UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ATMOSPHERIC VARIATIONS IN COLUMN INTEGRATED CO2 ON SYNOPTIC AND SEASONAL TIME SCALE OVER THE U.S.

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

QINGYU WANG
Norman, Oklahoma
2018

ATMOSPHERIC VARIATIONS IN COLUMN INTEGRATED CO2 ON SYNOPTIC AND SEASONAL TIME SCALE OVER THE U.S.

A THESIS APPROVED FOR
SCHOOL OF METEOROLOGY

BY

Dr. Petra Klein, Chair

Dr. Sean Crowell

Dr. Jason Furtado

Acknowledgements

I would like to express my special thanks of gratitude to my advisor, Dr. Sean Crowell, who gave me generous help and encouragement in last two and half years. Thank you for your patients whenever I have numerous questions on this research. Without your guidance and persistent help, this thesis would not have been possible. Thanks must also go to my committee members, Dr. Petra Klein and Dr. Jason Furtado, who greatly improved the quality of this work through timely and useful suggestions. Secondly I would also like to thank my parents and friends, especially Yuan Yuan, Yuwei Shi, who helped me a lot in finalizing this project within the limited time frame and kept encouraging me to insist on my choice. I would also like to show our gratitude to my friends in GeoCarb lab, Hélène Peiro and Jeff Nivitanont for sharing their pearls of wisdom with me in my thesis. Most importantly, I wish to thank my loving and supportive boyfriend, Yue Zhao, whose love and guidance are with me in whatever I pursue. And Thank you Zhuangzhuang for your company.

Finally, I must thank the Orbiting Carbon Obsevatory-2 mission by NASA for providing funding for this research.

Table of Contents

List of Tables

List of Figures

Abstract

Past studies have demonstrated that synoptic events play an important role in the spatial and temporal variations of carbon dioxide ($\mathrm{CO_2}$). In this study, in order to investigate whether cold fronts have impact on synoptic $\mathrm{CO_2}$ concentrations, we collect 83 cold frontal cases over United Sates, east Pacific Ocean and west Atlantic Ocean (or the contiguous United States: CONUS) from 2015 to 2017 with data from Orbiting Carbon Observatory-2 (OCO-2) and Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), calculate the Column-averaged carbon dioxide dry air mole fraction ($X_{\mathrm{CO_2}}$) difference anomalies across the fronts, and apply significance test in each season based on non-frontal days climatology to decide whether cold fronts relate to horizontal $\mathrm{CO_2}$ changes. The large day-to-day $\mathrm{CO_2}$ variability at the same location from Weather Research and Forecasting-Vegetation Photosynthesis and Respiration Model (WRF-VPRM) suggest that randomly selected orbits which are not crossing any fronts are a better reference than $X_{\mathrm{CO_2}}$ monthly mean on the same spatial resolution in one month.

Seeing that OCO-2 measures well on mesoscale and synoptic scale, we regard OCO-2 data as the truth and examine whether simulations from WRF-VRPM are simulating well over the contiguous United States (CONUS). Based on land cover classifications given by the Moderate Resolution Imaging Spectroradiometer (MODIS), we compute the difference between the model simulations and OCO-2 observations for each land cover type in each season. The evaluation reveals that the model agrees well with OCO-2 for some specific land

cover and season, like forests in the winter, while still relative high bias in summer for most surface types.

In summary, the result that $X_{\mathrm{CO_2}}$ frontal gradients have similar pattern with past studies in boundary layer demonstrates that OCO-2 is a good tool to see mesoscale or synoptic and seasonal variations over the CONUS. Accordingly, regarding OCO-2 data as the truth, outputs from WRF-VPRM are evaluated based on 7 land cover types, which was proved generally well in winter over some land cover types, but may not ideal for others.

Chapter 1 : Introduction

1.1 Background

1.1.1 Global warming situation and local measurements

The 30-year period from 1983 to 2012 has been reported to be the warmest period in the past 800 years (Pachauri et al. 2014). Also, sea level rise explained by about 75% by glacier mass loss and ocean thermal expansion, and CO_2 concentrations increased at the fastest observed decadal rate of change for 2002-2011 by IPCC Fifth Assessment Report (Pachauri et al. 2014). From 1970 to 2014, cumulative anthropological CO_2 emissions have tripled from fossil fuel combustion, cement production and flaring, and increased by about 40% from deforestation and other land use (IPCC Fifth Assessment Report, (Pachauri et al. 2014)). Before the Industrial Revolution, the amount of carbon moving between trees, soil, oceans and the atmosphere was relatively balanced. As greenhouse gases (GHGs), the increase of CO_2 concentrations and unbalanced carbon cycle drive global climate warming via trapping outgoing longwave radiation emitted from the Earth to keep the Earth warm. Global warming alters the carbon cycle as well, for example, releasing dissolved CO_2 from warmer oceans and lakes enhanced sources. The major components of the carbon cycle are shown in Fig 1.1. The large growth in global CO_2 emissions has had a significant impact on the concentrations of CO_2 in Earth's atmosphere. Local CO_2 emission along with transported CO_2, can be measured by several methods like airborne LIDARs, satellites such as the Orbiting CarbonObservatory-2 (OCO-2), the Total Carbon Column Observing Network (TCCON), which is a network of ground-based Fourier Transform Spectrometers recording direct solar spectra in the near-infrared spectral region (https://tccon-wiki.caltech.edu), and etc. With the in-situ or satellite measurement, we can better understand the carbon emissions, which plays an important role in global carbon cycle.

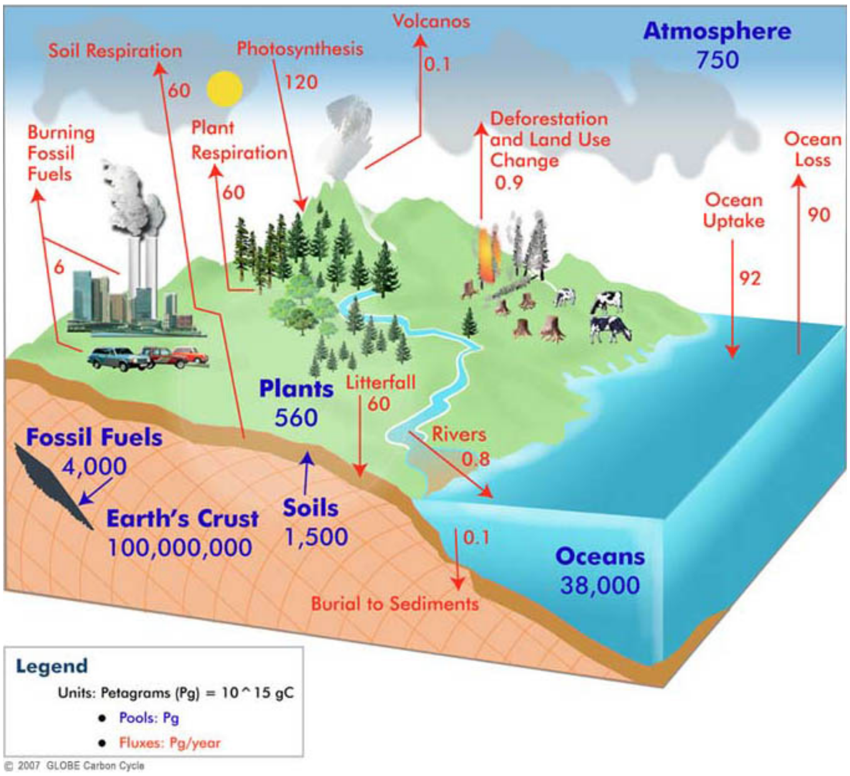


Figure1.1 - The global carbon cycle. The main sources of carbon are oceans, soils and vegetation respiration, fossil fuel& cement production. Carbon sinks include terrestrial biosphere, soils, oceans, and fossil fuels, http://globecarboncycle.unh.edu

Synoptic weather, like fronts, coupled with surface carbon dioxide flux, is related with day-to-day CO_2 variations in atmosphere (Geels et al. 2004; Gerbig et al. 2003; Lauvaux et al. 2008; Law et al. 2002; Lin et al. 2004; Parazoo et al. 2008; Peylin et al. 2005). TCCON data from Park Falls, Wisconsin, suggest that even on hourly timescales, CO_2 variations in a column is dominated by non-local effects (Keppel-Aleks et al. 2012). Keppel-Aleks et al. (2012) illustrated that more robust flux estimates can be inferred if looking into the CO_2 variability over the local or large-scale based on potential temperature. Fronts are believed to be related to temperature and moisture discontinuity from the earliest , simplest model developed by Margules (1906), may also be a starting point to assess the relationship between CO_2 variations and synoptic weathers, and furthermore, the examination of the precision of CO_2 measurements from OCO-2.

1.1.2 Fronts and CO_2 contrasts

Fronts are defined as the zones of transition between two different air masses (Hewson 1998; Holton and Hakim 2012; Renard and Clarke 1965). The scales of fronts are variable in space and time — surface fronts often last for about three days on time scale and span about 800 km horizontally, 3km vertically on spatial scale (Bluestein 1992), but from one side of a front to the other, one clearly would sense that the properties of an air mass had changed significantly (e.g., contrasts in temperature and dew point, wind direction, cloud cover, and sensible weather). There are four types of fronts defined: cold front, warm front, stationary front, and occluded front (vertical cross-section of cold fronts, warm fronts are shown in Figure 1.2). The type of front depends on both the direction in which the air mass is moving and the characteristics of the air mass. A cold front is a front in which cold air replaces warm air at the surface, whose typical frontal slope is 1:100 (vertical to horizontal). The temperature transition zone is a band between warmer and cooler air masses which drop by 8K over 50 miles (Miles 1962). A warm front is a front in which warm air replaces cooler air at the surface, with a typical frontal slope 1:200 (more gentle than cold fronts). The warm air is above the cold wedge, which is not only warmer but also moister than the cooler air in the wedge (Miles 1962; Renard and Clarke 1965), because of density differences. A stationary front is a front that does not move or barely moves, which typically forms when polar air masses are modified significantly so as to lose their character. If, as often happens, a warm front is overtaken by a cold front moving around a low-pressure center, then occluded fronts form, when cold fronts move faster than warm fronts, catch up and overtake their related warm front. In spite of the common density or the temperature criterion, many other features may distinguish a front, such as a pressure trough, wind direction changes, moisture discontinuities, and a band of clouds from a satellite imagery and precipitation. Fronts are often initiated by large-scale horizontal deformation field — the tendency of air parcels to change shape, and result in some sharp temperature contrasts and precipitation (Wallace and Hobbs 2006). Hurwitz et al. (2004) demonstrated that abrupt changes in CO_2 mixing ratio happened in the presence of inclement weather and low pressure system using water vapor mixing ratio, temperature, wind speed and wind direction data measured by flux tower. Also a lot of past studies (Bianchi et al. 2009; Boutin et al. 2008; Keppel-Aleks et al. 2012; Lee et al. 2012; Mahadevan and Archer 2000; Parazoo et al. 2008) indicate that spatial CO_2 variations are also related with fronts.

The research done by Keppel-Aleks et al. (2012) illustrates that the changes of vertically integrated CO_2 mixing ratio are primarily driven by non-local effects as seen from TCCON data at Park Falls, Wisconsin. And large-scale gradients of integrated CO_2 are correlated with synoptic-scale variations in potential temperature θ. Thus we can also regard CO_2 gradients as seen from TCCON as a reference. CO_2 frontal changes were observed by in-situ data in Parazoo et al. (2008), in which 96-hour in-situ data were applied to see the whole frontal passage. Frontal CO_2 changes as seen from in-situ continuous sites were found to have higher prefrontal CO_2 than postfrontal CO_2 at some sites like SGP (Southern Great Plains of North America, characterized by agriculture), WKT (Great Plains of North America in a region of strong moisture gradient, characterized by cattle grazing) and SBI (Sable Island, Island off the coast of Nova Scotia) in summer.

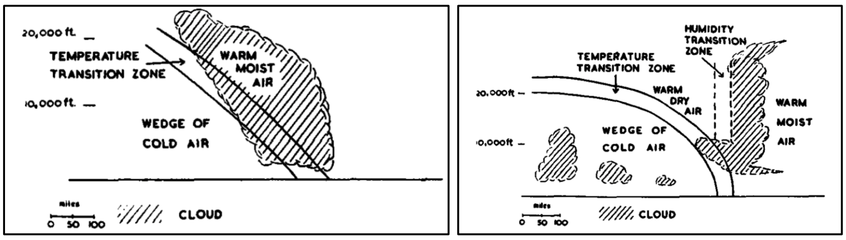


Figure 1.2 - From (Miles 1962), schematic cross-section through a front (a). A warm front (b). A cold front

In this thesis, we only look at the cases when there is a cold front because cold fronts are more easily spotted from satellite imagery, which often go along with cloudy zones. Parazoo et al. (2008) found that in midlatitudes, horizontal advection is responsible for 60-70% of CO_2 daily variations on average in boundary layer, along with moisture convection and surface flux, are three components of boundary layer CO_2 budget equation. Hence the surface flux is more important in days when advection and moist convection is weaker. However, whether the OCO-2 can see the variation in that scale is still unknown. With the work done before, we can compare results from OCO-2 to verify the accuracy of OCO-2 data.

1.1.3 ACT-America

Surface measurements like TCCON (site map is given in Fig 1.3) are made only where instruments can be positioned, while weather phenomena will not always pass where the instruments are located. The Atmospheric Carbon and Transport – America (ACT-America) (https://act-america.larc.nasa.gov) campaign has conducted five airborne campaigns across three regions in the eastern United States to study the transport and fluxes of atmospheric carbon dioxide and methane (Davis et al. 2017). Each 6-week campaign measures how weather systems transport these greenhouse gases. The objective of the study is to enable more accurate and precise estimates of the sources and sinks of these gases by better understanding atmospheric transport mechanisms. Better estimates of greenhouse gas sources and sinks are needed for climate management and for prediction of future climate. ACT-America addresses three primary sources of uncertainty in our ability to infer carbon dioxide and methane sources and sinks - transport error, prior flux uncertainty and limited data density.

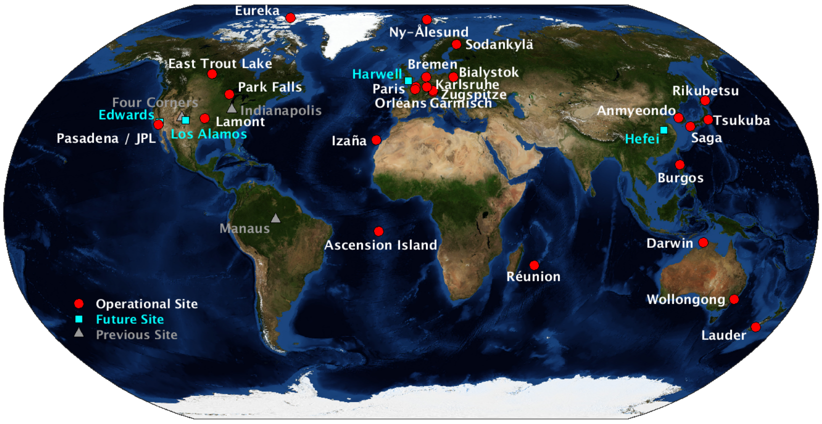
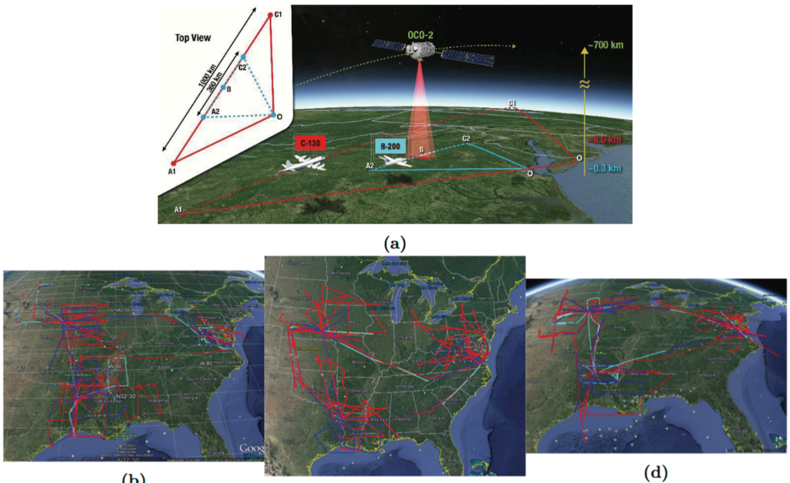


Figure 1.3 - TCCON site locations. Current operational sites are indicated with filled red circles, planned future sites are indicated with filled blue squares, and previous sites are indicated with grey triangles. From https://www.esrl.noaa.gov/gmd/annualco

According to flight experiments from ACT America, the seasonal variations of CO_2 concentration frontal gradients (CO_2 concentration at warm sector minus the one at cold sector) is apparent in the boundary layer in the mid-Atlantic, mid-west U.S., and south U.S. to Gulf of Mexico regions in summer 2016, winter 2017, fall 2017, and spring 2018 (shown in Fig 1.4). In the boundary layer, the average CO_2 concentration frontal gradient in summer for mid-west U.S. are demonstrated to be about 12.5 parts per million by volume (ppmv), -2.5 ppmv in winter, -5 ppmv in fall and -3 ppmv in spring (shown in Fig 1.5).



(a)

(b)

(c)

(d)

Figure 1.4 - (a) ACT-America OCO-2 underflight diagram, and flight tracks from the (b) Summer 2016, (c) Fall 2017, and (d) Winter 2017 campaigns, provided by Davis et al. (2017)

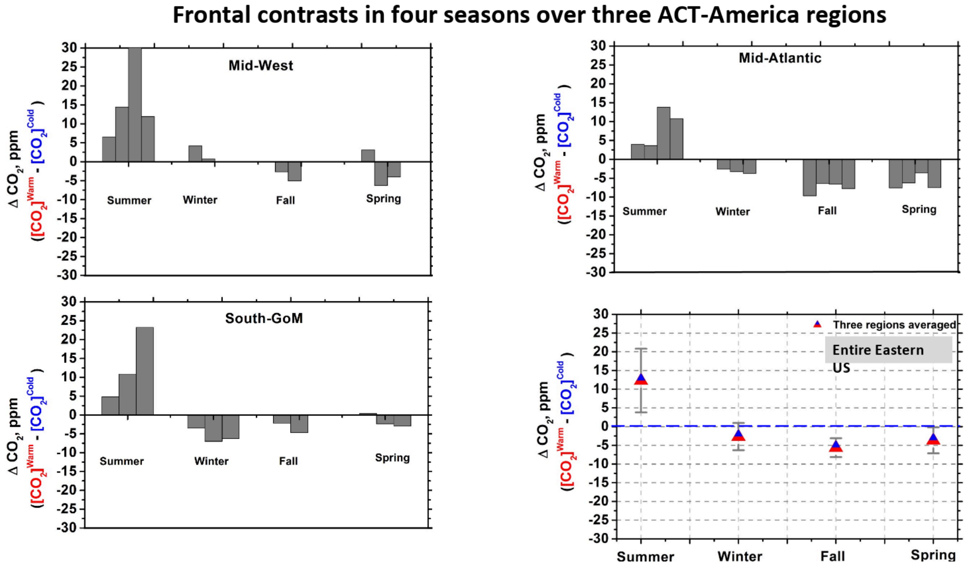


Figure 1.5 - CO_2 frontal gradient measured by ACT-America campaign. (a) is in mid-west US, the grey bars in each season are measured CO_2 frontal gradients in each flight; (b) is same with (a), but over mid-Atlantic; (c) is same with (a), but over Gulf of Mexico, (d) is the average of CO_2 frontal gradients in campaigns in (a)-(c) with respect to season. This figure is credit to Sandip Pal.

1.1.4 OCO-2

Most scientific satellites and many weather satellites are in a nearly circular, low Earth orbit. In order to enable coordinated science observations, the Earth Observations System has created the A-Train (or “afternoon train”) (Stephens et al. 2002). The satellites have low polar orbits 438 miles (705 km) above Earth at an inclination of 98 degrees. Together, their overlapping science instruments give a comprehensive picture of Earth weather and climate. The A-Train is so named because the lead satellite, Aqua, crosses the equator at the mean local time of approximately 1:30 pm. The A-Train satellite formation currently consists of six satellites flying in close proximity: Aqua, CloudSat, Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO), Aura, Global Change Observation Mission –

Water 1 (GCOM-W1) and OCO-2 (shown in Fig 1.6) (https://www.nasa.gov/mission_pages/a-train/a-train.html). Bakwin et al. (1998) suggested the mid-afternoon is the best time in a day for inferring the regional and synoptic influence on tracer concentrations in well-mixed atmospheric boundary layer because the effect is maximized at that time.

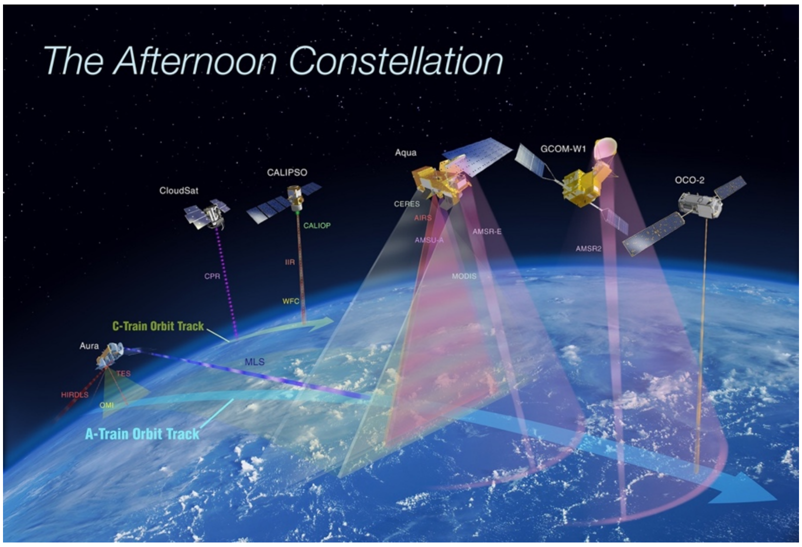


Figure 1.6 - The International Afternoon Constellation includes the A-Train satellites (OCO-2, GCOM-W1, Aqua, and Aura) as well as the C-Train satellites (CALIPSO and CloudSat), https://atrain.nasa.gov

OCO-2 is NASA's first dedicated satellite for monitoring atmospheric carbon dioxide distributions over the global scale. It is a replacement for The Orbiting Carbon Observatory (OCO), which was lost in a launch failure in 2009 (https://www.jpl.nasa.gov/news/news.php?release=2014-211). Introductions on OCO given by (Crisp et al. 2008; Crisp et al. 2004) are valid for OCO-2. OCO-2 is designed to provide space-based global measurements of atmospheric carbon dioxide (CO_2) with the precision and resolution needed to identify and characterize the processes that regulate this important greenhouse gas. With its three high spectral resolution grating spectrometers, data collected by

OCO-2 could be combined with meteorological observations and ground-based CO_2 measurement to help characterize CO_2 sources and sinks on regional scales at monthly intervals. CO_2 concentration frontal gradients detected by ACT-America flight can be seen as the reference to judge qualitatively if OCO-2 has sufficient precision to see changes in the atmosphere. If the variations from OCO-2 are similar to the results attained from flight, OCO-2 data will be considered as accurate. Especially in this thesis, the $X_{\mathrm{CO_2}}$ gradients across cold fronts are considered as a test for the accuracy of OCO-2 and whether OCO-2 is appropriate for looking at CO_2 variations in the atmosphere. Frontal contrasts observed from OCO-2 may be smaller because it collects the column integrated CO_2 from the surface to the top of atmosphere, while carbon dioxide concentration varies more in the boundary layer because of the respiration of vegetation, fuel burning etc. (Keeling et al. 1976; McClure et al. 2016; Thoning et al. 1989). OCO-2 measures the intensity of the sunlight reflected from the presence of CO_2 in a column of air rather than measuring CO_2 directly (shown in Figure 1.7). The intensity of 3 wavelength bands (Weak CO_2 at 1.61 microns, Strong CO_2 at 2.06 microns and Oxygen O_2 at 0.765 microns) from the spectrum is measured, each wavelength band is specific to one of the three spectrometers. OCO-2 will be able to track the variations in gas density in the atmosphere over time by repeatedly sampling the globe over days, seasons, and years.

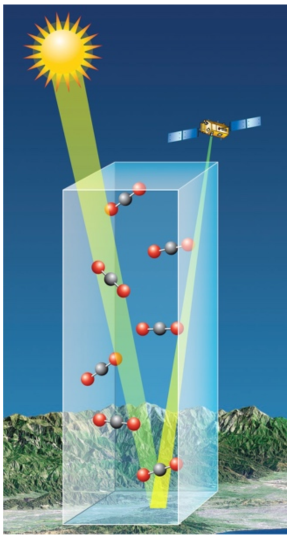


Figure 1.7 - A simple model for OCO-2 CO2 measurements in a column in glint mode. OCO-2 measures the sunlight reflected at the Earth surface. When passing through the atmosphere twice, energy in sunlight are absorbed by carbon dioxide and molecular oxygen molecules at some specific wavelengths (https://oco.jpl.nasa.gov/observatory/instrument/).

Nadir and glint are the two main viewing geometries of satellite observation for OCO-2. In nadir mode, the satellite points the instrument to the local nadir, in which the data is representative of the column of air between the satellite and the subsatellite point directly on the ground track. Each OCO-2 spectrometer will have 10 km – wide cross track field of view at nadir that is divided into eight cross-track elements, or “footprints”. While in the glint mode, the spacecraft points the instrument in the direction of the bright “glint” spot, where solar radiation is reflected from the surface. Glint measurements provide much higher signal-to-noise ratios (SNR). Originally, the observatory was set to exclusively operate in one mode – nadir or glint for each period in 16 periods. In order to reduce the error caused by thermal changes, since July 2nd, 2015, the measurement pattern was changed to switching by cycle (i.e., switching to nadir after one orbit of glint, followed by an orbit of glint, then nadir, and so on).

1.1.5 WRF-VPRM

WRF-VPRM simulates Gross Ecosystem Exchange with temperature data from the Weather Research and Forecasting (WRF) model and Land Use and Land Cover information from MODIS. WRF and the Vegetation Photosynthesis and Respiration Model (VPRM) (Mahadevan et al. 2008) were coupled (WRF-VPRM) to evaluate CO_2 fluxes and their subsequent atmospheric transport/dispersion (Ahmadov et al. 2007). WRF is a next-generation mesoscale numerical weather prediction system which is widely used by atmospheric research. It provides temperature and photosynthetically active radiation (PAR) information in WRF-VPRM. VPRM is based on the MODIS data, developing from the Vegetation Photosynthesis Model (VPM) (Xiao et al. 2004a; Xiao et al. 2004b). VPM was developed to estimate Gross Ecosystem Exchange GEE —the amount of chemical energy as biomass that primary producers create in a given length of time, with vegetation indices and environmental data from MODIS. VPRM additionally simulates Ecosystem Respiration (ER), the sum of all respiration occurring by the living organisms in the ecosystem, and then estimates Net Ecosystem Exchange (NEE) of CO_2 through (Mahadevan et al. 2008):

$$\mathrm{NEE} = \mathrm{GEE} + \mathrm{ER}\,, \qquad (1)$$

$$\mathrm{GEE} = \lambda \cdot \mathrm{T_{scale}} \cdot \mathrm{W_{scale}} \cdot \mathrm{P_{scale}} \cdot \mathrm{FAPAR_{PAV}} \cdot \mathrm{PAR} \cdot \frac{1}{1+\frac{\mathrm{PAR}}{\mathrm{PAR_0}}}\,, \ (2)$$

$$\mathrm{ER} = \alpha \times \mathrm{T} + \beta\,, \ (3)$$

where Ecosystem Respiration is linearly proportional to environmental temperature T in Equation (3). In Equation (3), α and β are two empirical parameters, which vary by land cover type. PAR in Equation (3) stands for Photosynthetically Active Radiation, which is the amount of light in the 400 to 700 nanometer wavelength range, provided by WRF, as well as T_{scale}, which is a parameter accounting for effects of temperature. $FAPAR_{PAV}$ is the Fraction of PAR, (μmol m^{-2} s^{-1}) absorbed by the photosynthetically active portion of the vegetation (PAV). MODIS provides Land Surface Water Index (LSWI) and Enhanced Vegetation Index (EVI) for VPRM, which are given as follows (Xiao et al., 2004a, 2004b):

$$\mathrm{EVI} = \mathrm{G} \cdot \frac{(\rho_{\mathrm{nir}} - \rho_{\mathrm{red}})}{\rho_{\mathrm{nir}} + (\mathrm{C}_1 \times \rho_{\mathrm{red}} - \mathrm{C}_2 \times \rho_{\mathrm{blue}}) + \mathrm{L}}, (4)$$

where ρ_{nir} and ρ_{red} are satellite-derived reflectance from the near infrared and red bands, G = 2.5, C_1 = 6, C_2 = 7.5 and $L = 1$.

$$\mathrm{LSWI} = \frac{\rho_{\mathrm{nir}} - \rho_{\mathrm{swir}}}{\rho_{\mathrm{nir}} + \rho_{\mathrm{swir}}}, (5)$$

where ρ_{swir} is short wave infrared reflectance. T_{scale} is a function of temperature, provided by WRF, and P_{scale} and W_{scale} are functions of LSWI, given by MODIS. Functions are as follows:

$$\mathrm{T_{scale}} = \frac{(\mathrm{T}-\mathrm{T_{min}})(\mathrm{T}-\mathrm{T_{max}})}{[(\mathrm{T}-\mathrm{T_{min}})(\mathrm{T}-\mathrm{T_{max}})-(\mathrm{T}-\mathrm{T_{opt}})^2]}, (6)$$

where $\mathrm{T_{min}}$, $\mathrm{T_{max}}$, and $\mathrm{T_{opt}}$ are minimun, maximum and optimal temperature (°C) for photosynthesis. For evergreen classes, $\mathrm{P_{scale}}$ is always set to be 1, while for deciduous vegetation and grasslands, $\mathrm{P_{scale}}$ is computed as a linear function of LSWI:

$$\mathrm{P_{scale}} = \frac{1+\mathrm{LSWI}}{2}. (7)$$

$\mathrm{W_{scale}}$ is computed as:

$$\mathrm{W_{scale}} = \frac{1+\mathrm{LSWI}}{1+\mathrm{LSWI_{max}}}, (8)$$

where $\mathrm{LSWI_{max}}$ is the maximum LSWI within the plant growing season for each site. Detailed explanations are available in Mahadevan et al. (2008).

In addition to vegetation contributions to CO_2 flux, WRF-VPRM also takes anthropogenic emissions into account using data from Emission Database for Global Atmospheric Research (EDGAR, Petrescu et al. (2012)) and monthly ocean CO_2 fluxes (Takahashi et al. 2009) for simulations over oceans. Atmospheric initial conditions and boundary conditions are taken from CarbonTracker (https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/). (https://www.ldeo.columbia.edu/res/pi/CO2/carbondioxide/pages/air_sea_flux_2000.html).

WRF-VPRM was evaluated over the South California Air Basin (Feng et al. 2016; Park et al. 2018), which is greatly affected by anthropogenic emissions. It demonstrated that compared with in atmospheric model , high spatial resolution in the anthropogenic CO_2 emissions, especially fossil fuel is more needed to capture CO_2 concentration variability across the LA megacity (Feng et al. 2016). However, WRF-VPRM hasn't been evaluated over the CONUS. In this study, WRF-VPRM simulations in 2016 over the CONUS are evaluated by comparing with OCO-2 $X_{\mathrm{CO_2}}$ if OCO-2 data are verified to be precise on the same scale WRF-VPRM is simulating.

