First look at the NOAA Aircraft-based Tropospheric Ozone Climatology in Colorado

Mark Leonard[1,2], Irina Petropavlovskikh[1,3], Meiyun Lin[4], Audra Mcclure-Begley[1,3], Bryan J. Johnson[1], Samuel J. Oltmans[1,3], David Tarasick[5]

[1]NOAA Earth System Research Laboratory / Global Monitoring Division, Boulder, CO, United States,
[2]Science and Technology Corporation, Boulder, CO, United States,
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, United States,
[4]Program in Atmospheric and Oceanic Sciences, Princeton University and NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, United States,
[5]Experimental Studies Research Division, Meteorological Service of Canada, Environment Canada, Ontario, Canada,

Abstract

The Global Greenhouse Gas Reference Network Aircraft Program at NOAA has sampled ozone and other atmospheric trace constituents in North America for over a decade (2005-present). The method to derive tropospheric ozone climatology from the light aircraft measurements equipped with the 2B Technology instruments is described in this paper. Since ozone instruments at most of aircraft locations are flown once a month, this raises the question of whether the sampling frequency allows for deriving a climatology that can adequately represent ozone seasonal and vertical variability over various locations. Here we interpret the representativeness of the tropospheric ozone climatology derived from these under-sampled observations using hindcast simulations conducted with the Geophysical Fluid Dynamics Laboratory chemistry-climate model (GFDL-AM3). We first focus on ozone measurements from monthly aircraft profiles over the Front Range of Colorado and weekly ozonesondes launched in Boulder, Colorado. The climatology is presented as monthly values separated in 5th, 25th, 50th, 75th, 95th percentiles, and averaged at three vertical layers: lower (1.6-3 km), middle (3-6 km), and upper (6-8 km) troposphere. The aircraft-based climatology is compared to the climatology derived from the nearest located ozonesondes launched from Boulder, Colorado, from GFDL-AM3 co-sampled in time with in-situ observations, and from GFDL-AM3 continuous 3-hour samples. Based on these analyses, we recommend the sampling frequency to obtain adequate representation of ozone climatology in the free troposphere. The 3-hour sampled AM3 model is used as a benchmark reference for the under-sampled time series. We find that the minimal number of soundings required per month for the all altitude bins (1.6-3, 3-6, and 6-8 km) to sufficiently match the 95% confidence level of the fully sampled monthly ozone means vary between 3 to 5 sounding per month, except in August with a minimum of 6 soundings per month. The middle altitude bin required the least number of samplings per month. We determine the reasonably good agreement between the ozonesondes and aircraft measurements near Boulder suggest that valuable climatologies could be developed from the aircraft sites where no ozonesondes exist even though the aircraft measurements are more limited in number than the ozonesondes. When averaged over a number of years the aircraft data provide valuable information. More frequent sampling could tell us more but the measurements given would indicate that they can provide interesting climatological results.

1. Introduction

The Tropospheric Ozone Aircraft Measurement Program is part of the NOAA Global Greenhouse Gas Reference Network. It collects information to analyze ozone vertical, temporal and interannual variability in the troposphere across North America. Ozone profile data were continuously collected at 7 locations over the last 10 years. Measurements taken by the in-situ instruments aboard aircraft are aimed to assess pollution events, boundary layer stability, tropospheric ozone changes, annual ozone cycles, and provide climatological information for model verification and initialization. These goals are similar to the former MOZAIC (Measurement of Ozone and Water Vapour on Airbus in-service Aircraft) and current IAGOS (In-service Aircraft for a Global observing System) program (Marenco et al., 1998), which collects atmospheric composition data from regularly calibrated in-situ instruments (TEI-49) aboard large commercial aircraft in human influenced regions since 1993. While the MOZAIC/IAGOS programs are linked to large international airport areas, where measurements near the surface can be prone to the urban pollution influence, ozone data collected by light commercial aircraft used in this investigation are representative of baseline conditions away from the large sources of air pollution.

Ozonesondes have a long record of measurements and data analysis suitable for comparing ozone vertical distribution and seasonality throughout the troposphere. Previous studies have compared the distribution and seasonal variability of ozonesondes, aircraft and models, along with discussions of measurement uncertainties (Thouret et al., 1998A; 1998B; Logan et al., 1999; Tilmes et al., 2012; Liu et al., 2013; Zbinden et al., 2013; Eskes et al., 2015). Recently, Lin et al., (2015a) used hindcast simulations with the GFDL-AM3 chemistry-climate model to interpret the representativeness of ozone interannual variability derived from weekly ozonesondes at Boulder and Trinidad Head. They found that the available weekly ozone profile measurements at these sites were too infrequent to capture the actual interannual variability of mean mid-tropospheric ozone in April-May when stratospheric ozone intrusions peak seasonally over the western US. Lin et al., (2015b) further investigated the representativeness of ozone trends derived from sparse in situ measurements during April-May over 1995-2008 in the western US free troposphere, originally reported by Cooper et al., (2010). They found that GFDL-AM3 co-sampled in space and time with observations reproduces the observed ozone trend (0.65 ± 0.32 ppb yr^{-1}), whereas the model "true average" with continuous daily sampling indicates an insignificant trend (0.25 ± 0.32 ppb yr^{-1}). These prior studies indicate that sampling deficiencies can substantially influence calculated ozone means, interannual variability, and long-term trends.

The main objective of this paper is to evaluate and compare ozone climatologies between active years of aircraft sampling with ozonesonde data. Since ozone instruments at most of aircraft locations are flown once a month, this raises the question of whether the sampling frequency allows for deriving a climatology that can adequately represent the actual ozone seasonal and vertical variability. Therefore, we use ozonesondes with weekly sampling and the GDFL-AM3 model forced with reanalysis winds to interpret the representativeness of the ozone climatology derived from aircraft measurements. Despite mean state biases, the GFDL-AM3 model captures the salient features of tropospheric ozone over western North America, including the influences from Asian pollution and stratospheric intrusion events (*Lin et al.*, 2012a, 2012b), as well as their variability on interannual to decadal time scales (*Lin et al.*, 2014, 2015a, 2015b), indicating that the model represents the underlying processes controlling ozone variability, and is thus a suitable tool for studying the representativeness of under-sampled observational records.

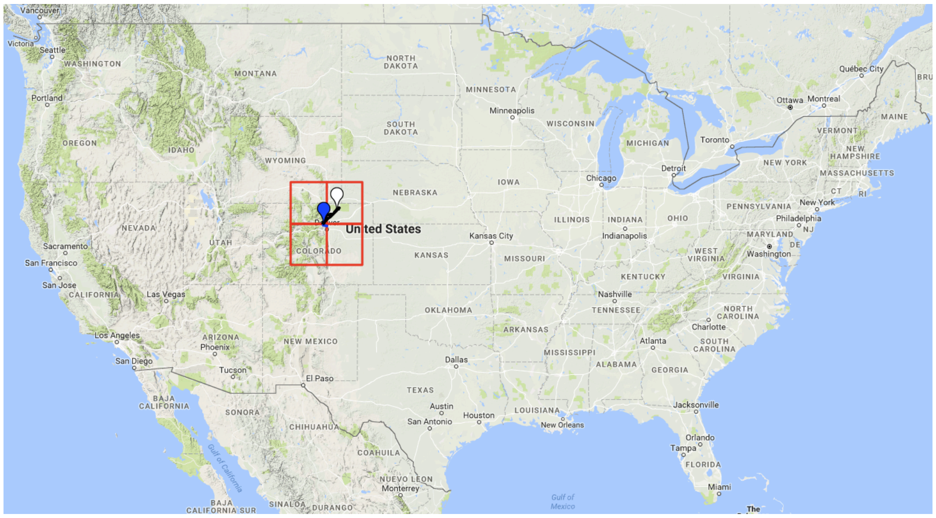
2. Ozone observations and the model

2.1 Light aircraft system for ozone measurements

Figure 1. Aircraft in-situ instrumentation package.
Model 205 Ozone Monitor repackaged in a suitcase for pilot convenience, instrument portability and durability. Additional ports were added for temperature and relative humidity probe as well as a GPS unit.

