

Validation of regional CO₂ concentrations in the ECMWF real-time analysis and Carbon-Tracker reanalysis with airborne observations from ACT-A field campaign

Hans W. Chen¹, Lily N. Zhang^{1,2}, Fuqing Zhang¹, Kenneth J. Davis¹, Thomas Luavaux¹,
Sandip Pal¹, and Joshua P. Digangi³

¹Department of Meteorology and Atmospheric Science, and Center for Advanced Data
Assimilation and Predictability Techniques, The Pennsylvania State University,
University Park, Pennsylvania, USA

²State College Area High School, State College, Pennsylvania, USA

³NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Corresponding author: Fuqing Zhang (fzhang@psu.edu)

To be Submitted to *Atmospheric Chemistry and Physics* for Publications

DRAFT DRAFT DRAFT

March 18, 2018

Abstract

Through verifying against hundreds of hours of airborne in-situ measurements from the NASA-sponsored Atmospheric Carbon and Transport – America (ACT-A) field campaign, this study systematically examines the regional uncertainties and biases of the carbon dioxide (CO₂) concentrations from two of the state-of-the-art global analysis products, namely the real-time analysis from the European Center (EC) for Medium Range Forecasting and NOAA’s near real-time Carbon Tracker (CT) reanalysis. It is found that both the EC and CT-NRT analyses agree reasonably well with the independent ACT-A flight-level CO₂ measurements in the free troposphere but the uncertainties are considerably larger in the boundary layer during both the summer months of 2016 and the winter months of 2017. There are also strong variabilities in accuracy and bias between seasons, and across three different subregions in the United States (Mid-Atlantic, Midwest and South). Overall, the analysis uncertainties of the EC and CT-NRT analyses in terms of root-mean square deviations against airborne data are comparable to each other, both of which are between 1-2 ppm in the free troposphere but can be as large as 10 ppm near the surface, which are grossly consistent with the difference between the two analyses. The current study not only provides systematic uncertainty estimates for both analysis products over North America but also demonstrated that these two independent estimates can be used to approximate the overall regional CO₂ analysis uncertainties. Both statistics are important in future studies in quantifying the uncertainties of regional carbon concentration and flux estimates, as well as in assessing the impact of regional transport through more refined regional modeling and analysis systems.

1. Introduction

The capabilities to monitor atmospheric carbon dioxide (CO₂) using in-situ and remote sensing observations combined with numerical models have rapidly evolved to meet the needs for more accurate climate projections, and to provide independent estimates of anthropogenic CO₂ emissions. Changes in atmospheric CO₂ can be used to infer uptake and release of CO₂ from land ecosystems and oceans through inversion methods, which in turn can help us understand how the natural carbon cycle responds to both natural and human-induced environmental changes, including climate disturbances such as droughts, increasing levels of CO₂ in the atmosphere, thawing permafrost in the Arctic, and human land-use management. Moreover, applied at regional to urban scales, inversions may become critical tools in the future to support policies aimed at limiting greenhouse gas emissions by providing independent data-driven checks on emissions accounting. To support these efforts, a wide variety of inversion systems have been developed over the past decade. Two state-of-the-art global near-real time CO₂ analyses are the European Center (EC) for Medium Range Forecasting real-time analysis and the National Oceanic and Atmospheric Administration (NOAA) CarbonTracker Near-Real Time (CT-NRT) reanalysis (both of which will be described in section 2). These analyses optimize the CO₂ state and CO₂ fluxes by assimilating CO₂ observations, and provide gridded estimates of global distributions of atmospheric CO₂ and CO₂ fluxes which are useful for analyzing changes in fluxes and as driver data for e.g. regional inversions.

The variability of CO₂ in the atmosphere depends on both the variability of CO₂ surface fluxes and the atmospheric transport. CO₂ is a relatively homogeneous trace gas

with small signals on the order of a few ppm, which imposes more stringent requirements for accurate modeling of the atmospheric transport compared with many other variables. Furthermore, CO₂ is a relatively long-lived gas, so small systematic transport errors can accumulate and have a large impact on the annual CO₂ budget. On the other hand, CO₂ surface fluxes are spatially heterogeneous and in-situ observations of CO₂ fluxes are typically representative of only a small area. Process-based vegetation models show a wide diversity of results on the regional scale, which reflect the large uncertainties in our current understanding of regional-scale biological CO₂ fluxes. It is therefore imperative to evaluate modeled CO₂ estimates against independent observations to detect systematic errors in atmospheric transport and CO₂ fluxes.

In-situ observations of CO₂ from instruments such as tall towers can provide long-term measurements of CO₂ at high precision (typically on the order of 0.01 ppm), but lack the spatial coverage to accurately capture spatial gradients in CO₂ created by varying weather systems. The new generation of satellites in sun-synchronous, polar low-Earth orbits that focus on CO₂ (GOSAT and OCO-2) provide column-average CO₂ measurements over the whole globe, but suffer in other areas such as long revisit times, lack of data in cloudy conditions and at high latitudes where the sunlight is insufficient, and no information about the vertical distribution of CO₂. Additional regular observations include surface and aircraft flasks and aircraft in-situ profiles, which also lack the spatial and temporal coverage to fully characterize the flux and transport uncertainties of CO₂.

Atmospheric Carbon and Transport – America (ACT-A) is a NASA-sponsored project that aims at filling the spatial gap in CO₂ observations by conducting five airborne campaigns over three regions spanning the eastern United States. The overall goal of ACT-

A is to improve the transport of CO₂ to obtain more accurate CO₂ flux estimates from inversions. During each intensive ACT-A field campaign, two weeks are spent in each region to sample the CO₂ spatial variability in fair weather conditions and across weather systems from two coordinated aircraft. The aircraft take in-situ measurements of CO₂ and other trace gases and meteorological variables along pre-designed flight patterns. The flight patterns include both long horizontal legs in the boundary layer and free troposphere, as well as several vertical profiles, and were designed to capture CO₂ where there are large uncertainties in transport and/or fluxes. Because of the large spatial extent and high spatial resolution of the flight measurements, ACT-A observations provide an ideal verification dataset to evaluate current inversion systems.

