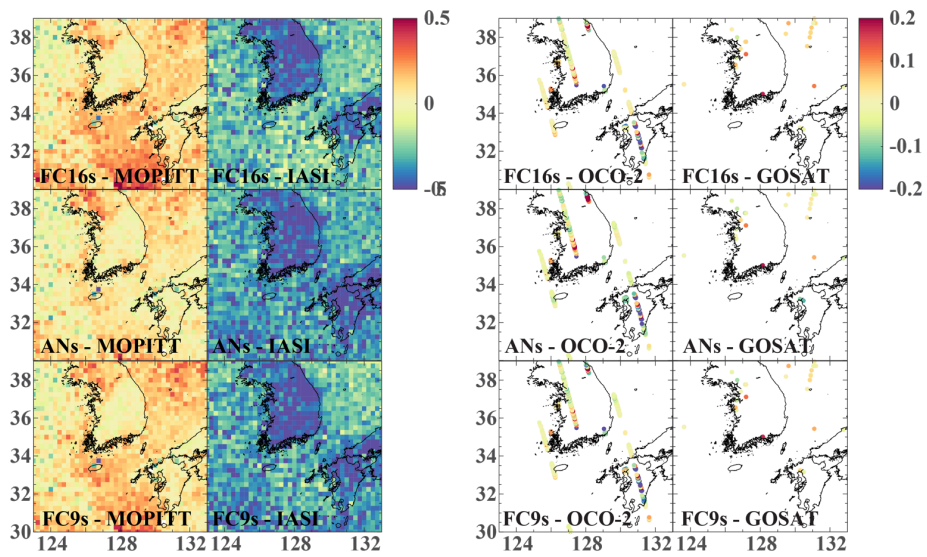


Figure 10. Spatial distributions of CAMS bias against satellite retrievals. For XCO_2 , the unit is 10^{18} molecules/ cm^2 while for XCO , the unit is 10^{21} molecules/ cm^2 .

Deleted: 1

5

10

15

Table 1. Configuration of CAMS global atmospheric composition products valid during the period of the KORUS-AQ Field Campaign (May to June 2016). The tracers evaluated in this paper are highlighted in bold face. Time availability is in number of days with respect to real time (N/A is used when this is not applicable).

CAMS product	Atmospheric composition tracers	Horizontal resolution	Number vertical levels	Initial conditions: Meteorology	Initial conditions: Atmospheric composition	Time availability observations/analysis of atmospheric composition	Time availability of product
AN_CHEM	Reactive gases (CO ,O3,NO ₂ ,etc) and aerosols	80 km	L60	Own analysis	Own analysis	<1day	<1day
FC_CHEM	Reactive gases (CO ,O3,NO ₂ ,etc) and aerosols	80 km	L60	AN_CHEM	AN_CHEM	<1day	0 days (real time)
AN_GHG	CO₂ , CH ₄	40 km	L137	Own analysis	Own analysis	2-4 days	4 days
FC16s	CO₂ , CH ₄ and linCO	16 km	L137	ECMWF operational analysis	Previous 1-day forecast	N/A	1 day
FC9s	CO₂ , CH ₄ , linCO and tagged tracers	9 km	L137	ECMWF operational analysis	AN_GHG 4-day fc for CO ₂ /CH ₄ and AN_CHEM for linCO	4 day for AN_GHG; <1 day for AN_CHEM	1 day

10

15

20

Table 2. Measurements during KORUS-AQ.