Light commercial aircraft, as part of the Global Greenhouse Gas Reference Network's aircraft program, are equipped with modified compact and lightweight 2B Technologies, Inc. Model 205 Ozone Monitors. The 2B ozone instruments use an ultraviolet (UV) absorption method to measure volume ozone mixing ratios following the beer-lambert law. These instruments are calibrated to the NOAA GMD surface ozone network standard instrument, which is regularly calibrated against the NIST Standard located in Gaithersburg, Maryland, USA. The instrument specifications are listed in **Table 1**. Air, after traveling into the inlet and through a particle filter, is “zero” referenced between two alternating cells. A mercury lamp illuminates these cells and a detector measures light not absorbed by ozone. A flowmeter is installed to maintain instrumental flow at approximately 1L/min. Flow measurements are taken continuously during flight. If the internal pump were to lose the required flow variable the data would be flagged for removal. Instrument measurements are taken at 10-

second intervals to provide a dynamical vertical resolution. Included with the ozone instrument, a Vaisala temperature and humidity sensor, and a Garmin GPS receiver are flown as a package (**Fig. 1**). Information about GPS latitude, longitude and geometric altitude is collected simultaneously with ozone measurements. More information on the aircraft program can be found at http://www.esrl.noaa.gov/gmd/ccgg/aircraft/. This study will be focused on aircraft flight data taken from flights based out of a Briggsdale, Colorado (CAR) (40.6347˚N, 104.3269˚W). CAR aircraft were flown bi-weekly to monthly from 2004-2014. Typical flight times range between 15-21 UTC (900-1500 MDT) and take place along the Front Range north of Denver, Colorado (**Fig. 2**). Ozone samples are collected over flat, mixed plains and agricultural land near Briggsdale in northeastern Colorado. A quality data filter was applied to each flight sampling to screen for instrumental errors. These include in-cabin air sampling, instrument warm up requirements, and uncertainty in measured pressure or temperature. This ensures that outlier values and periods of data contamination are removed from the record. All data recorded during flight was collected while the instrument was functioning correctly (within bounds of temperature, pressure, flow, etc).



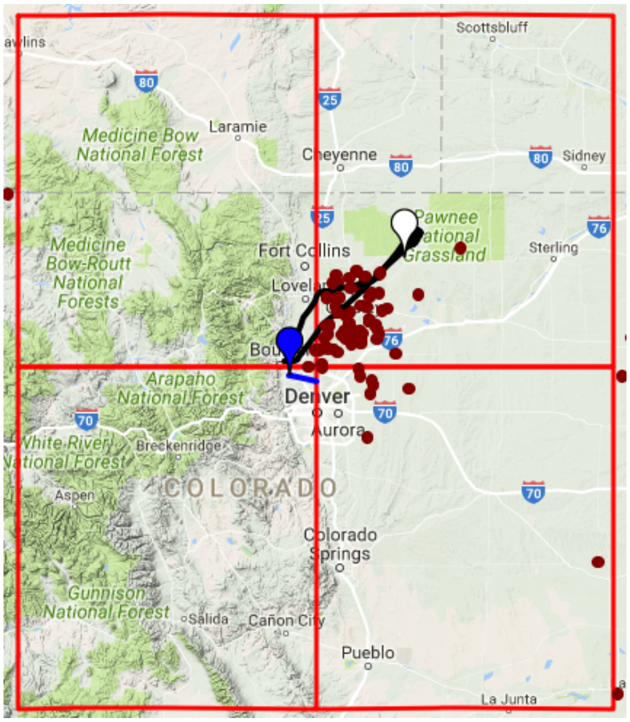


Figure 2. Locations of Observations from Front Range, Colorado. (Top) A large-scaled map of AM3 grid cells (red cells) location covering the Front Range of Colorado and launch location of ozonesondes in south Boulder, Colorado (blue marker) and aircraft at CAR station located in Briggsdale, Colorado (white marker). (Bottom) A regional map provides a typical flight trajectory of CAR aircraft (black) and trajectory of ozonesondes (blue) along the Front Range of Colorado (June 14-15, 2011). Included is the location of oil and gas wells (maroon) within the state of Colorado. These trajectories are analyzed further in Figure 4 where grid cells are labels Cell 1-4 and shown from upper right and counter-clockwise to lower right. Locations of oil and gas wells are courtesy of Colorado Oil and Gas Converation Commissions.

2.2 Ozonesondes

Vertical ozone profiles collected by CAR flights are compared to profiles taken with the electrochemical concentration cell (ECC) ozonesondes launched on balloons east of Marshall Lake in South Boulder, Colorado (39.949072˚N, -105.197083˚W). ECC ozonesondes are produced by Environment Scientific ENSCI 2ZV7 and are interfaced with an iMet radiosonde, which measures pressure, temperature, relative humidity, and GPS geopotential altitude along with ozone data from the ozonesonde. Ozonesondes have been occasionally equipped with GPS latitude and longitude measurements since 2004.

There were changes in ozonesonde operating procedures affecting the Boulder record. August 1997 was the transition from 1% Full Buffer Solution to 2% No Buffer Solution. Also, September 2005 had another transition from 2% No Buffer to the current solution 1% 1/10th Buffering. For more information, refer to NOAA Global Monitoring Division 10-year Summary Report No.28 2004-2013. Previous sets of ozonesonde datasets used to create climatologies (later discussed in Sections 3.2) have been processed in a way to fix some of the inconsistencies due to changes in the operating procedures. The data homogenization paper by Sterling et al., 2017 (in preparation) outlines the corrections used to adjust for the sensor solution changes, which are based on field and laboratory experiments.

The CAR aircraft and Boulder ozonesonde data can be downloaded from the NOAA GMD interactive data viewer http://www.esrl.noaa.gov/gmd/dv/iadv/.

ECC ozonesondes are launched weekly and have multi-decadal records. They provide ozone mixing ratios at high vertical resolution. Analyzed ozonesonde data in this study is based on the ascending portion of measured profiles because it provides the best vertical resolution and is less prone to measurement errors. Further information on ECC ozonesonde instrumentation can be found i.e. Komhyr et al., (1969); Johnson et al., (2002). Each profile is converted from geopotential altitude to geometric altitude for comparisons with CAR aircraft measurements.

When comparing ozonesondes and aircraft data, it is worth mentioning spatial and temporal differences that may be found within the planetary boundary layer (PBL) along the Front Range, Colorado. Data collection at the lowermost altitudes is affected by geographically diverse sources of anthropogenic and natural emissions of ozone precursors. Ozone measurements may vary within the PBL depending on the observational location along the Front Range, for instance, influenced by the daily buildup of surface ozone in the late summer months (e.g., Pfister et al., 2014). The site for the ECC ozonesonde launches is located 13 km south of Boulder, Colorado, whereas CAR aircraft flights originate from Briggsdale, Colorado that is characterized as a background site with no strong influence of urban pollution (Fig.2). Between these sites that are roughly 110 km apart there is an abundance of oil and gas wells. According to the U.S. Energy Information Administration, Top 100 U.S. Oil & Gas Fields (March 2015), Colorado is within the top 10 largest oil and gas fields in the United States, residing mostly along the Front Range. Studies have shown oil and gas wells emit large quantities of raw natural gas, including ethane and propane into ambient air, which act as a source of ozone precursor pollutants (Gilman, et al., 2013). Ozonesondes are launched southwest of most oil and gas wells, but during summer months, high concentrations of ozone are transported towards the ozonesonde site with prevailing winds detected from the northeast of Boulder (Fig.2 and **Supplementary Fig. S1**). This high concentration of ozone during summer months may influence climatological comparisons between ozonesonde and aircraft observations.