In this study, we provide a first evaluation of the EC analysis and CT-NRT reanalysis of CO₂ against ACT-A airborne measurements from the summer 2016 and winter 2017 field campaigns. We present comprehensive summary statistics from comparisons with hundreds of hours of airborne in-situ measurements to assess how reliable are the global re(analysis). Furthermore, we compare the model-observation mismatches with the differences between the two (re)analyses to investigate if the difference between the two independent CO₂ estimates can be used to quantify the uncertainties in the estimates. This study is unique because it provides the first inter-comparison with both the EC analysis and CT-NRT reanalysis against independent in-situ observations of CO₂ with high spatial coverage. Ongoing and subsequent studies to be reported elsewhere will further examine the transport uncertainties associated with different weather systems such as fronts and cyclones. The paper is structured as follows.

Section 2 describes the EC analysis, CT-NRT reanalysis, and ACT-A field campaigns. The inter-comparison results are presented in Section 3, and Section 4 concludes the paper.

2. Data and methodology

2.1 ECMWF analysis and Carbon Tracker reanalysis

The EC global analysis of CO₂ concentration (Massart et al. 2016) is generated by assimilating the XCO₂ products from the Japanese Greenhouse gases Observing Satellite (GOSAT; Kuze et. al, 2009) into the CO₂ forecast model (Agustí-Panareda et al. 2014) based on the ECMWF four-dimensional variational (4DVar) system (Engelen et al. 2009). The GOSAT satellite has a revisiting frequency of two weeks. Both the forecast model and the data assimilation system are based on those used in the Integrated Forecasting System (IFS), the leading operational global weather prediction system by the European Centre for Medium Range Weather Forecasts (ECMWF). A cycling system with a 12-h assimilation window is used for the CO₂ analysis with the background estimate derived from the short-term forecast initialized from the previous analysis cycle while the meteorology initial conditions at each forecast cycle comes from the ECMWF operational analyses. The real-time analysis and forecasts are available on a Gaussian grid of the ECMWF global model (~16 km x 16 km and 91 vertical levels). The surface fluxes from the terrestrial biogenic emissions in this carbon analysis and forecast modeling system at ECMWF are directly modeled but the other sources and sinks of CO₂ surface fluxes are prescribed from different inventory sources. These fluxes are not directly updated by the observations assimilated. The details of the EC analysis products are available online at http://www.gmes-atmosphere.eu/news/co2_forecasts, and the data assimilation system is described in Massart et al. (2016).

NOAA's CarbonTracker system assimilates in-situ CO₂ observations to optimize CO₂ fluxes using an Ensemble Kalman Filter system. In this study we used CarbonTracker Near-Real Time (CT-NRT) v2017, which is an extension of the CarbonTracker system to provide more timely CO₂ analyses. One difference from the standard CarbonTracker system is that CT-NRT uses a statistical prior flux model because the terrestrial biosphere model used by CarbonTracker, the Carnegie-Ames Stanford Approach (CASA) biogeochemical model, is not available in near-real time. Furthermore, for the CT-NRT product fewer observations are assimilated with less quality control. Observations assimilated by CT-NRT include in-situ tower, flask, and aircraft measurements. The atmospheric transport is simulated using the TM5 offline global chemical transport model driven by meteorology from the ERA-Interim reanalysis. TM5 is run globally at a resolution of 3° longitude × 2° latitude horizontal resolution and 25 vertical layers, and in a nested grid over North America at 1° × 1° resolution. The results presented in this study are from the regional nested high-resolution grid available as 3-hourly averages. In CarbonTracker and CT-NRT, a set of scaling factors are optimized for different pre-defined regions based on ecosystem types and ocean basins. Only biological and ocean fluxes are optimized. Prior biological fluxes are as mention from CASA for CarbonTracker and derived based on statistics for CT-NRT, prior ocean fluxes are derived from air-sea difference in partial pressure of CO₂ from ocean inversions or direct measurements, and the fossil fuel and wildfire emissions are based on inventories. A long assimilation window is used in the optimization, prior to the v2017 release the assimilation window length was 5 weeks, and in the current version 12 weeks are used to better capture the strong 2015–2016 El Niño. More information about the CarbonTracker system can be found in Peters

et al. (2004, 2007) and in the CarbonTracker documentation, which is available online at https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2016_doc.php.

2.2 ACT-A field campaigns

To achieve the three mission goals of ACT-America (i.e., reduce atmospheric transport uncertainties, improve regional-scale estimates of CO₂ and CH₄ fluxes, and evaluate the sensitivity of Orbiting Carbon Observatory-2 (OCO-2) column CO₂ measurements to regional variability in tropospheric CO₂), five field campaigns in four seasons over three regions (Mid-Atlantic, Mid-West, and South) are being conducted. In this paper, we report on the findings of first two field campaigns, i.e. summer-2016 and winter-2017 ACT-A field experiments (**Table 1**). The ACT-A flights encompassed most of the eastern US with an average of 25 research flights (RFs) during each season.

Table 1: A brief overview on the ACT-America summer -2016 and winter- 2017 field campaigns providing details on the research flight (RF) missions performed by NASA’s B-200 and C-130 aircraft over three ACT-A subregions, namely, mid-Atlantic (MA), Mid-west (MW) and South. Shown are the amount of observations in hours and total flight legs, number of profiles made along with brief information on some severe weather events during the two campaigns. RF: Research flight, legs: Horizontal/straight level flight legs, profiles: Obtained via spirals, on route ascents, descents, take-off and landing.