			CO ₂	CO
Airborne measurements	NASA DC-8 aircraft	Instrument	LI-COR	DACOM
		Time Response	1 second	1 second
		Precision	< 0.1 ppmv	< 1% or 0.1 ppbv
		Accuracy	0.25 ppmv (Vay et al., 2003)	2% (Warner et al., 2010)
Ground site measurements	Baengnyeong (37.97N,124.63E)	Instrument	/	Teledyne Gas analyzer
		Data intervals	/	1 hour
		Instrument	/	<u>Thermo</u> 48C
		Data intervals	/	1 hour
	Fukue (32.75N,128.68E)	Instrument	/	KENTEK CO analyzer
		Data intervals	/	5 minutes
	Olympic Park (37.52N,127.12E)	Instrument	/	
		Data intervals	/	
		Instrument	LI-COR LI-7500	Thermo 48i
		Data intervals	1 hour	1 hour
	Taehwa (37.31N,127.31E)	Instrument	G2201-I CO ₂ /CH ₄ carbon stable isotope analyzer	/
		Data intervals	30 minutes	/
Ship measurements	R/V Jangmok	Instrument	/	Thermo 48i-TLE
		Data intervals	/	1 minute
	R/V Onnuri	Instrument	/	Thermo Scientific, Inc., Model 48C
		Data intervals	/	1 minute
Satellite measurements	OCO-2	Date product	OCO-2 Level 2 v7 Full Product	/
		Resolution	XCO ₂ 2.25x1.29-km	/
		Global coverage	~16 days	/
		Revisit time	1:18 - 1:33 pm	/
		Uncertainty	1-2 ppm XCO ₂ (<u>Wunch et al. 2017; Osterman et al., 2015</u>)	/
		Date product	Level 2 V02	/
	GOSAT	Date product		

Formatted: Font:9 pt

Deleted: Boesch et al., 2011 and references therein

	MOPITT	Resolution	10.5 x 10.5 km ~12 days	/
		Revisit time	~1:00 pm	/
		Uncertainty	2 ppm for retrieval errors of XCO ₂ Griffith et al. 2011; Crisp et al. 2012	/
		Date product	/	TIR/NIR Level 2 v6 XCO
	IASI	Resolution	/	22 x 22 km ~3-4 days 10:30 am
		Revisit time	/	0.09e18 molecules/cm ² for total column retrieval; (Deeter et al., 2014)
		Uncertainty	/	Level 2 FORLI XCO
		Date product	/	12 km x 12 km twice a day
		Resolution	/	
		Revisit time	/	
		Uncertainty	/	<13% for FORLI (Wachter et al., 2012)

5

10

15

Table 3. Enhancement ratios of CO to CO₂ (ppbv/ppmv), CO and CO₂ correlations, and bias of CO to bias of CO₂ correlations from airborne measurements, CAMS FC16s, ANs, and FC9s.

		Seoul	Tachwa	the West Sea	Seoul-Jeju jetway	Seoul-Busan jetway	All
$\Delta\text{CO}/\Delta\text{CO}_2$ (ppbv/ppmv)	DC-8	9.09±0.48	15.3±0.56	28.17±0.75	10.37±0.31	15.86±0.73	13.29±0.2
	measurement						1
	FC16s	9.84±0.29	14.31±0.40	30.86±1.64	13.00±0.27	13.39±0.51	12.28±0.1
							5
	ANs	8.21±0.45	13.71±0.48	30.60±1.73	14.98±0.45	12.68±0.47	12.60±0.2
	FC9s	11.56±0.62	16.06±0.57	32.44±1.77	11.68±0.35	13.87±0.54	12.52±0.2
Correlation of CO and CO ₂	DC-8	0.78	0.68	0.89	0.62	0.60	0.66
	measurement						
	FC16s	0.94	0.83	0.42	0.83	0.74	0.82
	ANs	0.77	0.71	0.25	0.61	0.76	0.63
	FC9s	0.78	0.70	0.36	0.60	0.73	0.65
Correlation of Bias _{CO} and Bias _{CO₂}	FC16s	0.90	0.61	0.80	0.46	0.55	0.61
	ANs	0.66	0.59	0.82	0.36	0.63	0.51
	FC9s	0.64	0.52	0.82	0.33	0.54	0.49

- Formatted: Subscript
- Formatted: Font:9 pt
- Formatted: Font:9 pt, Subscript
- Formatted: Font:9 pt
- Formatted: Font:9 pt, Subscript
- Formatted: Font:9 pt
- Deleted: Statistics of CAMS performance evaluated against satellite observations.
- Formatted Table
- Deleted: Δ
- Deleted: Δ

Table 4. Statistics of CAMS performance compared against satellite observations.