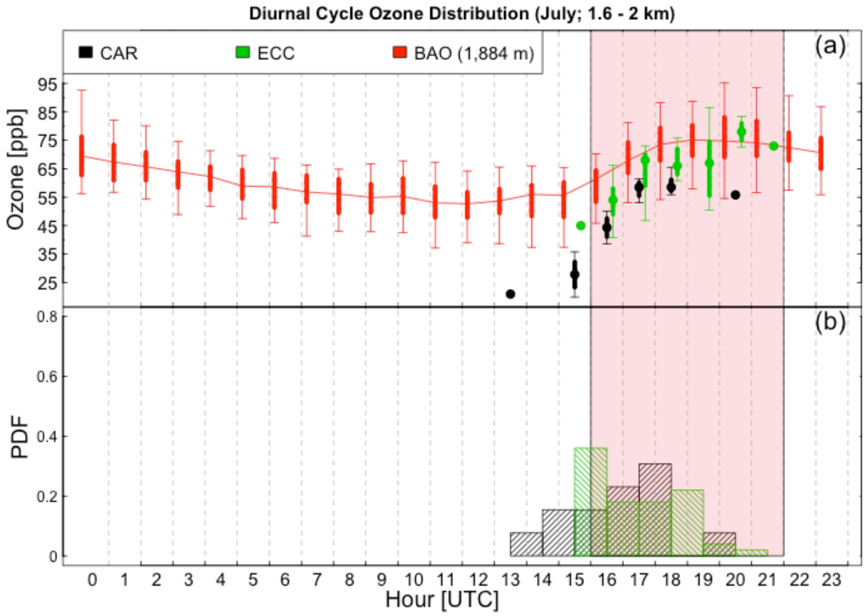


Figure 3. Diurnal distribution of measurements in July. (Top) Diurnal cycle of ozone in July from CAR aircraft measurements at **1.6-2 km altitude** (black), ECC ozonesonde (green) at 1.6-2 km altitude, and surface monitoring site at the Boulder Atmospheric Observatory (BAO; 1.8 km altitude) located in Erie, Colorado from years 2012-2014. The box-and-whiskers plots represent the 5th, 25th, 50th, 75th, and 95th percentiles of vertical ozone distribution taken from 2004 to 2014. (Bottom) Probability distribution of the sampling hours of aircraft (black) and ozonesonde measurements (green). The pink shaded area represents the hours used in this investigation.

Other differences may also be attributed by the later ozonesonde launching times compared to aircraft flight times in July (Fig.3). The Boulder Atmospheric Observatory (BAO) in Erie dataset is based on continuous ozone measurements taken by 2B instruments (similar to the aircraft instruments described in Section 2.1) located at 300 meters above the surface. The data is aggregated into hourly averages from July of 2012-2014 and shown in Fig. 3a as a reference for comparisons between ozonesondes and aircraft.

According to the diurnal cycle derived from ozone values averaged between 1.6 to 3 km, ozonesondes on average show higher ozone enhancements than aircraft observations (**Fig. 3a**). This may reflect the fact that ozonesondes are more frequently launched in later hours than CAR aircraft flights, as illustrated in the probability distribution function (PDF) of observational recording times from 2004 to 2014 time period (**Fig. 3b**). Late afternoon ozone measurements in July derived from ozonesondes are more consistent with in-situ ozone observations collected at 300 meters on the BAO tower than aircraft measurements. This can be attributed to diurnal changes in the PBL that can extend above 300 meters by early afternoon thus affecting both the BAO measurements and ozonesondes whereas the aircraft measurements are sampled within or above the PBL depending on the time and location. Therefore there might be bias between aircraft and other measurements.

2.3 Model

A 10-year hindcast simulation with the GFDL-AM3 global atmospheric chemistry-climate model, nudged to NCEP/NCAR winds, is used in this study to guide observation comparisons and evaluate any temporal and spatial differences found between CAR aircraft and ECC ozonesonde samplings. This model has a horizontal cubed-sphere grid resolution of about 200 x 200 km^2 and vertical resolution of about 70 m near the earth's surface to 1-1.5 km near the tropopause (Donner et al., 2011; Lin et al., 2012b; 2014).

The AM3 model has been used in comparison with ozonesonde measurements in California and in analysis of stratospheric intrusions. Lin et al., (2012a) showed that intrusions of ozone from the stratosphere affect air quality in the western US. The ability of the GFDL-AM3 model to capture tropospheric ozone variability was also verified in comparisons with ozonesonde measurements. The model was able to explain 50-90% of observed daily O_3 variability in Trinidad Head, Point Reyes, Point Sur, and Joshua Tree ozonesonde sites (see Supplementary Figs.S1-S2 in Lin et al., (2015a);http://www.nature.com/ncomms/2015/150512/ncomms8105/extref/ncomms8105-s1.pdf). Therefore, this model is suited for the assessment of the representativeness of ozone variability and trends derived from weekly ozonesonde and monthly aircraft observations in Colorado.

One of the limitations of the model is its coarse spatial resolution (~200x200 km^2). It is hard to detect the spatial differences between the aircraft and ozonesonde locations at this resolution. In addition, 3-hour averaged ozone fields archived from the model are used in this study, but ozonesonde vertical profiling of the troposphere is usually limited to less than 30 minutes. The ozone comparisons between the model and ozonesonde profiles might differ in the troposphere due to temporal or spatial averaging. For instance, the timing or spatial location of the stratospheric intrusions can vary; they may reach the free troposphere over Boulder an hour after an ozonesonde launch. Nevertheless, we are using the model to capture seasonal variability that is representative of the intra-annual variability. At the same time, we expect the model to capture the ozone variability in association with synoptic-scale meteorology as the model is relaxed towards assimilated meteorological fields.

3. Method of comparing datasets

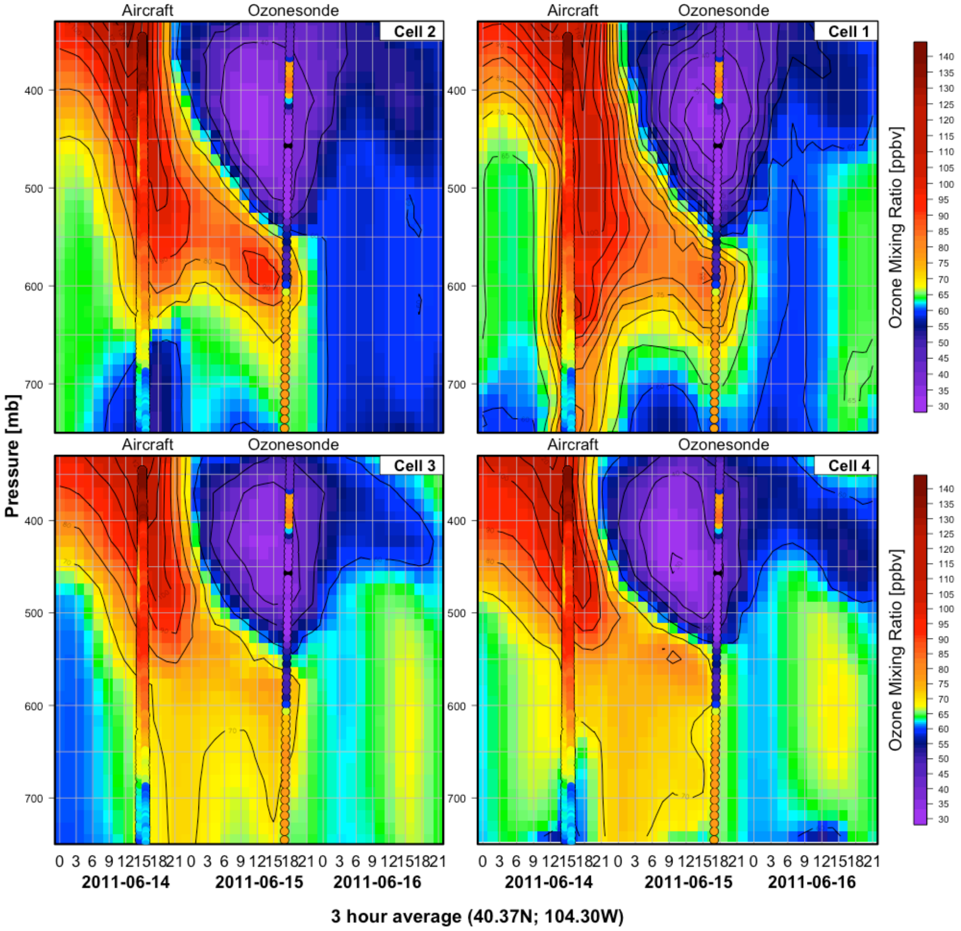


Figure 4a. Case Study of June 14-15, 2011.