	Summer – 2016			Winter- 2017		
Overview	25 RFs, 225 flight hours, 300 Vertical profiles			27 RFs, 216 flight hours, 220 Vertical profiles		
Regions	MA	MW	South	South	MW	MA
Dates	18-27 Jul	1-14 Aug	16-28 Aug	1-13 Feb	13-27 Feb	1-10 Mar
# RFs	7	9	9	8	9	8
# legs	51	62	69	48	67	42
# Profiles	90	105	95	64	86	66
Severe weather events	Catastrophic flooding in Louisiana from 11-14 Aug 2016			3 named winter storms: “Quid” over MW, “Reggie” and “Stella” over MA		

The ACT-A field experiments aim to obtain in-situ and remote sensing measurements of state variables, GHGs and numerous other trace gases (e.g. CO, O₃, OCS) covering four seasons over three regions of the Eastern US. Two instrumented NASA aircraft were operated out of NASA's Langley Research Center, NASA's Wallops Flight Facility for the Mid-Atlantic part of the mission, Lincoln, Nebraska for Mid-West and Shreveport, Louisiana for South (**Figure 1**). The B-200 aircraft was mainly equipped with in-situ sensors measuring state variables and GHGs while the C-130 aircraft was equipped with both remote sensing (e.g. MFL, CPL, radar) and in-situ sensors measuring both tracers and meteorological variables including temperature, wind, humidity and pressure. During summer 2016 and winter 2017 field campaigns, we deployed both aircraft for 2 weeks period in each region spanning 6 weeks for each campaign. Three main types of RF patterns were performed: (1) frontal weather RFs aiming to obtain GHG structures across frontal boundary within and above the ABL, (2) fair weather RFs for obtaining GHG variability inside and above ABL over large regions during typical anti-cyclonic synoptic settings yielding fair weather, steady horizontal wind regimes, (3) OCO-2 underflights to compare CO₂ columnar content measurements obtained with MFL and OCO-2.

Some detailed information about the ACT-A field campaigns in summer 2016 and winter 2017 are provided in **Table 1**. For instance, the summer field campaign was conducted for a six-week period between 15 July and 31 August 2016 involving 25 RFs with 7 over Mid-Atlantic, 9 over Mid-West, 9 over South. During the campaign, we designed daily flight plans primarily based on forecasts of meteorological conditions and source-sink distributions of CO₂ and CH₄ and classified the RFs into frontal or fair-weather flights; few RFs were hybrids of both.

In addition to the horizontal flight legs often over a distance of more than 600 km, vertical variability of GHGs and thermodynamic variables were sampled via spirals, on route ascents and descents in both warm and cold sectors, allowing horizontal and vertical variability of GHGs and other tracers. These general flight patterns of both B-200 and C-130 were guided by both short-term weather forecasts, usually on the day before of the RF and the now casting on the day of RF. For instance, for fair weather RFs, we selected days with no precipitation, few clouds, low-moderate wind speed, and homogeneous atmospheric properties – so changes in CO₂ are predominantly functions of the local surface fluxes and PBL height while for frontal RFs we selected days with passage of cold and/warm fronts in the three regions. The speed of B-200 and C-130 aircraft were on average 120 ms⁻¹ and 100 ms⁻¹ and we used quality controlled 5 s averaged PICARRO Cavity Ring Down Spectrometer (CRDS) measurements of atmospheric CO₂ for the results presented in this paper so that a horizontal resolution of the measurements presented were around 500 m.

2.3 Inter-comparison strategy and metrics

We first linearly interpolate both the EC and CT-NRT analyses of CO₂ concentration both in time and space to each of the valid ACT-A airborne measurements at the observed time and location. Our emphasis is on the overall differences, uncertainties and biases over different subregions and different seasons as a function of altitude, which guides our calculation and characterization of the error and difference statistics accordingly. It is beyond the scope of this study to examine the variability of such statistics for each mission or for different weather patterns such as across the frontal boundaries.

3. Results

3.1 Summer 2016

Figure 3 shows the comparison of the mean CO₂ concentration profiles averaged over all flight measurements along with each of the 3 subregions of the 2016 summer phase of the ACT-A field campaign. Averaged over all flights and over all subregions, the mean CO₂ profiles simulated by both the EC and CT-NRT global analyses interpolated to the flight-level positions are very consistent with the mean profile of the ACT-A measurements. All estimates show much reduced CO₂ concentrations in the boundary layer below 2 km with the lowest values at around 1.25 km. Above the boundary layer, the mean CO₂ increase gradually with height from a mean value of around 400 ppm at 3 km for both analyses and ACT-A to between 403-405 ppm at the highest level of aircraft measurements.

Nevertheless, there are some systematic biases in both analyses (**Figure 4**): the CT-NRT reanalysis has a systematic low bias of 1-3 ppm mostly in the boundary layer (below 2.5 km) while the EC analysis has a slight high bias of around 1 ppm above the boundary layer. Further examination shows that most of such biases in both the EC and CT-NRT analyses come from the Mid-Atlantic region: The CT-NRT analysis has a boundary layer low bias of as much as 10-12 ppm below 2 km but reduces to below 0.5 ppm at 3 km and above. The EC analysis has a high bias of 12 ppm near the surface but quickly reduces to around 2-3 ppm at 1-6 km, and below 1 ppm above 6 km. The mean difference between each of the analyses and ACT-A observations are rather small (less than 1 ppm) for the flights across the other two subregions (Midwest and South) except for a value of 1-4 ppm below the boundary layer. The enhanced mean difference in the boundary layer in all

subregions is at least partially due to larger spatiotemporal variabilities near the surface (where there are sources and sinks of CO₂) that will likely lead to larger sampling and interpolation errors both in time and space.

Figure 5 shows the root-mean square deviations (RMSD) between both analyses and the ACT-A measurements as well as between the two analysis products (CT-NRT and EC). Due to enhanced biases, and larger spatiotemporal variabilities, the RMSDs of both analyses are as high as above 10 ppm in the boundary layer but reduce to around 2 ppm or less above 3 km averaged across all the 2016 summer ACT-A flight-level CO₂ measurements. The RMSDs are the smallest (from 6 ppm near the surface to ~1 ppm in the free troposphere for both CT-NRT and EC) over the South subregion which has the smallest spatiotemporal variabilities, and are the largest in the mid-Atlantic region (from as high as 10-18 ppm in the boundary layer to around 2-4 ppm for EC and 1-2ppm for CT-NRT above 2.5km) which has the largest variabilities (**Figure 1**).