1.2 Thesis overview

In this thesis, we study the $X_{\mathrm{CO_2}}$ frontal gradient with data from OCO-2 to compare what was seen in the studies with TCCON data by Parazoo et al. (2008), and S.Pal et al in ACT-America to quantitatively evaluate OCO-2 data in the CONUS scale. All of these studies found gradients of ~10 ppmv in summer.

We will compare WRF-VPRM simulations with the observations from OCO-2 in terms of land cover types and different seasons, preliminary evaluations of factors of simulation biases are also included. This thesis has 4 chapters: Chapter 1 gives a brief introduction for global carbon cycle, OCO-2, and WRF-VPRM. Chapter 2 describes Data and Methods, where detailed OCO-2 data are discussed, and the method to locate and define a cold front on a weather map, satellite orbiting map and potential temperature at 700 hPa. This chapter also describes how we compute the $X_{\mathrm{CO_2}}$ gradient across a cold front and the method for significance test for frontal gradients. Chapter 3 gives results of this research and Chapter 4 incorporates discussion and summary.

Chapter 2 : Data and Methods

2.1 Detailed discussion of OCO-2 data

The Orbiting Carbon Observatory 2 (OCO-2), launched in 2014 by NASA (The National Aeronautics and Space Administration), flies in a sun-synchronous, near-polar orbit over a 16-day (233-revolution) repeat cycle and crosses the equator at about 1:30 PM Mean Local Time (MLT) (https://oco.jpl.nasa.gov/mission/quickfacts/). OCO-2 does not measure $\mathrm{CO_2}$ directly, but the intensity of the sunlight reflected from the presence of $\mathrm{CO_2}$ in the column of air. The OCO-2 level 2 data product includes column-averaged carbon dioxide ($\mathrm{CO_2}$) dry air mole fraction ($X_{\mathrm{CO_2}}$), surface pressure, surface-weighted estimates of the column-averaged water vapor, atmospheric temperature and other diagnostic products from mostly cloud-free scenes. $X_{\mathrm{CO_2}}$ is the primary product delivered by OCO-2, and is defined as the average concentration of carbon dioxide in a column of dry air extending from Earth's surface to the top of the atmosphere. Estimates of CO_2 are derived by taking the ratio of the column integrated number densities of carbon dioxide and molecular oxygen along the optical path between the Sun, the surface footprint, and the instrument, and then multiplying these results by the global mean column-averaged oxygen concentration (0.20935) (function is given in Equation (9)).

$$X_{\mathrm{CO_2}} = \frac{\mathrm{CO_2^{col}}}{(\mathrm{O_2^{col}}/\mathrm{O_2^{mf}})}, \quad (9)$$

where $\mathrm{CO_2^{col}}$ is retrieved absolute CO_2 column (in molecules/cm2), $\mathrm{O_2^{col}}$ is retrieved absolute O2 column (in molecules/cm2), and assumed (column-averaged) mole fraction of O_2 used to convert the O_2 column into a corresponding dry air column and is equal to 0.20935 (Schneising et al. 2008).

The OCO-2 spacecraft carries a single instrument that incorporates 3 spectrometers collecting high spectral resolution spectra of reflected sunlight, one in the molecular oxygen (O_2) A band, centered near 765 nm, and other twos in the CO_2 bands centered near 1610 and 2060 nm. Collecting 24 spectra per second, each spectrometer yields about a million observations each day over the sunlit hemisphere. Radiance measurements are made from the three spectrometers and then produce the estimates of X_{CO_2} with a "full-physics" retrieval algorithm, with O_2 A band and CO_2 bands as prescreeners before processing (O'Dell et al. 2018). Two prescreeners are applied: the first is A-band Preprocessor (ABP), which uses the oxygen A band and assumes no clouds or aerosol throughout the detection. This pre-processor identifies where surface pressure differences between the retrieved and a priori greater than 25 hPa are and screens these scenes assuming cloud and/or aerosol contamination. The second is the Iterative Maximum A Posteriori-Differential Optical Absorption Spectroscopy preprocessor (IMAP-DOAS, or IDP), which performs fast, clear-sky fits to the two CO_2 bands centered near 1610 and 2060 nm separately, and screens based on excessive differences in retrieved CO_2 between the two bands. Clouds and optically thick aerosols preclude observations of the full atmospheric column in many regions, especially where there is a weather phenomenon like fronts, storm and snow. Retrievals of X_{CO_2} also fail when the solar zenith angle is too high, or when there are is low surface albedo such as in the case of snow and ice, which causes the Signal to Noise Ratio to be too low. The fraction of soundings passing in the tropics is larger than at higher latitudes mainly due to sub-solar latitudes—regions between 23.5°S and 23.5°N are brighter than surface in higher latitudes, where surface albedos are stronger (O'Dell et al. 2018) and solar zenith angles are smaller. With this approach, over 100,000 cloud-free full-column $X_{\mathrm{CO_2}}$ observations are collected by OCO-2 each day.

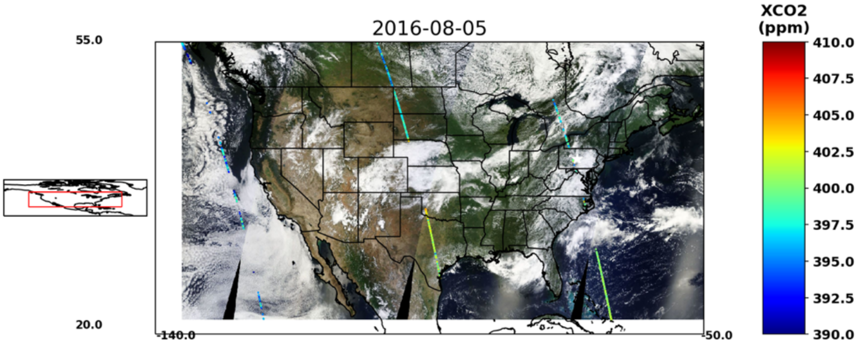


Figure 2.1 - An example for OCO-2 tracks in August 5th, 2016. The satellite flies from east to west, south to north. The colors of the soundings vary with $X_{\mathrm{CO_2}}$.

In this study, we use Version 8 $X_{\mathrm{CO_2}}$ from OCO-2 Level 2 (L2) dataset (http://oco2.gesdisc.eosdis.nasa.gov/data/s4pa/OCO2_DATA/OCO2_L2_Lite_FP.8r/) (O'Dell et al. 2018) from 2015 to 2017. Before Version 9, which is released in October, 2018, Version 8 (also referred to as build 8 or B8) is the latest version of OCO-2 data after version 7, which has the lowest biases and highest throughput of any version so far. It has been processed with an updated version of the OCO-2 Level 1 (L1) and Level 2 algorithms. The latest versions of the algorithms include updated radiometric calibration for the L1b product, updated spectroscopic parameters, the addition of a stratospheric aerosol type, and a more realistic treatment of surface reflectance in the L2 algorithm. For large scale measurements, a previously reported positive bias with respect to models over southern hemisphere mid-latitude oceans is nearly removed in B8, while regionally coherent biases still remain at a significant level (~1 ppmv) (O'Dell et al. 2018). Data in X_{CO_2} is marked with quality flag including number 0 and 1. X_{CO_2} marked with number 1 are determined as bad data due to reasons like ‘calibration door not open for science’, ‘invalid attitude data received’, ‘algorithmic errors detected in the frame’… and those marked with number 0 are in good quality ('xco2_quality_flag’=0). Bad quality data are removed when analyzing X_{CO_2} gradients across cold fronts in our study.

2.2 The method to locate and define a front

3-hourly weather maps of surface analysis, mostly at 18UTC or 21UTC from Weather Prediction Center in NOAA (Figure 2.2) are applied to locate cold fronts. In this example at August 5th, 2016, the cold front was over Nebraska and Kansas, where it was cloudy according to the satellite map and there was cold front sign in the surface analysis. Because OCO-2 overpass times range from 15 UTC over the western Atlantic Ocean to 21 UTC over west coast of the U.S. on each day. We examine the weather map at 18UTC or 21 UTC over most regions in U.S.: reading time information at each sounding along tracks over U.S., find the closer time between 18UTC and 21UTC. Here we also make use the “OCO-2 MODIS Vistool” (https://github.com/hcronk/oco2_modis_vistool) developed by Heather Cronk from Colorado State University, which is to pull MODIS Aqua RGB images from Worldview using the NASA GIBS API and overlay various OCO-2 data fields for case study analysis in support of OCO-2 cloud and aerosol screening validation, filtering with warning level or quality flag in OCO-2 Lite files. We zoom in the satellite track images to our interested domain in this study, and retain soundings marked with quality flag ‘0’. In this domain (50˚W-140˚W, 20˚N-55˚N), we can find at most 4 tracks, because the ground track repeat cycle of OCO-2 is 16 day/233 orbit, approximately 25˚ of Longitude (Figure 2.1). We compare the Weather Prediction Center 3-hourly surface analysis with satellite images we produced with the OCO-2 MODIS Vistool(https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/), find the same location where the front is in the surface analysis, and determine when the frontal case meets the requirement that satellite tracks are ahead and behind the fronts.

In this thesis, we utilize an updated version of WRF-VPRM, in which VPRM parameters was calibrated by (Hilton et al. 2013) using covariance tower data over North America, to simulate CO_2 over CONUS with a resolution of 12 km for year 2016 in a continuous run using an optimal downscaling configuration justified in (Miles 1962) and then integrate CO_2 on 47 pressure levels up to 10 hPa to compute $X_{\mathrm{CO_2}}$. CO_2 between 10 hPa to 0hPa (i.e., the top of atmosphere) is neglected because it is less than 0.1 ppmv, which will contribute even less in $X_{\mathrm{CO_2}}$.

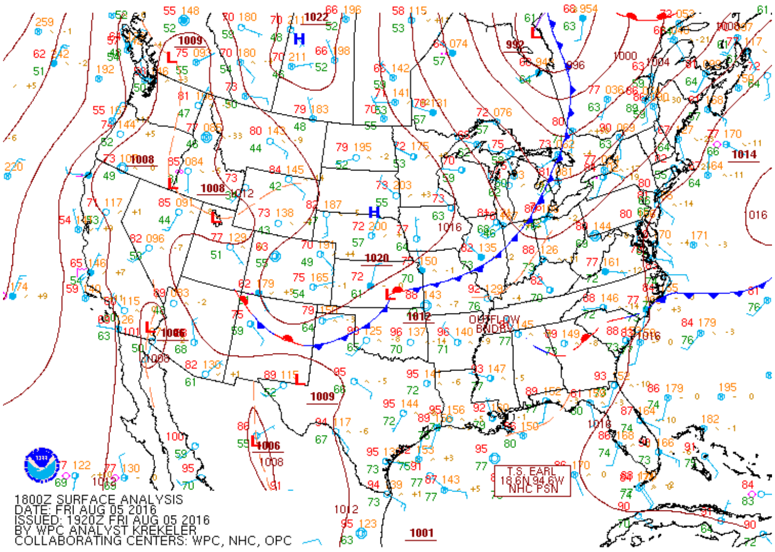


Figure 2.2 - An example for weather prediction center surface analysis in August 5th, 2016, 18 UTC. This surface analysis is used to compare with OCO-2 tracks in Figure 2.1 to locate the cold front.

(https://www.wpc.ncep.noaa.gov/archives/web_pages/sfc/sfc_archive_maps.php?arcdate=08/05/2016&selmap=2016080518&maptype=namussfc)

Frontal zones, usually marked by sharp horizontal gradients and outright discontinuities in wind and temperature, are often observed with a cloud band and precipitation. A potential problem is that $X_{\mathrm{CO_2}}$ frontal anomalies are hidden under clouds for satellite observations (Corbin and Denning 2006). Moisture varies at the frontal area, as seen from the strong gradients in dew point and equivalent temperature, especially when the cold air is of continental origin and the warmer air is of marine origin, as is often the case over the eastern United States. Water vapor in the warm air mass ascends to saturation due to cold air approaching, and is lifted to higher altitudes to form the clouds. The strong temperature gradient and moisture gradient across the front account for clouds over the frontal zone. In the process that we are looking for frontal zones corresponding to instantaneous satellite images, the frontal zones are often characterized by clouds parallel to front lines. However, the 3-hourly surface analysis cannot provide the front location at the exact time when the satellite passing across the front, we consider the band-like cloudy areas near the front line from the satellite image are all frontal zones, and the track to the north of the cloudy area is crossing the cold sector, and similarly the one to the south of the cloudy area is crossing the warm sector.

The potential temperature of a parcel of a fluid at pressure P is the temperature that the parcel would attain if adiabatically brought to a standard reference pressure P_0, usually 1000 millibars (Wallace and Hobbs 2006). The potential temperature is denoted θ and, for a gas well-approximated as ideal, is given by

$$\theta = T\left(\frac{P_0}{P}\right)^{R/C_p}, \quad (10)$$

where T is the current absolute temperature (in K) of the parcel, R is the gas constant of air, and C_p is the specific heat capacity at a constant pressure. $R/C_p = 0.286$ for air.

Potential temperature at 700 hPa over the CONUS (20°N~55°N latitude, 50°W~140°W longitude) from The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) (Gelaro et al. 2017) is used to define the boundaries of cold fronts. The MERRA-2 data product is 3-hourly, instantaneous, at 0.5° latitude by 0.625° longitude resolution with 42 vertical pressure levels. Following the criteria from Wunch et al. (2011), we overlap satellite tracks from OCO-2 on 700 hPa potential temperature map from MERRA-2, and verify the length of fronts both at warm sector and cold sector. Due to the missing OCO-2 X_{CO_2} data at cloudy areas, we find the first soundings from blank area at both sectors (A, B) and the first sounding (C, D) which has 2 K potential temperature difference from the former soundings (A, B) (i.e., θ_c= $\theta_A - 2K$, θ_D= $\theta_B + 2K$, θ_c is the potential temperature at sounding C), take the average of $X_{\mathrm{CO_2}}$ at soundings between the start sounding A (or B) and the end sounding C (or D) . Also 3-sounding running mean is used to "smooth" the soundings — reduce outliers which are apparent larger or smaller than nearby soundings. For example, assuming there are N+1 soundings between A and C, the original soundings of $X_{\mathrm{CO_2}}$ between A and C are:

$$[\mathrm{P_A},\ \mathrm{P_{A+1}}, \mathrm{P_{A+2}}, \mathrm{P_{A+3}}, \mathrm{P_{A+4}}, \ldots, \mathrm{P_{A+N-3}}, \mathrm{P_{A+N-2}}, \mathrm{P_{A+N-1}}, \mathrm{P_{A+N}}], \ (11)$$

where $\mathrm{P_A}$ is $X_{\mathrm{CO_2}}$ at soungding A, $\mathrm{P_{A+1}}$ is $X_{\mathrm{CO_2}}$ at the sounding next to A towards C, $\mathrm{P_{A+3}}$ is $X_{\mathrm{CO_2}}$ at the sounding next to A+1 towards C, similar for $\mathrm{P_{A+4}}$, $\mathrm{P_{A+N-3}}$, $\mathrm{P_{A+N-2}}$, $\mathrm{P_{A+N-1}}$, $\mathrm{P_{A+N}}$. $\mathrm{P_{A+N}}$ is $X_{\mathrm{CO_2}}$ at the sounding C. If applying 3-sounding running mean on that sounding series, (11) turns out to be:

$$[\frac{\mathrm{P_A}+\mathrm{P_{A+1}}+\ \mathrm{P_{A+2}}}{3}, \frac{\mathrm{P_{A+1}}+\mathrm{P_{A+2}}+\ \mathrm{P_{A+3}}}{3}, \frac{\mathrm{P_{A+2}}+\mathrm{P_{A+3}}+\ \mathrm{P_{A+4}}}{3}, \ldots$$

$$\ldots, \frac{\mathrm{P_{A+N-3}}+\mathrm{P_{A+N-2}}+\ \mathrm{P_{A+N-1}}}{3}, \frac{\mathrm{P_{A+N-2}}+\mathrm{P_{A+N-1}}+\ \mathrm{P_{A+N}}}{3}], \ (12)$$

where there are N-1 soundings in (12). $X_{\mathrm{CO_2}}$ at warm sector are operated similarly from (11) to (12) between B and D.

In this way, the moving-averaged mean — the average of elements in (12) $X_{\mathrm{CO_2}}$ between A and C ($\overline{\mathrm{XCO_2}}_{\mathrm{AC}}$) is the average $X_{\mathrm{CO_2}}$ at cold sector and the one between B and D ($\overline{\mathrm{XCO_2}}_{\mathrm{BD}}$) is at warm sector. Then $X_{\mathrm{CO_2}}$ frontal gradient in this case is:

$$\Delta X_{\mathrm{CO_2}} = \overline{X_{\mathrm{CO_2}}}_{\mathrm{BD}} - \overline{X_{\mathrm{CO_2}}}_{\mathrm{AC}}, \ (13)$$

where $\Delta \mathrm{XCO_2}$ is the $X_{\mathrm{CO_2}}$ frontal gradient, $\overline{X_{\mathrm{CO_2}}}_{\mathrm{BD}}$ is the mean $X_{\mathrm{CO_2}}$ at warm sector, $\overline{X_{\mathrm{CO_2}}}_{\mathrm{AC}}$ is the mean $X_{\mathrm{CO_2}}$ at cold sector.

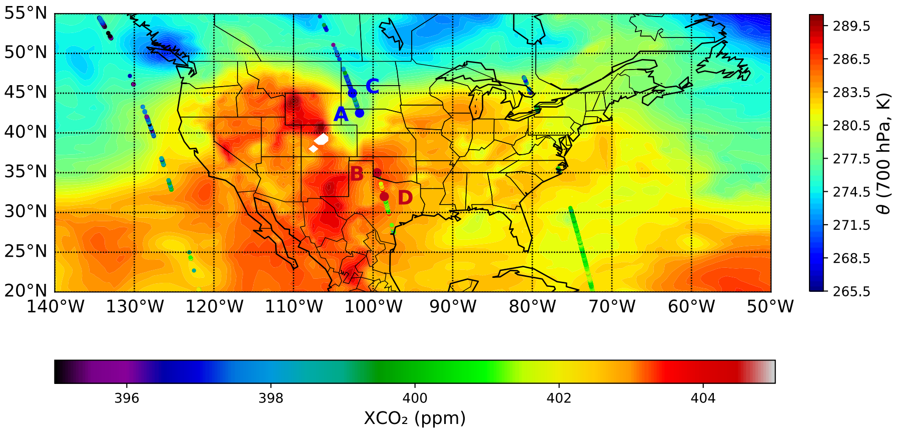


Figure 2.3 - An example for OCO-2 tracks over the potential temperature at 700 hPa. We can find the boundary on this map according to the potential temperature differences along the tracks.

We apply above method in the cases from January 2015 to December 2017, in which there was a cold front and an instantaneous satellite track from OCO-2 over the U.S. and nearby oceans, and classify them into four seasons.

2.3 Significance test

We utilize a statistical method to explore the significance of the frontal gradients over and above the day to day local variability of X_{CO_2} as well as the scatter of the OCO-2 data itself.

A null hypothesis is required in significance tests, and is denoted H_0. The null hypothesis is a general statement or default position that the value of a population parameter is equal to some claimed value. In this research, the null hypothesis is the relationship between X_{CO_2} frontal gradients we collected and the climatology. The method we use to attain the climatology is: (1) Compute the mean length of gap between warm and cold sector (i.e., where the cloud band is and data are not available) and the length of warm sector and cold sector in latitudes; (2). Randomly select an orbit when there is not a front if the length of that orbit is longer than [length of gap + length of warm sector + length of cold sector], and compute the X_{CO_2} “frontal” gradient with those lengths; (3). Repeat the random selection and computing for 1500 times for each season.

A significance test (P<0.1) is then applied to disprove the null hypothesis, which is that X_{CO_2} gradients across fronts are no different than non-frontal spatial gradients in X_{CO_2}: if X_{CO_2} frontal gradients lie in top 5% or bottom 5% of random samples, they are statistically significant. A two-tailed test will test both if the X_{CO_2} frontal gradient is significantly greater than non-frontal gradients and if it is significantly less than non-frontal gradients. To test whether the real X_{CO_2} frontal gradients are significantly different from the non-frontal gradients, we apply two-tailed significance test with a significance level of 0.1 (P<0.1). Details are as follows: We randomly select OCO-2 tracks over the CONUS when there is not a front along the selected satellite track, and then randomly set a gap with the length of the mean length of collected cases according to the real weather map before and find the boundary of both sectors according to our ‘2K-Criterion’, then take the average of both sides of X_{CO_2}, and the difference as well. For each season, we choose 1500 random samples and sort them from small to large. The X_{CO_2} frontal gradient is considered significant if the test statistic is in the top 5% (top 75 samples) or bottom 5% (bottom 75 samples) of its probability distribution, resulting in a P-value less than 0.1. Test when P<0.3 can also be applied to find which one interpreting better.

Another method to test the significance is that we calculate the mean $\bar{x}$ and standard deviation s of 1500 sampling assumed frontal gradient for each season, and consider gradients greater than $\bar{x} + s$ or less than $\bar{x} - s$ as significant and those falling in between as insignificant. These two methods are equivalent for a Gaussian distribution of samples.

The statistical results can be qualitatively compared with cases collected by flight in ACT-America mission. The main concerns are: (1) $X_{\mathrm{CO_2}}$ frontal gradients are significantly different from assumed gradients in non-frontal orbits; (2) $X_{\mathrm{CO_2}}$ frontal gradients in summer are greatly positive, while other seasons' are slightly positive or slightly negative; (3) $X_{\mathrm{CO_2}}$ frontal gradients in summer are statistically significantly different from other seasons; if all three of these criteria are met, we can consider $X_{\mathrm{CO_2}}$ from OCO-2 are precise on synoptic scale and mesoscale and we can verify whether WRF-VPRM simulates over the CONUS well or not based on OCO-2 data.

The assumed 'frontal' gradients of non-frontal orbits over the CONUS from 2015 to 2017 are applied to calculate the climatology and be compared with real $X_{\mathrm{CO_2}}$ frontal gradients.

2.4 Evaluate WRF-VPRM with OCO-2 $\mathrm{X_{CO_2}}$

2.4.1 Evaluate on seasonal time scale

Both monthly mean of $\mathrm{X_{CO_2}}$ from OCO-2 and WRF-VPRM are gridded into 1° Latitude×1° Longitude, and difference between WRF-VPRM and OCO-2 refer as the model bias, in which OCO-2 soundings correspond to instantaneous model outputs. The biases for each season are evaluated as follows: based on land cover types, compute the difference between X_{CO_2} from OCO-2 and WRF-VPRM. Evaluate the bias of WRF-VPRM X_{CO_2} on each surface type. For example, over the U.S. domain, (1). Grid X_{CO_2} from OCO-2 and WRF-VPRM in January into 1° Latitude×1° Longitude, and acquire time from soundings measured by OCO-2 on each grid box; (2). Match WRF-VPRM X_{CO_2} grid box with gridded X_{CO_2} from OCO-2 according to time; (3). Take the difference of WRF-VPRM X_{CO_2} grid box in (2) and OCO-2 X_{CO_2} grid box in (1); (4). Divide grid boxes in (3) with respect to land covers (shown in Figure 2.4), evaluate on seasonal time scale. Methods are the same for other land cover types in other months. Months can also be combined to evaluate WRF-VPRM seasonally. Land cover classifications are plotted according to MODIS data cover dataset.

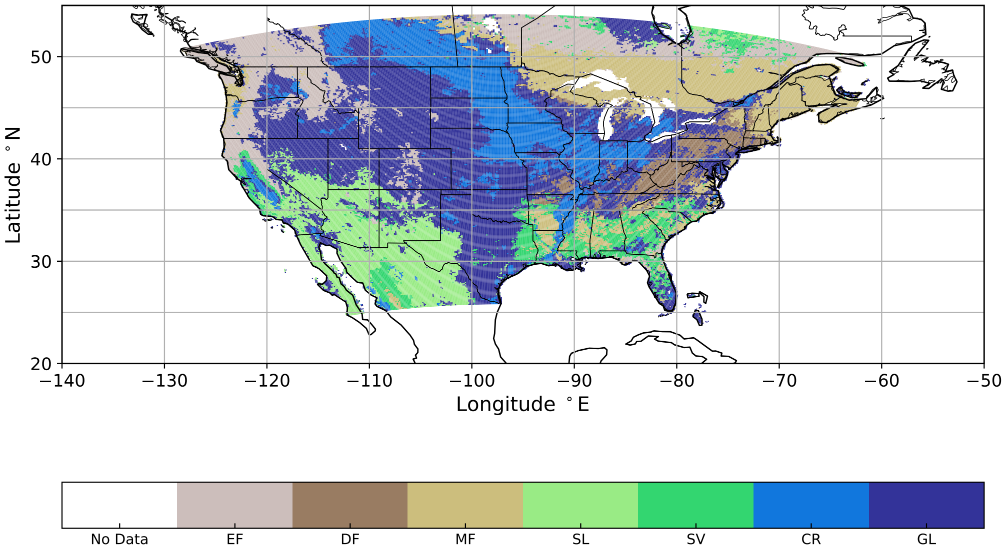


Figure 2.4 - The land covers over the CONUS. There are seven types: Evergreen Forest, Deciduous Forest, Mixed Forest, Savannah, Shrubland, Cropland, Grassland.

Overall, once confirm the OCO-2 X_{CO_2} is an appropriate measurement via comparing X_{CO_2} frontal gradients with previous work done in ACT-America, X_{CO_2} from OCO-2 can be seen as a metric to evaluate how well the WRF-VPRM simulates over the CONUS.

2.4.2 Separate factors of NEE

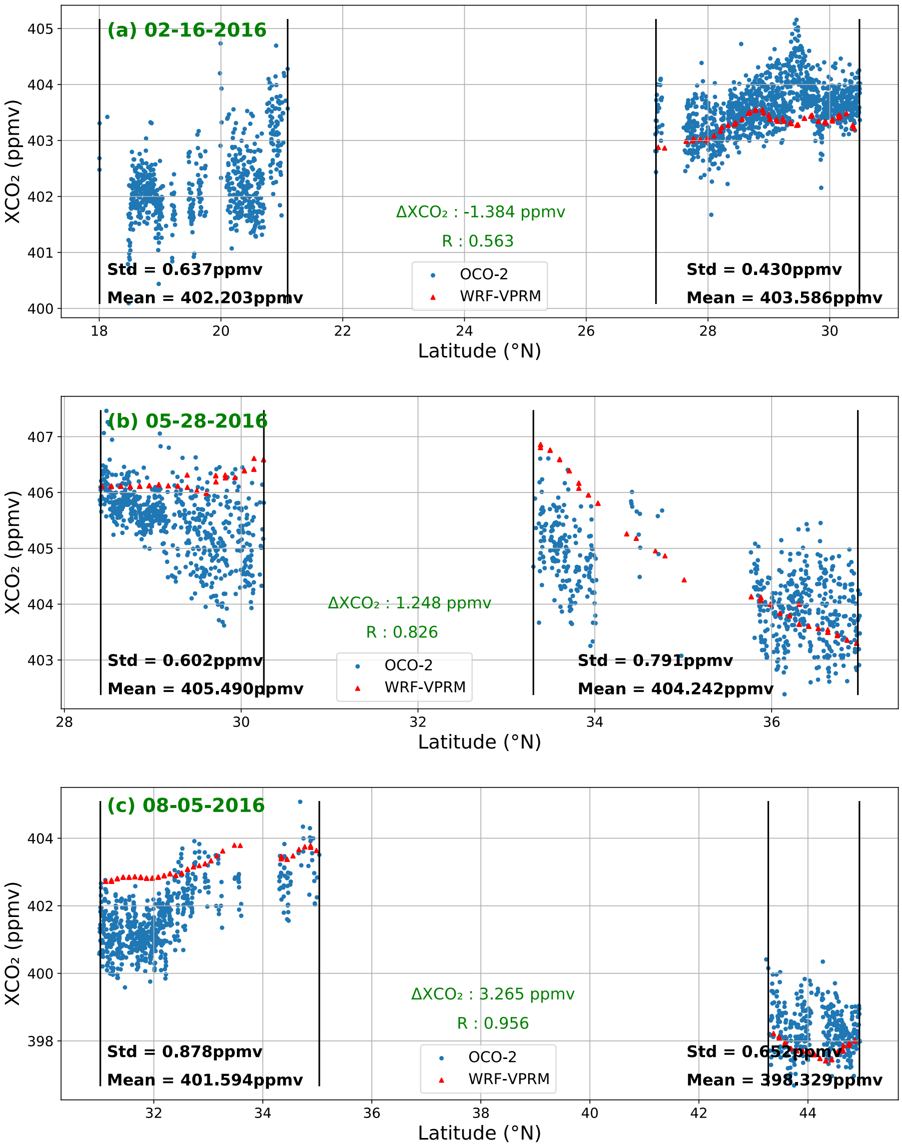
As stated in Equation (1) - (3), the estimation of NEE is the sum of GEE and Ecosystem Respiration. we can extend WRF-VPRM analysis using OCO-2 monthly or seasonal SIF (Solar-Induced chlorophyll Fluorescence) (Baker 2008; Baret et al. 2007; Meroni et al. 2009; Papageorgiou 2007) data correlations to isolate Gross Primary Production (GPP) and Ecosystem Respiration (ER) parameterizations, and continue to refine WRF-VPRM for regional scale study, where GPP is the amount of chemical energy as biomass that main energy producers of ecosystem create in a given length of time (https://en.wikipedia.org/wiki/Biomass_(ecology)), which equals to $(-1) \cdot \mathrm{GEE}$. SIF, which can be directly measured from space, can serve as a physiological indicator or proxy for GPP (Cui et al. 2017). SIF is related to vegetation growth, which is greater when there is a stronger vegetation growth and corresponding greater GPP (or $(-1) \cdot \mathrm{GEE}$). We can preliminary compare SIF from OCO-2 (at 757 nm) with GPP simulated by WRF-VPRM to evaluate the two factors in Equation (1). For example, when there is a strong bias in a specific season between WRF-VPRM and OCO-2, and SIF from both has a strong correlation, the bias may be caused by ecosystem respiration simulation.

Chapter 3 : Results

3.1 X_{CO_2} gradients across cold fronts

3.1.1 Cases and significance tests

We collect 83 cases from 2015 to 2017 in which a cold front was observed by OCO-2: 22 in spring, 31 in summer, 16 in fall, and 14 in winter. Per year, there are 17 frontal cases in 2015, 37 in 2016, 29 in 2017.



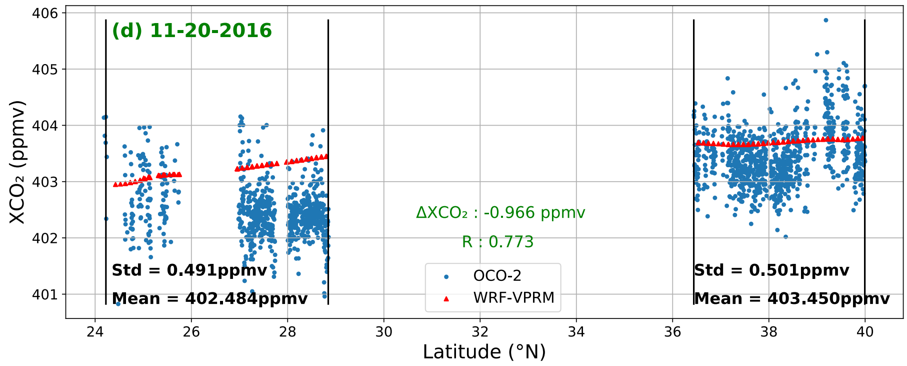


Figure 3.1 - Four X_{CO_2} frontal gradient in seasons. Each panel has three sections separated by black vertical lines: the first section is the warm sector in a cold front, blue dots in this section are the satellite soundings in warm sector, and the red triangles are corresponding WRF-VPRM simulations along the track. Similar for the third section, but in the cold sector. The blank second section is invalid soundings because of clouds or other reasons. The gradients marked in the figures are the X_{CO_2} in warm sector minus that in cold sector and 'R' is the correlation coefficient between OCO-2 observations and WRF-VPRM simulations. (a) is for Febuary 16th, 2016, (b) is for May 28th, 2016, (c) is for August 5th, 2016, (d) is for November 20th, 2016. (Other days are given in Appendix I.)