Time series of three hourly averaged tropospheric ozone profiles for June 14 to 16 2011 from the GFDL-AM3 model (color shading) sampled at the four grid boxes closest to the observation locations (confined to coordinates 103.75-106.25˚W and 39-41˚N). Color-coded circles denote observations from the aircraft profile over the Colorado Front Range on June 14 and Boulder ozonesonde on June 15.

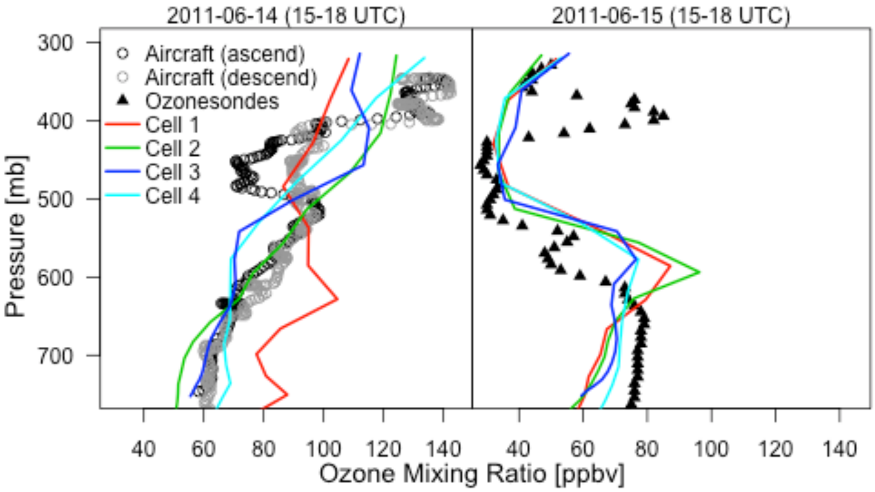


Figure 4b. Ozone profiles from aircraft (black and grey circles) on June 14 and ozonesonde at Boulder (black triangles) on June 15, along with the model results sampled at the four closest grid cells.

3.1 Observations and Model

For comparison with aircraft data, ECC ozonesondes and the GDFL-AM3 model are sampled from years 2004-2014 and for hours 15-21 UTC (900-1500 MDT) of active CAR flight data. This range of hours is found in 98% of all ozonesonde records. The number of aircraft and ozonesonde profiles used for deriving the monthly climatology is shown in **Table 2**. The sample size for each month ranges from 4 to 15 for aircraft flights and 38 to 91 for ozonesondes. The AM3 model provides a continuous record (3-hour averages) of tropospheric ozone profiles during active years of aircraft flights over the Colorado Front Range. For spatial proximity to observations, the model is confined to four grid cells at coordinates (103.75-106.25˚W) and (39-41˚N). The model uses atmospheric pressure as a vertical parameter and can be compared to observational measurements using a conversion from pressure to altitude by a derivation based on a subset of the International Standard Atmosphere (ISA) model formulated by the International Civil Aviation Organization (ICAO). For illustrative purposes, **Figure 4a** shows time series of the three hourly averaged ozone profiles from the GFDL-AM3 model at 4 grid cells during June 14-15, 2011 (see Figure 2 for location of grid cells relative to the Colorado Front Range). The curtain plots show descent of the enhanced stratospheric ozone air to the troposphere. The four panels show spatial variability of the ozone filament descending over the Front Range. The vertical ozone distributions sampled by aircraft on June 14, 2016 and by ozonesonde on June 15, 2011 are over-plotted in four panels to provide an easy comparison with the modeled results. The model captures the large-scale ozone enhancements at 400-600 hPa observed by the aircraft on June 14 as well as the rapid decrease on the following day observed by the ozonesonde at Boulder. The model missed the small-scale ozone filament at approximately 400 hPa recorded by the sonde.

In order to match spatial variability of the modeled ozone fields with aircraft sampling location, a distance-weighted average from four AM3 grid cells is calculated with respect to GPS coordinates taken on each flight. This method is used because the ozonesonde launch location and CAR station are located near the intercept of four individual grid cells. This model-averaging approach is illustrated in the June 14-15, 2011 case study, briefly discussed above. **Figure 4b** shows detailed comparison of the aircraft and ozonesonde profiles (black symbols in respective left and right panels) with AM3 model co-sampled with observations in time and space. The spatial ozone variability in AM3 model is represented by profiles selected from 4 matching grid cells (profile colors match colors of the cells shown in Figure 2). The large spatial tropospheric ozone variability across Colorado Front Range on June 14, 2011 is captured by the model (left panel of Figure 4). None of the individual cell profiles closely match the aircraft sampled ozone vertical distribution, but the weighted average profile compares well with observations.

Since the model, with roughly 0.7-1 km vertical resolution in the middle to lower troposphere cannot reproduce a fine vertical resolution of observations, we bin both observations and models into in the following layers: 1.6-3 km, 3-6 km, and 6-8 km. The lowest altitude (1.6 km ASL) is roughly the elevation along the Colorado Front Range. These altitude bins are chosen to represent the mixed planetary layer (MPL), free troposphere, and upper troposphere altitudes. Averaging vertical profiles within these altitude bins helps to minimize the impact of non-representative tropospheric variability due to poor temporal sampling of the aircraft record, and the impact from the ozone filaments that impact the climatologies, predominantly in spring season. We construct a vertical ozone seasonal climatology at these layers.

Within each altitude section, daily median ozone values were aggregated to produce monthly median climatologies for aircraft, ozonesondes, and the continuously sampled AM3 “true” model (referred to as AM3-True). The model was also co-sampled separately to match the dates of each aircraft flight and ozonesonde. These daily median values were then aggregated to a monthly median climatology. Also included are standard deviation and standard error associated with monthly variability. Monthly median values were used because under-sampled aircraft flights and ozonesonde soundings may produce a wide range of variability that is best described using a median rather than an arithmetic mean.

For further comparisons, vertical profiles were made using 1 km intervals. Within each altitude interval, daily ozone distribution for both aircraft and ozonesonde measurements were compared using a 5th, 25th, 50th, 75th, and 95th percentiles. This same comparison method was used to help differentiate the vertical variability in the model. In this comparison, the model-based AM3-True and corresponding co-sampled climatologies were compared to the respective aircraft and ozonesonde climatologies. This comparison between observation and the model also frames the discussion on limitations in the observation sampling frequency and its influences on the ozone climatology development.

3.2 Other climatologies

In this study we also create monthly mean climatology based on ozonesondes data measured at Boulder from 2004-2014. This climatology (further referred as Leonard climatology) uses mean instead of median ozone values to match the historical climatology, while it also averages highly resolved ozonesonde profiles into three thick altitude layers (1.6-3 km, 3-6 km, and 6-8 km). We compare Leonard climatology with the previously published ozonesonde climatology for the Front Range of Colorado published in Logan et al., (1999), Tilmes et al., (2012) and Liu et al., (2013).

The Logan climatology is available as monthly means of ozonesonde soundings launched near Boulder, Colorado from 1985 to 1993. Ozone values are given in partial pressure, but converted to a mixing ratio. Monthly means are provided from 1000 to 10 hPa in 100 hPa intervals, but converted to altitude (km) for comparisons with other climatologies. This data is available at http://www.people.fas.harvard.edu/~logan/research7.html.