The overall RMSD between the two analyses, though slightly larger, are comparable with respective analysis uncertainties at different vertical levels verifying against the ACT-A data (**Figure 5**). This result suggests that these two independent analysis estimates can be used to quantify the overall regional CO₂ analysis uncertainties over North America, which can be used as a baseline reference for future studies in quantifying the uncertainties of regional carbon concentration and flux estimates, as well as in assessing the impact of regional transport through more refined regional modeling and analysis systems.

3.2 Winter 2017

Figure 6 shows the comparison of the mean CO₂ concentration profiles averaged over all flight measurements along with each of the 3 subregions of the 2017 winter phase of the ACT-A field campaign. First, the mean profiles by both analyses and ACT-A show a reversal of the vertical CO₂ gradient from increasing with height during the summer of 2016 to decreasing with height during the winter of 2017, which is expected as there is a net loss of CO₂ at the surface during the summer plant growing season over North America, while there is a net gain from the surface during winter, which is due primarily to plant respiration and fossil fuel emissions. The net change over this 6-month period is about 5 ppm in the free troposphere average of all ACT-A flight-level measurements, but the mean net change in the boundary layer is as high as 10-20 ppm.

The mean CO₂ profile interpolated from the CT-NRT reanalysis agrees exceptionally well with the ACT-A 2017 winter measurements across all subregions except for maybe missing a peak value by 2-3 ppm at around 1 km in the boundary layer (**Figure 7**). The EC analysis on the other hand has a systematic high bias across all subregions and at all vertical levels ranging from around 5 ppm near the surface to a near persistent 1-2 ppm high bias above 2 km. Given that the EC analysis only assimilates infrequent remote sensing CO₂ measurements from polar-orbiting satellites (without including surface-based tower observations as used in the CT-NRT reanalysis), and that the bias is persistent throughout the whole atmospheric column, it is possible such a bias is at least partially inherited from a bias in the satellite measurements. The exact cause of this high bias in the EC analysis is beyond the investigation of the current study but the ACT-A measurements provide excellent independent observations to identify such a high and persistent bias in the EC analysis during the winter months that will certainly be valuable to product

developers in future improvement of the model and/or data assimilation including enhanced bias correction. Also worth noting that some of the ACT-A flight patterns were designed to fly directly under the OCO-2 satellite path so that the direct airborne measurements can be used to verify satellite measurements, and possibly correct the biases in satellite retrievals.

The RMSEs for both analyses are considerably smaller during winter 2017 (2-6 ppm) than those during summer 2016 (6-12ppm) in the boundary layer below 2 km but have overall similar level of analysis uncertainties (1-2ppm) in the free troposphere (**Figure 8** versus **Figure 4**). Due largely to the persistent positive bias in the EC analysis, the analysis uncertainties in the CT-NRT estimates are noticeably smaller than those of the EC analysis. The RMSEs are rather comparable across all subregions for both analyses except in the boundary layer.

It is also worth noting that if a simple bias correction is applied to both the EC and CT-NRT analyses (by subtracting the overall mean differences from ACT-A observations above 1 km across all flight-level measurements for each region from the respective raw analysis estimates), the bias corrected EC analysis would have analysis uncertainties comparable to those of the CT-NRT analysis (**Figure 9**).

Note that since the XCO₂ product assimilated by the real-time EC analysis is bias corrected based on TCCON data from the previous year (not available for the current year due to about 6-month TCCON data latency) while assuming a 2-ppm nominal annual changes, such a 2-ppm bias in the EC analysis during the winter is well within the range of the uncertainties in the annual changes (Massarat et al. 2016).

4. Concluding remarks

Through verifying against hundreds of hours of airborne in-situ measurements collected during the summer 2016 and winter 2017 phases of the ACT-A field campaign, this study systematically examines the regional uncertainties and biases of the CO₂ concentrations from two of the state-of-the-art global analysis products. One is the experimental real-time global carbon analysis produced by the European Center for Medium Range Forecasting using a 4DVar system and the other one is the near real-time reanalysis generated by the NOAA's Carbon Tracker system that is based on the ensemble Kalman filter techniques.

It is found that both the EC and CT-NRT CO₂ analyses agree reasonable well with the independent ACT-A flight-level measurements in particular above the boundary layer, although there are strong variabilities in accuracy and bias between seasons, and across different sub-regions of the field campaign. During the summer months of 2016, the analysis uncertainties of the EC and CT-NRT analyses against airborne data are comparable to each other, both of which are within 2 ppm from above 3 km but increase to about 3 ppm at 2 km reaching as large as 10 ppm near the surface. The overall biases are small for both analysis products while the differences between the two analyses is comparable to the analysis uncertainties verifying against flight measurements. Similar analysis uncertainties are found to be true also during the winter months of 2017, except for the need for removing a ~2 ppm systematic high bias from the EC carbon analysis.

The current study not only provides systematical uncertainty estimates for both analysis products over North America but also demonstrated that these two independent estimates can be used to quantify the overall regional CO₂ analysis uncertainties, both of

which are important in future studies in quantifying the uncertainties of regional carbon concentration and flux estimates, as well as in assessing the impact of regional transport through more refined regional modeling and analysis systems.

The current study provides the first overall uncertainty analysis of the EC and CT-NRT CO₂ concentration estimations compared with the ACT-A airborne field measurements. Ongoing and planned future studies to be reported elsewhere when completed will look at the CO₂ concentration and transport uncertainties under different large-scale weather patterns. This includes but not limited to how well the analyses compare for fair-weather (during which CO₂ uncertainties are thought to be predominantly flux-driven) and stormy weather (mostly transport-driven errors), and for different regions, to investigate potential flux errors due to anthropogenic vs biological fluxes. Other future studies could also investigate if the global analyses can capture the horizontal structure of CO₂, e.g. the strong frontal gradients that were observed during ACT-A. The CO₂ (re)analysis and ACT-A field measurements can also be further used to assess the fidelity and uncertainties in the EC real-time forecasts of CO₂. Ultimately these uncertainty estimates and subsequent modeling experiments will help to address the uncertainties in the CO₂ flux estimates from different sources and sinks as well as from atmospheric transport, as the ultimate goal of the ACT-A field campaign.