The average length of gaps between warm and cold sectors for those cases are 4.959° in latitudes. The average X_{CO_2} frontal gradients for all three years in the spring is -0.2495 ppmv, the fall average is -0.5241 ppmv, -0.7657 ppmv for winter frontal gradients, and +2.0957 ppmv for summer cases. Most cases are at central or eastern U.S.. From Table 3.1, we see that the average of summer X_{CO_2} frontal gradients is much greater than other seasons', and most summer gradients are greater than 2 ppmv (20 in 31), indicating that in terms of average, absolute X_{CO_2} frontal gradients in summer is greater than all other seasons. Also, cases in spring fall and winter span over smaller ranges (-3.020 ppmv to 1.720 ppmv for spring, -3.064 ppmv to 1.592 ppmv for fall, and -2.757 ppmv to 0.536 ppmv for winter), which means X_{CO_2} frontal gradients in spring, fall and winter vary less case by case. The pattern can also be seen in Figure 3.2.

	Case Counts	Average (ppmv)	Standard Error (ppmv)	Median (ppmv)	Min (ppmv)	Max (ppmv)
Spring	22	-0.3835	0.2722	-0.4020	-3.0197	1.7205
Summer	31	2.5788	0.3360	2.5729	-2.3166	6.6914
Fall	16	-0.6075	0.3238	-0.3379	-3.0636	1.5922
Winter	14	-1.2591	0.2094	-1.2736	-2.7570	0.5365
Seasons except Summer	52	-0.6882	0.1698	-0.7600	-3.0636	1.7205
All Seasons	83	0.5320	0.2391	0.1640	-3.0636	6.6914

Table 3.1 - Counts, mean and standard error, median, minimum and maximum of $\boldsymbol{X}_{\mathbf{CO_2}}$ frontal gradients in each season

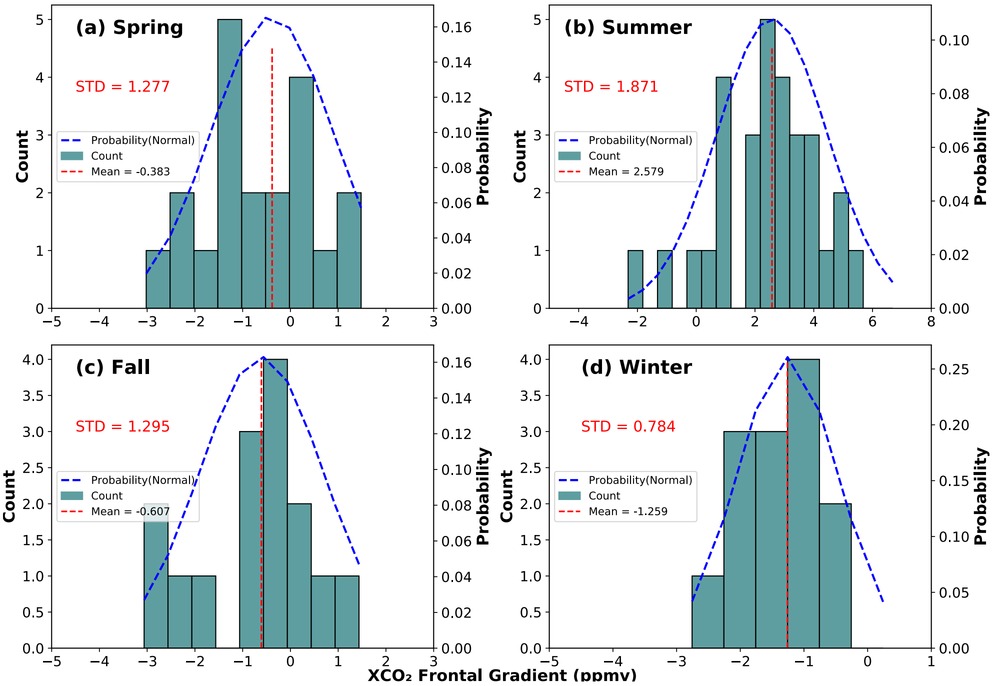


Figure 3.2 - Histograms for $X_{\mathrm{CO_2}}$ gradients across fronts. Bars are counts of $X_{\mathrm{CO_2}}$ frontal gradients, red dashed lines are mean of $X_{\mathrm{CO_2}}$ gradients in each season, and blue dashed lines are normal distribution lines. (a) is for spring, (b) is for summer, (c) is for fall, (d) is for winter.

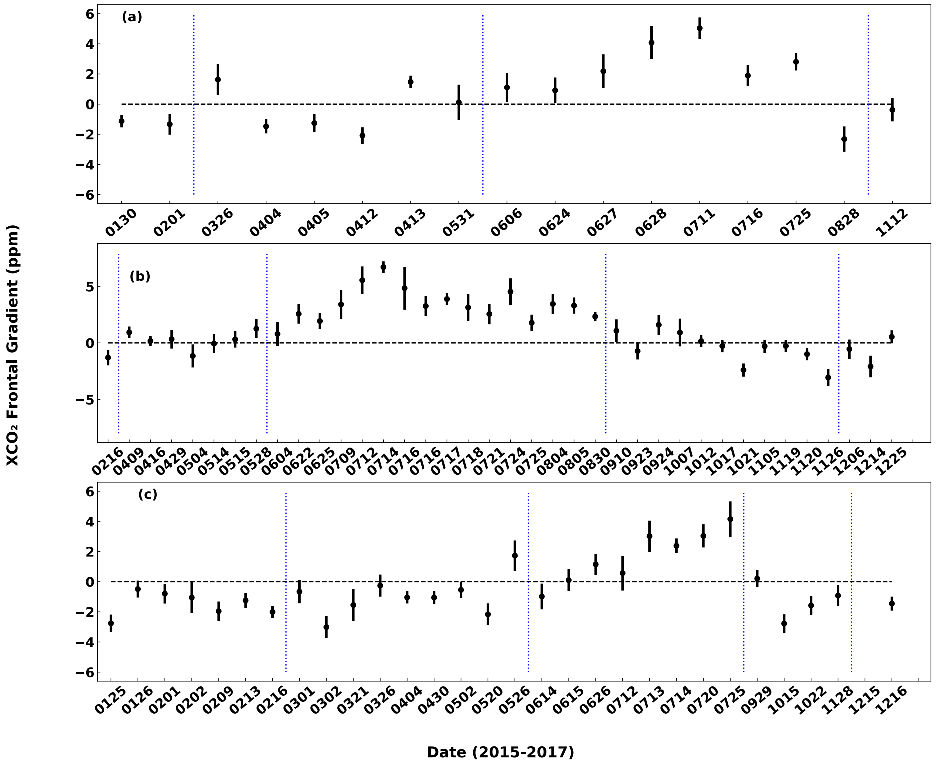


Figure 3.3 - Error bars for single X_{CO_2} frontal gradients from OCO-2. Black dots are X_{CO_2} frontal gradients and the range of black lines are the average of two standard deviations: one at warm sector and the other at cold sector. Seasons are divided by blue dashed lines. (a), (b), (c) are for 2015,2016, and 2017.

However, in spite of the effect from cold fronts, surface fluxes are also another big part of the size of the gradient. For example, the X_{CO_2} OCO-2 observations in January at latitudes more than 40°N and in February at latitudes more than 45°N are rare because data marked with 'xco2_quality_flag'=1 is filtered — large solar zenith angles in higher latitudes in winter resulting in less sunlight for good retrieval, which indicates that most cold front cases we collect locate between $20^{\circ}\mathrm{N} - 40^{\circ}\mathrm{N}$. Smaller the domain we work on, fewer land cover type there will be, so the CO_2 horizontal distribution is less affected by plants' respiration and photosynthesis. For example, in the winter, the land cover types that impact the X_{CO_2} frontal gradients are mainly shrubland in southwest U.S., grassland in central U.S., and savannah in southwest and southeast U.S. However, in the summer, there are more good quality soundings between $40^{\circ}\mathrm{N} - 55^{\circ}\mathrm{N}$, which results in more active biosphere over the domain, for example, lower $X_{\mathrm{CO_2}}$ over corn belt in the mid-western U.S. and higher $X_{\mathrm{CO_2}}$ over shrublands and savannahs in south and central U.S., especially in north Mexico, New Mexico and south Colorado. Also, active plants in the summer take effect on CO_2 distributions, because strong photosynthesis and respiration varies a lot from daytime to nighttime for forests, but for land covers like grassland and shrubland, the CO_2 concentration varies not that much. Since the OCO-2 flies over the equator at about 1:30 pm local time, in the summer, $X_{\mathrm{CO_2}}$ over evergreen forests, deciduous forests, and mixed forests is less than other surface types, i.e., in July (shown in Fig 3.7 (a)), the monthly mean $X_{\mathrm{CO_2}}$ from OCO-2 over forests is ~396 ppmv, while ~409 ppmv over shubland at southwest U.S. The land cover type difference from north to south is also different at east and west: in Figure 3.4 (b), in summer in 2016, from the simulation of $X_{\mathrm{CO_2}}$ from WRF-VPRM, along the longitude 105°W, the gradient is about 10 ppmv from 50° latitude to 30° latitude and along the longitude 80°W, the gradient is about 6 ppmv from 50° latitude to 30° latitude. As shown in Figure 3.4, the $X_{\mathrm{CO_2}}$ gradient in the summer is generally northeast-southwest, approximately from 395 ppmv to 407 ppmv, while $X_{\mathrm{CO_2}}$ south-to-north gradient is smaller in winter. Our frontal cases are located in western and central U.S., where there are smaller seasonal average $X_{\mathrm{CO_2}}$ gradients, but $X_{\mathrm{CO_2}}$ frontal gradients are still significant in these regions.

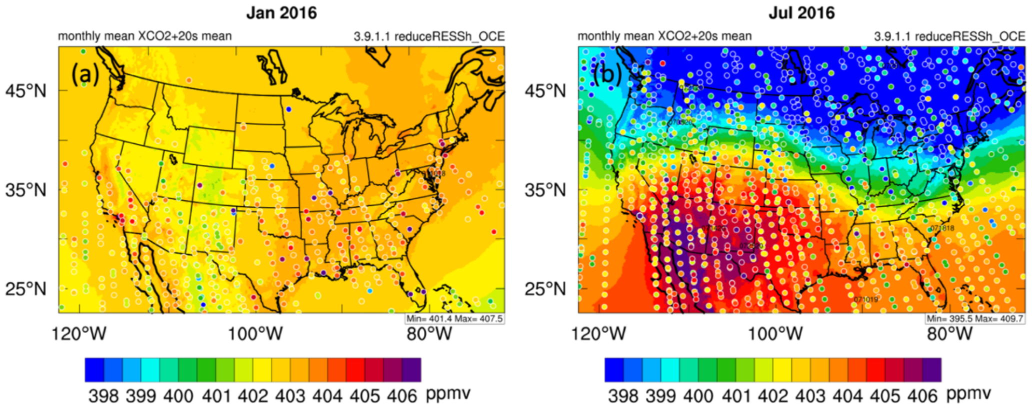


Figure 3.4 – (a) Monthly mean of X_{CO_2} in January 2016 over the CONUS. Filled circles indicate very 20 seconds mean of X_{CO_2} from OCO-2. (b) is for July. This figure is credit to Xiao-Ming Hu.

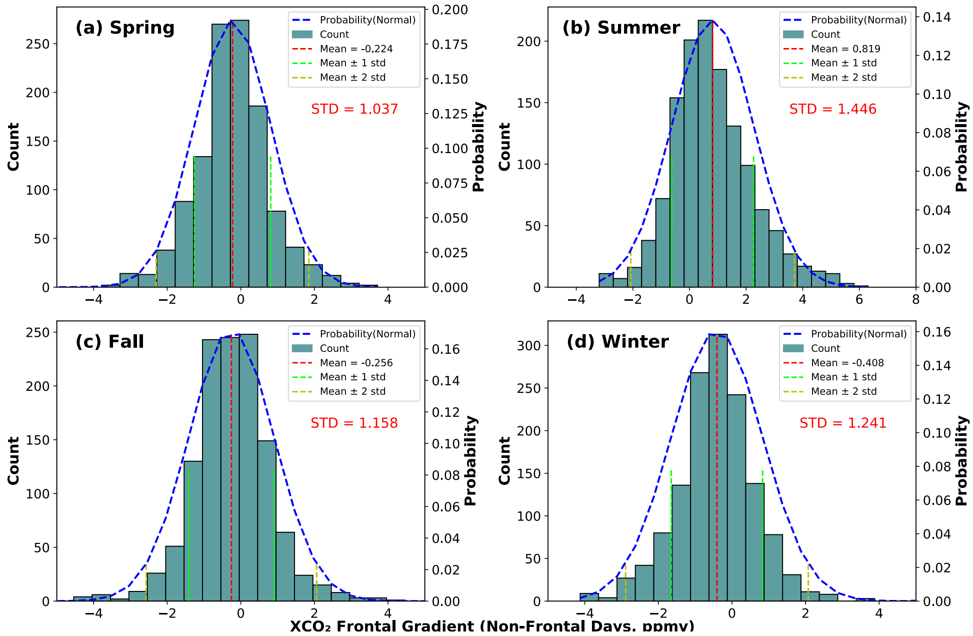


Figure 3.5 - Assumed 'frontal' gradients based on random non-frontal days over the CONUS. (a) is for spring, the blue dashed line is normal distribution, red dashed line indicates the mean of 1500 samples in spring, the left (right) green dashed line is the mean of samples in spring minus (plus) their standard deviation, the left (right) yellow dashed line is the mean of samples in spring minus (plus) their doubled standard deviation; (b) is for summer, (c) is for fall and (d) is for winter.

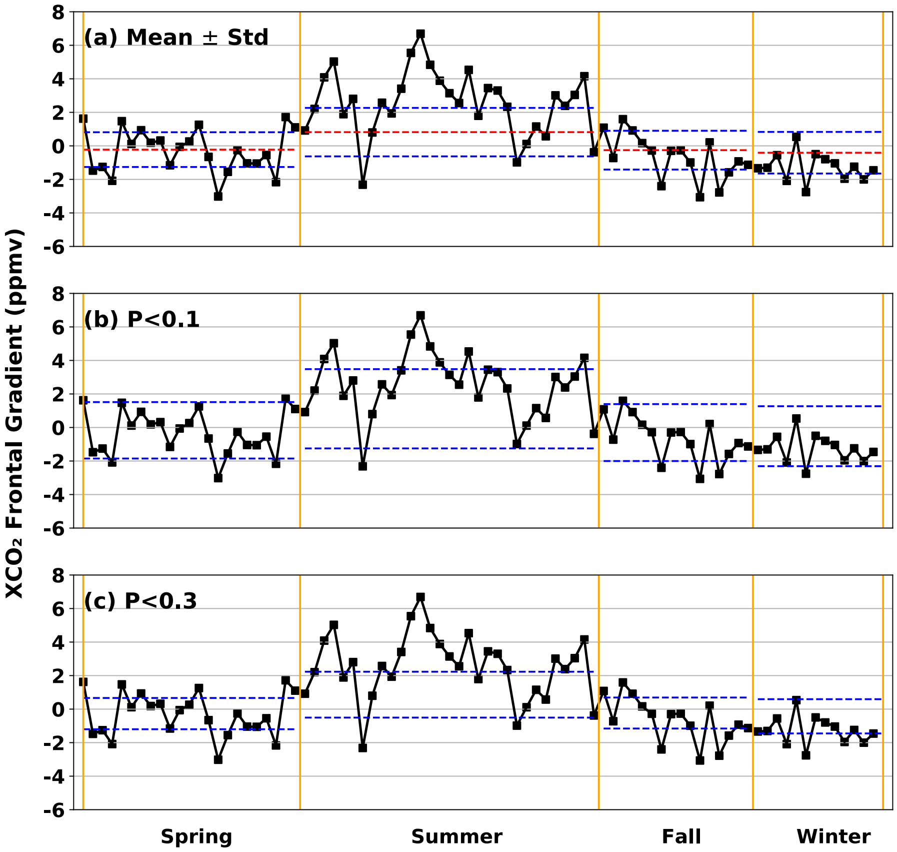


Figure 3.6 - 2015-2017 $X_{\mathrm{CO_2}}$ frontal gradients (black dots) grouped by season. In (a), the mean (red) and a single standard deviation (blue) of assumed frontal gradients on non-frontal days are included. In (b), the upper blue dashed lines for each season is the top 5% value and the lower blue dashed line is for bottom 5% of samples ($P < 0.1$). And (c) is using significance test with $P < 0.3$: cases greater than top 15% samples or less than bottom 15% samples are defined to be significant.

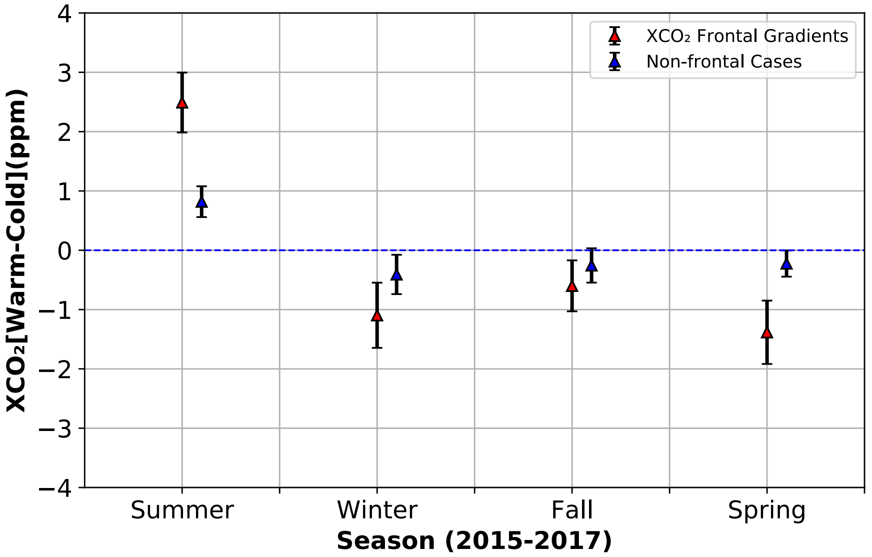


Figure 3.7 - Seasonal mean frontal gradients and associated uncertainties. The red triangle is the average of $X_{\mathrm{CO_2}}$ frontal gradients in specific seasons, and the blue triangle is the computed “frontal” gradients based on non-frontal samples. The error bars on the triangles are averages of standard error of soundings at warm sector and cold sector, where the numbers of cases are: summer: 31; winter: 14; fall: 16; spring: 22.

As we wrote in Section 2.4, we randomly select 1500 non-frontal orbits and calculate the assumed frontal gradients (shown in Fig 3.5). The significance test results are shown in the Fig. 3.6: most samples fall in between mean ± 1 standard deviation. In our random sampling significance test, we compare frontal gradients with computed gradients from non-frontal orbits in Fig 3.6. In the middle panel, $X_{\mathrm{CO_2}}$ gradients across fronts greater than 95% of climatological gradients or less than 5% of them are considered significant. In Table 3.2, in spring, 2 cases are greater than the upper limit, and 3 less than 5% of samples— 5 in 22 in spring are significant. 4 cases in fall are significant: 1 greater than top 5% and 3 less than bottom 5%. Only one in 14 winter cases is significant—less than bottom 5%. For summer, there are 8 in 31 cases are significant: 7 greater than top 5% and 1 less than 5%. Except for winter, other fractions are all about $\frac{1}{4}$. If lower confidence is applied — P value less than 0.3, i.e., gradients falling within 15% or bottom 15% are significant, there are more cases should be considered to be significant: 21 in summer, 11 in spring, 7 in fall and 5 in winter. If we test the significance with the mean and standard deviation of sampled assumed frontal gradients, ie., for each season, cases greater than $\mathrm{\bar{x}} - \mathrm{s}$ or less than $\mathrm{\bar{x}} + \mathrm{s}$ are significant, where $\mathrm{\bar{x}}$ is the mean of 1500 samples and s is the standard deviation in corresponding season. With this testing method, the fractions turn out to be: $\frac{10}{22}, \frac{21}{31}, \frac{7}{16}$, and $\frac{4}{14}$ — only summer cases are greater than 1/2, which is very similar to the results of the method with mean and standard deviation of non-frontal samples. Due to a lot of reasons like local fuel burning, plants' respiration, $\mathrm{X}_{\mathrm{CO_2}}$ retrieval bias, as well as strength of the frontal gradient, the anomalies between climatological warm sector and cold sector may vary between 1 standard deviation of samples. The climatological $\mathrm{X}_{\mathrm{CO_2}}$ gradient is due to atmospheric "standing waves" in $X_{\mathrm{CO_2}}$, which might be the result of land cover difference, polar vortex or other reasons. However, some of the measurements (outstanding gradients greater than 1 standard deviation in each season) have such great difference in $\mathrm{X}_{\mathrm{CO_2}}$ so that it can not be explained by this "standing wave" pattern. In Fig 3.7, the $\mathrm{X}_{\mathrm{CO_2}}$ frontal gradients and their spread in summer and spring out of the range of non-frontal samples, which indicates the $\mathrm{X}_{\mathrm{CO_2}}$ frontal gradients in those two seasons are beyond the average spatial variations of $\mathrm{X}_{\mathrm{CO_2}}$ distributions when there is not a cold front. If compared with CO_2 frontal contrast in ACT-America campaign in Figure 1.5, we can find similar patterns: positive gradients in summer, negative in others, and the magnitudes in summer are about 3 times of any other season's. The thought is that these gradients are caused by land cover difference and large-scale CO_2 transport, which causes climatological X_{CO_2} gradient. But the reason why fronts causing strong X_{CO_2} gradient in summer and weaker ones in other seasons is still unknown.

	Spring	Summer	Fall	Winter
$\bar{x} \pm s$	10	21	7	4
P<0.1	5	8	4	1
P<0.3	11	21	7	5
Total	22	31	14	16

Table 3.2 - Numbers of significant cases with three standards: 1. Cases greater than mean of non-frontal samples plus their standard deviation or less than mean minus standard deviation are significant; 2. Cases in top 5% or bottom 5% of samples are significant (P<0.1); 3. Cases in top 15% or bottom 15% of samples are significant (P<0.3). The first row is seasons. Second row indicates numbers of significant cases with standard 1; third row for standard 2; fourth row for standard 3. The last row indicates the number of all cases collected in each season.

The simulated differences of assumed cold and warm air mass on 3-year monthly climatological X_{CO_2} may vary between ±standard deviation, but those falling outside of ±standard deviation are not reasonable only because of occasional gradient, and we should consider those cases are influenced by other reasons like fronts, or other weather phenomena.

3.1.2 The relationship between the strength of cold fronts and X_{CO_2} gradients

Surface fronts are coincident with cyclones at upper levels, probably evolving from baroclinic waves, which tend to be strongest over the ocean, but can develop over land (Wallace and Hobbs 2006). The strength of the upper-level cyclone is associated with the scale and strength of surface fronts, which is also relative to surface temperature gradient, pressure gradient, wind speed and direction. Considering wind speed, temperature and specific humidity, Parazoo et al. (2008) illustrated that frontal CO2 is related to deformational compression and strong advection along the front, which is also sensitive to locations and seasons. In order to find the relationship between the strength of cold fronts and the magnitude of $X_{\mathrm{CO_2}}$ frontal gradient, we select several cases when there was a strong temperature gradient across the cold front from NOAA weather prediction center's surface analysis. For example, at July 14th, 2016, 18 UTC, the temperature difference between two nearby surface stations near Chicago across the front is 12 K, which is a very strong temperature gradient, whereas the normal temperature gradient is ~5K. This front also caused a strong $X_{\mathrm{CO_2}}$ gradient (7.696 ppmv), which is the largest one in the summer in this three years. Another example is at December 25th, 2016, the temperature gradient is nearly 20 K for two surface stations in Oklahoma, which also came with a surface cyclone in Colorado. Coincidentally, the $X_{\mathrm{CO_2}}$ gradient across that cold front is the only one which is significant in that winter, which may indicate that frontal strength in terms of the temperature gradient is related to the strength in the $X_{\mathrm{CO_2}}$ gradient.

In short, results of above two significance testing method suggest the $X_{\mathrm{CO_2}}$ frontal gradient is generally statistically significant in the summer, but less so in other seasons. Compared with the airborne measurements from ACT-America, seasonal difference of OCO-2 $X_{\mathrm{CO_2}}$ gradients across the frontal have similar pattern: great positive in summer and small negative in other seasons, even though the magnitude is smaller. Thus we can draw the conclusion that the OCO-2 data has sufficient precision to see synoptic scale and mesoscale changes in the atmosphere. In this way, we can evaluate WRF-VPRM simulations using OCO-2 data.

3.2 Evaluating the precision of WRF-VPRM with OCO-2 data

In this section, we aim to evaluate the accuracy of WRF-VPRM with respect to observations from OCO-2 in 2016 based on the difference of gridded monthly mean of $X_{\mathrm{CO_2}}$ from WRF-VPRM and from OCO-2. In Figure 3.5, in July, WRF-VPRM overestimates $X_{\mathrm{CO_2}}$ in areas from 20° to 40° latitudes and underestimates from 40° to 55°. However, few OCO-2 soundings above 40° latitudes are valid after filtered with 'xco2_quality_flag'. In the winter, resulting in less bias grid boxes in winter (like Fig 3.8 (c)), which may be the reason of relatively better simulations. Also, grid boxes of some land cover types reduced because of our method to decide the land cover type of a grid box: Consider the type taking up most fractions in a grid box is the land cover type for this grid box. For example, in Figure 2.4, in central Colorado, there are some regions are combined of Evergreen Forest and Grassland, but when gridding, more fraction of land cover of grassland than Evergreen Forest resulting in more grassland grid boxes in this region. The accuracy of WRF-VPRM varies from regions not only with latitudes, for instance, the model performs better in northwest than in northeast U.S. in July, even at the same latitudes. If land cover types are considered, we find that WRF-VPRM underestimates OCO-2 in the summer over the forested land cover (Fig 2.4, 3.9 and 3.10).