The Tilmes climatology is available as monthly means of ozonesonde soundings launched near Boulder, Colorado from 1995 to 2009 using altitude grid with 0.5 km resolution. Data is available at http://www.acom.ucar.edu/gctm/download.shtml. The Liu et al., (2013) analyses provide a globally and vertically gridded dataset of monthly averaged ozone. While the typical sampling of the ozone by balloon-borne sounding at a station is limited to once a week launches, the method described in Liu et al., (2013) provides additional data to the gridded ozone record by means of a long-range-transport of the ozone air masses driven by meteorological reanalysis. It redistributes ozone from various geographical places where it was measured by ozonesondes to the remote grid point and effectively increases the number of “sampling” days available for the gridded record in each month. Liu et al., (2013) dataset will be further called TOST (Trajectory-mapped Ozonesonde dataset for Stratosphere and Troposphere). TOST dataset is available as monthly means ozone time series from 1985 to 2013. In order to spatially match Boulder ozonesonde record, TOST was sampled within the 5x5 degree box based on the longitude and latitude that is collocated to represent the Front Range of Colorado. TOST is available at ftp://es-ee.tor.ec.gc.ca/pub/ftpdt/ under “Tropospheric Climatology”. In order to compare TOST with the modern-day climatologies (i.e Tilmes and Leonard), the 2004-2013 subset of time series was averaged into 12 monthly mean values (henceforth called TOST (2013)).

To help link any changes between the Logan time period and more recently derived climatologies, a historical subset of available monthly mean data from Liu is selected between years 1985 and 1993 (further called TOST (1985)). These previously published climatologies were combined into 3 altitude layers (1.5-3km, 3-6km and 6-8km) to compare with the Leonard climatology.

4. Results

4.1 Comparison with previously published climatologies

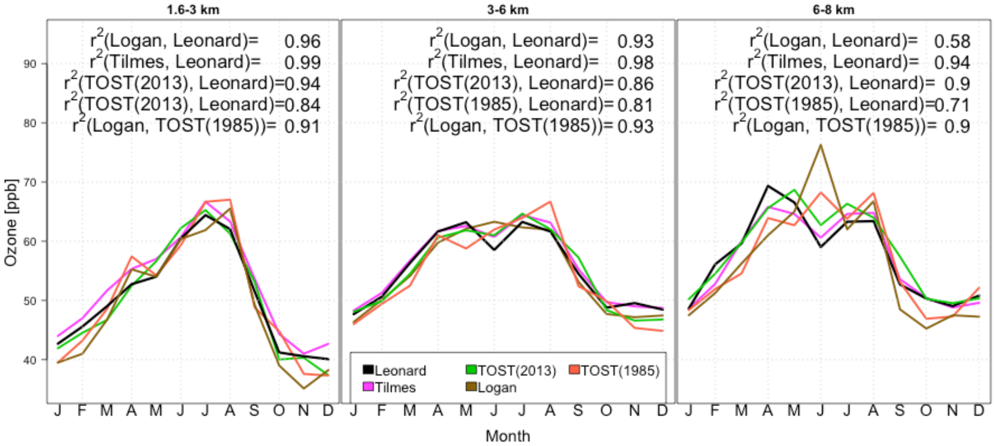


Figure 5. Comparisons of monthly mean ozone climatology at Boulder ozonesonde site derived from the present study for the period 2004-2014 (black) with those from the published literature: Logan et al., (1999) for the period 1985-1993 (brown), Tilmes et al., (2012)

for the period 1995-2009 (magenta), Liu et al., (2013) for the period 2004-2012 (green) (called TOST(2013)) and for the period 1985-1993 (called TOST(1985)) (orange). Note that Liu spatially interpolates ozonesonde profile sampling using forward and backward trajectories of meteorological reanalysis between 2004-2012 and again between 1985-1993.

When comparing the Leonard climatology to past studies, it was found that the best agreement occurred within the lower (1.6-3 km) and middle (3-6 km) layers of the troposphere (see Supplemental Table 1). It was also found that the seasonal distribution in the climatologies may have reformed over the time span of years 1985-2014.

Beginning with the lower tropospheric layer (1.6-3 km), we note that all climatologies are in agreement in regards to the amplitude of the seasonal cycle, while offsets are found for individual months. In comparison to the more recent climatologies (i.e. derived over the more recent time-period of Boulder ozonesonde record, after 1995), the two historical climatologies of TOST(1985) and Logan, contain lower monthly ozone values throughout the late fall and winter months. These climatologies also exhibit higher ozone anomalies in April and August, whereas more recent climatologies find lower ozone levels in April and summer ozone maximum in July.

In the middle layer (3-6 km) of the troposphere, all discussed climatologies are in fairly close agreement. The historical TOST(1985) climatology exhibits the largest deviations in month-to-month variability compared to the other climatologies, featuring a high ozone anomaly in April and the lowest ozone levels in winter and early spring. Even though Logan spans the same time period (1985-1993) as TOST(1985), it does not follow the same seasonal distribution, instead it resembles the more recent climatologies. Since Logan and TOST(1985) have large deviations in their respected seasonal ozone distributions no conclusion can be made about decadal changes in the 3-6 km tropospheric ozone presented in the more recent climatologies in the 3-6 km layer. Still withstanding, the Leonard climatology is found in reasonable agreement with all other climatologies.

The largest variability between the climatologies is found in the upper layer (6-8 km) of the troposphere. In this layer, it is more apparent that the shape of the seasonal cycle may have changed over the last 3 decades. The two historical climatologies have a comparable seasonal cycle that is different from the three more recent climatologies. The difference between these climatologies in the upper troposphere may be attributed to a possible decadal change in the position of the sub-tropical jet stream relative to Boulder, CO, which could affect stratosphere-to-troposphere ozone transport (Lin et al., 2012) and atmospheric composition in the upper troposphere/lower stratosphere (UTLS) layer in the vicinity of the jet (Manney et al., 2015).

4.2. Comparison of ozonesonde and aircraft climatology

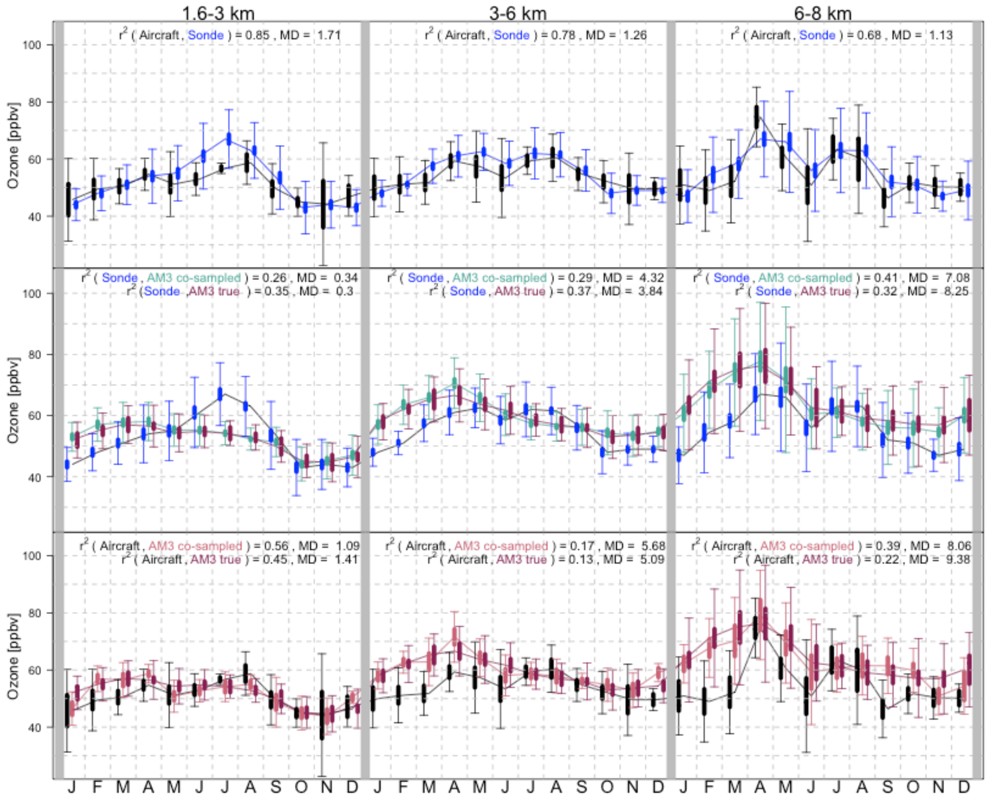


Figure 6. Monthly ozone climatology along Front Range, Colorado. Shown are monthly median climatologies derived from CAR aircraft observations (black), ECC ozonesonde observations (blue), and simulated results of the continuously daily sampled (AM3-True) (purple) and the GDFL AM3 model co-sampled to aircraft (pink) and ozonesondes (light blue). The model is co-sampled to match the dates of aircraft flights and ozonesondes. The box and arrow plots represent the standard error and standard deviation for each monthly median value. Climatologies are taken at altitudes of 1.6-3 km, 3-6 km, and 6-8 km.