Acknowledgements: [to be added]

References: [to be added]

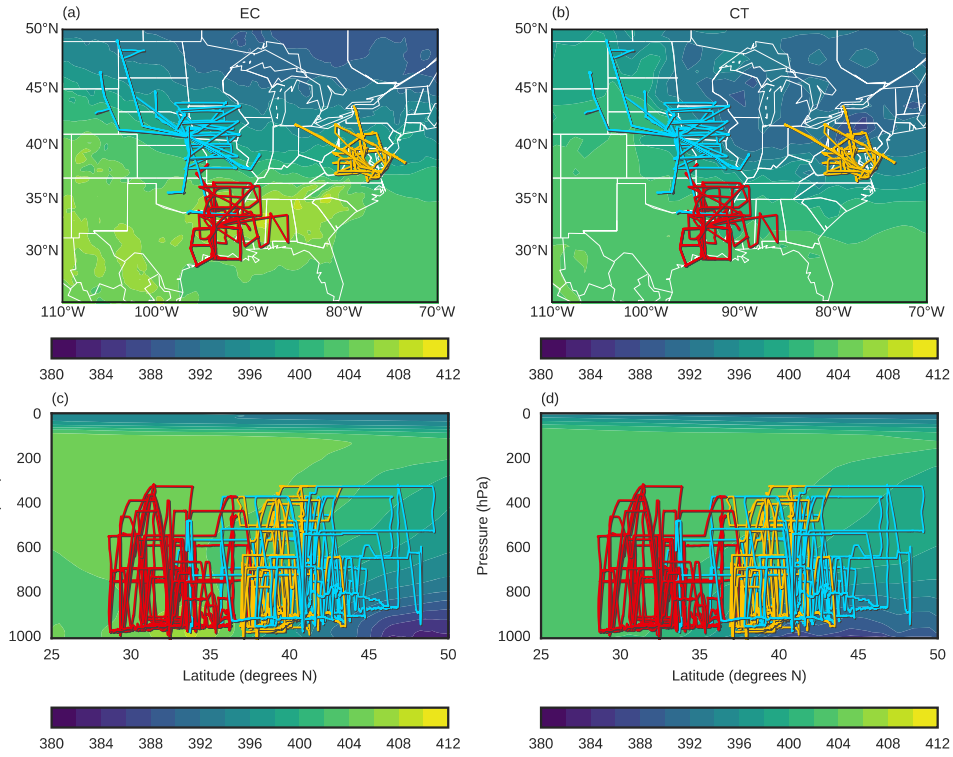


Figure 1. Flight tracks during the ACT-A summer 2016 campaign in Midwest (blue lines), Mid-Atlantic (yellow lines), and South (red lines). (a) and (b) show the flight tracks with respect to longitude and latitude, and (c) and (d) show the same tracks with respect to latitude and elevation. The shadings in (a) and (b) show the July—August mean CO₂ concentration at around 850 hPa for the EC analysis and the CT-NRT reanalysis, respectively. Similarly, the shadings in (c) and (d) show the zonally averaged July—August mean CO₂ concentration for the EC analysis and the CT-NRT reanalysis.

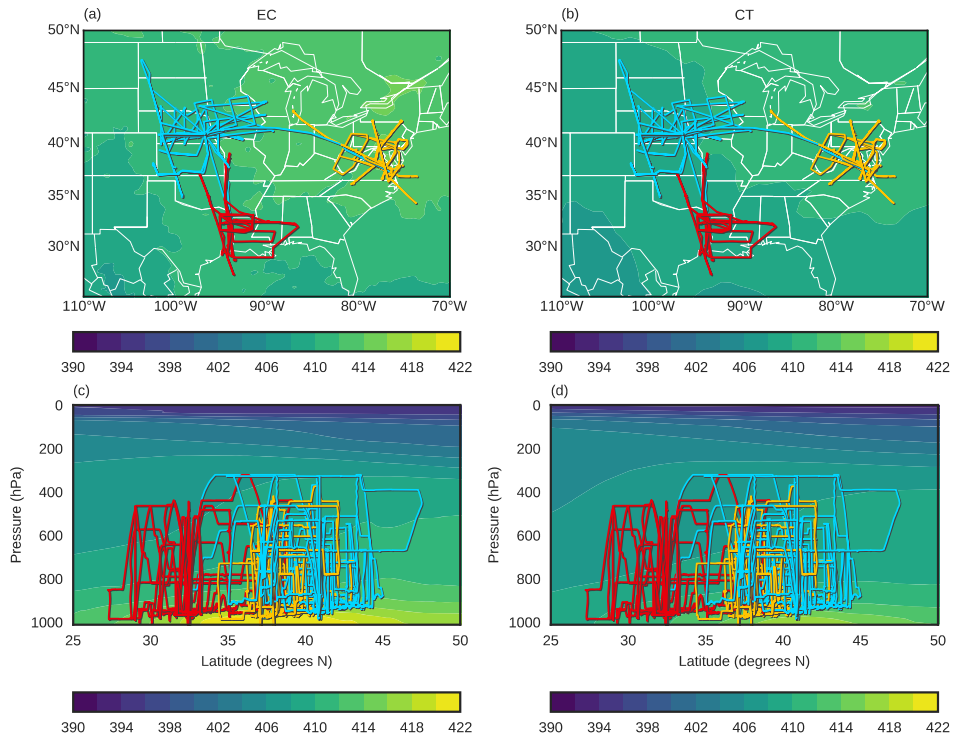


Figure 2. Same as Fig. 1 but for the ACT-A winter 2017 campaign.