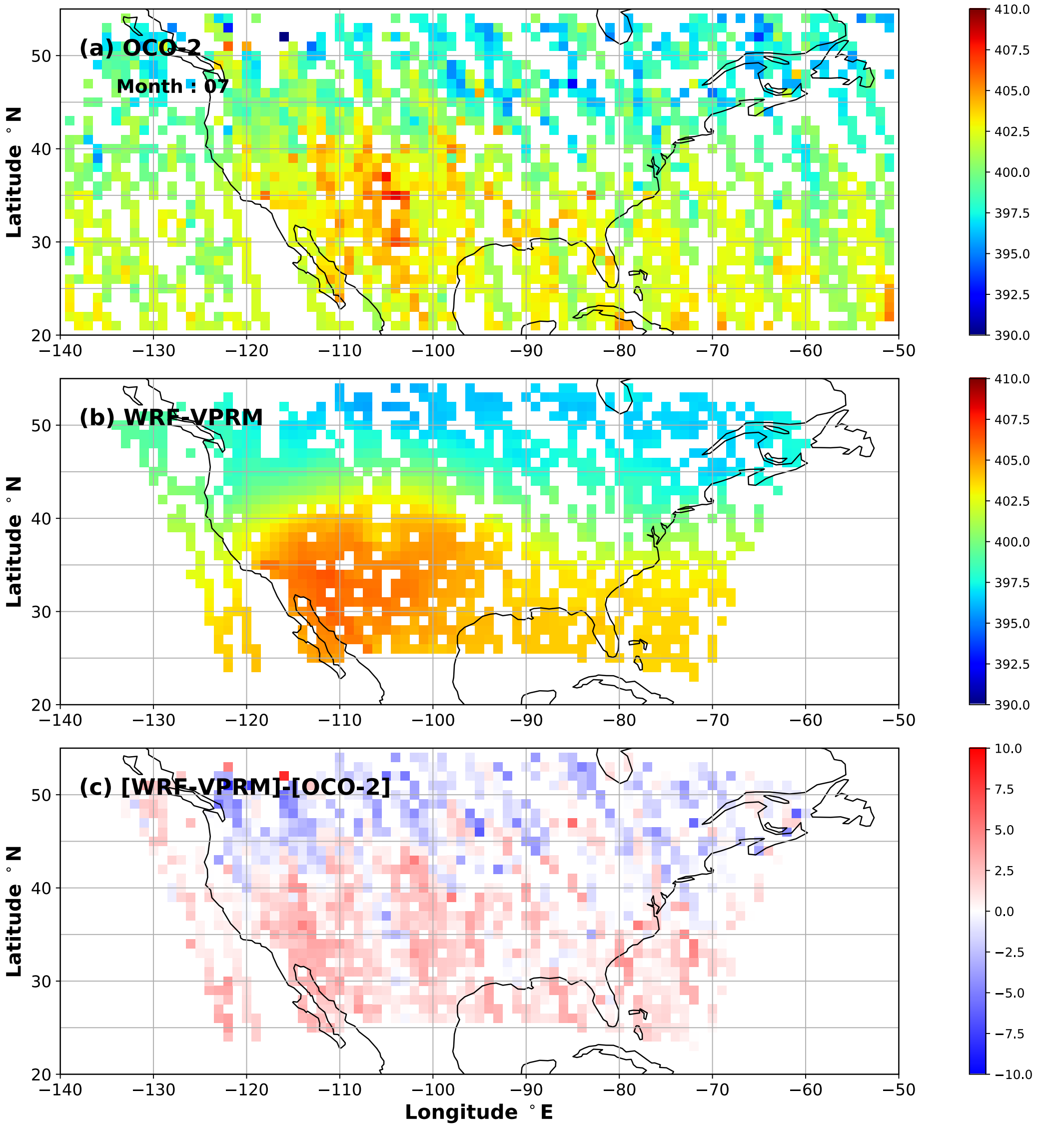
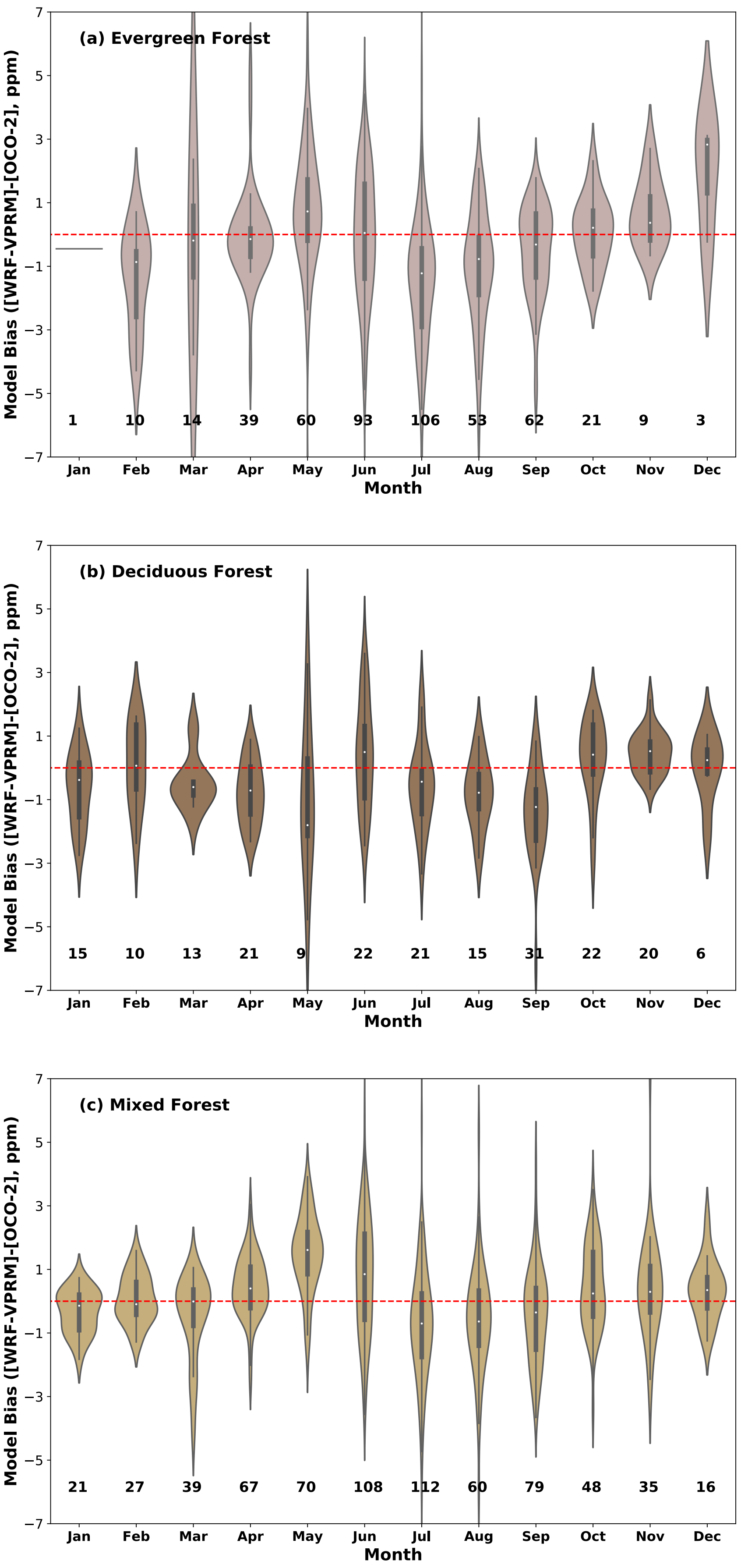
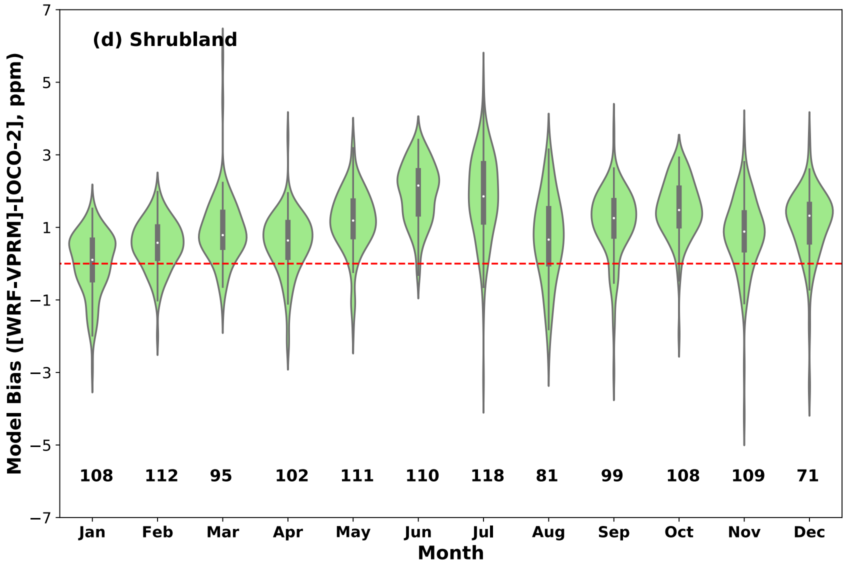
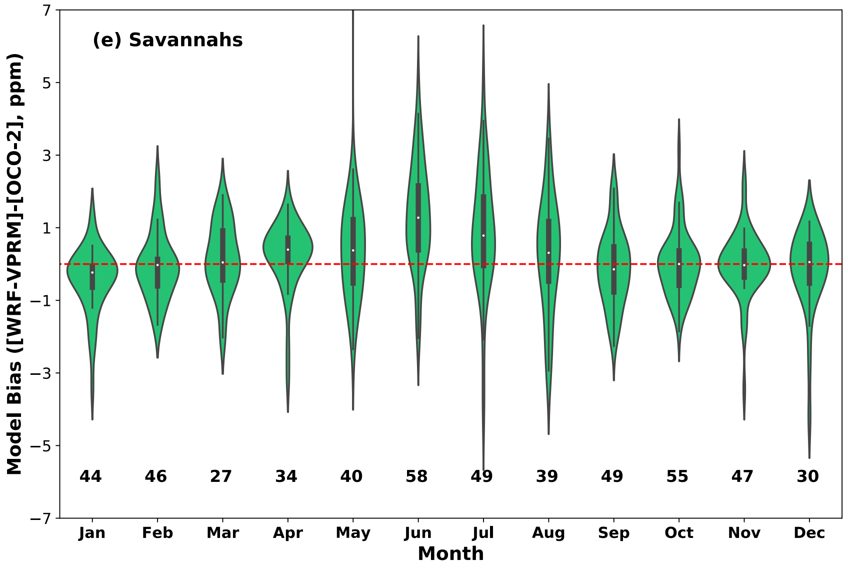
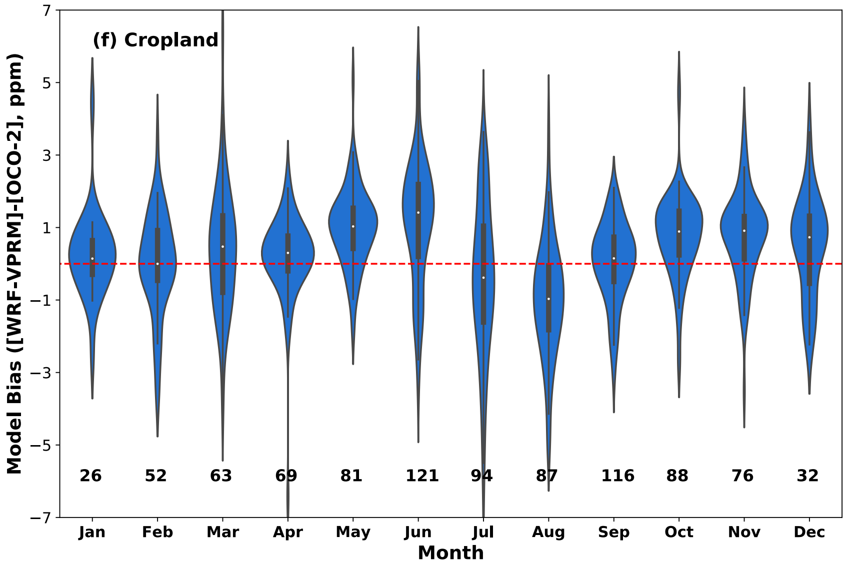


Figure 3.8 - (a) is monthly mean of OCO-2 $X_{\mathrm{CO_2}}$ in 2016 July, filtered with quality flag and gridded into 1°×1°. (b) is corresponding simulations from WRF-VPRM to grid boxes in (a), and each grid box in (b) matches the time of that in (a). (c) is the difference of (b) and (a). (Other months are given in Appendix II.)


(a) Evergreen Forest
(b) Deciduous Forest
(c) Mixed Forest
Model Bias ([WRF-VPRM]-[OCO-2], ppm)
Month
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
7
5
3
1
−1
−3
−5
−7
1
10
14
39
60
93
106
53
62
21
9
3
15
10
13
21
9
22
21
15
31
22
20
6
21
27
39
67
70
108
112
60
79
48
35
16


(d) Shrubland
Model Bias ([WRF-VPRM]-[OCO-2], ppm)
7
5
3
1
−1
−3
−5
−7
108
112
95
102
111
110
118
81
99
108
109
71
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


(e) Savannahs
Model Bias ([WRF-VPRM]-[OCO-2], ppm)
7
5
3
1
−1
−3
−5
−7
44
46
27
34
40
58
49
39
49
55
47
30
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


(f) Cropland
Model Bias ([WRF-VPRM]-[OCO-2], ppm)
7
5
3
1
−1
−3
−5
−7
26
52
63
69
81
121
94
87
116
88
76
32
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

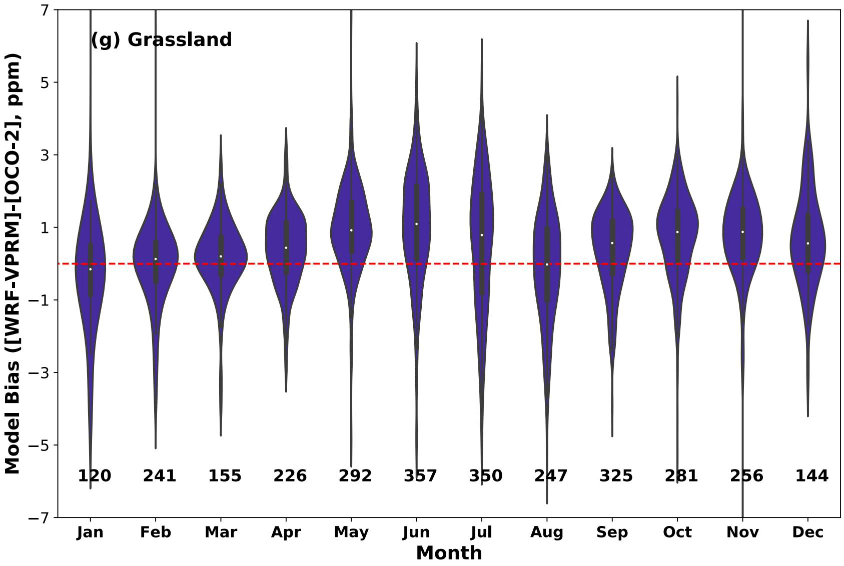


Figure 3.9 - The model bias (X_{CO_2} in WRF-VPRM minus in OCO-2, example is given in Figure 3.7 (c)) with respect to surface types. The box plot in 'violins' indicating median and quartile, the widths of the violins indicate the probability of different bias values and the length is the range of bias including outliers. The numbers at the bottom are the counts of grid boxes in this land cover types.

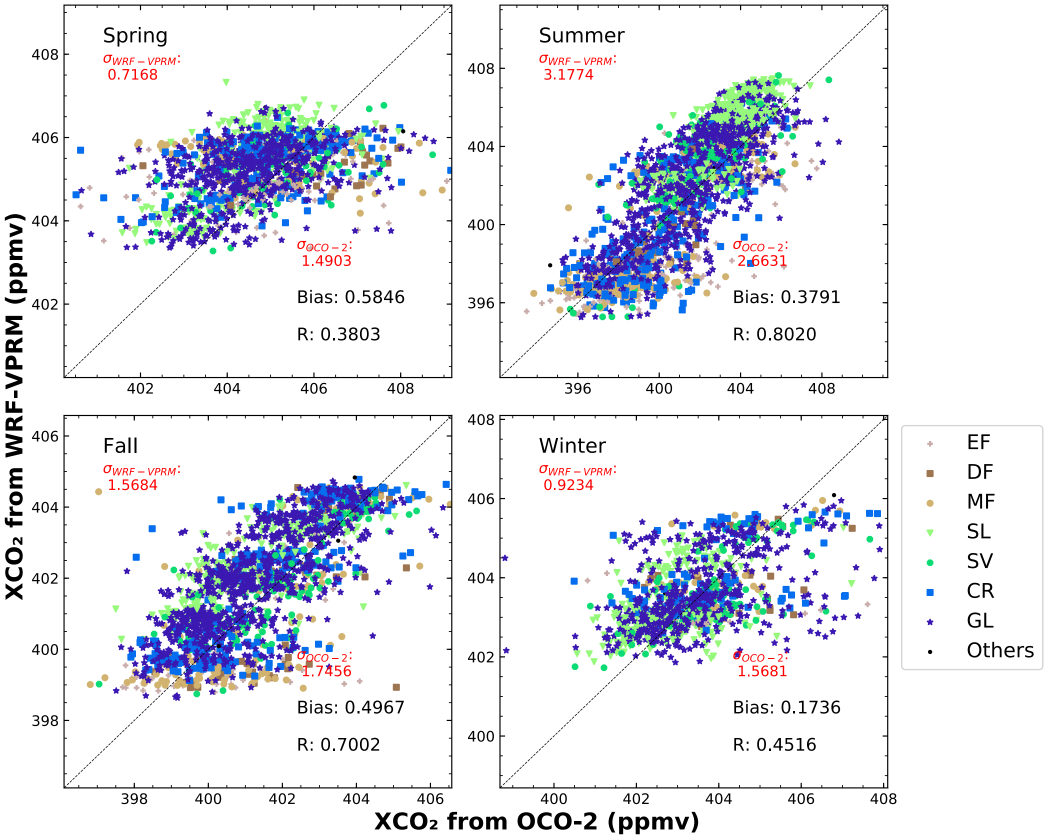


Figure 3.10 - Scatters of bias with respect to season and land cover types. Colors are respect to land cover types in Fig 2.4.

	Spring					Summer				
	Sampl	$\sigma_{WRF-VPRM}$	σ_{OCO-2}	Bias	R	Samples	$\sigma_{WRF-VPRM}$	σ_{OCO-2}	Bias	R
Evergreen	109	0.638	1.519	0.253	0.152	249	2.736	2.818	-0.921	0.737
Deciduous	42	0.460	1.204	-0.629	0.308	58	2.134	1.931	-0.257	0.741
Mixed	176	0.443	1.224	0.645	-0.187	279	2.942	2.683	-0.176	0.784
Shrubland	306	0.910	1.224	0.866	0.678	309	2.086	1.778	1.582	0.821
Savannah	100	0.672	1.169	0.235	0.511	146	2.638	2.354	0.884	0.797
Cropland	211	0.607	1.332	0.545	0.421	302	2.843	2.397	0.047	0.711
Grassland	671	0.698	1.175	0.584	0.422	954	2.794	2.452	0.547	0.791
All	1618	0.716	1.248	0.562	0.402	2299	3.177	2.629	0.376	0.807

	Fall					Winter				
	Sampl	$\sigma_{WRF-VPRM}$	σ_{OCO-2}	Bias	R	Samples	$\sigma_{WRF-VPRM}$	σ_{OCO-2}	Bias	R
Evergreen	92	1.448	1.629	-0.278	0.610	14	0.525	1.825	-0.629	-0.171
Deciduous	73	2.014	1.640	-0.385	0.697	31	0.878	1.329	-0.294	0.481
Mixed	159	1.897	1.974	-0.032	0.707	64	0.900	1.049	-0.002	0.667
Shrubland	314	1.057	1.420	1.173	0.764	291	0.802	0.930	0.471	0.449
Savannah	150	1.395	1.564	-0.014	0.819	119	0.914	1.206	-0.197	0.676
Cropland	279	1.814	1.814	0.486	0.790	109	0.915	1.437	0.163	0.494
Grassland	855	1.432	1.571	0.632	0.735	501	0.933	1.364	0.073	0.437
All	2299	3.177	2.629	0.376	0.807	1131	0.924	1.319	0.132	0.510

Table 3.3 - Standard deviation, bias, and correlation coefficient for each land cover type for seasons in 2016.

Statistics of time-matched $X_{\mathrm{CO_2}}$ data pairs between WRF-VPRM and OCO-2 are calculated for each season in 2016 (Fig. 3.10) over 7 land cover types over the U.S. domain, i.e., evergreen forest (EF), deciduous forest(DF), mixed forest (MF), shrubland (SL), savannah (SV), cropland (CR), grassland (GL) (see the spatial distribution of land covers in Fig 2.4). In Fig. 3.10, each marker stands for data aggregated in a $1^{\circ}\times1^{\circ}$ grid box. Outliers falling outside of 3 standard deviations of biases are removed, which might be due to errors in the OCO-2 data, e.g., contaminated by cloud. Overall, as seen from Figure 3.10, WRF-VPRM performs the best in winter in terms of bias and correlations with OCO-2 data followed by fall, during which seasons $X_{\mathrm{CO_2}}$ spans over larger ranges than in other seasons, especially over cropland (396.57-406.44 ppmv) and grassland (395.26-406.22 ppmv). Particularly, WRF-VPRM simulations are biased low relative to OCO-2 in growing seasons (month 4-9) over forests. For different land categories WRF-VPRM performs the best over cropland and grassland while it overestimates X_{CO_2} during all year long over shrubland which dominates in the southwestern U.S. where temperature is high (Table 3.3). The positive bias of X_{CO_2} over shrubland may be due to model error associated with the current simple parameterization of respiration (eq. (3)), which simply linearly depends on air temperature. High correlation between Gross Primary Production (GPP) simulated by WRF-VPRM and Solar Induced Chlorophyll fluorescence (SIF) from OCO-2 (in Figure 3.10) also suggests that bias in summer is more due to respiration simulation because GPP is simulated well. However, Savannahs are spreading mainly at three regions over the CONUS: southwest U.S., southeast U.S. and northeast U.S. (marked in Figure 3.11) — most grid boxes standing for savannahs (60%-86.67% with respect to month) are located at southeast U.S., Region #1-#3 are marked with red, blue, and green squares. Simulation precisions may vary with not only land cover types, but also locations because to climate and albedo differences. So it is also needed to separate savannahs into different locations because they are far apart. Results are shown in Figure 3.12 Are grid boxes in Region #2 has similar pattern with Figure 3.8 (e): WRF-VPRM overestimates X_{CO_2} in growing season (April to August). In Region #1 at west coast U.S., X_{CO_2} are also overestimated in spring. Grid boxes in Region #3 in July, August and September show that X_{CO_2} is largely underestimated, but uncertainty remains due to few grid boxes in this region. Differences by separating savannahs with respect to locations reveal that WRF-VPRM overestimates X_{CO_2} in grow season in all savannahs, and those in west coast U.S. in spring as well. Estimates for savannahs in north can be taken for granted because too few grid boxes too be considered as statistical results.

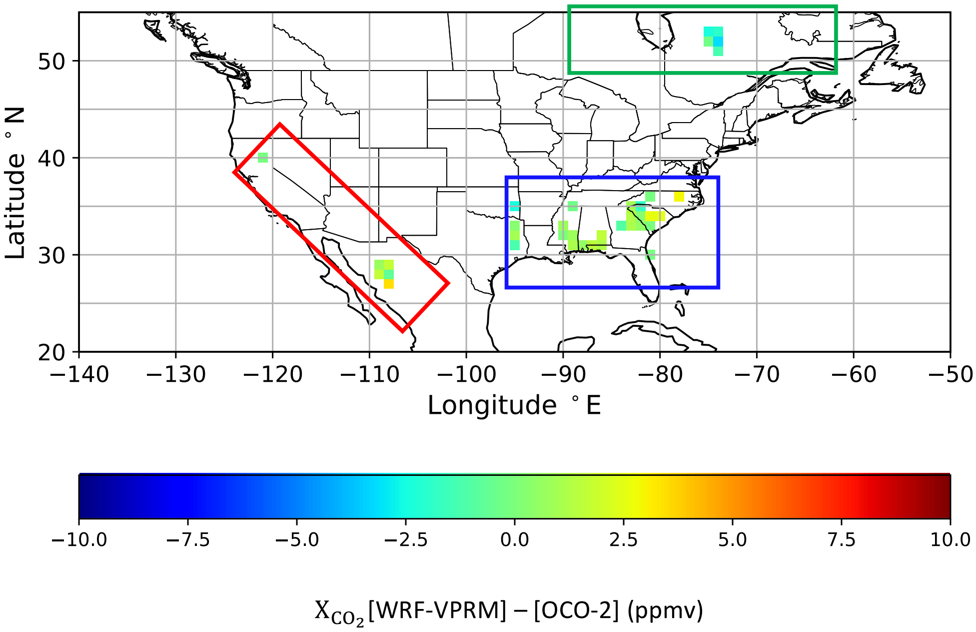
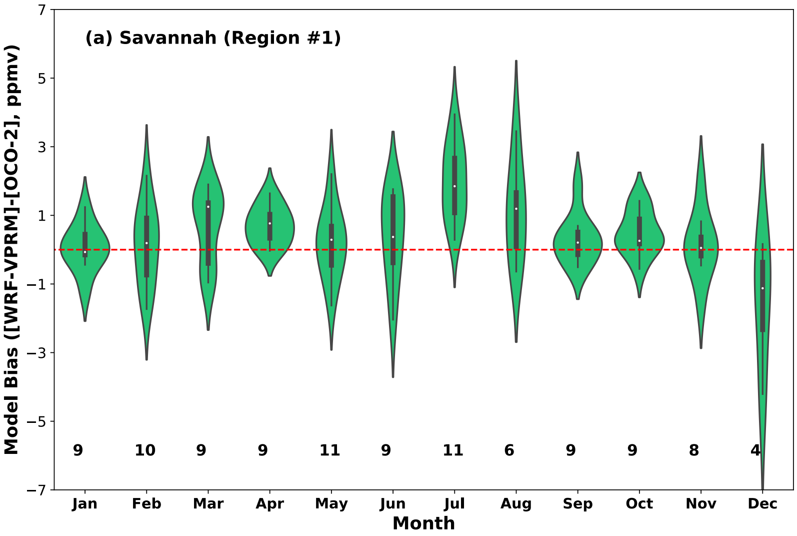


Figure 3.11 - An example to divide Savannah Regions into three (marked with squares) in August, 2016, filtered with quality flag and gridded into 1° Lat ×1° Lon. Each grid box is the average of difference between WRF-VPRM outputs and OCO-2 observations within this grid box. Region #1 is denoted by red square, Region #2 is denoted by blue square, and #3 is denoted by green square.



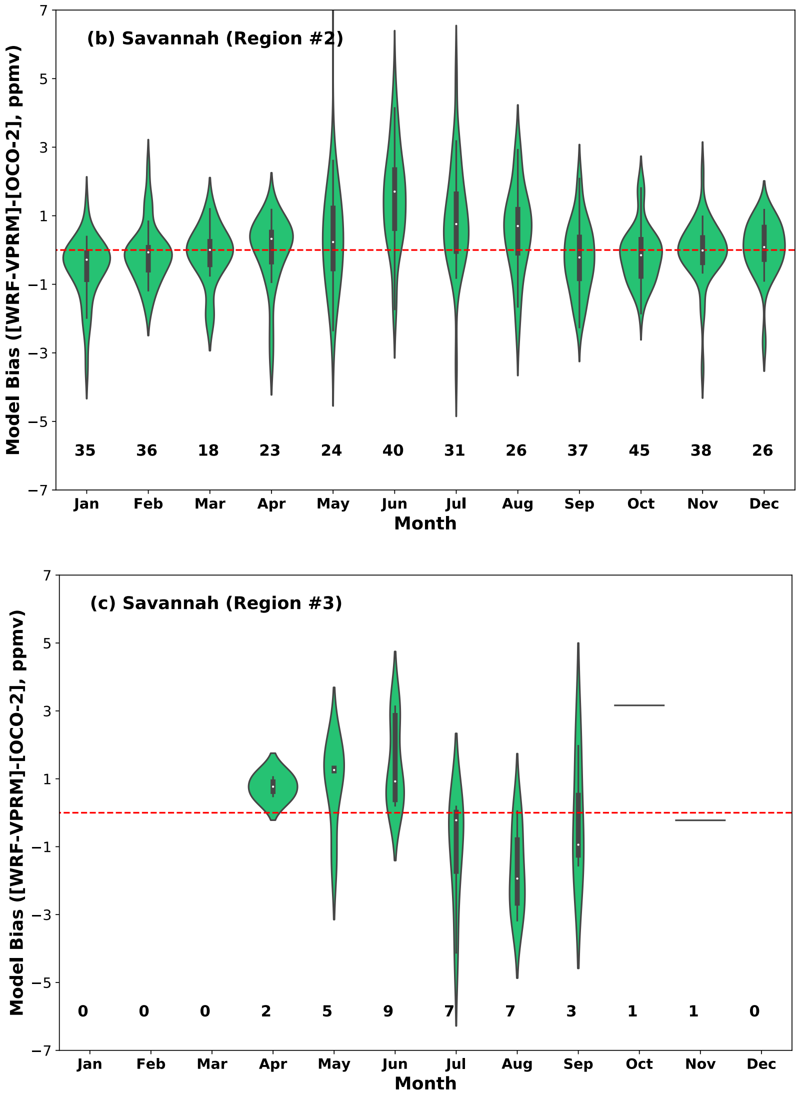


Figure 3.12 - Difference of grid boxes between WRF-VPRM simulations and OCO-2 data on Savannahs in 2016, which is separated by regions marked in Figure 3.11 The numbers at the bottom are the counts of grid boxes in this land cover types.

Generally, the WRF-VPRM simulates X_{CO_2} over the CONUS well in winter — small biases between simulations and observations (as seen from Figure 3.10), but not ideal in summer since the ratios span a greater range in summer. In terms of land cover types, the model overestimate X_{CO_2} all year long over shrubland and in most months except January over grassland, which may be caused by strong respiration or high temperature in Equation (3) and overestimated NEE in Equation (1). In growing season from April to September, WRF-VPRM underestimates $X_{\mathrm{CO_2}}$ in Evergreen Forest, Deciduous Forest, Mixed Forest and Cropland due to strong GPP — dependent on chlorophyll content of the plant.

As stated in Section 2.4.2, we can also separate the factors of model estimation from GEE and ER. Thus we directly compare SIF from measured by OCO-2 and GPP simulated by WRF-VPRM, and results are given in Figure 3.13. Each dot or marker in other shapes stand for one grid box in specific season over specific surface type, and values falling out of mean ± 3 standard deviation (of all grid boxes in one season for each panel) are removed. In Figure 3.13, GPP and SIF in spring, summer and fall have strong correlations, but those in winter not, which indicates that GPP in spring, summer and fall are evaluated well by WRF-VPRM, thus respiration biases are more responsible for overall bias of NEE in these three seasons. For example, $X_{\mathrm{CO_2}}$ is overestimated over the grassland by about 1 ppmv in the summer according to Figure 3.9 (g), however GPP in Figure 3.13 is strongly correlated. Thus respiration over the grassland in summer may be overestimated more, which can also be explained by high temperature and Equation (3). But the correlation of GPP and SIF is bad in winter, which illustrates that GPP estimation in winter can be further improved to modify the overall estimation of NEE.

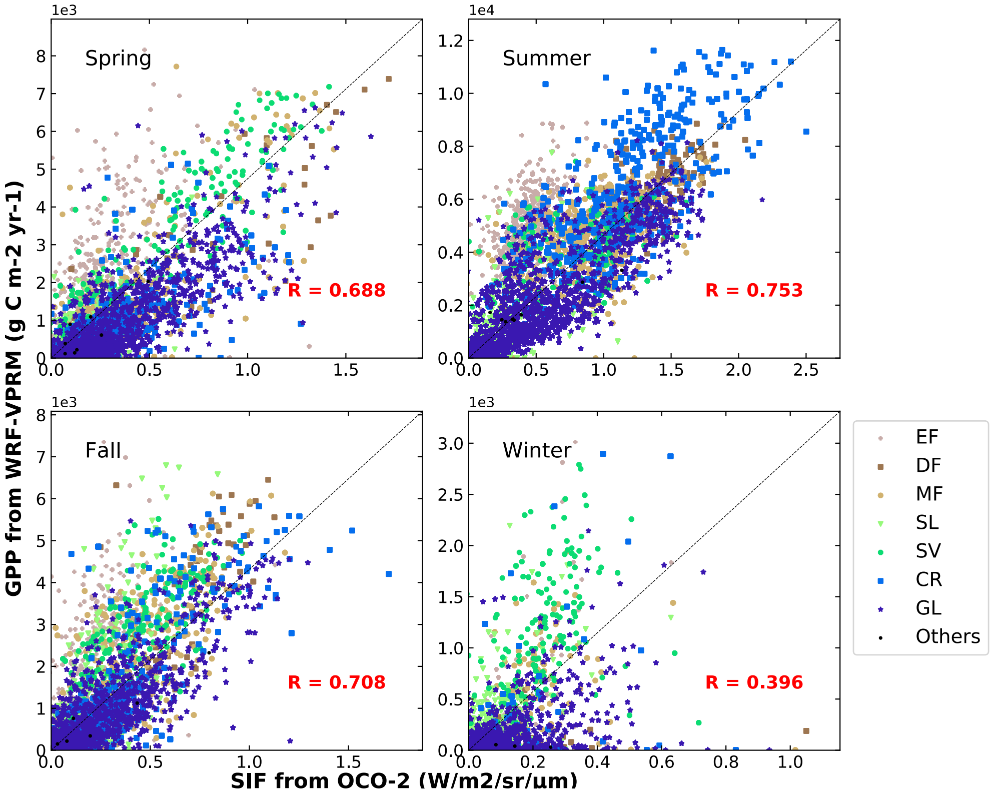


Figure 3.13 - Comparison between GPP from WRF-VPRM and SIF from OCO-2 with respect to season. Colors represent land cover types as in Figure 3.10.

Chapter 4 : Summary and Conclusions

CO_2 gradients across cold fronts in Atmospheric boundary layer was found vary with seasons in a similar qualitative fashion to Parazoo et al. (2008), etc.. In their work, they found the gradients in summer might be 10-30 ppmv more than other season — gradients in spring, fall and winter are slightly positive (~5ppmv) or negative (3-10 ppmv), while those in summer are between 15-30 ppmv, because CO_2 concentration varies less in upper atmosphere, the magnitude of X_{CO_2} gradients across cold fronts may be smaller than the contrast in boundary layer. Whether OCO-2 can detect the variability in a smaller scale like a front was still an open question. Through our work on collecting X_{CO_2} gradients across cold fronts over the CONUS, we find X_{CO_2} from OCO-2 can detect a similar pattern: great positive contrast in summer and small negative contrast in spring, fall and winter.

After testing the significance via random sampling in each season, we can conclude that gradients are statistically significant in summer and spring, but other seasons not, especially in winter. Therefore, it's not sufficient to say that cold fronts are responsible for large scale X_{CO_2} variability for all seasons.

Since the OCO-2 data has sufficient precision to see synoptic scale changes in the atmosphere, we can use OCO-2 to evaluate whether WRF-VPRM simulates X_{CO_2} well on the CONUS scale. With regard to land cover types provided from MODIS, we draw the conclusion that WRF-VPRM simulates generally well over the CONUS in winter — small bias, but for some surface types like shrubland, WRF-VPRM overestimates CO_2 concentration all year long. For all seven land cover types, the bias is high in growing season, when the plants' respiration and photosynthesis change rapidly. Either strong temperature or strong GPP may reduce the quality of WRF-VPRM simulations, because the Ecosystem Respiration is linearly proportional to temperature in the simple model, while in nature respiration may be also relative to CO_2 concentration, water content and light (Davidson et al. 1998). The correlation between simulated GPP and SIF from OCO-2 helps to find whether the GEE or respiration parameterization need be more modified. Results show GPP of shrubland and savannah are overestimated in spring, fall and winter, while overall high bias indicates that respiration bias and fossil fuel emission may be causing strong model bias.

The model simulations can also be verified on a smaller scale: computing the $X_{\mathrm{CO_2}}$ frontal gradients. Compared with $X_{\mathrm{CO_2}}$ frontal gradients collected from OCO-2. WRF-VPRM gradients are of a similar magnitude. After all the simulations have similar bias at both sides of the cold front, for example, WRF-VPRM overestimates $X_{\mathrm{CO_2}}$ by 1 ppm in warm sector, as well as in cold sector, while the overestimates at both sides cancel when calculating the gradients.

Chapter 5 : Future Work

X_{CO_2} and SIF data can help to separate flux and transport and evaluate their bias when simulated by WRF-VPRM. We look forward to improve flux model in VPRM with X_{CO_2} and SIF, and furthermore, coupled with temperature and radiation information, WRF-VPRM will be able to simulate NEE better.

In our definition for the length of cold fronts, they often span about 10° latitudes, in which case the fronts may happen over both ocean and land. Concerning vegetation respiration and photosynthesis, fossil fuel burning and carbon exchange through ocean surface, the gradients of X_{CO_2} may vary if the one sector of the front is over the ocean but the other is over the land. We should take this condition into consideration in our future work.