		CO		CO ₂	
		MOPITT	IASI	OCO-2	GOSAT
total observations		13612	25509	4591	42
during campaign					
bias (molecules/cm ²)	FC16s	-1.13E+17	8.28E+16	9.30E+18	-2.64E+19
	ANs	-6.42E+16	1.36E+17	4.48E+19	1.05E+19
	FC9s	-1.01E+17	7.52E+16	-1.31E+19	-1.28E+19
RMSE (molecules/cm ²)	FC16s	2.47E+17	4.19E+17	7.11E+19	5.67E+19
	ANs	2.31E+17	4.12E+17	8.48E+19	6.42E+19
	FC9s	2.56E+17	4.19E+17	8.29E+19	5.49E+19
correlation	FC16s	0.65	0.44	0.88	0.78
	ANs	0.66	0.52	0.85	0.63
	FC9s	0.61	0.45	0.85	0.75

Formatted Table

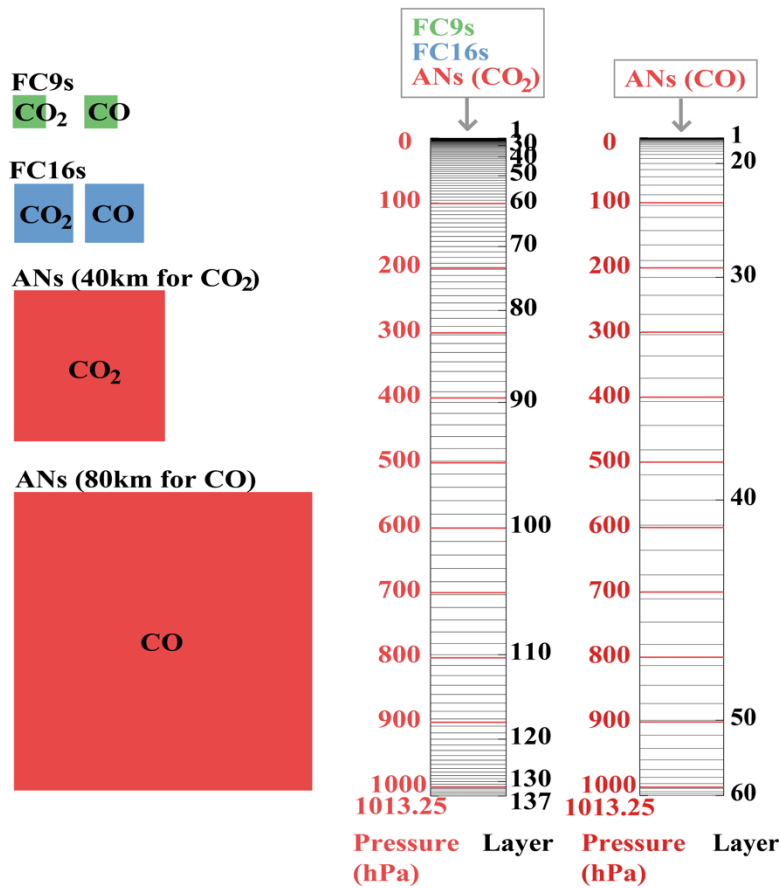


Figure 1. Model grid sizes of the CAMS and vertical structures of the model layers assuming the surface pressure being 1013.25hPa. FC9s, FC16s, and ANs for CO_2 (40 km) have 137 vertical layers. ANs for CO (80 km) have 60 vertical layers.

Evaluating High-Resolution Forecasts of Atmospheric CO and CO₂ from a Global Prediction System during KORUS-AQ Field Campaign