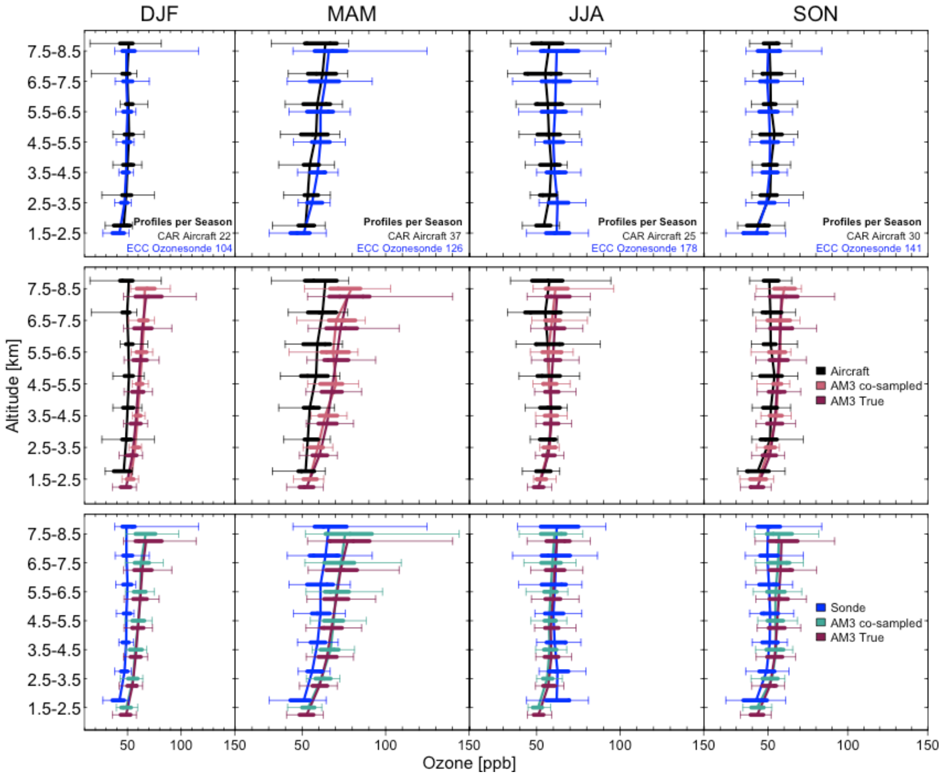


Figure 7. Seasonal median ozone vertical profiles along Front Range of Colorado.
Comparison of seasonal ozone vertical profiles: (top panels) between CAR aircraft (black) and ECC ozonesonde observations (blue); (middle panels) among aircraft data (black), AM3 co-sampled on aircraft launch days (pink), and AM3 "true" average with continuous daily sampling (brown); (bottom panels) among sonde data (blue), AM3 co-sampled on sonde launch days (cyan) and AM3 true average (brown). The box-and-whiskers plots represent the 5th, 25th, 50th, 75th, and 95th percentiles of vertical ozone distribution taken from 2004 to 2014. The number of observational samplings per meteorological season is labeled above.

Since ozonesonde profiles collected in Boulder, CO from 2004 to 2014 provide an adequate climatological representation of ozone variability in the troposphere (i.e Tilmes et al., 2010; Logan et al., 1999; Liu et al., 2013), the CAR aircraft-based climatology is compared to the ozonesonde climatology (top panels in Fig. 6 and Fig.7). The sonde and aircraft climatologies show similar ozone seasonal cycle that feature the late summer maximum below 3 km altitude and the bimodal maximum (spring and late summer) at 3-6 km and 6-8 km altitudes. The summertime afternoon buildup of ozone may be a cause of higher ozone values recorded by ozonesonde data in July as compared to ozone data from the earlier aircraft flights. Below 3 km altitudes, sonde measurements show 5-10 ppbv greater ozone values than aircraft data from May to September – the ozone pollution season in the boundary layer, consistent with the comparison for July presented in Fig.3. This reflects the fact that ozonesondes are launched at locations closer to urban pollution in Boulder and Denver (Fig.2) and more frequently in later afternoon hours when boundary layer ozone pollution peaks in this region (Fig.3), as discussed in Section 2.2. The monthly correlations between the two datasets range from R^2= 0.85 at 1.6-3 km altitudes, R^2= 0.78 at 3-6 km altitudes, R^2= 0.68 at 6-8 km altitudes (top panels in Fig.6). The vertical profiles of ozone derived from the two datasets are similar for winter (DJF) and fall (SON), but ozonesondes show ~5 ppb greater ozone values for spring (MAM) and summer (JJA) (top panels in Fig.7), likely for reasons noted above.

4.3 Model and observation comparisons

The middle and lower panels in Figs.6 and 7 compare sonde and aircraft measurements with the GFDL-AM3 model co-sampled with observations in time and the “true average” with continuous daily sampling (see also Supplemental Table 2 for the differences). AM3 captures the seasonal cycle of ozone measured by aircraft below 3 km altitude (r^2= 0.56 for AM3 co-sampled and r^2=0.45 for AM3-True) and the springtime maximum of observed ozone at 6-8 km altitude (r^2= 0.4 for co-sampled results). The simulated monthly variation of ozone is weakly correlated with aircraft measurements in the middle troposphere (r^2=0.13-0.17). With regard to the magnitude, AM3 shows overall positive biases during the non-summer months. The positive biases are found in both the AM3 co-sampled results and the model “true average” with continuous sampling (Fig.6 and Supplemental Table 3). The biases in winter and spring increase with altitude (Fig.7), indicating that excessive stratosphere-to-troposphere transport may be responsible for the biases (see more detailed discussions in Lin et al., 2012a). In the boundary layer (1.6-3 km altitude), the model agrees better with aircraft data than ozonesonde measurements (left panels in Fig.6). The model is biased low by ~10 ppbv compared to sonde measurements during the summer months, reflecting that the model has difficulty resolving observed urban pollution due to its coarse resolution (~200x200 km^2).

The continuously sampled model is included for comparisons with the observational climatology and vertical profiles. We find that monthly variations of ozone in sonde measurements correlates better with the model “true average” (R^2=0.35 to 0.37) than its respective co-sampled model climatology (R^2=0.26 to 0.29) within the lower and middle troposphere (1.6-6 km). In the upper troposphere (6-8 km), however, AM3 co-sampled with sonde launch days shows greater seasonal correlations with observations than the model “true average” (R^2=0.41 vs. 0.32). The model co-sampled results are also better correlated with aircraft measurements in the lower and upper troposphere, implying that sampling deficiency influences the calculated ozone climatology by aircraft in the lower and upper troposphere and by ozonesondes in the upper troposphere.

4.4 Sampling studies

We next evaluate the minimum number of vertical ozone profiles per month that would be required to obtain representative monthly mean ozone climatology within each discussed altitude bin. We use the 2004-2014 time series of the daily median AM3-true ozone simulations as a benchmark for the comparisons. This AM3-true time series, sub-sectioned within each altitude bin, was sequentially sampled by different frequencies of daily ozone values per each month throughout each year. Daily median time series sequentially selected for particular frequencies (e.g. 5 possibilities of 1 profile per weekday time series) were then aggregated by month to find the mean values for each different samplings time series.