Weather condition is complicated in the frontal zone, which can help understand the weather changing process and improve weather forecast. While satellite data cannot provide since retrievals cannot be received due to clouds.

References:

Ahmadov, R., C. Gerbig, R. Kretschmer, S. Koerner, B. Neininger, A. Dolman, and C. Sarrat, 2007: Mesoscale covariance of transport and CO2 fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model. *Journal of Geophysical Research: Atmospheres*, **112**.

Baker, N. R., 2008: Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol.*, **59,** 89-113.

Bakwin, P. S., P. P. Tans, D. F. Hurst, and C. Zhao, 1998: Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program. *Tellus B: Chemical and Physical Meteorology*, **50,** 401-415.

Baret, F., V. Houles, and M. Guérif, 2007: Quantification of plant stress using remote sensing observations and crop models: the case of nitrogen management. *Journal of Experimental Botany*, **58,** 869-880.

Bianchi, A. A., and Coauthors, 2009: Annual balance and seasonal variability of sea-air CO2 fluxes in the Patagonia Sea: Their relationship with fronts and chlorophyll distribution. *Journal of Geophysical Research: Oceans*, **114**.

Bluestein, H. B., 1992: *Synoptic-dynamic Meteorology in Midlatitudes: Observations and theory of weather systems.* Vol. 2, Taylor & Francis.

Boutin, J., L. Merlivat, C. Hénocq, N. Martin, and J. Sallée, 2008: Air-sea CO2 flux variability in frontal regions of the Southern Ocean from Carbon Interface Ocean Atmosphere drifters. *Limnology and Oceanography*, **53,** 2062-2079.

Corbin, K. D., and A. S. Denning, 2006: Using continuous data to estimate clear-sky errors in inversions of satellite CO2 measurements. *Geophysical research letters*, **33**.

Crisp, D., C. E. Miller, and P. L. DeCola, 2008: NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. *Journal of Applied Remote Sensing*, **2,** 023508.

Crisp, D., and Coauthors, 2004: The orbiting carbon observatory (OCO) mission. *Advances in Space Research*, **34,** 700-709.

Cui, Y., and Coauthors, 2017: Temporal consistency between gross primary production and solar-induced chlorophyll fluorescence in the ten most populous megacity areas over years. *Scientific reports*, **7,** 14963.

Davidson, E. A., E. Belk, and R. D. Boone, 1998: Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. *Global change biology*, **4,** 217-227.

Davis, K. J., and Coauthors, 2017: Progress toward improving regional atmospheric inversions using airborne measurements: Results from ACT-America. *AGU Fall Meeting Abstracts*.

Feng, S., and Coauthors, 2016: Los Angeles megacity: a high-resolution land–atmosphere modelling system for urban CO 2 emissions. *Atmospheric Chemistry and Physics*, **16,** 9019-9045.

Geels, C., S. Doney, R. Dargaville, J. Brandt, and J. Christensen, 2004: Investigating the sources of synoptic variability in atmospheric CO2 measurements over the Northern

Hemisphere continents: a regional model study. *Tellus B: Chemical and Physical Meteorology*, **56,** 35-50.

Gelaro, R., and Coauthors, 2017: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, **30,** 5419-5454.

Gerbig, C., and Coauthors, 2003: Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *Journal of Geophysical Research: Atmospheres*, **108**.

Hewson, T. D., 1998: Objective fronts. *Meteorological Applications*, **5,** 37-65.

Hilton, T., K. Davis, K. Keller, and N. Urban, 2013: Improving North American terrestrial CO 2 flux diagnosis using spatial structure in land surface model residuals. *Biogeosciences*, **10,** 4607-4625.

Holton, J. R., and G. J. Hakim, 2012: *An introduction to dynamic meteorology.* Vol. 88, Academic press.

Hurwitz, M. D., D. M. Ricciuto, P. S. Bakwin, K. J. Davis, W. Wang, C. Yi, and M. P. Butler, 2004: Transport of carbon dioxide in the presence of storm systems over a northern Wisconsin forest. *Journal of the atmospheric sciences*, **61,** 607-618.

Keeling, C. D., R. B. Bacastow, A. E. Bainbridge, C. A. Ekdahl Jr, P. R. Guenther, L. S. Waterman, and J. F. Chin, 1976: Atmospheric carbon dioxide variations at Mauna Loa observatory, Hawaii. *Tellus*, **28,** 538-551.

Keppel-Aleks, G., and Coauthors, 2012: The imprint of surface fluxes and transport on variations in total column carbon dioxide. *Biogeosciences*, **9,** 875-891.

Lauvaux, T., and Coauthors, 2008: Mesoscale inversion: first results from the CERES campaign with synthetic data. *Atmospheric Chemistry and Physics*, **8,** 3459-3471.

Law, R., P. Rayner, L. Steele, and I. Enting, 2002: Using high temporal frequency data for CO2 inversions. *Global biogeochemical cycles*, **16,** 1-1-1-18.

Lee, T. R., S. F. De Wekker, A. E. Andrews, J. Kofler, and J. Williams, 2012: Carbon dioxide variability during cold front passages and fair weather days at a forested mountaintop site. *Atmospheric environment*, **46,** 405-416.

Lin, J., and Coauthors, 2004: Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO2 Budget and Rectification Airborne (COBRA) study. *Journal of Geophysical Research: Atmospheres*, **109**.

Mahadevan, A., and D. Archer, 2000: Modeling the impact of fronts and mesoscale circulation on the nutrient supply and biogeochemistry of the upper ocean. *Journal of Geophysical Research: Oceans*, **105,** 1209-1225.

Mahadevan, P., and Coauthors, 2008: A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM). *Global Biogeochemical Cycles*, **22**.

Margules, M., 1906: Uber Temperaturschichtung in stationar bewegter und ruhender Luft. *Meteorologische Zeitschrift.*, **23,** 243-254.

McClure, C. D., D. A. Jaffe, and H. Gao, 2016: Carbon dioxide in the free troposphere and boundary layer at the Mt. bachelor observatory. *Aerosol Air Qual. Res*, **16,** 717-728.

Meroni, M., M. Rossini, L. Guanter, L. Alonso, U. Rascher, R. Colombo, and J. Moreno, 2009: Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sensing of Environment*, **113,** 2037-2051.

Miles, M., 1962: Fronts. *Weather*, **17,** 45-47.

O'Dell, C., and Coauthors, 2018: Improved Retrievals of Carbon Dioxide from the Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm.

Pachauri, R. K., and Coauthors, 2014: *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change.* IPCC.

Papageorgiou, G. C., 2007: *Chlorophyll a fluorescence: a signature of photosynthesis.* Vol. 19, Springer Science & Business Media.

Parazoo, N., A. Denning, S. Kawa, K. Corbin, R. Lokupitiya, and I. Baker, 2008: Mechanisms for synoptic variations of atmospheric CO 2 in North America, South America and Europe. *Atmospheric Chemistry and Physics*, **8,** 7239-7254.

Park, C., and Coauthors, 2018: CO2 Transport, Variability, and Budget over the Southern California Air Basin Using the High-Resolution WRF-VPRM Model during the CalNex 2010 Campaign. *Journal of Applied Meteorology and Climatology*, **57,** 1337-1352.

Petrescu, A., R. Abad-Viñas, G. Janssens-Maenhout, V. Blujdea, and G. Grassi, 2012: Global estimates of carbon stock changes in living forest biomass: EDGARv4. 3–time series from 1990 to 2010. *Biogeosciences*, **9,** 3437-3447.

Peylin, P., P. Rayner, P. Bousquet, C. Carouge, F. Hourdin, P. Heinrich, and P. Ciais, 2005: Daily CO 2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology. *Atmospheric Chemistry and Physics*, **5,** 3173-3186.

Renard, R. J., and L. C. Clarke, 1965: Experiments in numerical objective frontal analysis.

Schneising, O., and Coauthors, 2008: Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite-Part 1: Carbon dioxide.

Stephens, G. L., and Coauthors, 2002: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, **83,** 1771-1790.

Takahashi, T., and Coauthors, 2009: Climatological mean and decadal change in surface ocean pCO2, and net sea–air CO2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, **56,** 554-577.

Thoning, K. W., P. P. Tans, and W. D. Komhyr, 1989: Atmospheric carbon dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC data, 1974–1985. *Journal of Geophysical Research: Atmospheres*, **94,** 8549-8565.

Wallace, J. M., and P. V. Hobbs, 2006: *Atmospheric science: an introductory survey.* Vol. 92, Elsevier.

Wunch, D., and Coauthors, 2011: A method for evaluating bias in global measurements of CO 2 total columns from space. *Atmospheric Chemistry and Physics*, **11,** 12317-12337.

Xiao, X., D. Hollinger, J. Aber, M. Goltz, E. A. Davidson, Q. Zhang, and B. Moore III, 2004a: Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote sensing of environment*, **89,** 519-534.

Xiao, X., and Coauthors, 2004b: Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of Environment*, **91,** 256-270.

Appendix I: Figures for all collected 83 cases of X_{CO_2} from 2015-2017

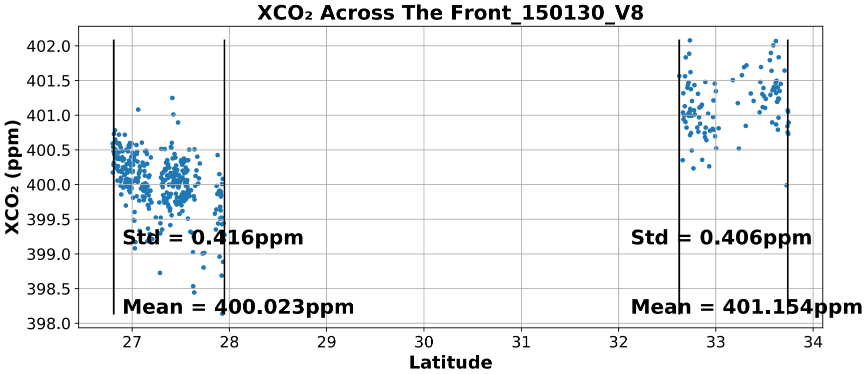


Figure I.1 – 3-sounding moving averaged X_{CO_2} gradient across cold front on Jan 30th, 2015, blue dots in left sections are X_{CO_2} in warm sector and right ones are in cold sector. Means and standard deviations are given at both sectors.