Wenfu Tang et al., 2017

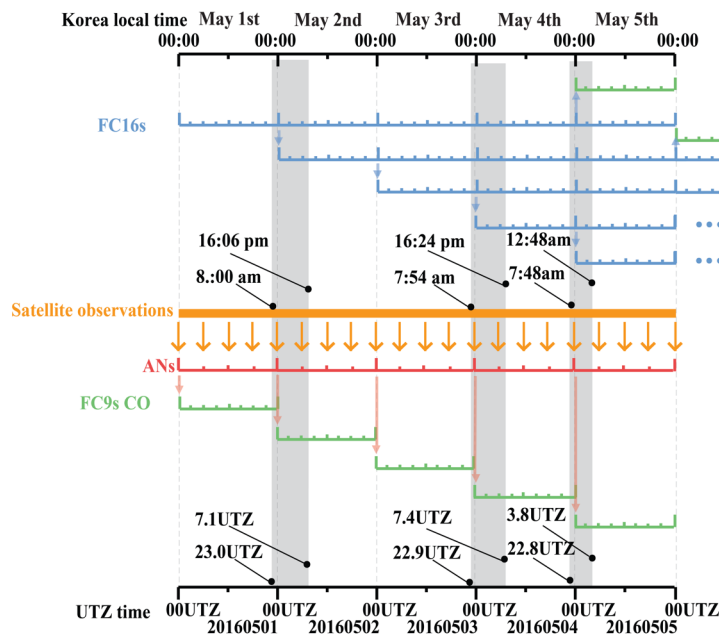
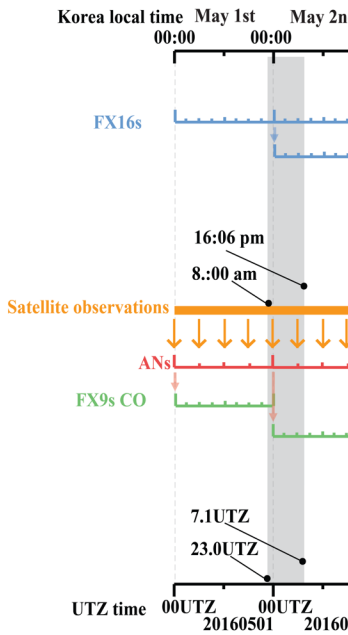


Figure S1. CAMS configuration. Left panel corresponds to the time configuration of the CAMS CO and CO₂ evaluated in this study. The black lines represent Korea local time (on the top) and UTC time (on the bottom). The blue lines represent CAMS 5-day FC16s. FC16s are initialized with forecasts from the previous day. The orange line represents satellite observations (i.e., CO from MOPITT and IASI, CO₂ from GOSAT) assimilated in CAMS (ANs). Gray shade denotes campaign time of the DC-8 aircraft. A typical DC-8 flight starts at 8am Korea time (23 UTC of previous day) and ends at 4pm Korea time (7 UTC).



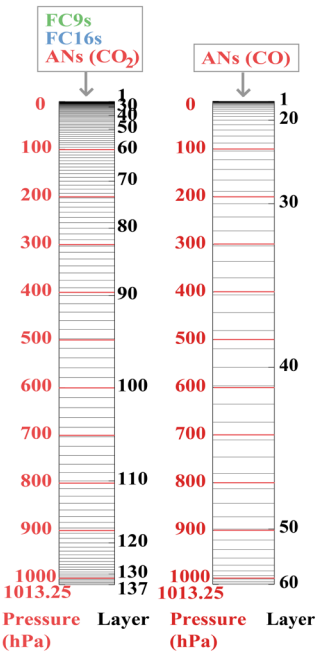
Deleted:
Formatted: Font:(Default) Times New Roman, (Asian)
Times New Roman, 9 pt
Deleted: 1

FC9s
CO₂ CO

FC16s
CO₂ CO

ANs (40km for CO₂)
CO₂

ANs (80km for CO)
CO



Formatted: Font:9 pt, Font color: Text 1

Figure S2. Model grid sizes of the CAMS and vertical structures of the model layers assuming the surface pressure being 1013.25hPa. FC9s, FC16s, and ANs for CO₂ (40 km) have 137 vertical layers. ANs for CO (80 km) have 60 vertical layers.