The sequentially selected time series is a realistic method for testing how many soundings should be launched per week to capture the variability of any daily and continuously sampled month. In this method, each sampling frequency time series was remade with all possible combinations for number of days per weekday. For instance, a 2 days-per-weekday (Monday to Friday) frequency consists of 10 possible combination (e.g. Tuesday-Thursday, Monday-Wednesday) time series. These time series are aggregated by month to find their respective mean values.

We next evaluate these monthly mean values by comparing them to the continuously sampled 'true' climatology. For each month, the mean values produced by each combination of weekday time series is averaged and divided by the 'true' monthly means. Furthermore, this ratio of subsampled and 'true' monthly means is compared at each altitude bin. If the number of samplings per weekday provided a mean value that is within ±CI of the true mean value, it is considered adequate for that particular month.

Months	1.6-3 km	3-6 km	6-8 km
Jan	4 m.	4 m.	3 m.
Feb	4 m.	5 m.	4 m.
Mar	5 m.	3 m.	3 m.
Apr	4 m.	3 m.	4 m.
May	4 m.	4 m.	5 m.
Jun	4 m.	3 m.	4 m.
Jul	4 m.	4 m.	4 m.
Aug	6 m.	4 m.	4 m.
Sep	4 m.	4 m.	4 m.
Oct	4 m.	4 m.	4 m.
Nov	4 m.	5 m.	5 m.
Dec	4 m.	4 m.	4 m.

Table 3: Minimal number of samplings required for all possibilities of sequentially sampled monthly means to average within a 95% confidence interval (CI) of the 'true' monthly means. The brackets include the minimal number of samplings required for the CI of the sub-sampled monthly means to intercept with the CI of the continuously sampled 'true' monthly means. Abbreviation is m. (per month)**.**

We find that the minimal number of soundings required for the all altitude bins (1.6-3, 3-6, and 6-8 km) vary between 3 to 5 sounding per month to (see **Table 3.**), except in August with a minimum of 6 soundings per month. The middle altitude bin required the least number of samplings per month.

5. Conclusion

Observations of ozone from ozonesondes and aircraft were compiled and analyzed from different sections of the troposphere. Vertical ozone climatology and distribution profiles were compared along the Colorado Front Range from years 2004-2014. The GDFL-AM3 model ozone profile dataset was sampled to match each observation with the best spatial and temporal proximity that is limited to 3 hour and 200 by 200 km^2 resolution of the model. This report provided daily and monthly ozone values for Boulder, Colorado ozonesondes and CAR station aircraft along with various percentiles, standard deviations, error, and number of profiles used.

Seasonally and regionally aggregated ozonesonde climatologies were found to be similar to those of Tilmes et al., (2012) and Liu et al., (2013), even though years used in the present study were different. The climatologies derived from aircraft and ozonesonde data are similar in representing tropospheric ozone seasonal cycle and variability in the Colorado Front range area, except in the lower troposphere. Sonde measurements are shown to have ~10 ppb greater values in lower tropospheric ozone during summer months in comparison with aircraft measurements likely related to the greater impact of urban related pollution impacts at the ozonesonde site with its closer proximity to the Denver metropolitan area.

The comparison of median ozone climatologies between observations and GDFL-AM3 model simulations identified both similarities and differences in the seasonal cycles. The model captures the observed day-to-day variations of ozone in the free troposphere associated with tropopause folds (Fig.4) and the prominent seasonal cycle of ozone in the upper troposphere (Fig.6). The model has difficulty replicating small-scale differences, e.g., the weak seasonal cycle of ozone at 3-6 km altitude and the summertime urban pollution measured by ozonesondes at Boulder. Future work should explore model hindcasts at higher horizontal resolution (e.g., 50x50 km^2).

The present-day tropospheric ozone climatology was derived using aircraft and ozonesonde observations along Front Range Colorado. While CAR aircraft do not have an abundant sampling frequency, they provide vertical and seasonal variability in tropospheric ozone that is comparable to variability derived from 2004-2014 ozonesonde observations. Using bi-monthly sequential sampling of the AM3 true time series for analysis of the sampling limitations of aircraft measurements to derive tropospheric ozone climatology we find that sub-sampled and AM3-true climatologies agree within their respective 95% CI.

The tropospheric ozone climatology can be also derived from yet another NOAA location with co-located aircraft and ozonesondes monitoring, such as Trinidad Head, California (not presented in this study). The use of the high resolution AM3 model (i.e. 50x50 km^2 or better) can help to improve the geographical matching criteria with observations sampled near the ocean. Other NOAA light-aircraft sampling locations (provided at http://www.esrl.noaa.gov/gmd/ccgg/aircraft/) are stand-alone operations and do not have co-located ozonesonde measurements, and thus these climatologies are more difficult to validate. The aircraft based climatologies are useful for validation of the tropospheric/surface ozone predicted by regional and sub-scaled chemistry models because it represents ozone variability under the influence from local meteorological conditions and chemistry specific to aircraft locations. The limitation of the aircraft measurements is due to very limited temporal sampling. However, if the sampling is increased, these datasets could be useful for studies of the inter- and intra-annual ozone variability across the US. This includes validation of tropospheric ozone derived by current satellites (i.e. combined product from Aura OMI and MLS, GOME/SCHIAMACHI, Aura/TES, IASI, etc.) to refine their product for small-scale pollution events. In the future, tropospheric products from the JPSS/OMPS, and geostationary TROPOMI and Sentinel series satellites can be validated by a fleet of small commercial aircraft that can provide information on the temporal and spatial ozone variability with regards to the small geographical scales.

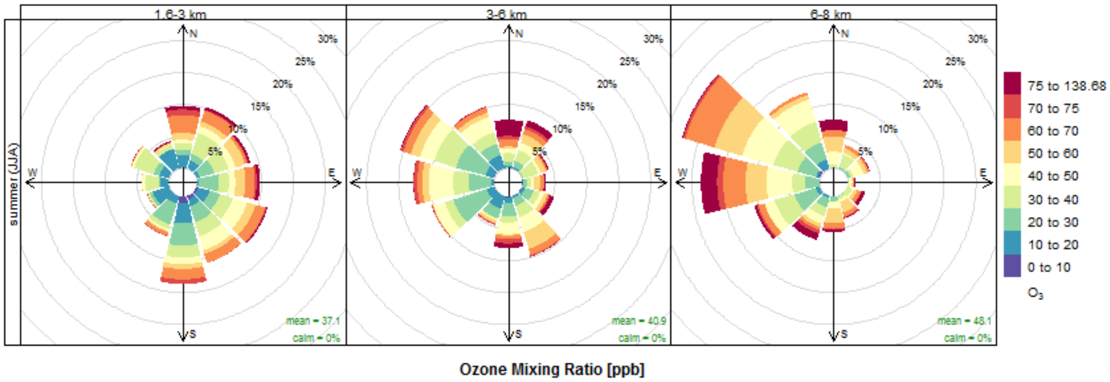
The reasonably good agreement between the ozonesondes and aircraft measurements near Boulder suggest that valuable climatologies could be developed from the aircraft sites where no ozonesondes exist even though the aircraft measurements are more limited in number than the ozonesondes. When averaged over a number of years the aircraft data provide valuable information. The aircraft measurements also have measurements from other species that could be useful in telling something about the source of ozone being measured. More frequent sampling could tell us more but the measurements given would indicate that they can provide interesting climatological results.

2B Technologies 205	Specifications
Range	1-250 ppmv
Zero Noise	.25 ppbv
Lower limit	2 ppbv
Drift: Zero	Less than 1 ppbv
Drift: Span	Less than 1 % full scale
Response time	10 seconds
Precision	Greater of 1 ppbv or 2%
Flow Rate	1.8 L/min
Calibration	NIST Traceable 0-200 ppbv

Table 1. Specifications of Model 205 Ozone Monitor.
Further information can be found at http://www.twobtech.com/model-205-ozone-monitor.html.

Month	CAR Aircraft Flights	ECC Ozonesonde Flights
1	7	44
2	11	38
3	9	51
4	15	52
5	15	47
6	11	61
7	8	50
8	7	91
9	11	52
10	12	48
11	7	46
12	4	39
Total	117	619

Table 2. Summary of observations taken per month during 2004-2014.