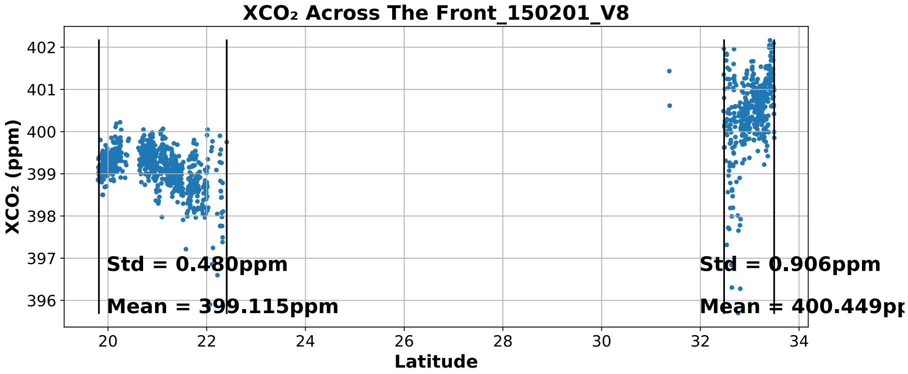


Figure I.2 – Same with Figure I.1 on Feb 1st, 2105

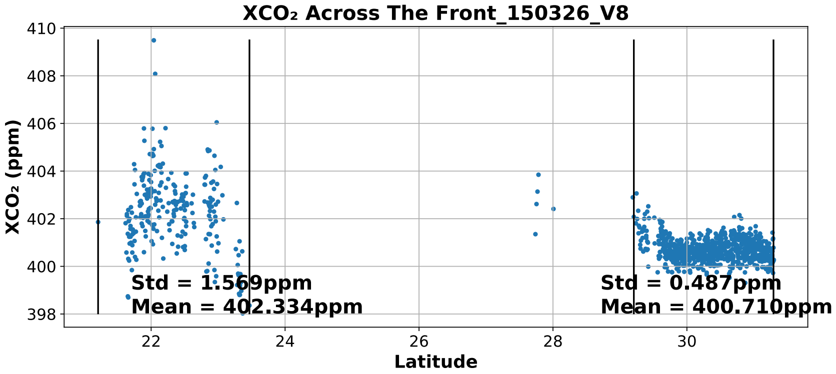


Figure I.3 – Same with Figure I.1 on Mar 26th, 2015

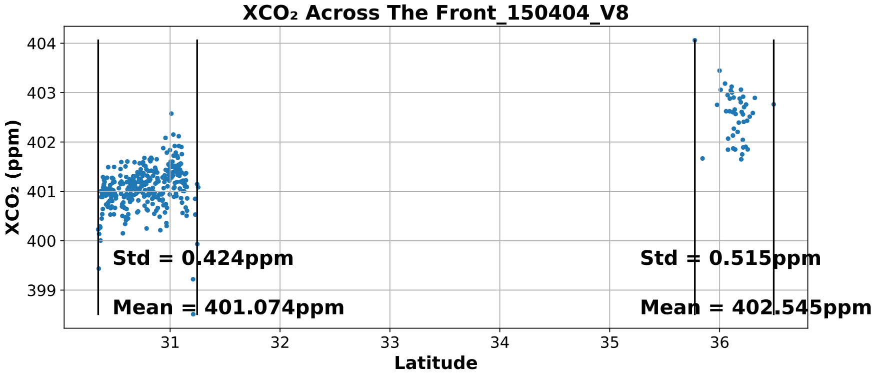


Figure I.4 – Same with Figure I.1 on Apr 4th, 2015

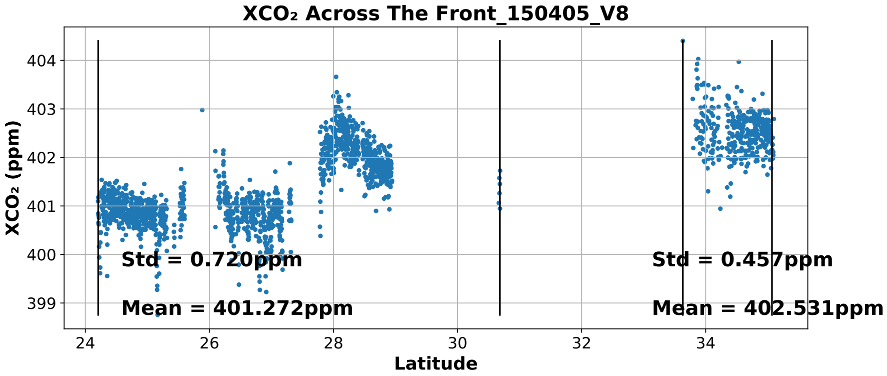


Figure I.5 – Same with Figure I.1 on Apr 5th, 2015

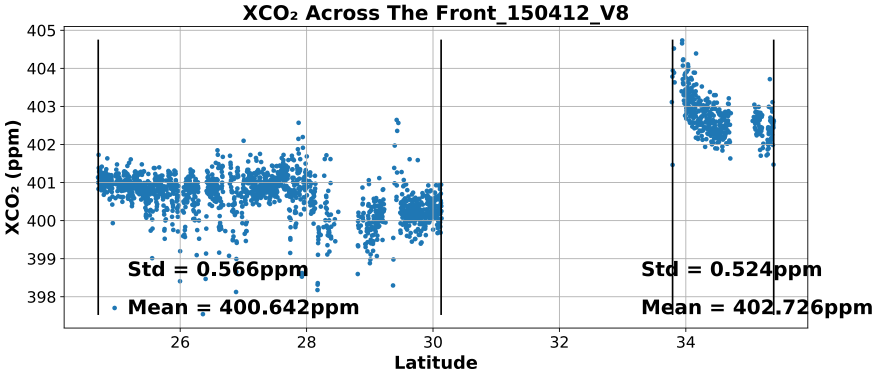


Figure I.6 – Same with Figure I.1 on Apr 12th, 2015

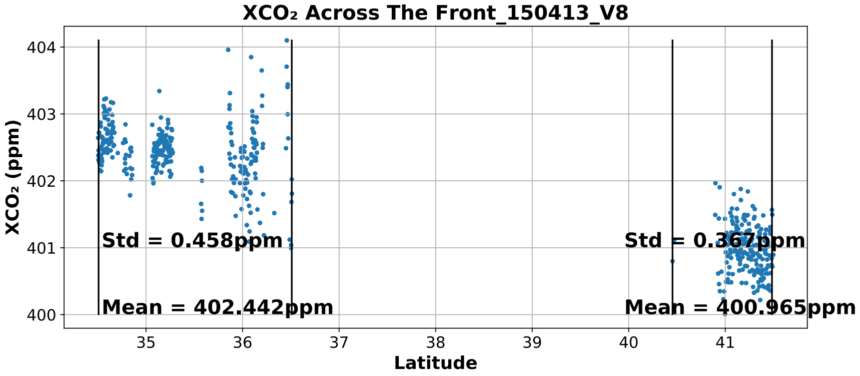


Figure I.7 – Same with Figure I.1 on Apr 13th, 2015

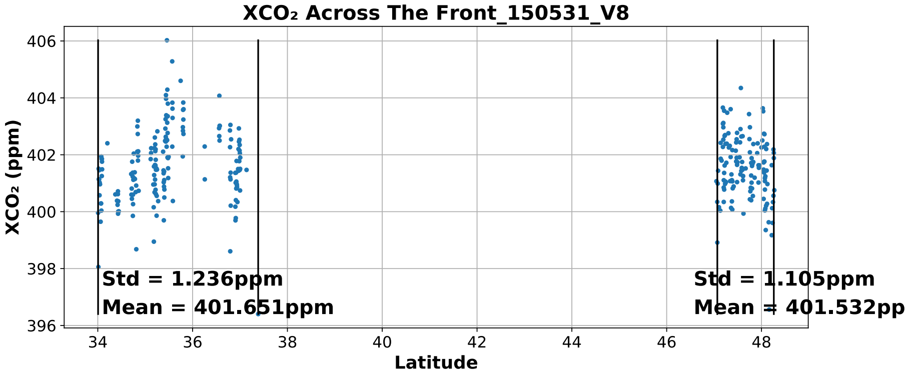


Figure I.8 – Same with Figure I.1 on May 31st, 2015

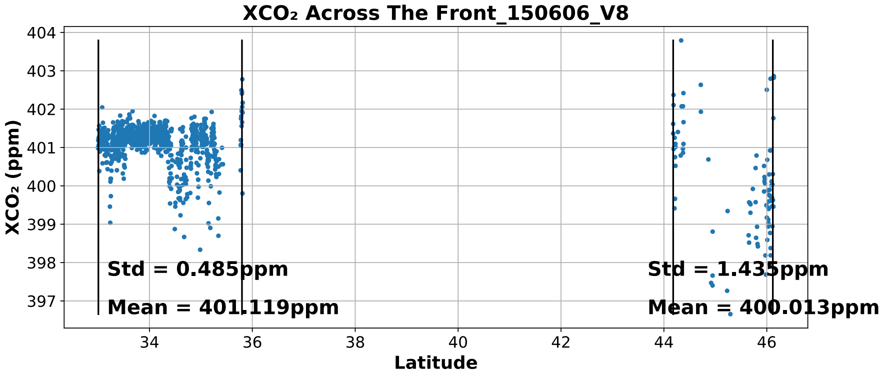


Figure I.9 – Same with Figure I.1 on Jun 6th, 2015

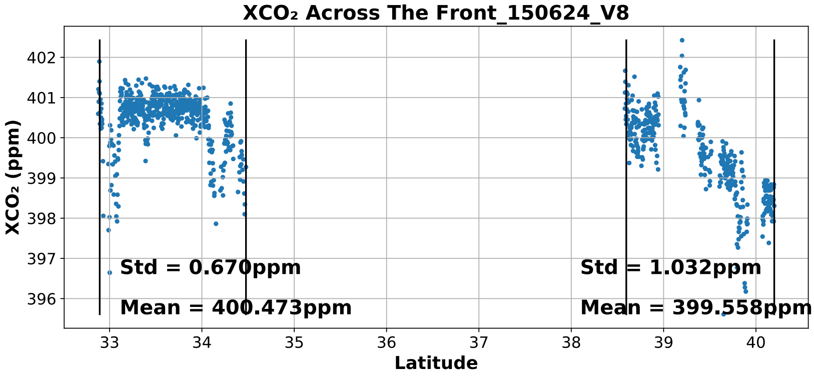


Figure I.10 – Same with Figure I.1 on Jun 24th, 2015

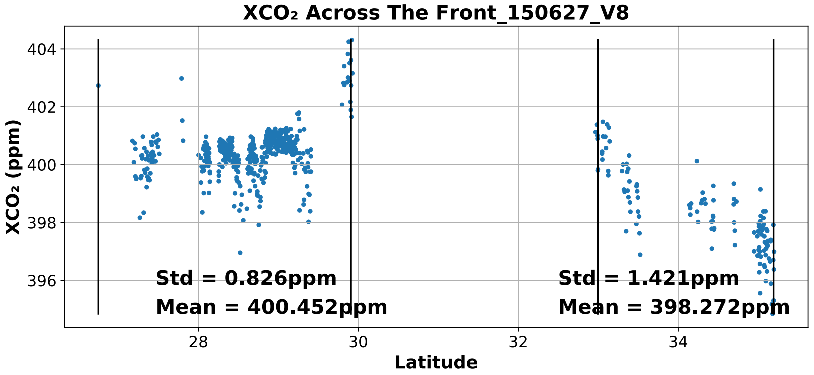


Figure I.11 – Same with Figure I.1 on Jun 27th, 2015

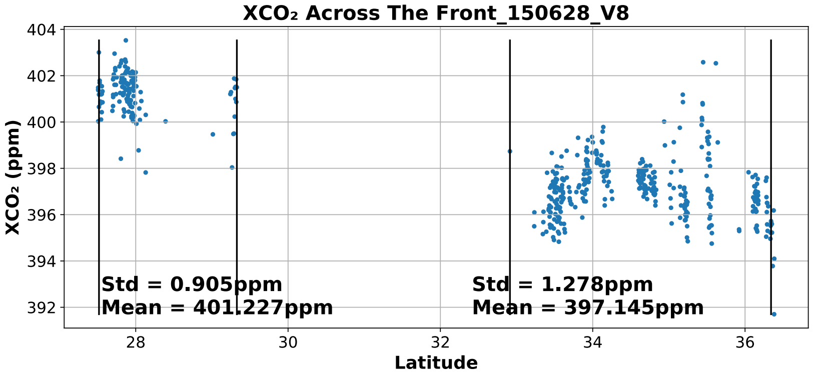


Figure I.12 – Same with Figure I.1 on Jun 28th, 2015

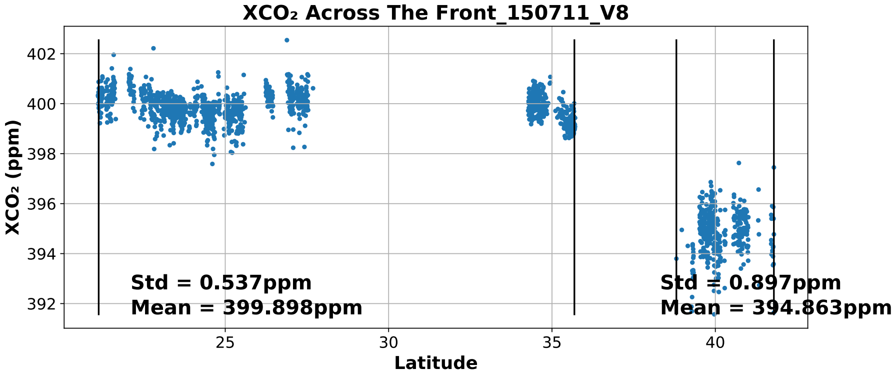


Figure I.13 – Same with Figure I.1 on Jul 11th, 2015

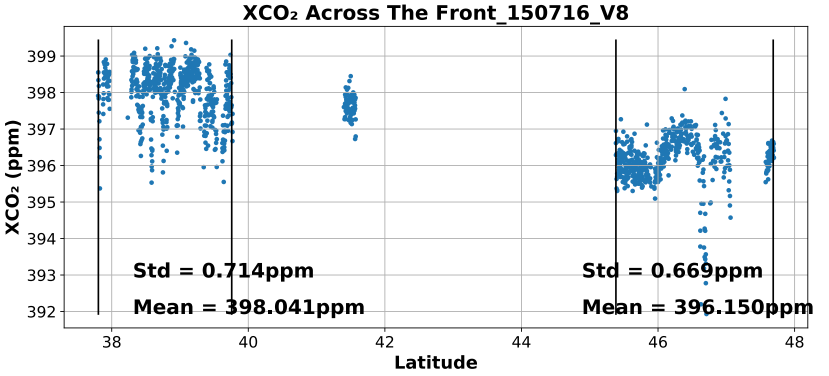


Figure I.14 – Same with Figure I.1 on Jul 16th, 2015

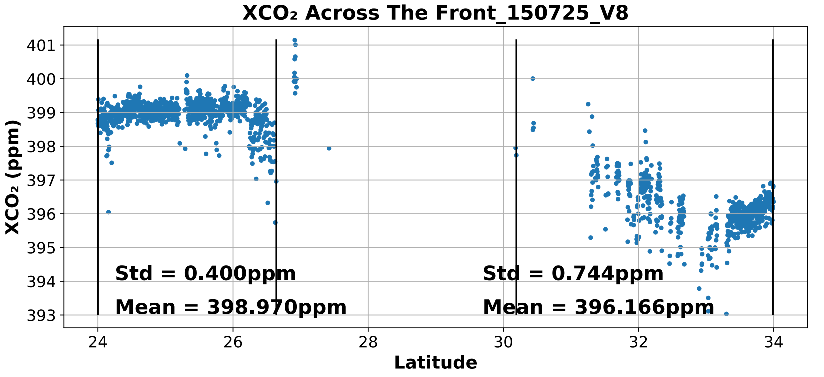


Figure I.15 – Same with Figure I.1 on Jul 25th, 2015

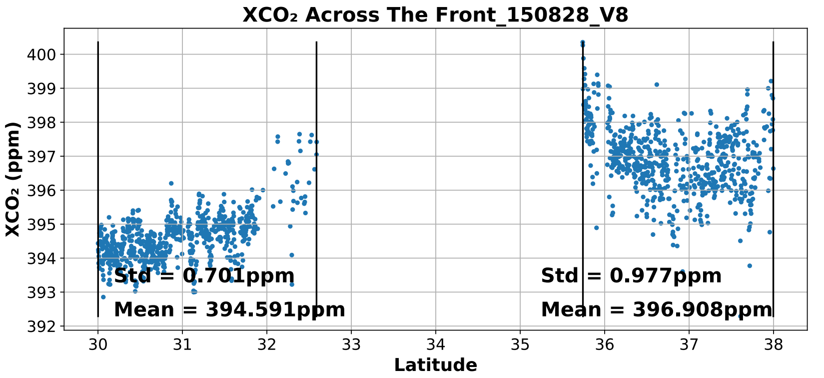


Figure I.16 – Same with Figure I.1 on Aug 28th, 2105

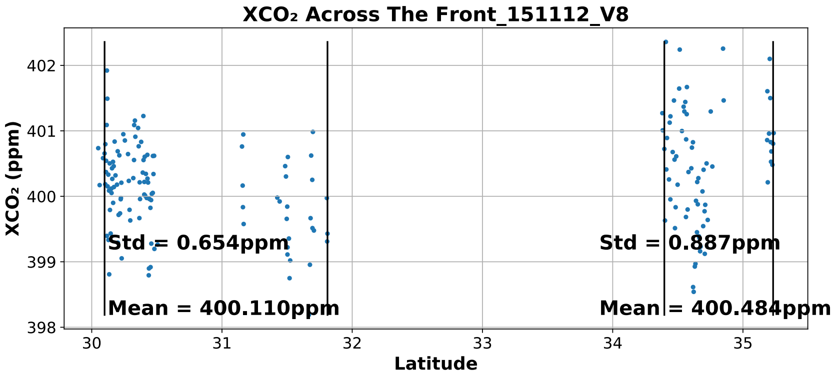


Figure I.17 – Same with Figure I.1 on Nov 12th, 2015

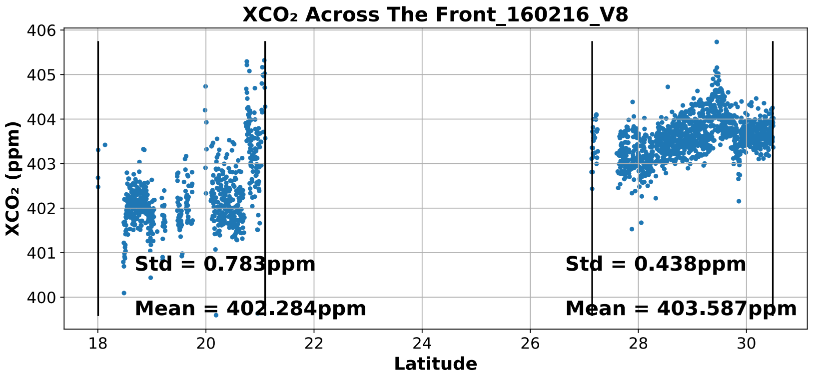


Figure I.18 – Same with Figure I.1 on Feb 16th, 2016

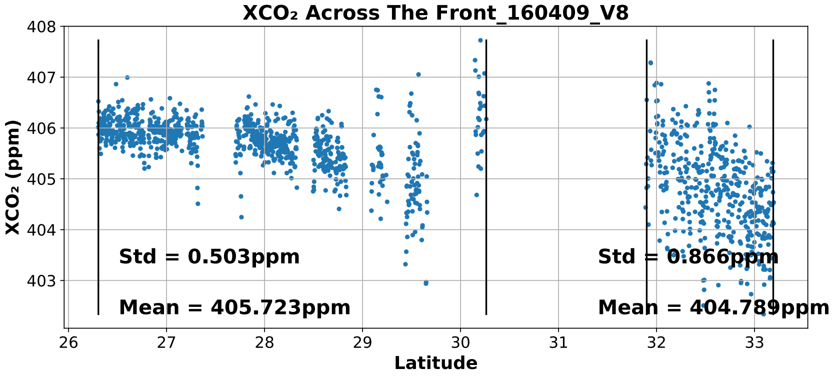


Figure I.19 – Same with Figure I.1 on Apr 9th, 2016

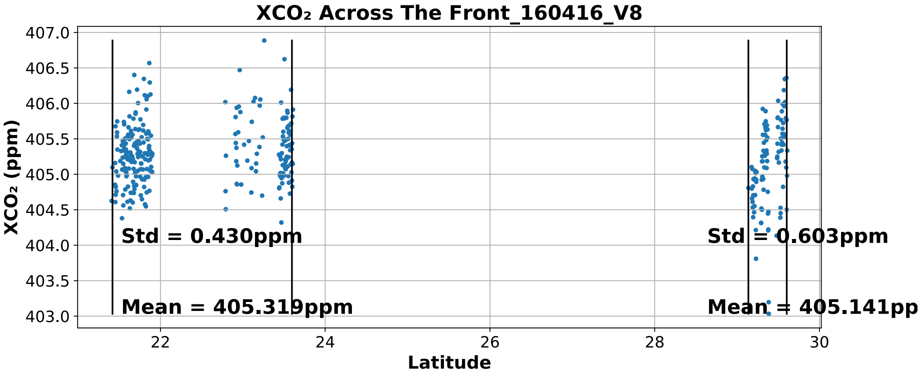


Figure I.20 – Same with Figure I.1 on Apr 16th, 2016

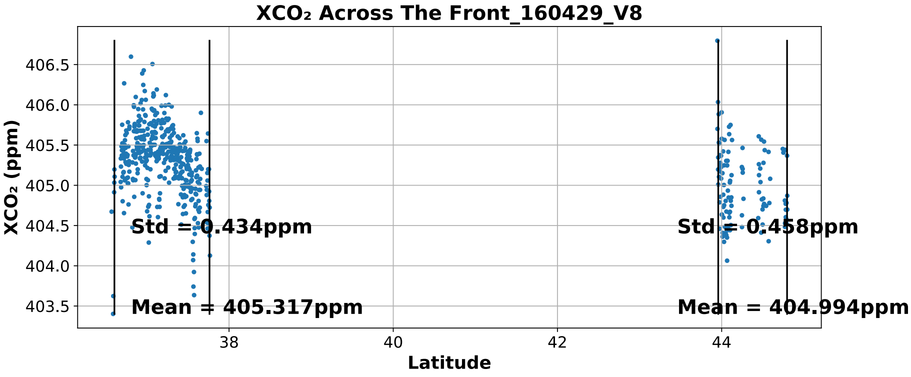


Figure I.21 – Same with Figure I.1 on Apr 29th, 2016

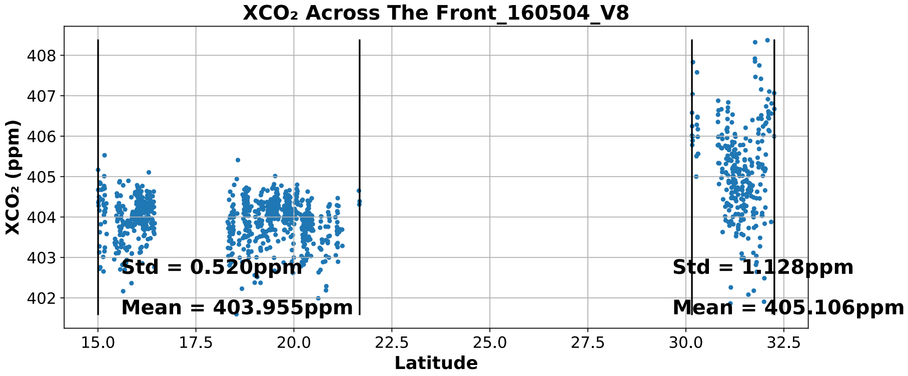


Figure I.22 – Same with Figure I.1 on May 4th, 2016

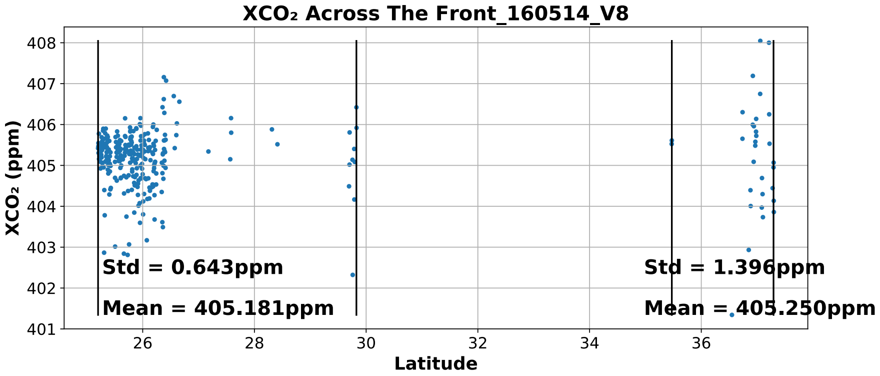


Figure I.23 – Same with Figure I.1 on May 14th, 2016

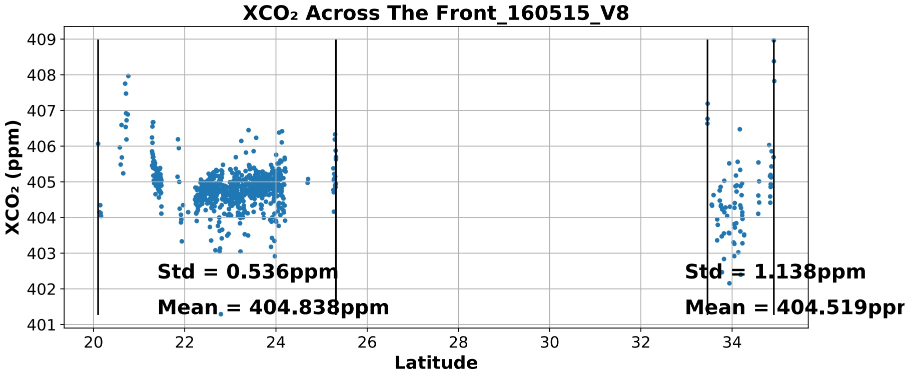


Figure I.24 – Same with Figure I.1 on May 15th, 2016

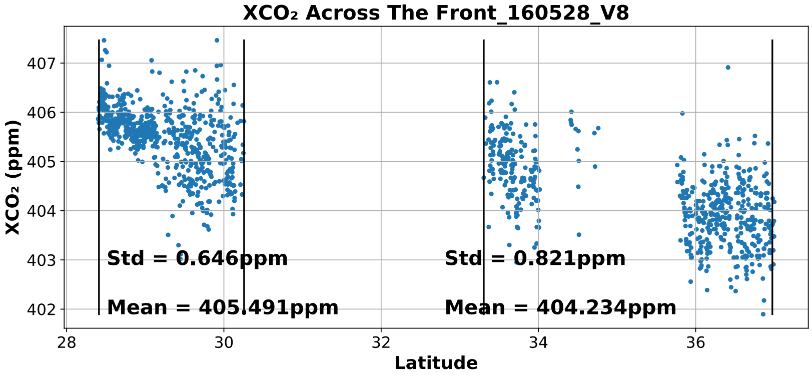


Figure I.25 – Same with Figure I.1 on May 28th, 2016

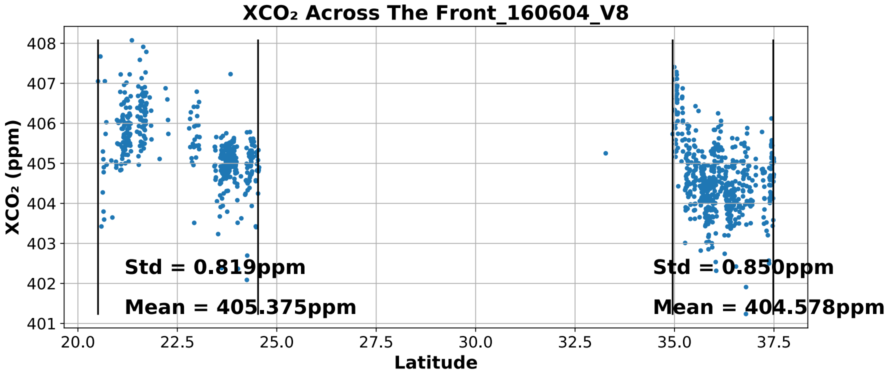


Figure I.26 – Same with Figure I.1 on Jun 4th, 2016

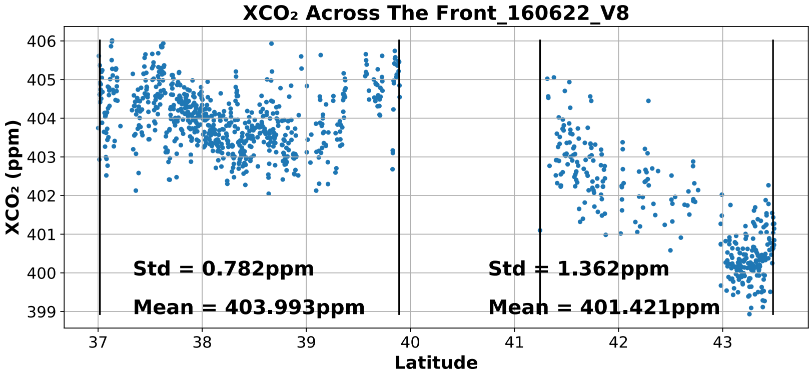


Figure I.27 – Same with Figure I.1 on Jun 22nd, 2016

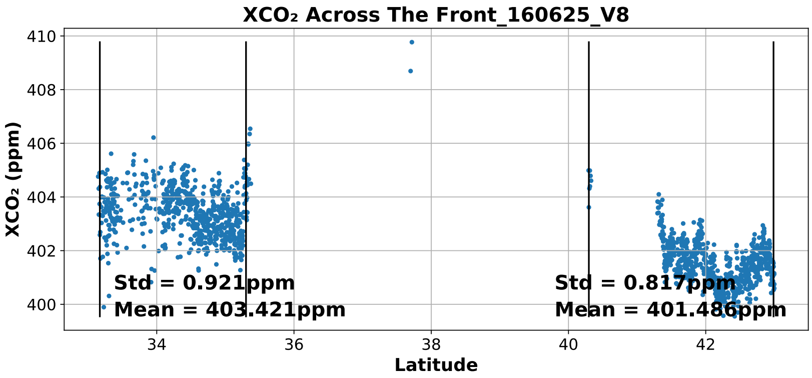


Figure I.28 – Same with Figure I.1 on Jun 25th, 2016

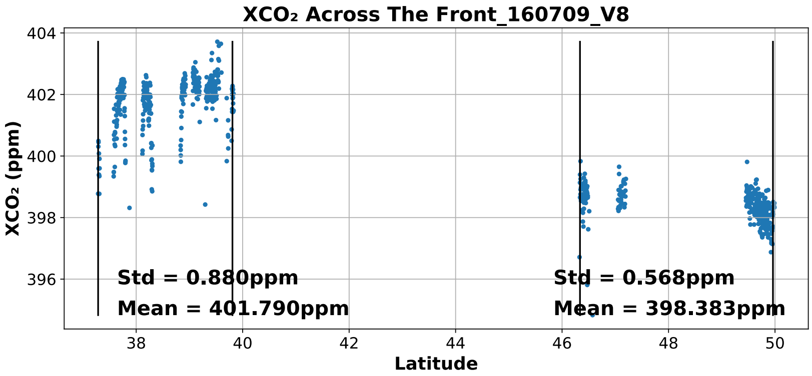


Figure I.29 – Same with Figure I.1 on Jul 9th, 2016

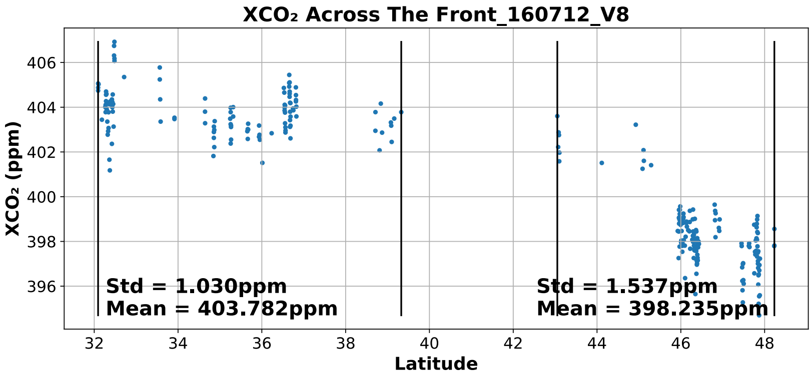


Figure I.30 – Same with Figure I.1 on Jul 12th, 2016

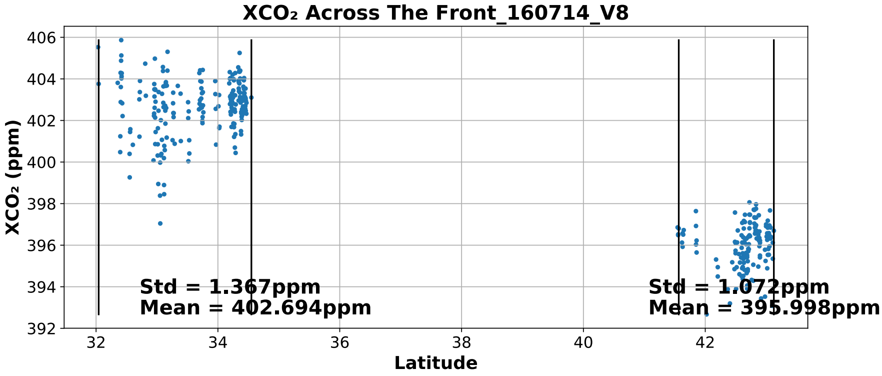


Figure I.31 – Same with Figure I.1 on Jul 14th, 2016

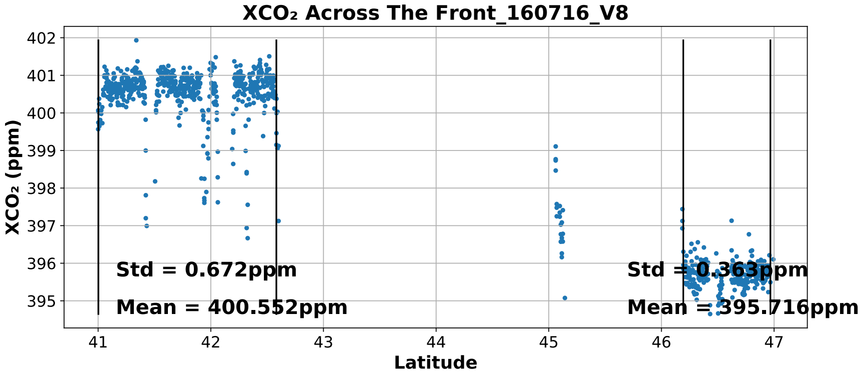


Figure I.32 – Same with Figure I.1 on Jul 16th, 2016

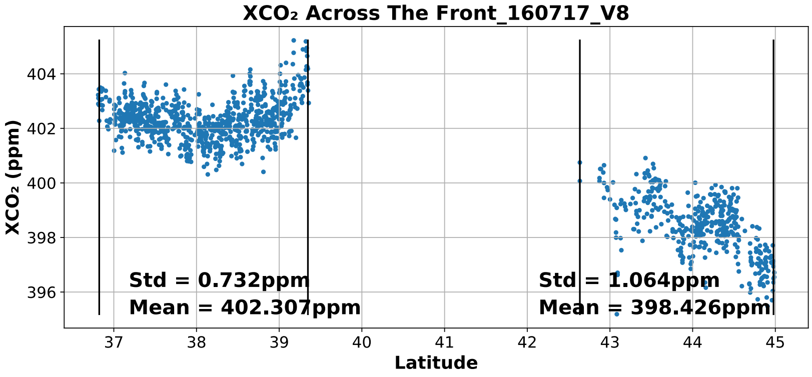


Figure I.33 – Same with Figure I.1 on Jul 17th, 2016

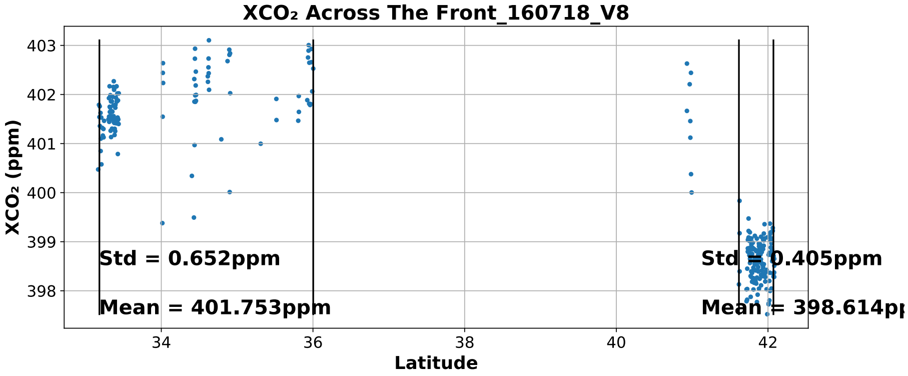


Figure I.34 – Same with Figure I.1 on Jul 18th, 2016