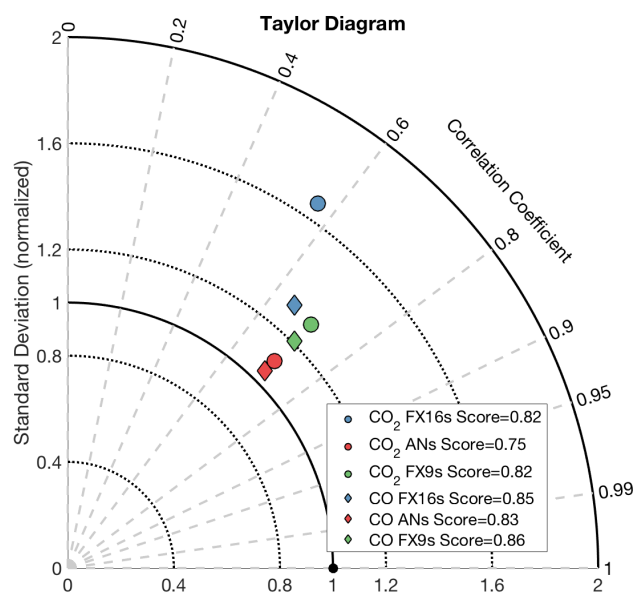


Figure S3. Taylor diagram for CAMS CO (diamonds) and CO₂ (circles) from FC9s (green), FC16s (blue), and ANs (red). Also shown are the Taylor scores.

Deleted: 2

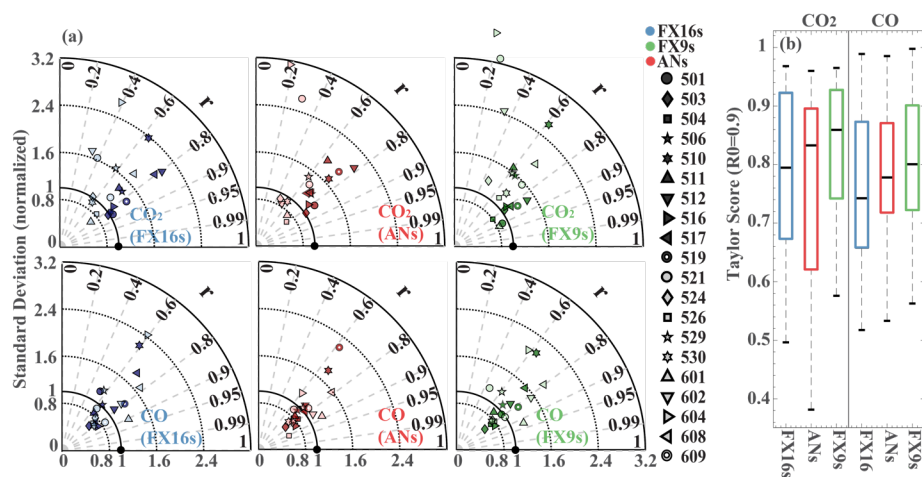


Figure S4. (a) Taylor diagrams for CAMS CO_2 (first row) and CO (second row) from 16-km forecasts (FC16s, left column), analyses (ANs, middle column), and 9-km forecasts (FC9s, right column) for individual flights (different symbols). (b) Boxplot of Taylor scores for CAMS CO_2 (left panel) and CO (right panel) from FC16s (blue), ANs (red), and FC9s (green).