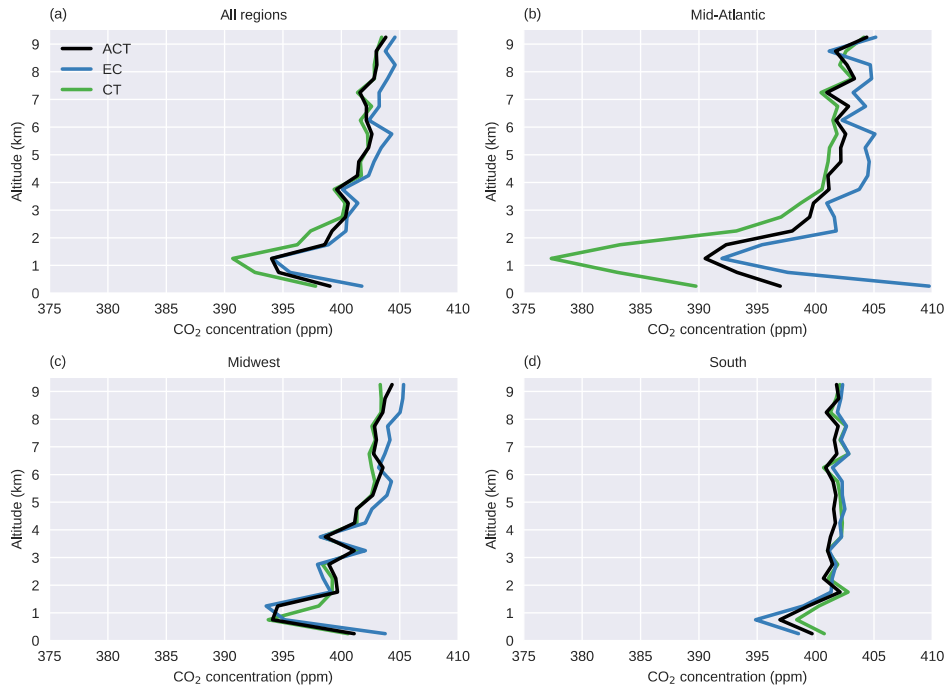


Figure 3. Vertical distribution of CO₂ concentration averaged over all flights during the ACT-A summer 2016 campaign for (a) all regions, (b) Mid-Atlantic, (c) Midwest, and (d) South. The EC and CT-NRT CO₂ products were linearly interpolated in time and space to match the ACT-A flight tracks.

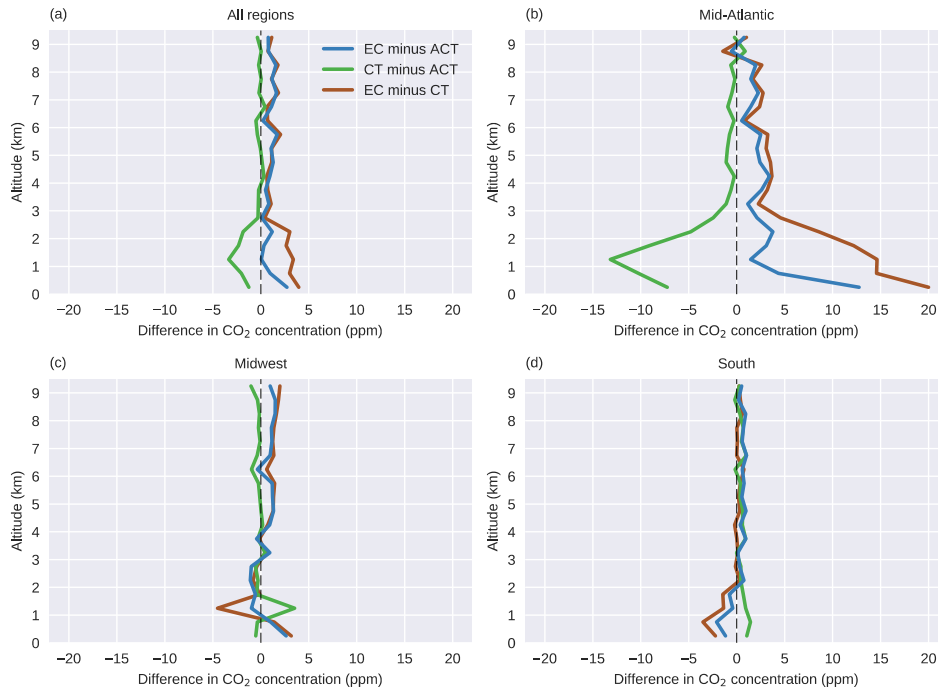


Figure 4. Vertical bias of CO₂ concentration averaged over all flights during the ACT-A summer 2016 campaign for (a) all regions, (b) Mid-Atlantic, (c) Midwest, and (d) South. The blue lines show the bias in EC with respect to ACT-A observations, the green lines show the bias in CT-NRT with respect to ACT-A observations, and the brown lines show the difference between the two CO₂ (re)analysis products.