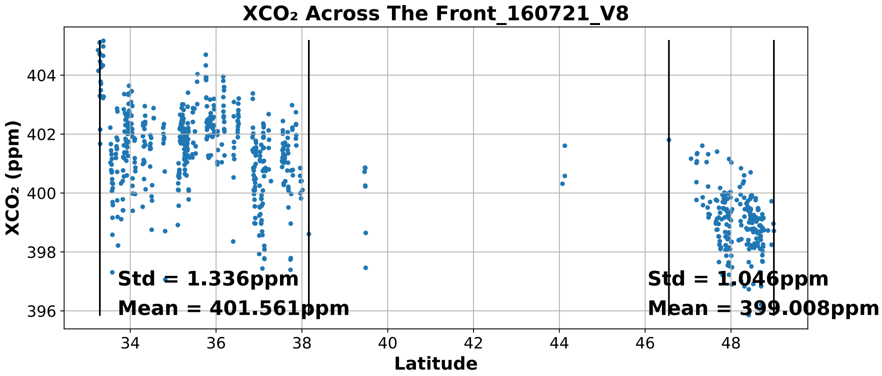


Figure I.35 – Same with Figure I.1 on Jul 21st, 2016

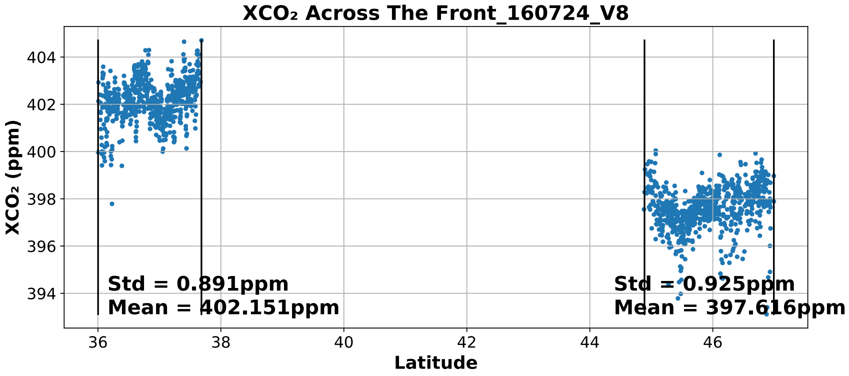


Figure I.36 – Same with Figure I.1 on Jul 24th, 2016

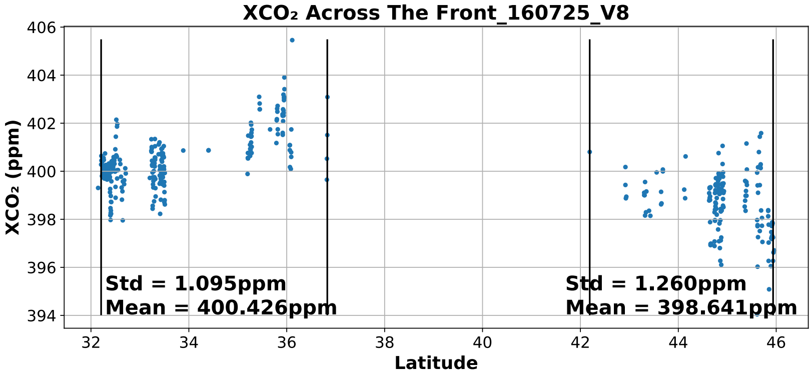


Figure I.37 – Same with Figure I.1 on Jul 25th, 2016

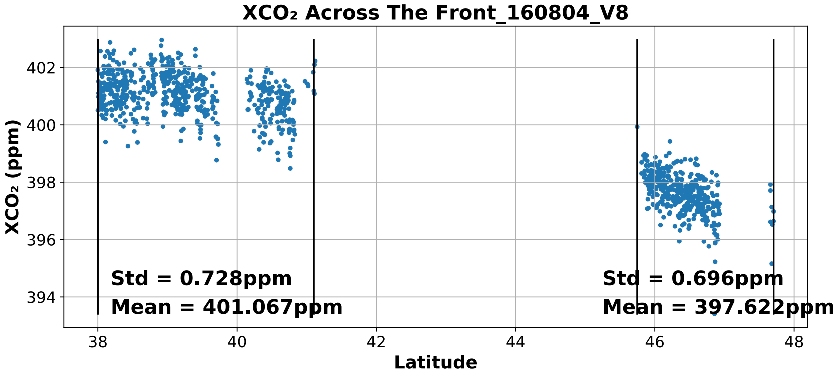


Figure I.38 – Same with Figure I.1 on Aug 4th, 2016

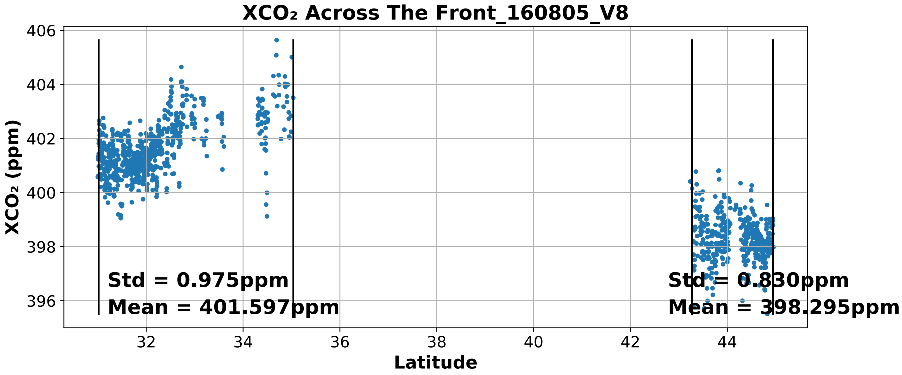


Figure I.39 – Same with Figure I.1 on Aug 5th, 2016

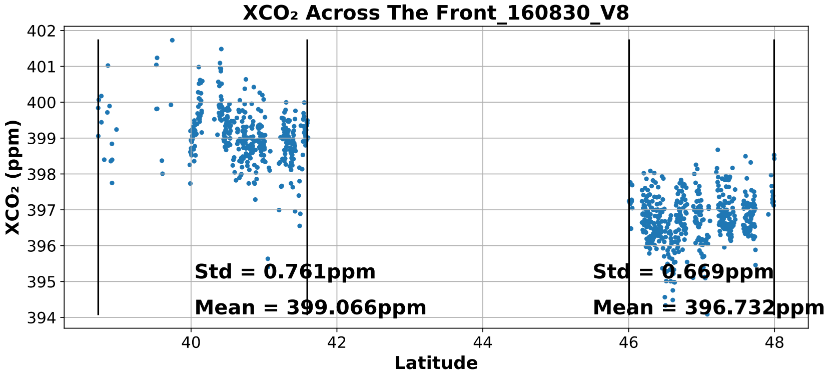

XCO₂ Across The Front_160830_V8

Figure I.40 – Same with Figure I.1 on Aug 30th, 2016

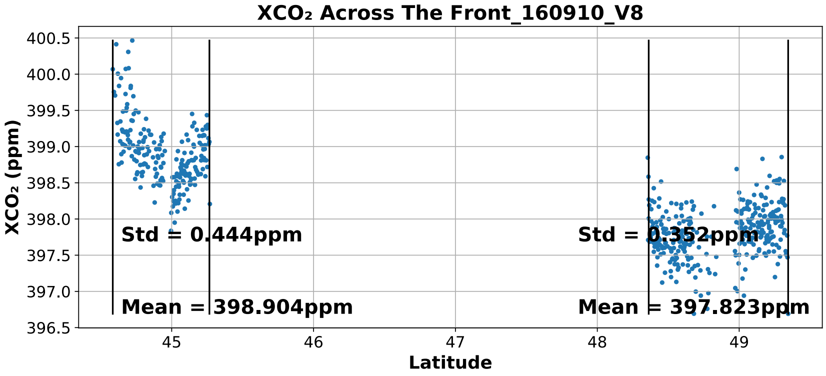

XCO₂ Across The Front_160910_V8

Figure I.41 – Same with Figure I.1 on Sep 10th, 2016

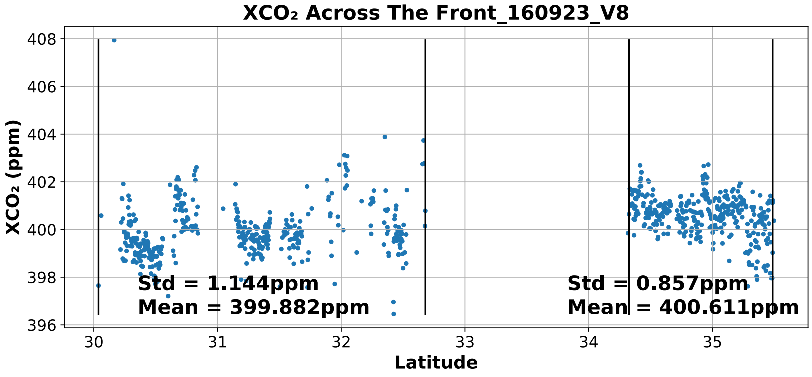

XCO₂ Across The Front_160923_V8

Figure I.42 – Same with Figure I.1 on Sep 23rd, 2016

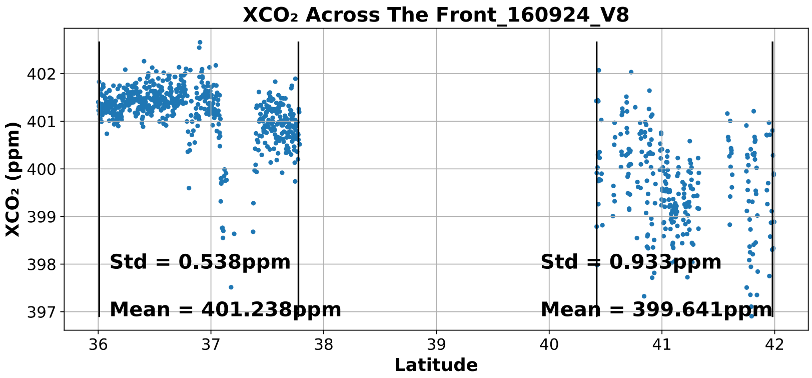


Figure I.43 – Same with Figure I.1 on Sep 24th, 2016

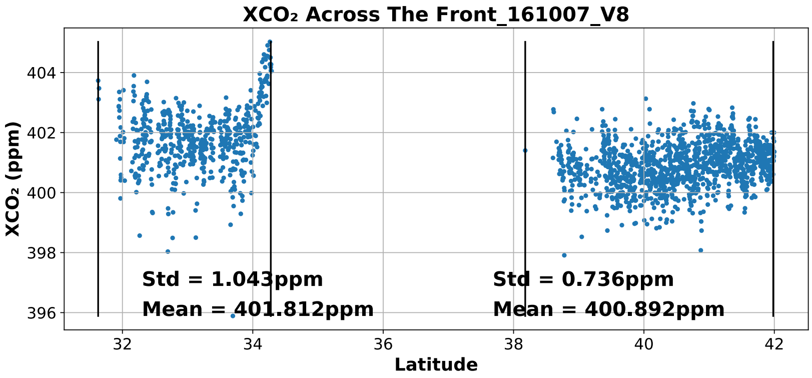


Figure I.44 – Same with Figure I.1 on Oct 7th, 2016

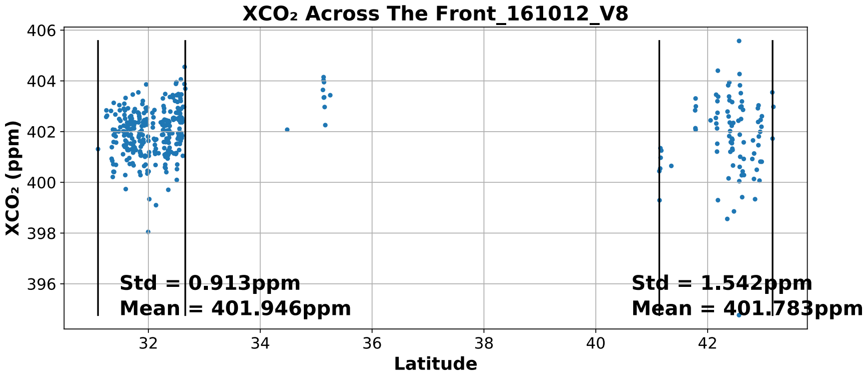


Figure I.45 – Same with Figure I.1 on Oct 12th, 2016

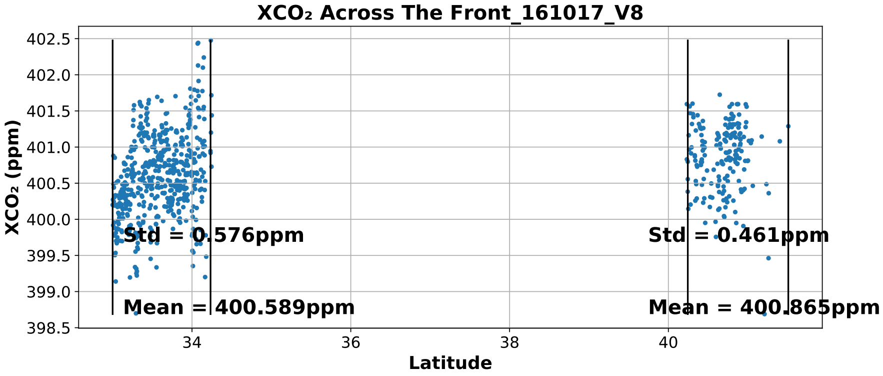


Figure I.46 – Same with Figure I.1 on Oct 17th, 2016

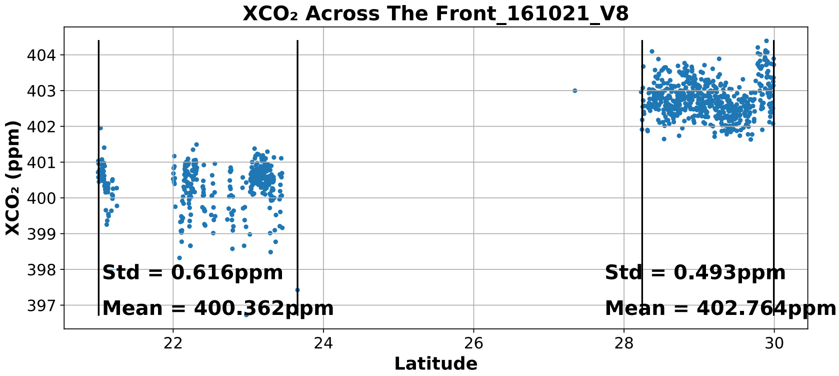


Figure I.47 – Same with Figure I.1 on Oct 21st, 2016

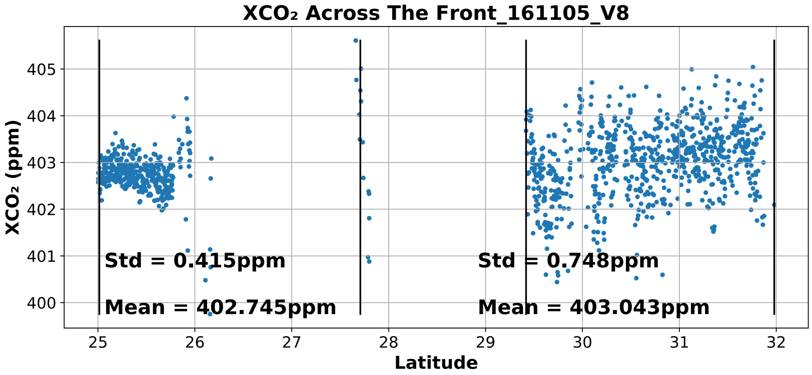


Figure I.48 – Same with Figure I.1 on Nov 5th, 2016

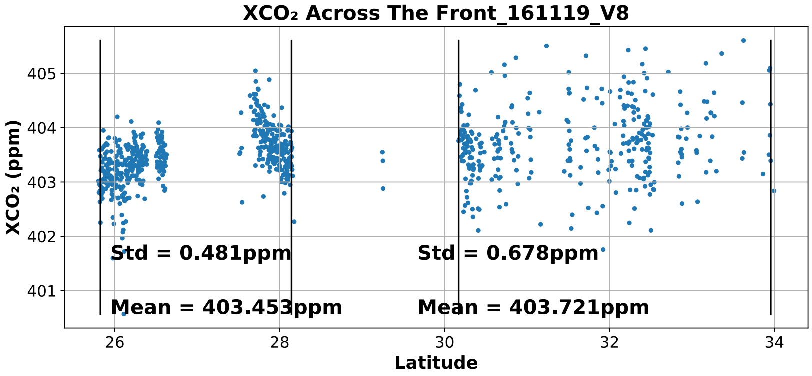


Figure I.49 – Same with Figure I.1 on Nov 19th, 2016

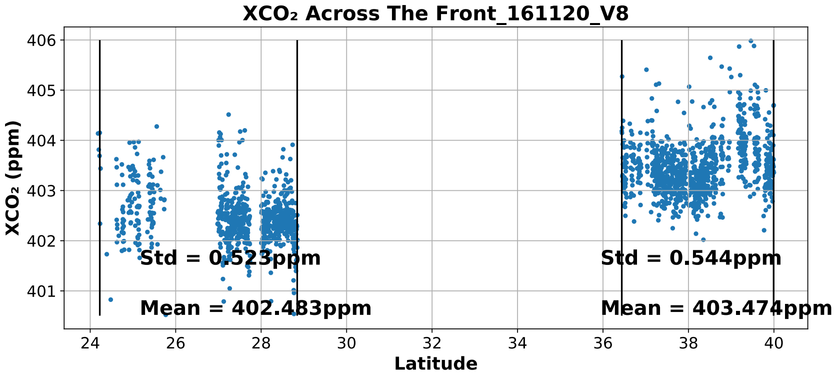


Figure I.50 – Same with Figure I.1 on Nov 20th, 2016

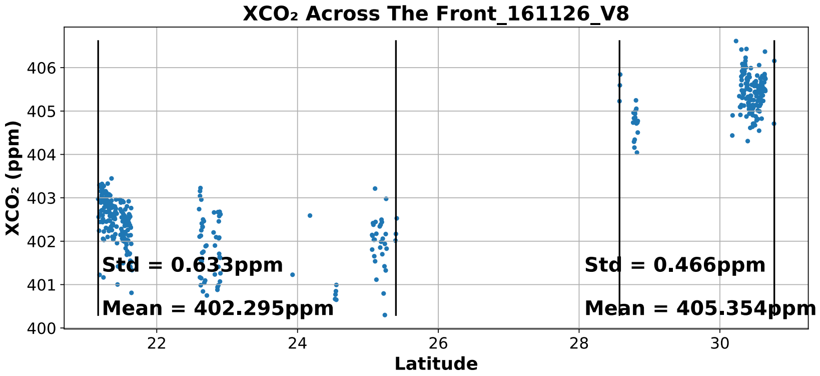


Figure I.51 – Same with Figure I.1 on Nov 26th, 2016

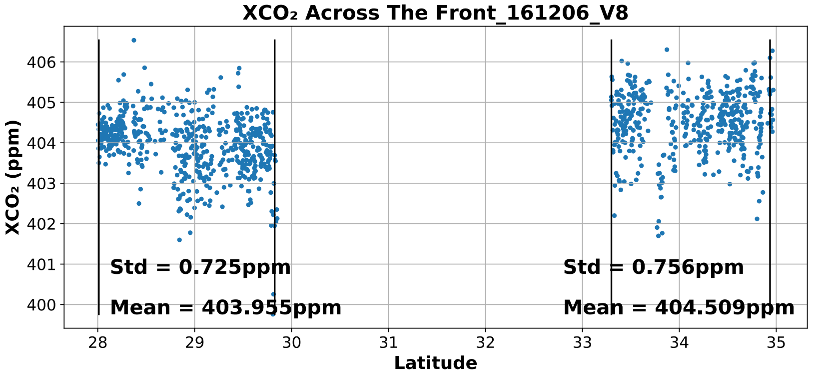


Figure I.52 – Same with Figure I.1 on Dec 6th, 2016

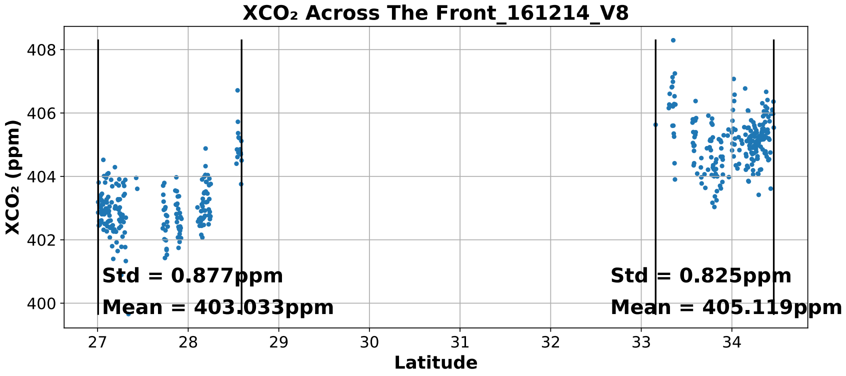


Figure I.53 – Same with Figure I.1 on Dec 14th, 2016

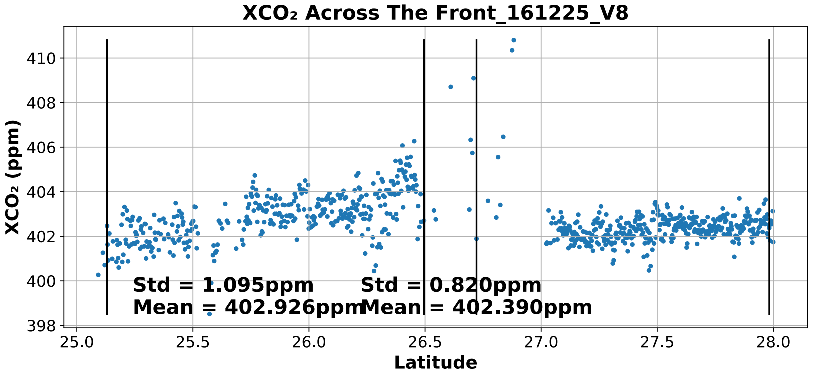


Figure I.54 – Same with Figure I.1 on Dec 25th, 2016

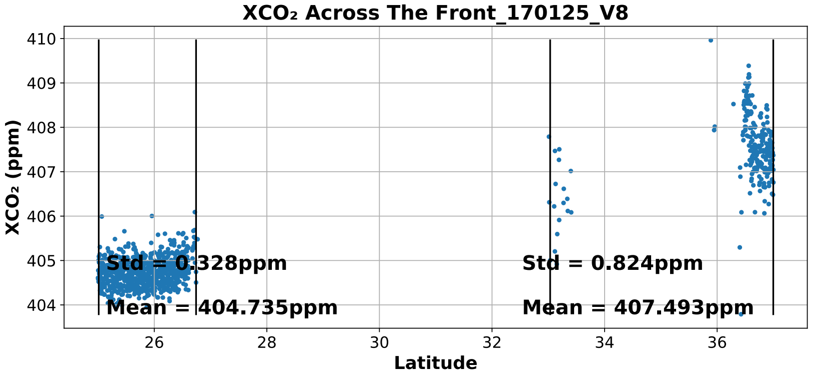


Figure I.55 – Same with Figure I.1 on Jan 25th, 2017

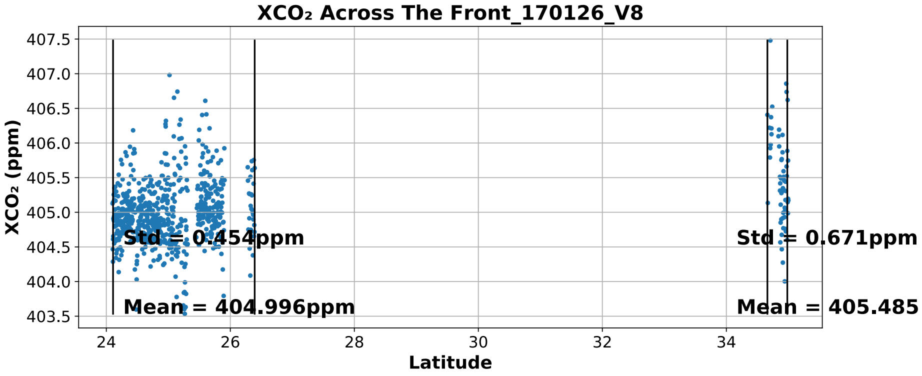


Figure I.56 – Same with Figure I.1 on Jan 26th, 2017

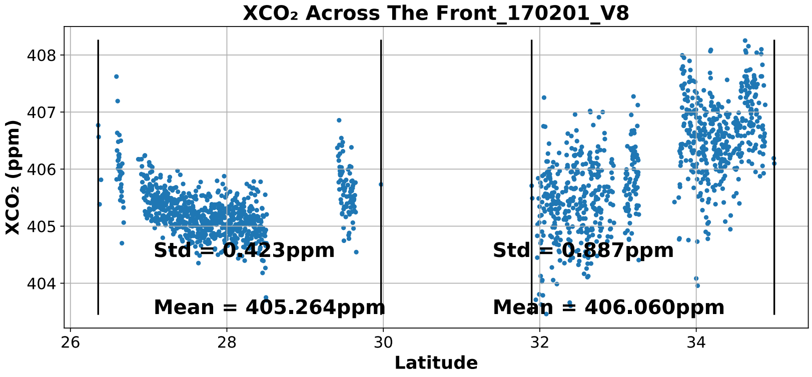


Figure I.57 – Same with Figure I.1 on Feb 1st, 2017

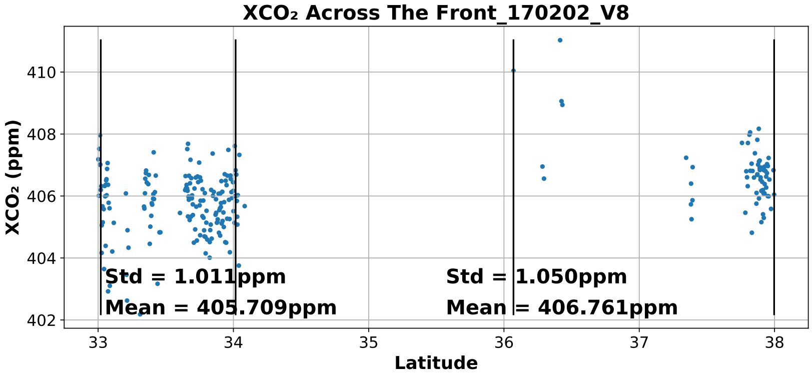


Figure I.58 – Same with Figure I.1 on Feb 2nd, 2017

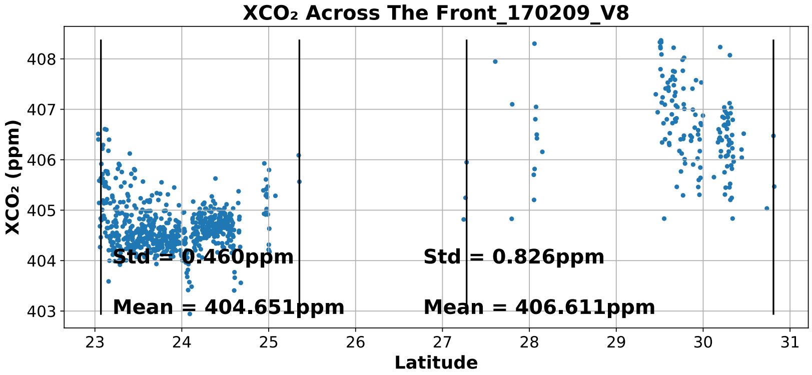


Figure I.59 – Same with Figure I.1 on Feb 9th, 2017

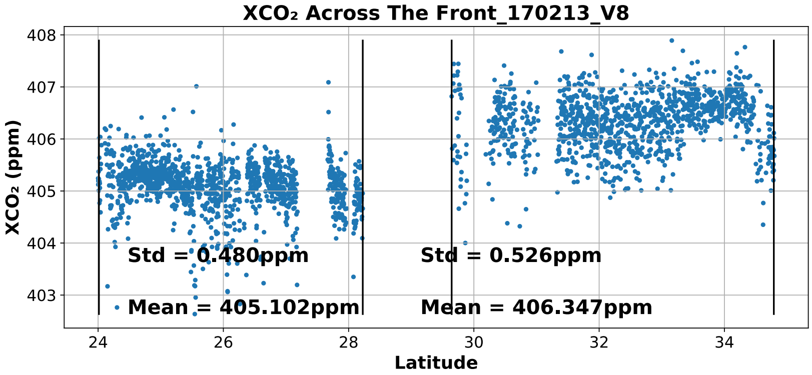


Figure I.60 – Same with Figure I.1 on Feb 13th, 2017

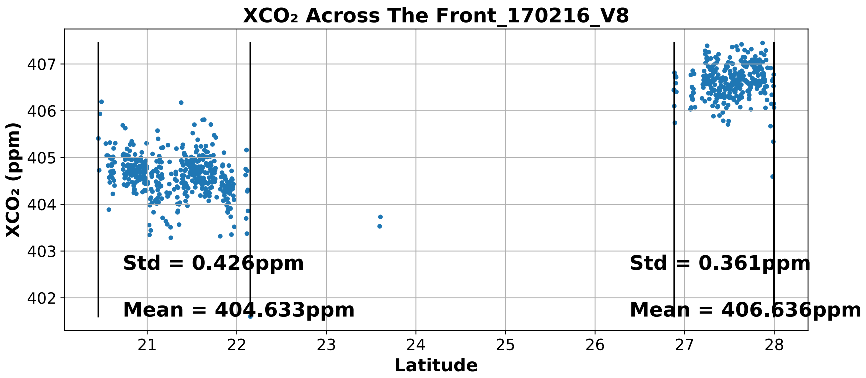


Figure I.61 – Same with Figure I.1 on Feb 16th, 2017

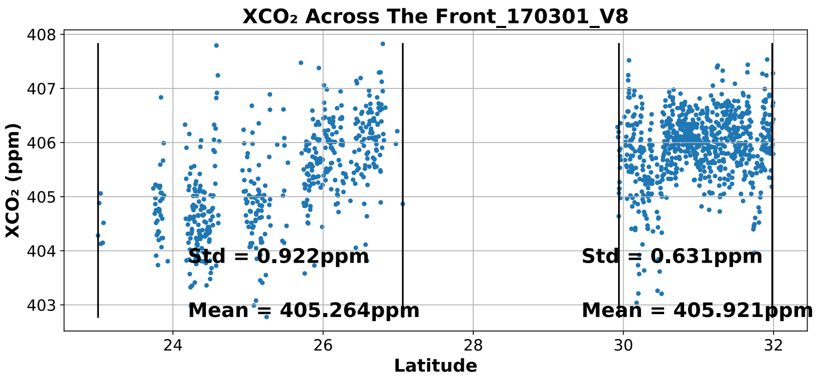


Figure I.62 – Same with Figure I.1 on Mar 1st, 2017

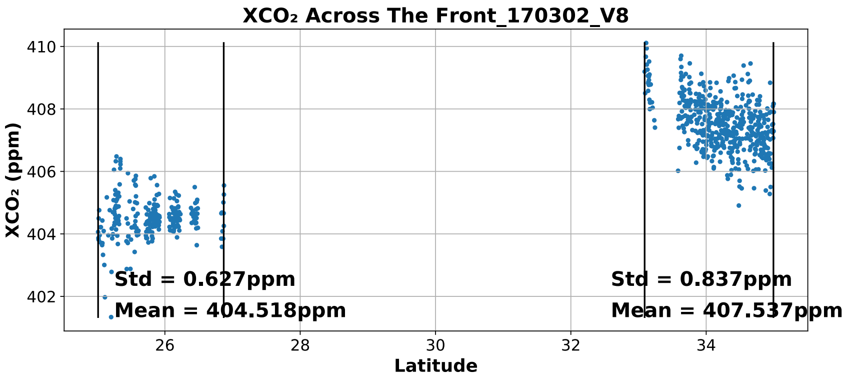


Figure I.63 – Same with Figure I.1 on Mar 2nd, 2017

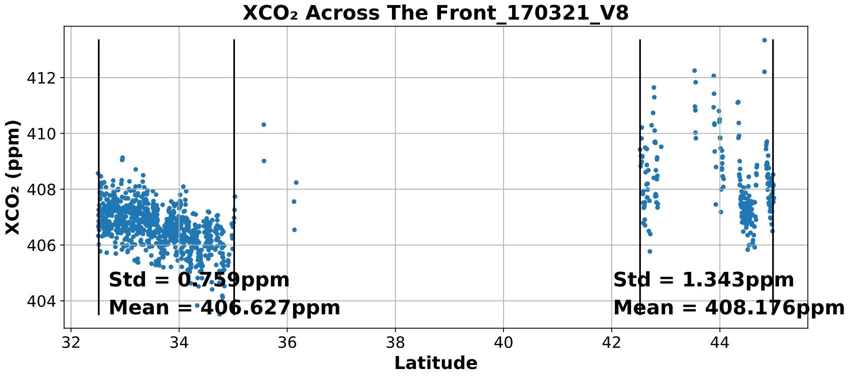


Figure I.64 – Same with Figure I.1 on Mar 21st, 2017

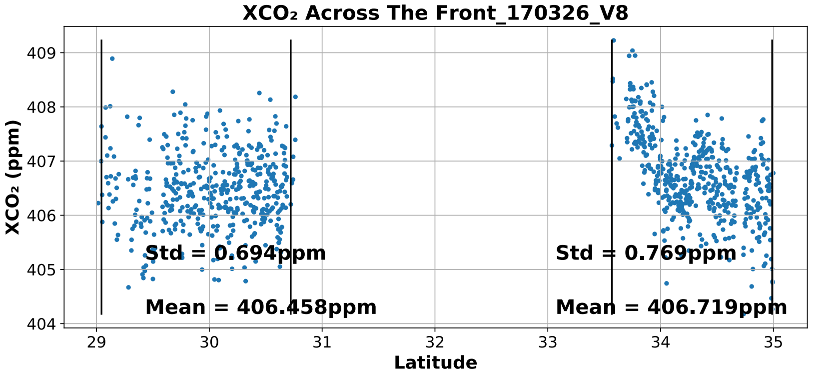


Figure I.65 – Same with Figure I.1 on Mar 26th, 2017

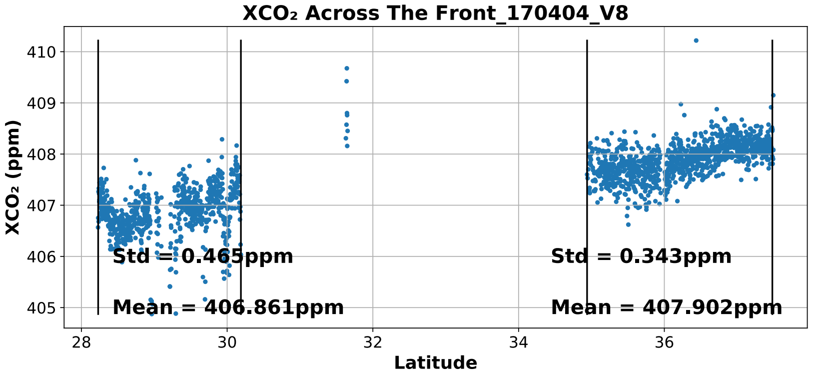


Figure I.66 – Same with Figure I.1 on Apr 4th, 2017