Deleted: 3

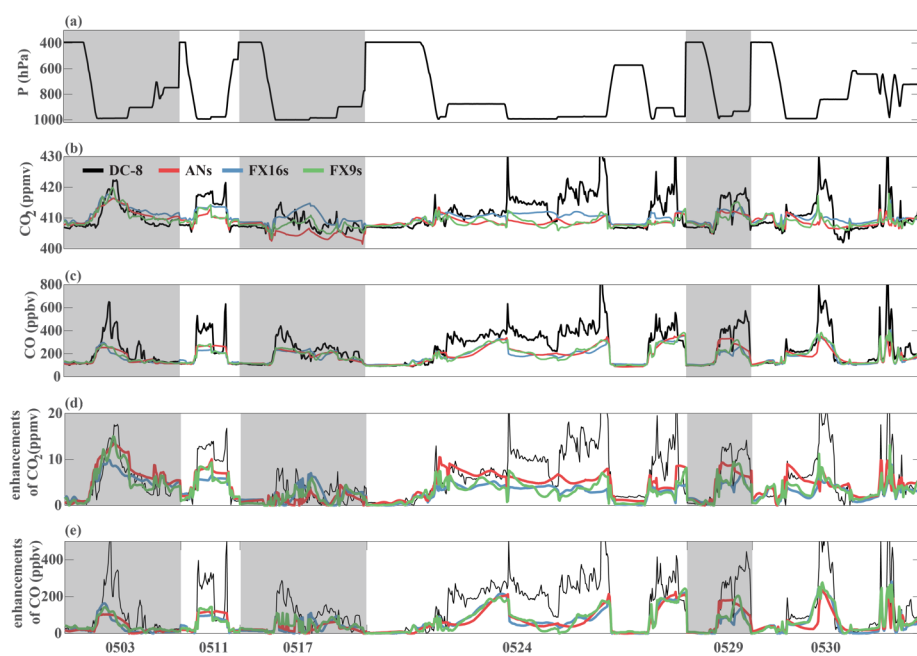


Figure S5. Time series of (a) pressure levels, (b) CO_2 concentrations and (d) their enhancements relative to background values, (c) CO concentrations and (e) their enhancements relative to background values along DC-8 aircraft tracks over the West Sea from measurements (black), 16-km forecasts (FC16s, blue), analyses (ANs, red), and 9-km forecasts (FC9s, green).

Deleted: 4

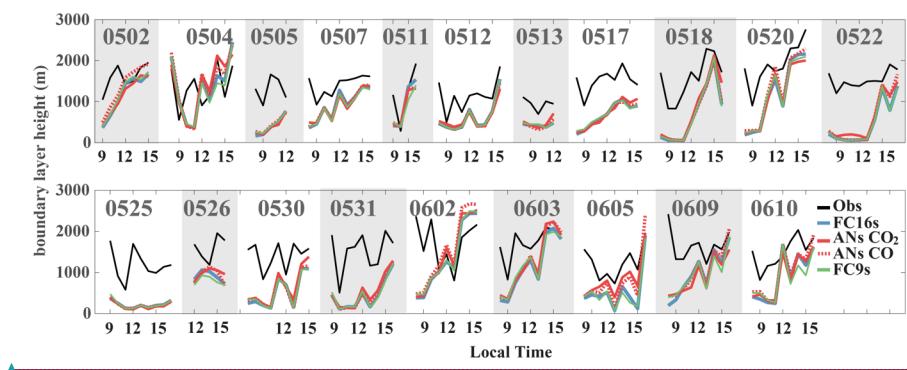


Figure S6. Hourly-average time series of the mixing layer heights derived from the airborne DIAL-HSRL measurements of aerosol backscatter (black) and corresponding boundary layer heights from the four CAMS configurations (colored) along the DC-8 flight track during KORUS-AQ.

Formatted: Font:(Default) Times New Roman, (Asian)
Times New Roman, 9 pt

Formatted: Font:Bold

Formatted: Justified

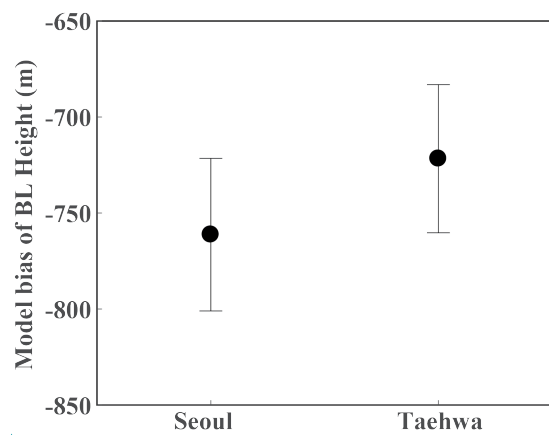


Figure S7. Model bias of boundary layer heights against boundary layer heights derived from the airborne DIAL-HSRL measurements of aerosol backscatter. Error bars represent standard deviations among the four CAMS configurations (i.e., FC16s, 80km ANs for CO, 40km ANS for CO₂, and FC9s).