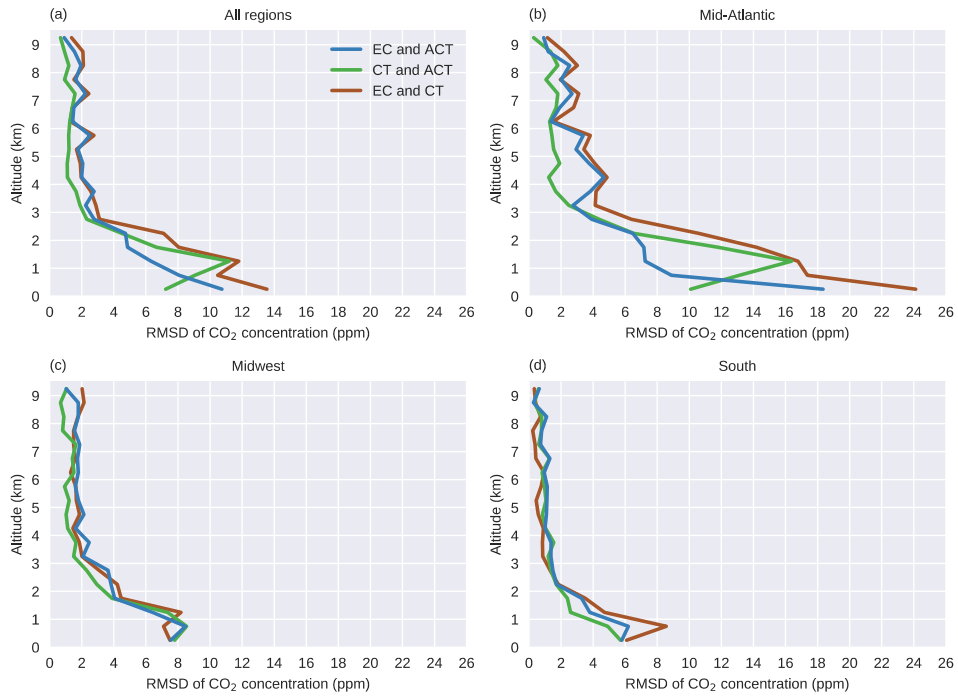


Figure 5. Vertical distribution of RMSD of CO₂ concentration averaged over all flights during the ACT-A summer 2016 campaign for (a) all regions, (b) Mid-Atlantic, (c) Midwest, and (d) South. The blue lines show the RMSD between EC and ACT-A observations, the green lines show the RMSDE between CT-NRT and ACT-A observations, and the brown lines show the RMSD between the two CO₂ (re)analysis products.

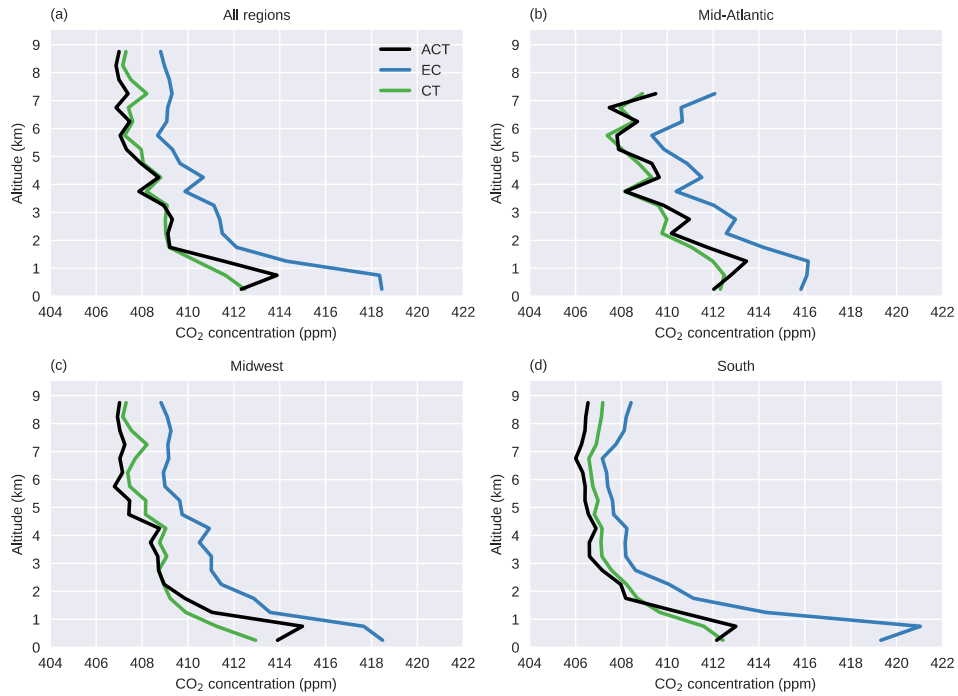


Figure 6. Vertical distribution of CO₂ concentration averaged over all flights during the ACT-A winter 2017 campaign for (a) all regions, (b) Mid-Atlantic, (c) Midwest, and (d) South. The EC and CT-NRT CO₂ products were linearly interpolated in time and space to match the ACT-A flight tracks.

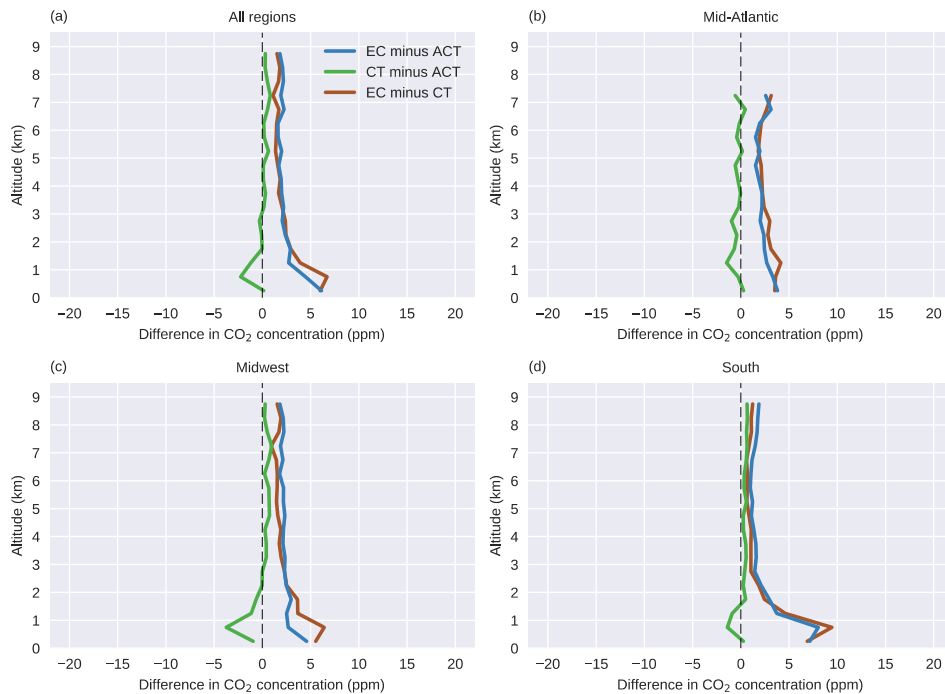


Figure 7. Vertical bias of CO₂ concentration averaged over all flights during the ACT-A winter 2017 campaign for (a) all regions, (b) Mid-Atlantic, (c) Midwest, and (d) South. The blue lines show the bias in EC with respect to ACT-A observations, the green lines show the bias in CT-NRT with respect to ACT-A observations, and the brown lines show the difference between the two CO₂ (re)analysis.