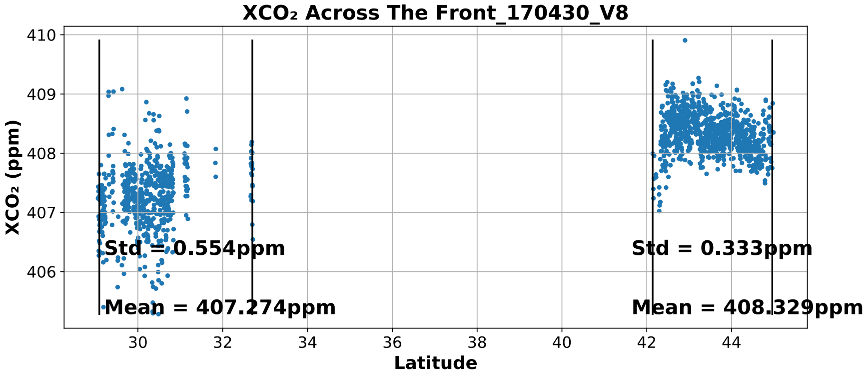


Figure I.67 – Same with Figure I.1 on Apr 30th, 2017

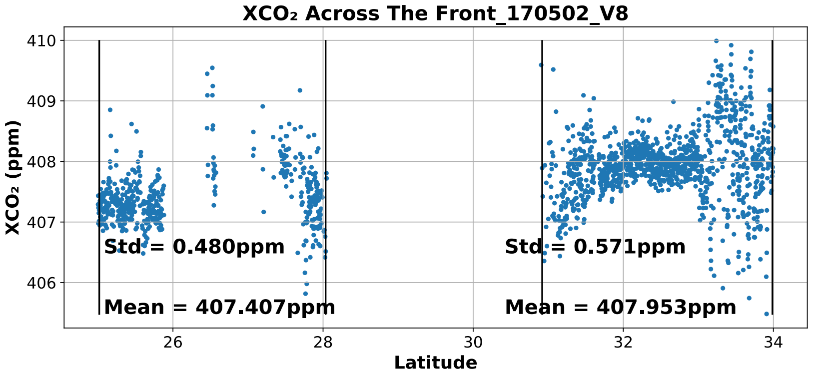


Figure I.68 – Same with Figure I.1 on May 2nd, 2017

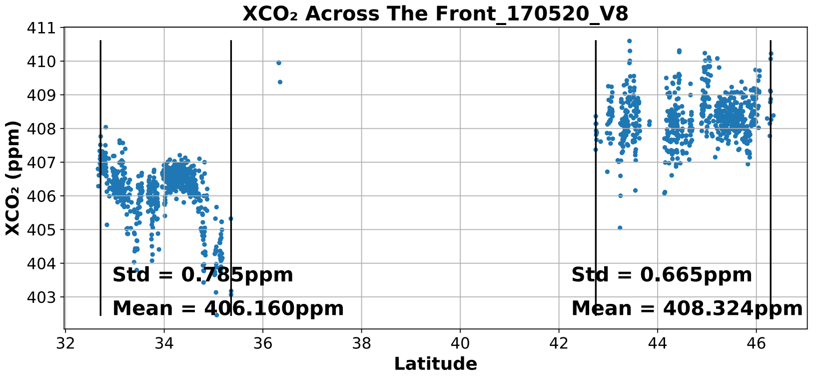


Figure I.69 – Same with Figure I.1 on May 20th, 2017

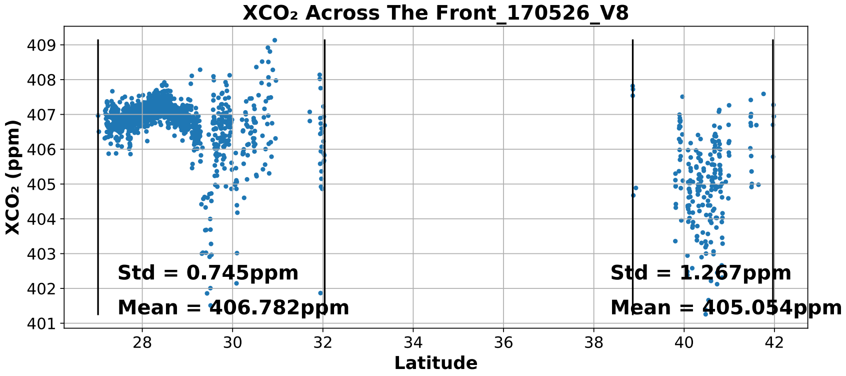


Figure I.70 – Same with Figure I.1 on May 26th, 2017

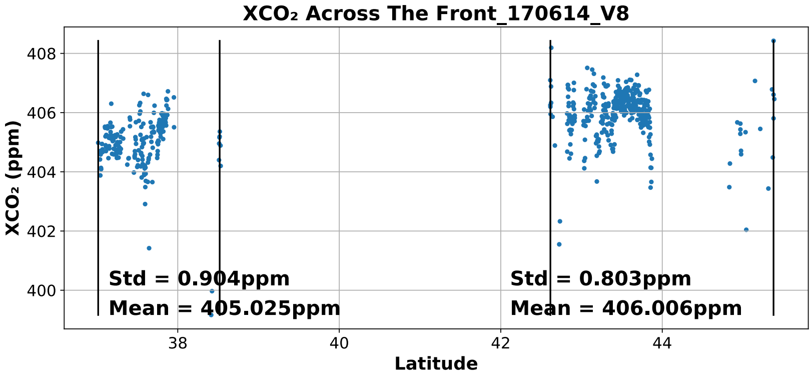


Figure I.71 – Same with Figure I.1 on Jun 14th, 2017

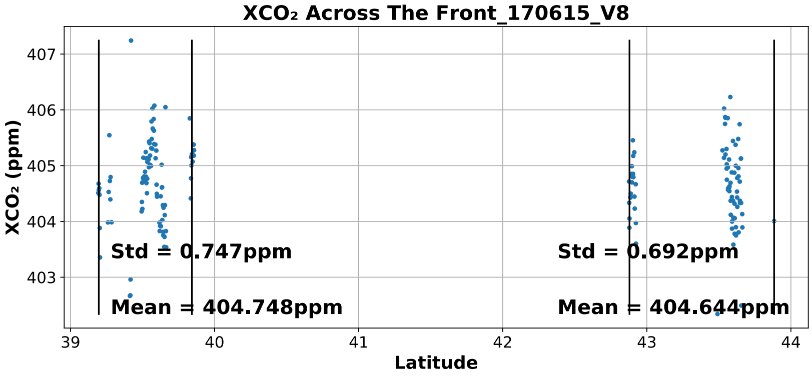


Figure I.72 – Same with Figure I.1 on Jun 15th, 2017

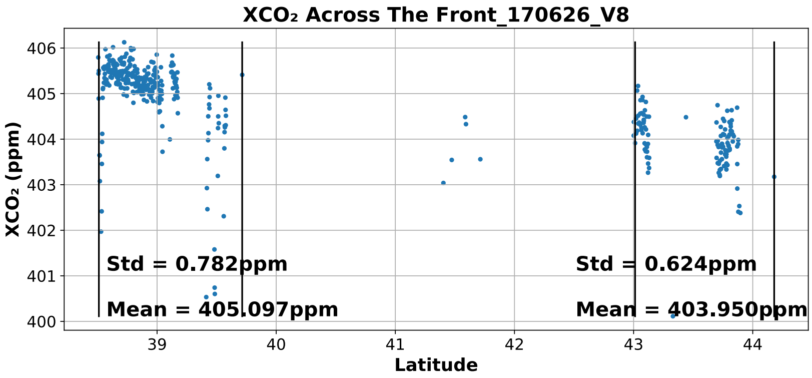


Figure I.73 – Same with Figure I.1 on Jun 26th, 2017

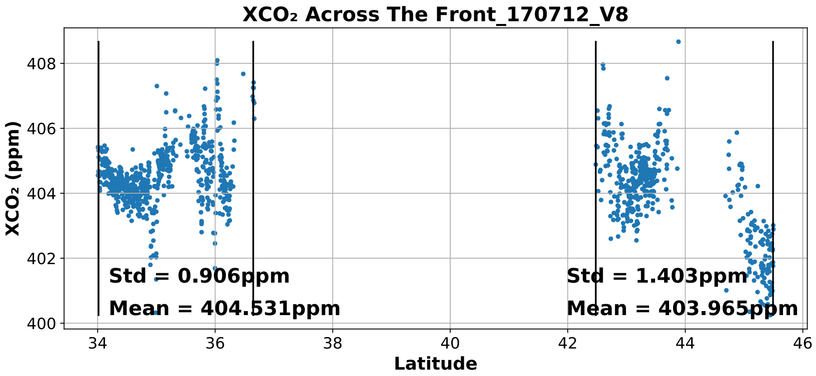


Figure I.74 – Same with Figure I.1 on Jul 12th, 2017

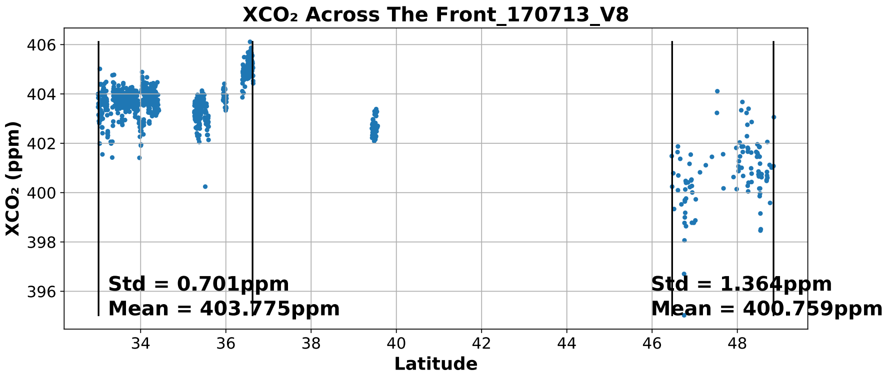


Figure I.75 – Same with Figure I.1 on Jul 13th, 2017

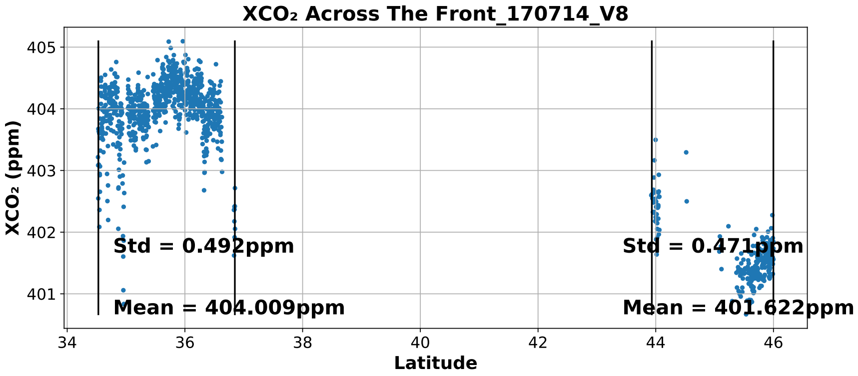


Figure I.76 – Same with Figure I.1 on Jul 14th, 2017

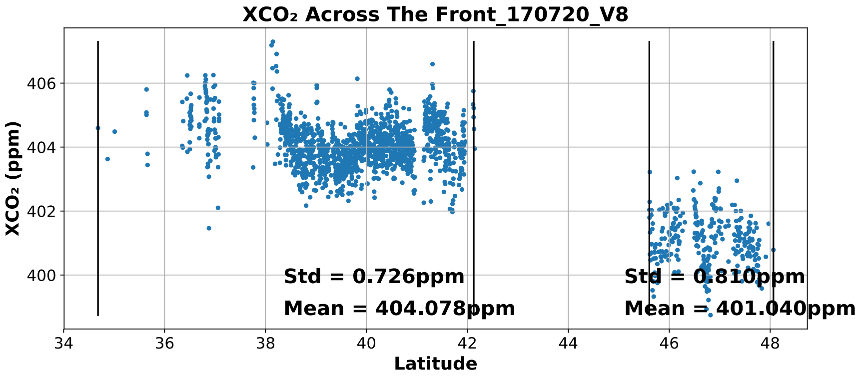


Figure I.77 – Same with Figure I.1 on Jul 20th, 2017

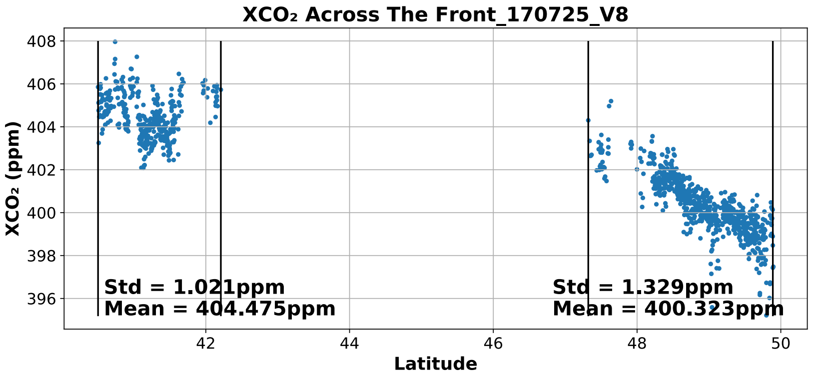


Figure I.78 – Same with Figure I.1 on Jul 25th, 2017

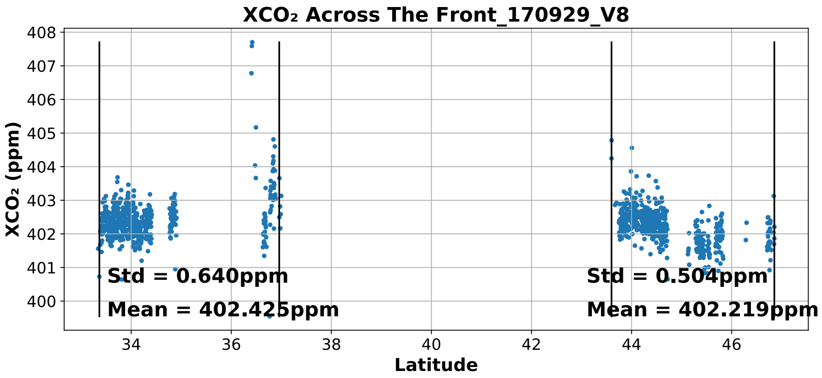


Figure I.79 – Same with Figure I.1 on Sep 29th, 2017

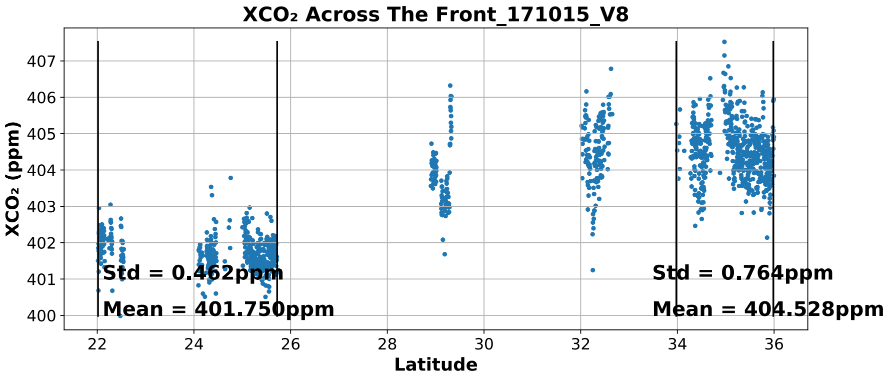


Figure I.80 – Same with Figure I.1 on Oct 15th, 2017

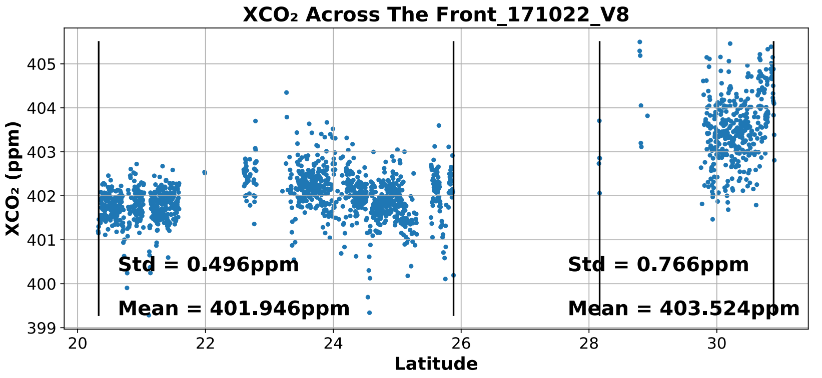


Figure I.81 – Same with Figure I.1 on Oct 22nd, 2017

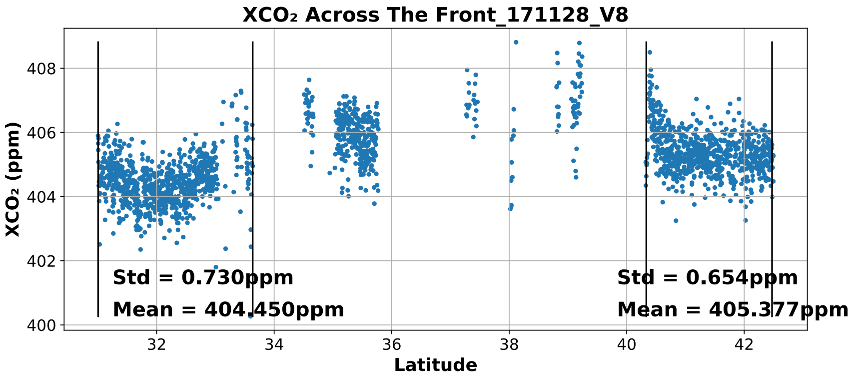


Figure I.82 – Same with Figure I.1 on Nov 28th, 2017

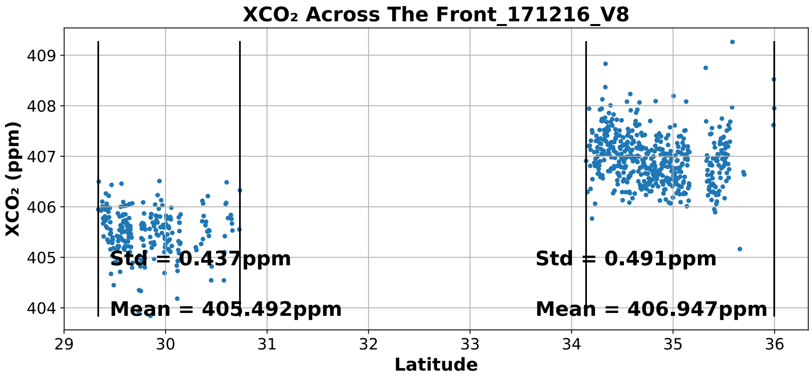


Figure I.83 – Same with Figure I.1 on Dec 16th, 2017

Appendix II: Figures for monthly distribution and difference between $X_{\mathrm{CO_2}}$ from WRF-VPRM and OCO-2 in 2016

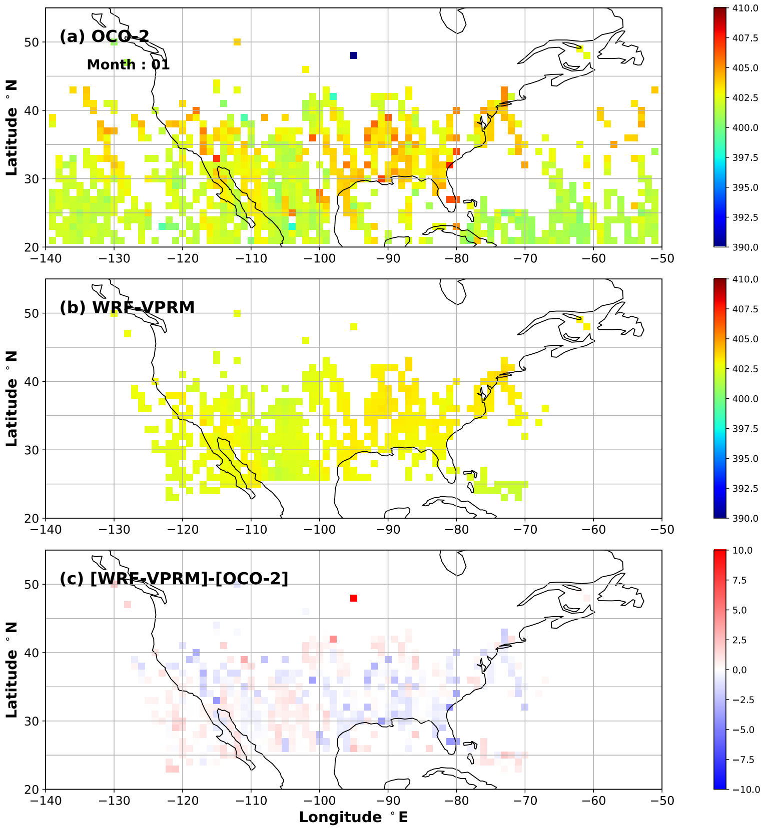


Figure II.1 – (a) is monthly mean of OCO-2 $X_{\mathrm{CO_2}}$ in January 2016, filtered with quality flag and gridded into 1°×1°. (b) is corresponding simulations from WRF-VPRM to grid boxes in (a), and each grid box in (b) matches the time of that in (a). (c) is the difference of (b) and (a).

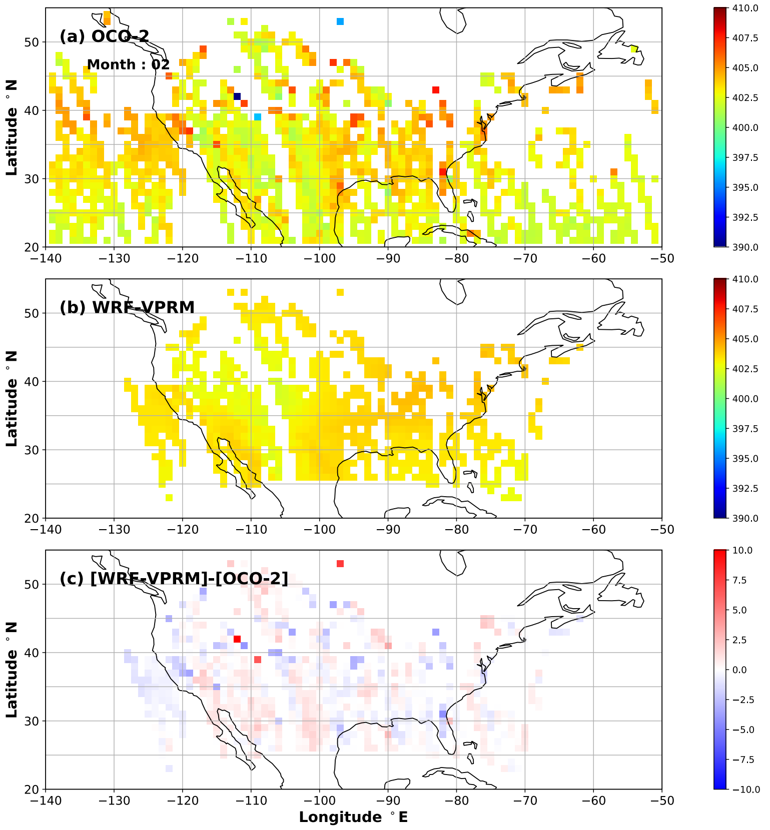


Figure II.2 – Same with Figure II.1 in February, 2016

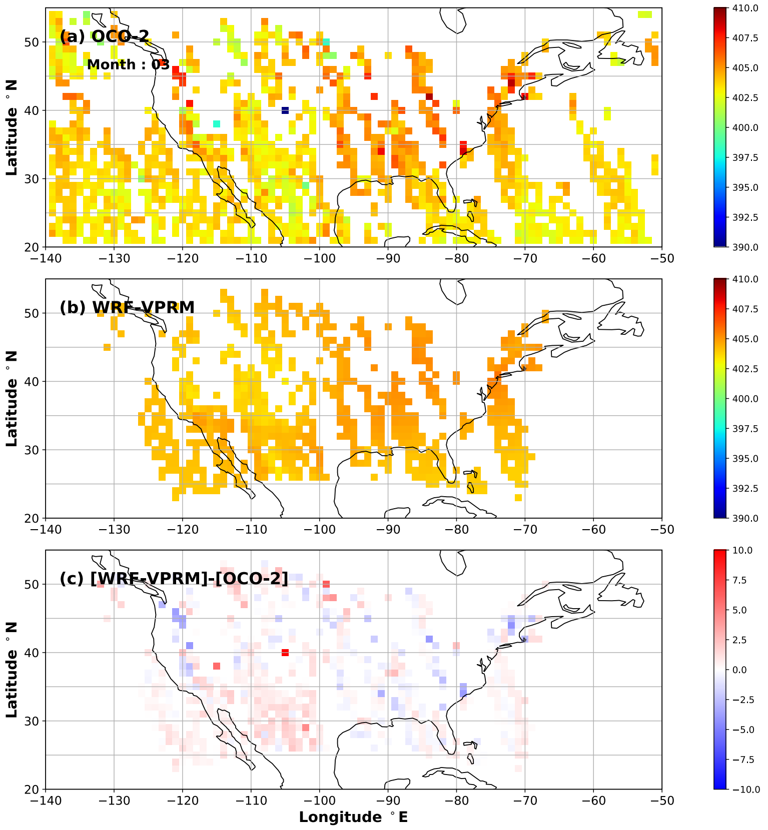


Figure II.3 – Same with Figure II.1 in March, 2016

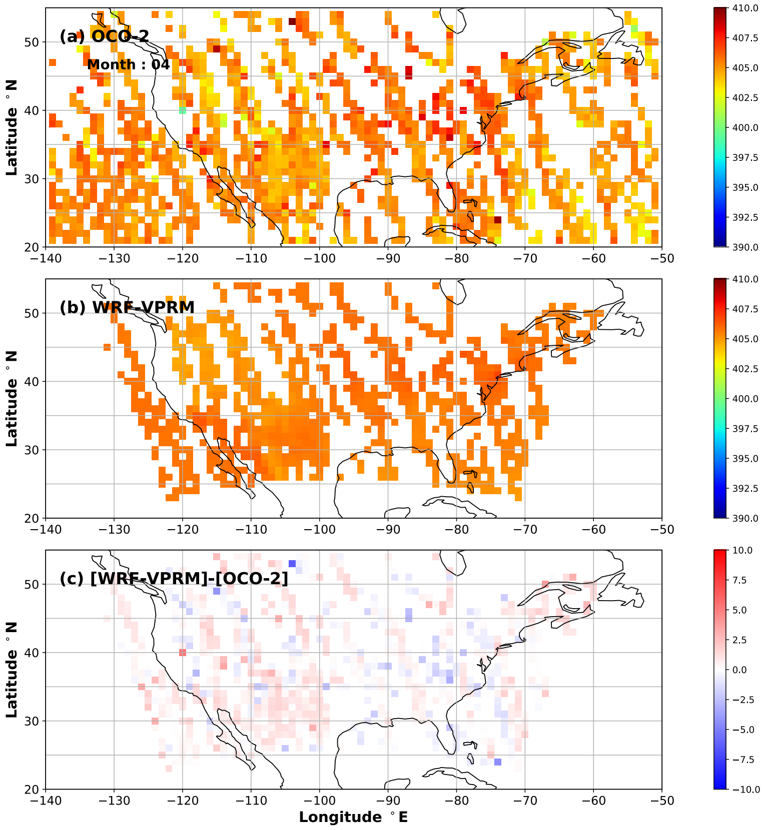


Figure II.4 – Same with Figure II.1 in April, 2016

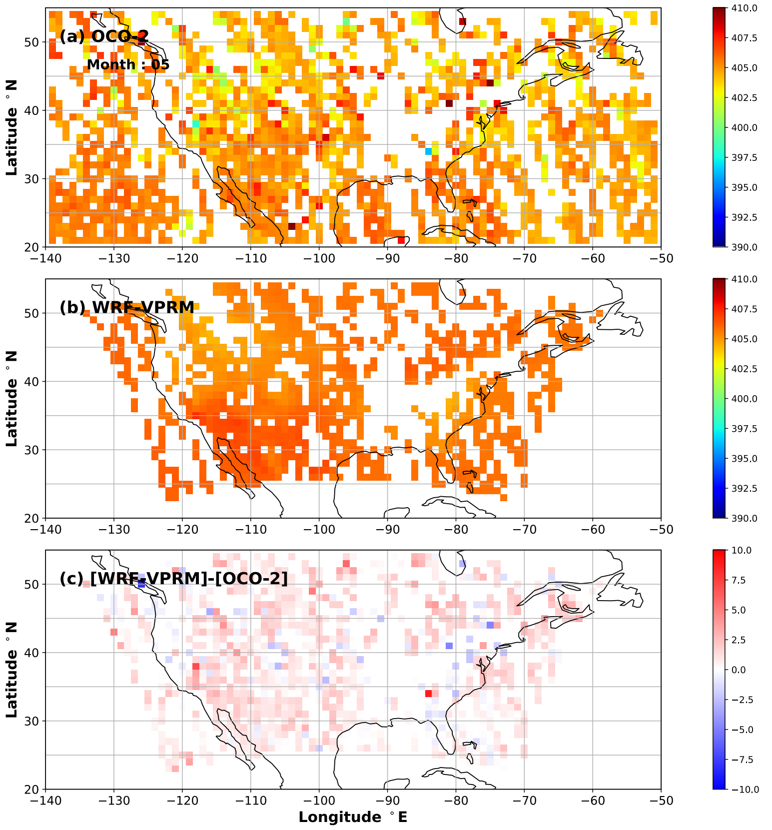


Figure II.5 – Same with Figure II.1 in May, 2016

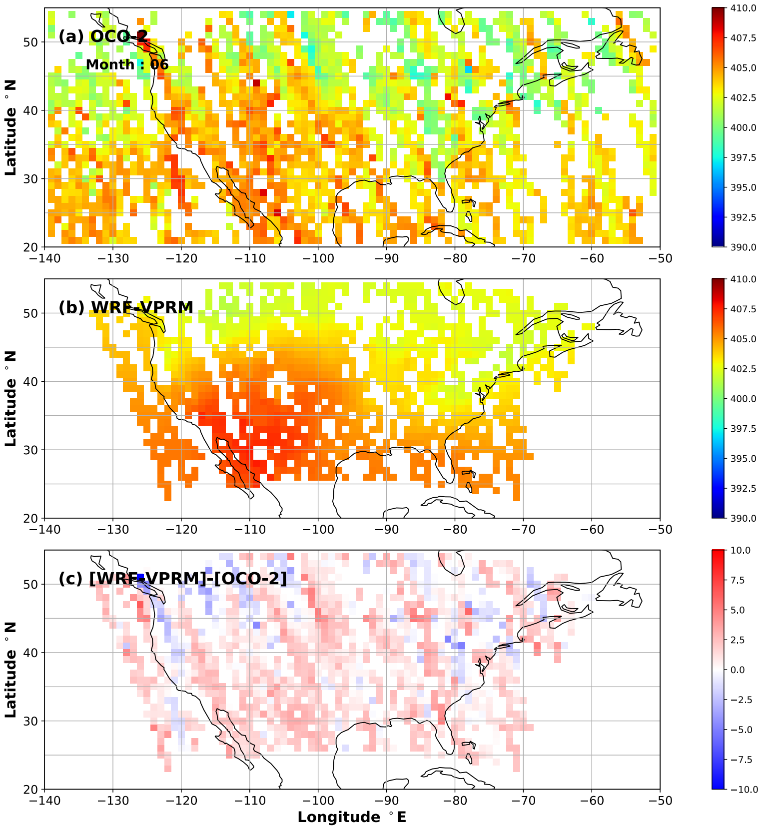


Figure II.6 – Same with Figure II.1 in June, 2016

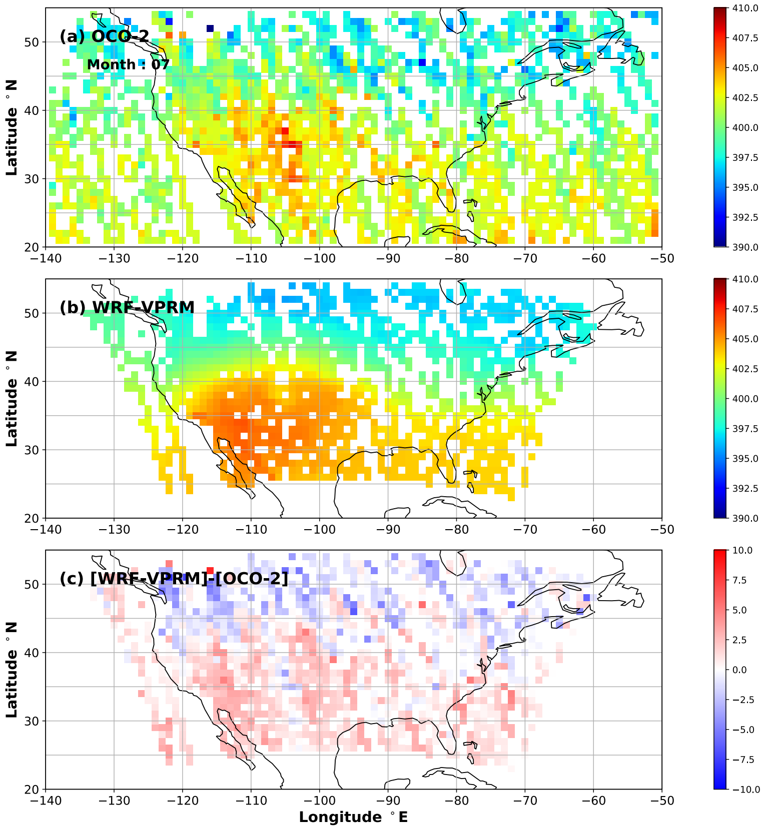


Figure II.7 – Same with Figure II.1 in July, 2016

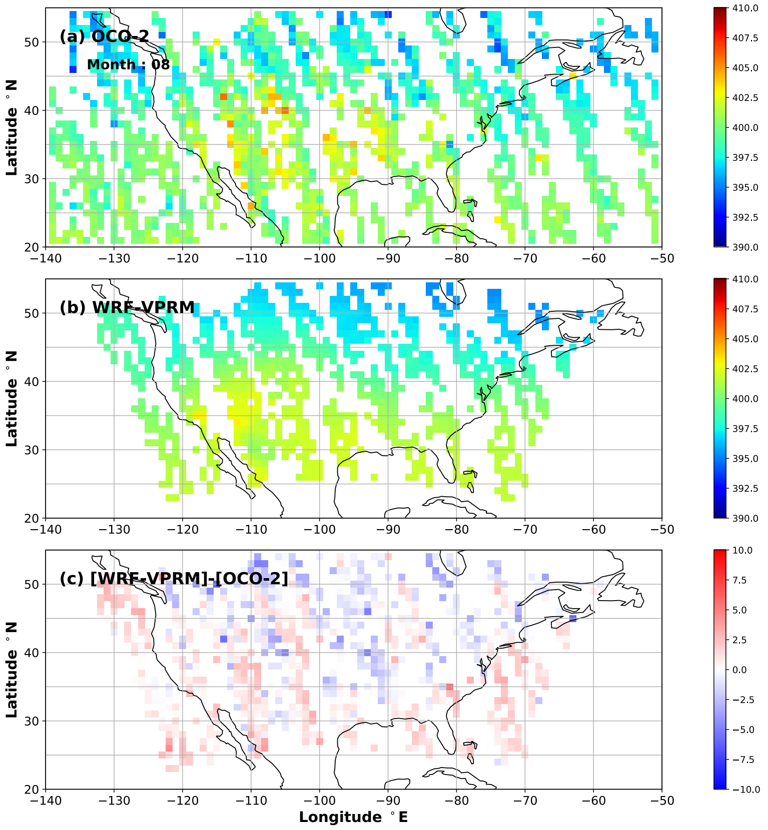


Figure II.8 – Same with Figure II.1 in August, 2016

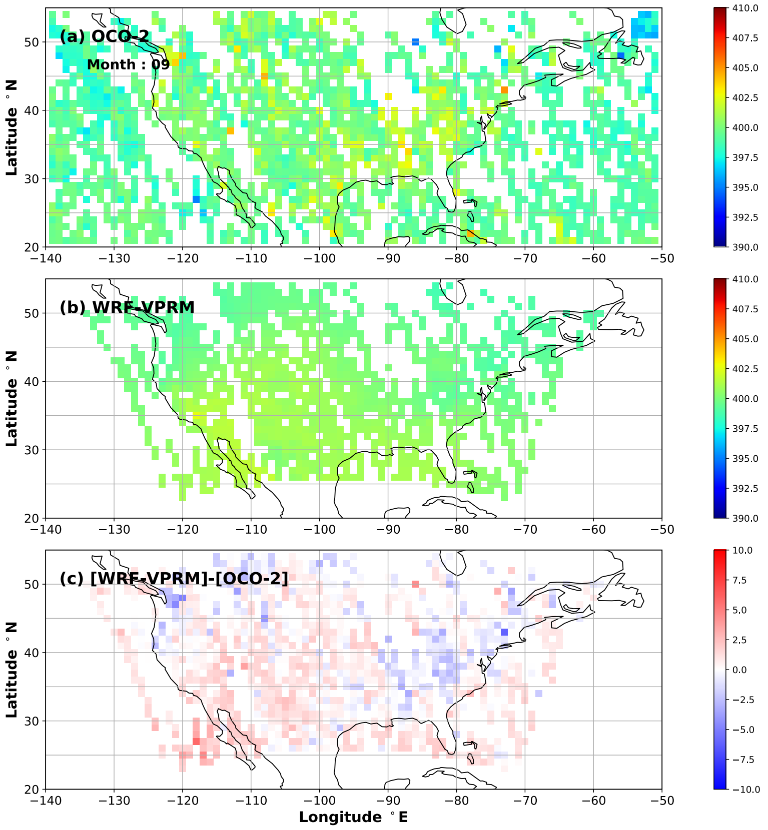


Figure II.9 – Same with Figure II.1 in September, 2016

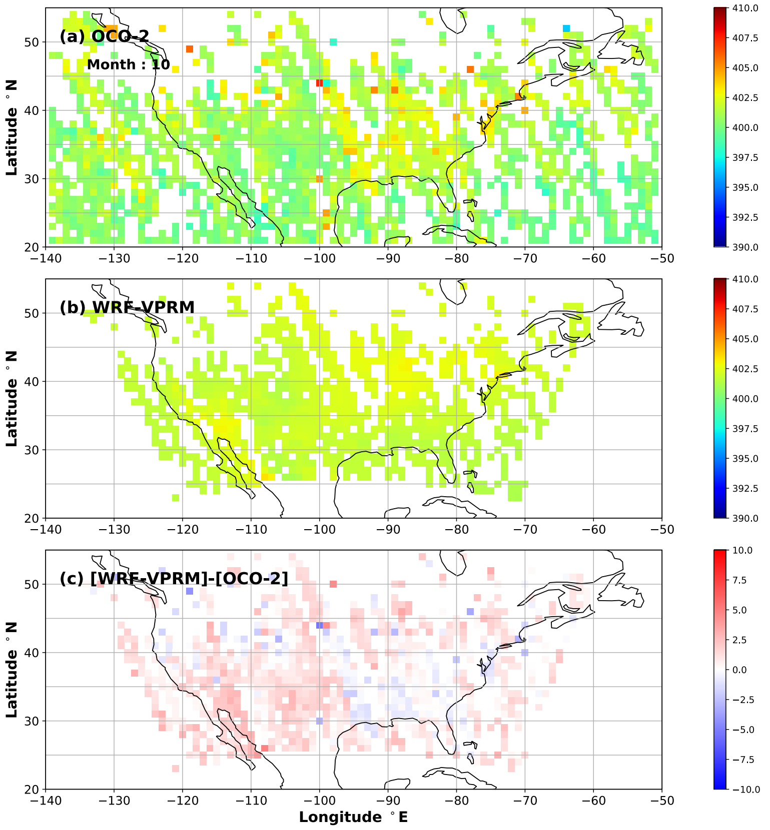


Figure II.10 – Same with Figure II.1 in October, 2016

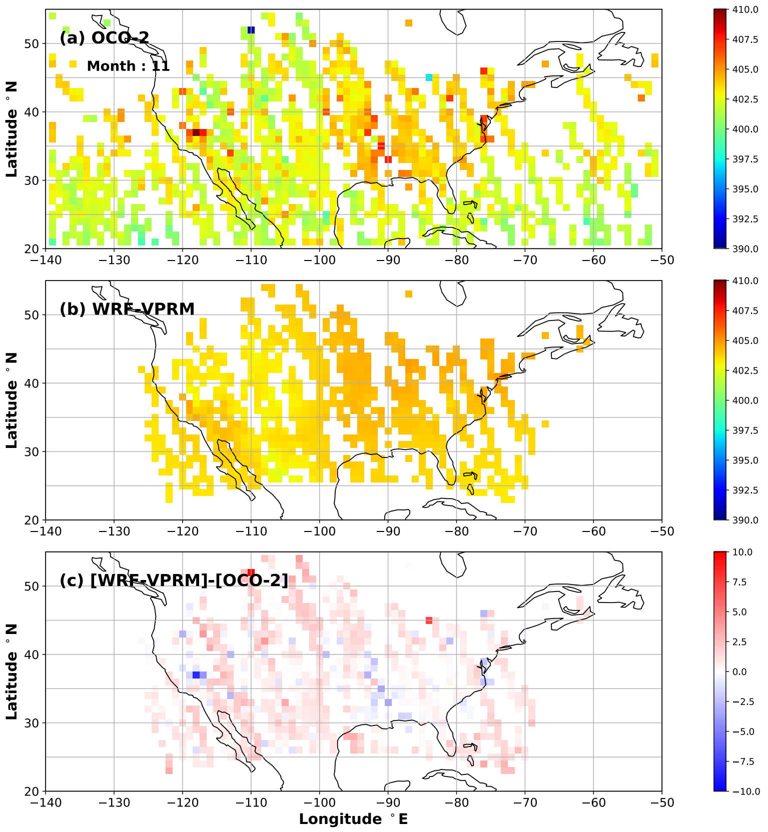


Figure II.11 – Same with Figure II.1 in November, 2016

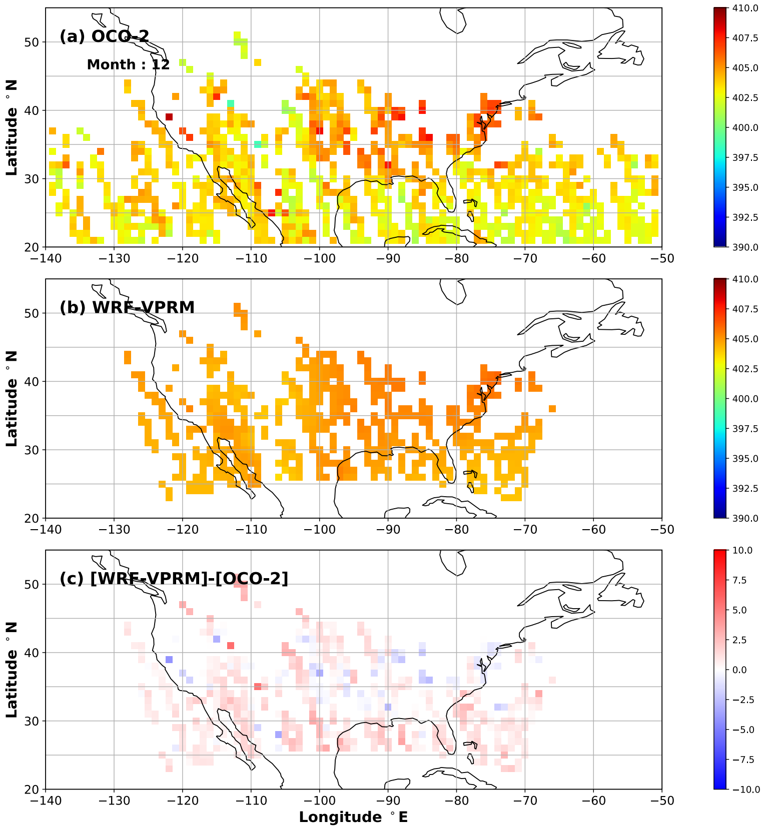


Figure II.12 – Same with Figure II.1 in December, 2016