Formatted: Font:(Default) Times New Roman, (Asian)
Times New Roman, 9 pt

Formatted: Subscript

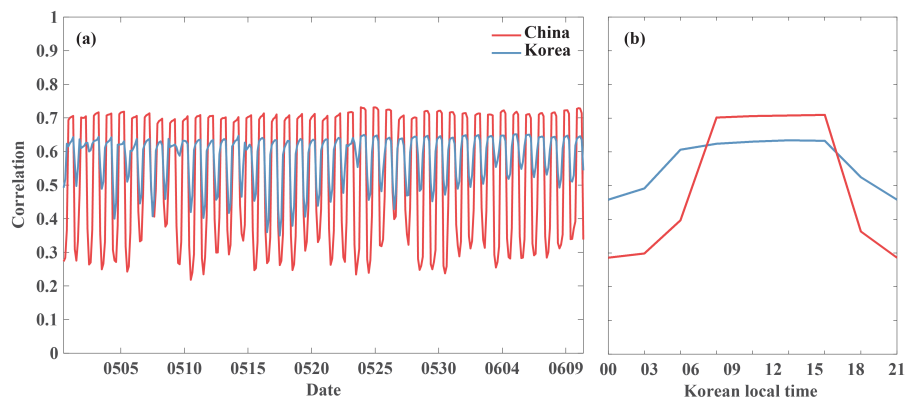


Figure S8. (a) Time series of spatial correlations between CO emissions and CO2 fluxes in CAMS over East China (which dominates Chinese contribution to the West Sea (Tang et al., 2018)) and Korea. (b) Averaged diurnal cycle of spatial correlations between CO emissions and CO2 fluxes in CAMS over East China and Korea.

Formatted: Font:(Default) Times New Roman, (Asian)
Times New Roman, 9 pt

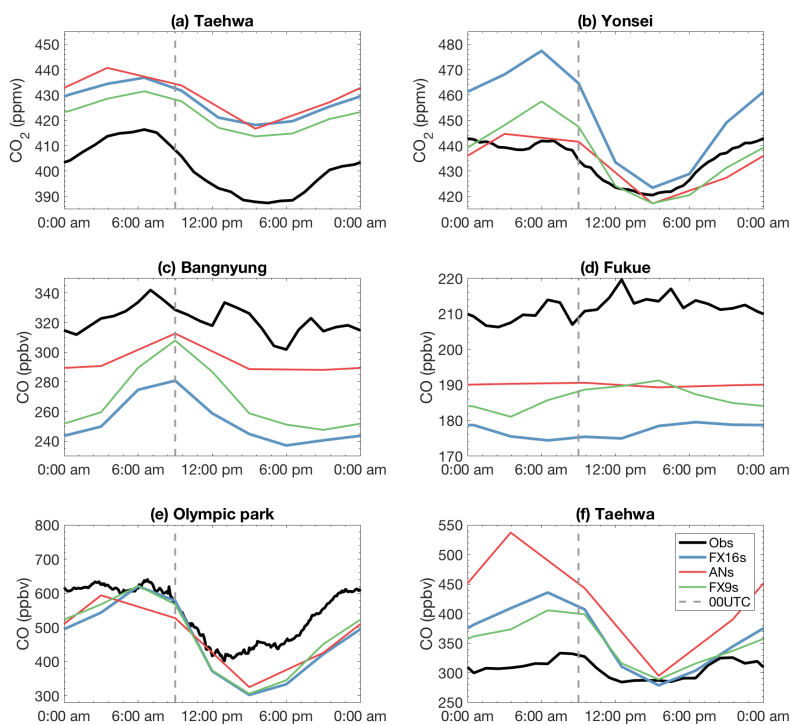


Figure S9. Diurnal cycles of CO₂ (a–b) and CO (c–f) concentrations averaged over days with available data during the KORUS-AQ period from observations (black), 16-km forecasts (FX16s, blue), analyses (ANs, red), and 9-km forecasts (FX9s, green) at fix ground sites, including (a, f) Taehwa, (b) Yonsei, (c) Bangnyung, (d) Fukue, and (e) Olympic park. CAMS values are averages across layers with pressure higher than 95% of the surface pressure.