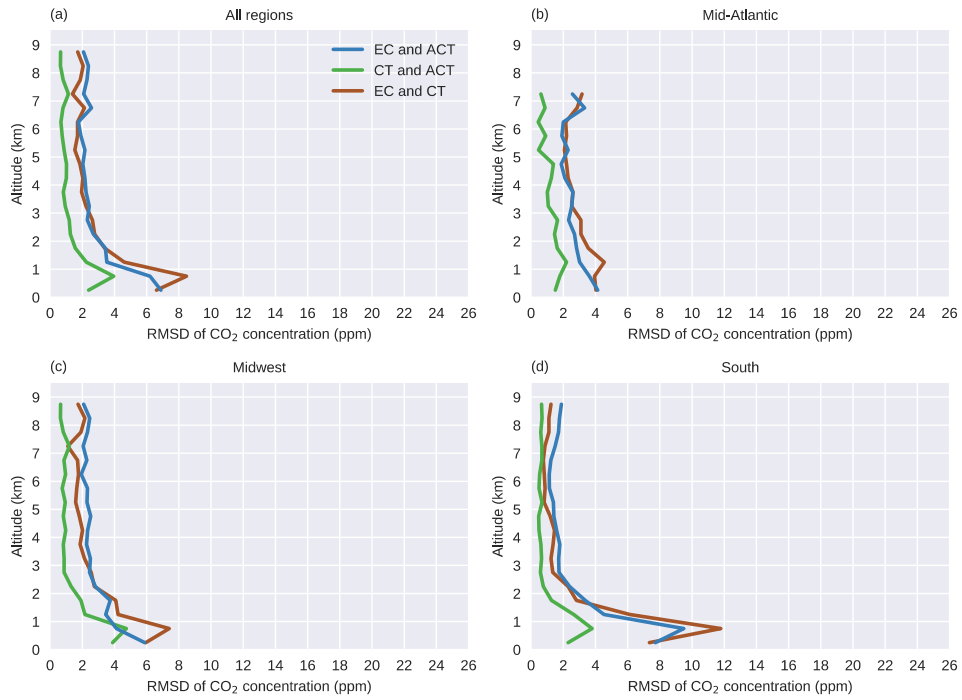


Figure 8. Vertical distribution of RMSD of CO₂ concentration averaged over all flights during the ACT-A winter 2017 campaign for (a) all regions, (b) Mid-Atlantic, (c) Midwest, and (d) South. The blue lines show the RMSD between EC and ACT-A observations, the green lines show the RMSDE between CT-NRT and ACT-A observations, and the brown lines show the RMSD between the two CO₂ (re)analysis products.

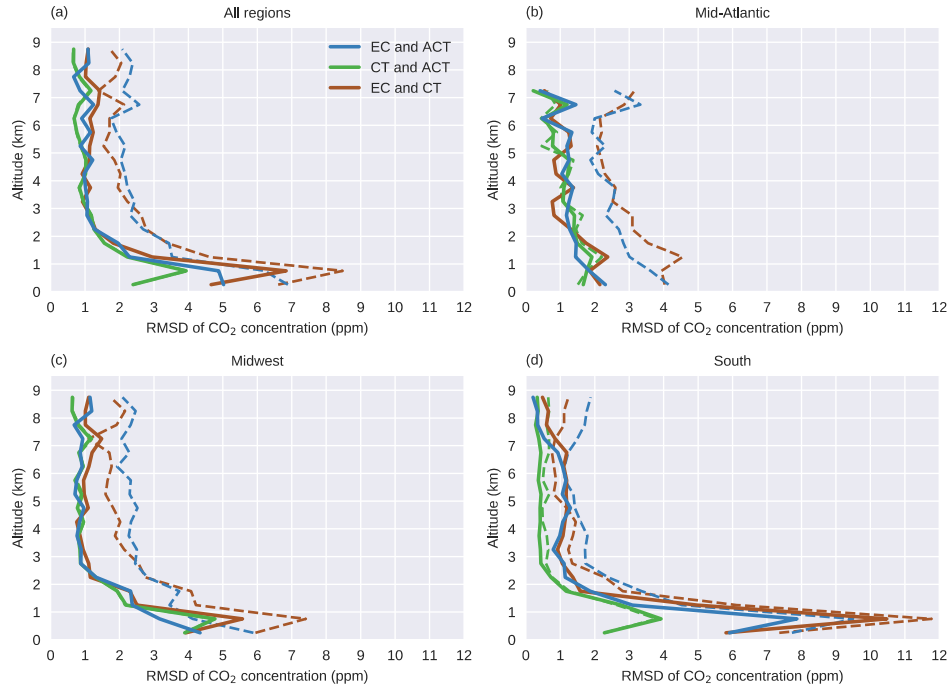


Figure 9. Vertical distribution of RMSD of CO₂ concentration during the ACT-A winter 2017 campaign after removing the biases in the EC and CT-NRT products with respect to ACT-A observations. A single bias was calculated for each region by averaging the differences above 1 km between the CO₂ (re)analysis products and ACT-A observation (shown in Fig. 7). This bias was then subtracted from all levels before calculating new RMSDs between EC/CT-NRT and ACT-A observations. Dashed lines show the RMSDs before the bias correction (same as the lines in Fig. 8), and solid lines show the RMSDs after the bias correction.