Deleted: 5

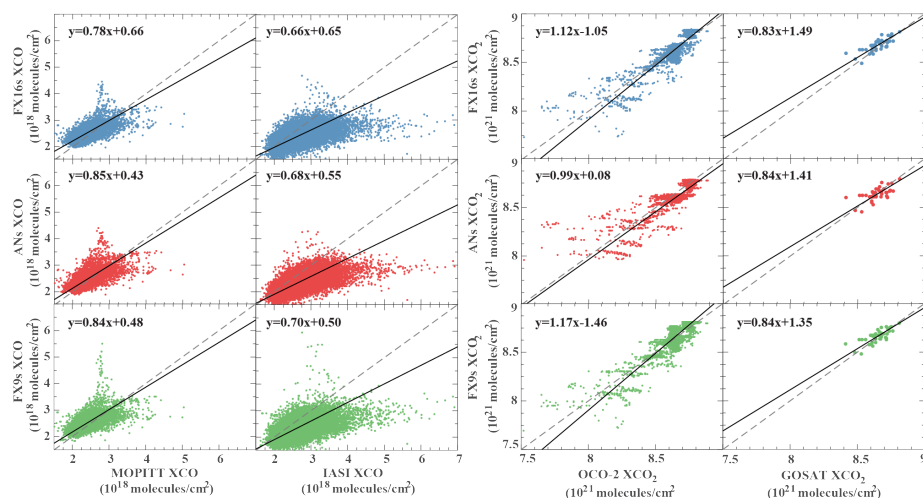


Figure S10. Comparisons between satellite observations and CAMS XCO and XCO₂ from 16-km forecasts (FC16s, blue, 1st row), analyses (ANs, red, 2nd row), and 9-km forecasts (FC9s, green, 3rd row). The columns from left to right correspond to MOPITT XCO, IASI XCO, OCO-2 XCO₂, and GOSAT XCO₂, respectively.

Deleted: 6

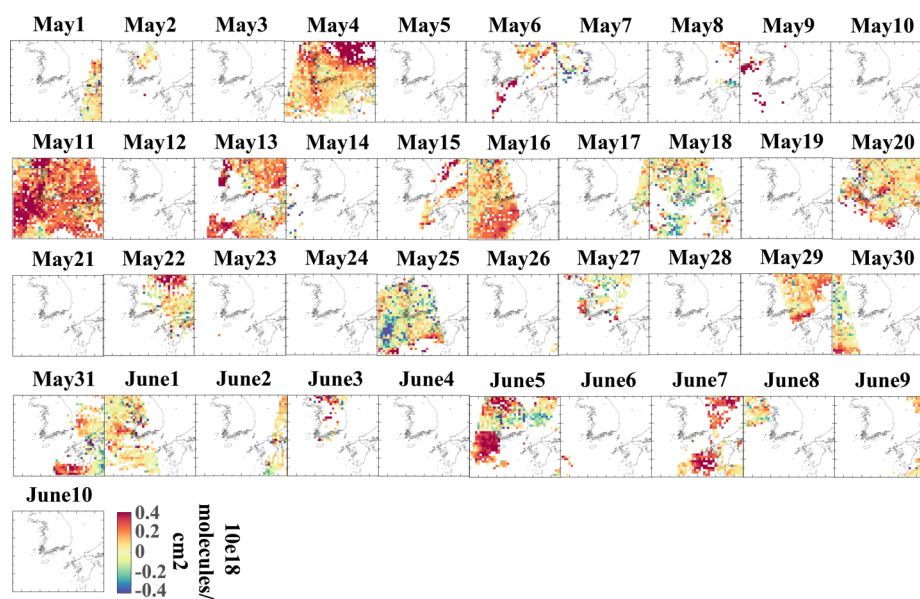


Figure S11. Daily spatial distributions of XCO biases in CAMS 9-km forecasts compared with MOPITT observations during the KORUS-AQ period (May 1 to June 10, 2016).

Deleted: 7