Supplementary Figure 1. Summer ozone concentration wind rose. Shown is a collection of summertime wind and ozone measurements collected from ozonesondes in south Boulder, Colorado. Altitudes are sectioned in from 1.6-3km, 3-6km, and 6-8 km. These altitude sections represent the surface and plantary boundary layer (PBL), free troposphere air, and upper troposphere.

Altitude	Tilmes 1995-2009	Liu 2004-2012	Liu 1985-1993	Logan 1985-1993
1.6-3 km	1.24±1.19	0.02±1.49	0.26±3.01	-1.52±2.75
3-6 km	0.54±1.12	-0.08±1.78	-0.71±3.08	-0.7±1.96
6-8 km	1.09±1.02	1.01±2.41	-0.82±4.50	-0.98±6.05

Supplemental Table 1. Monthly mean climatology difference in mean and standard deviation of previous studies.

12-month mean climatologies of previous studies are compared using mean differences and variance to the 2004-2014 ozonesonde-derived Leonard climatology within 3 altitude bins. Previous studies compared are Logan et al., (1999), Tilmes et al., (2012) and two separate datasets from Liu et al., (2013).

Altitude	Sonde	AM3 co-sampled	AM3-true
1.6-3 km	1.71±4.36	1.09±3.39	1.41±3.65
3-6 km	1.26±3.01	5.68±4.97	5.09±4.96
6-8 km	1.13±4.62	8.06±6.84	9.38±7.88

Supplemental Table 2. Monthly median climatology mean difference and standard deviation.
Within 3 altitude bins, monthly median climatologies for CAR aircraft (called Aircraft) are compared with ozonesondes (called Sonde), the continuously sampled AM3 'true' model, and the model co-sampled to match the dates of each aircraft flight using a 12-month mean difference and standard deviation.

Altitude	AM3 co-sampled with aircraft	AM3 co-sampled with ozonesondes
1.6-3 km	-0.31±2.64	0.65±0.87
3-6 km	0.59±2.46	0.48±1.56
6-8 km	-1.32±3.45	-1.17±1.34

Supplemental Table 3. Monthly median climatology between model samplings.
Within 3 altitude bins, monthly median climatologies for the continuously sampled AM3 'true' model is compared to the model co-sampled to match the dates of each aircraft flight and ozonesondes using a 12-month mean difference and standard deviation.

Cooper, O. R., Parrish D. D., Stohl A., Trainer M., Nédélec P., Thouret V., J. P. Cammas, S. J. Oltmans, B. J. Johnson, D. Tarasick, T. Leblanc, I. S. McDermid, D. Jaffe, R. Gao, J. Stith, T. Ryerson, K. Aikin, T. Campos, A. Weinheimer and M. A. Avery (2010), Increasing springtime ozone mixing ratios in the free troposphere over western North America, Nature, 463, 344-348, doi:10.1038/nature08708.

Johnson B., Oltmans S., and Vo¨mel H., Electrochemical concentration cell (ECC) ozonesonde pump efficiency measurements and tests on the sensitivity to ozone of buffered and unbuffered ECC sensor cathode solutions. J. Geophys.
Res. 107, D194394 (2002).

Donner L.J., Hemler R.S., Horowitz L.W., Ming Y., Ming Z., Golaz J.-C., Ginoux P., Lin S.-J., Schwarzkopf M.D., Austin J., Alaka G., Cooke W.F., Delworth T.L., Freidenreich S.M., Gordon C.T., Griffies S.M., Held I.M., Hurlin W.J., Klein S.A., Knutson T.R., Langenhorst A.R., Lee H.-C., Lin Y., Magi B.I., Malyshev S.L., Milly P.C.D., Naik V., Nath, M.J., Pincus R., Ploshay J.J., Ramaswamy V., Seman C.J., Shevliakova E., Sirutis J.J., Stern W.F., Stouffer R.J., Wilson J.R., Winton M., Wittenberg A.T., Wyman B.L., Zeng F, The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL Global Coupled Model CM3. J. Clim. 24, 3484–3519 (2011).

Eskes H., Arola A., Benedictow A., Blechschmidt A.-M., Botek E., Boucher O., Bouarar I., Chabrillat S. , Cuevas E., Engelen R., Flentje H., Gaudel A., Griesfeller J., Huijnen V., Jones L., Kapsomenakis J., Katragkou E., Kinne S., Langerock B., Razinger M., Richter A., Schultz M., Schulz M., Sudarchikova N., Thouret V., Vrekoussis M., Wagner A., and Zerefos C., Validation of reactive gases and aerosols in the MACC global analysis and forecast system, Geophys. Model Dev., Volume: 8, Issue: 11, Pages: 3523-3543, DOI: 10.5194/gmd-8-3523-2015, 2015.

Gilman J., Gouw J., Kuster W., and Lerner B., Source Signature of Volatile Organic Compounds from Oil and Natural Gas Operations in Northeastern Colorado. Environ. Sci. Technol., 47 , 1297–1305 (2013).

Komhyr, W.D., Electrochemical concentration cells for gas analysis, Ann. Geophys. 25, 203-210, (1969).

Lin, M.Y., Cooper O.R., Fiore A.M., Horowitz L.W., Langford A.O., Levy H.II, Johnson B.J., Naik V., Oltmans S.J., and Senff C., Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions, Journal of Geophysical Research, 117, D00V22, (2012A), doi:10.1029/2012JD018151

Lin, M.Y., Cooper O.R., Fiore A.M., Holloway J., Horowitz L.W., Langford A.O., Levy H.II, Johnson B.J., Middlebrook A.M., Naik V., Oltmans S.J., Pollack I.B., Ryerson T.B., Warner J.X., Wiedinmyer C., Wilson J., and Wyman B., Transport of Asian ozone pollution into surface air over the western United States in spring, Journal of Geophysical Research,117, D00V07, (2012B), doi:10.1029/2011JD016961

Lin, M.Y., Horowitz L.W., Oltmans S.J., Fiore A.M., and Fan S., Tropospheric ozone trends at Manna Loa Observatory tied to decadal climate variability, Nature Geoscience, 7, 136-143, (2014), doi:10.1038/NGEO2066.

Lin, M.Y., Fiore A.M., Horowitz L.W., Langford A.O., Oltmans S.J., Tarasick D.W., and Rieder H., Climate variability modulates western US ozone air quality in spring via deep stratospheric intrusions. Nature Comm. 6, 7105, (2015a).

Lin, M.Y., Horowitz L.W., Cooper O.R., Conley S., Tarasick D.W., Iraci L.T., Johnson B., Leblanc T., Petropavlovskikh I., and Yates E.L., Revisiting the evidence of increasing springtime ozone mixing ratios in the free troposphere over western North America, Geophysical Research Letter, 42, doi:10.1002/2015GL065311(2015b)

Logan, J.A.: An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models, and development of a gridded climatology for tropospheric ozone, J. Geophys. Res. 104, 16115–16149, (1999)

Liu, G., Fioletov V.E., Jin J.J., Liu J., Liu X., Moeini O., Osman M., Sioris C.E., and Tarasick D.W., A global tropospheric ozone climatology from trajectory-mapped ozone sounding, Atmos. Chem. Phys., 13, 10659-10675, 2013, doi:10.5194/acp-13-10659-2013.

Marenco, A., Cook T., Helten M., Hume C., Karcher F., Kley D., Law K., Nédélec P., Poschmann G., Pyle J., Simon P., Smit H., Thouret V., and Von Wrede R., Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, An overview, J. Geophys. Res., 103, 25,631-25,642, (1998)

Manney, G., Hegglin M., Daffer W., Schwartz M., Santee M., and Pawson S., Climatology of Upper Tropospheric–Lower Stratospheric (UTLS) Jets and Tropopauses in MERRA. *J. Climate,* 27, 3248–3271, (2014) doi: 10.1175/JCLI-D-13-00243.1.

Pfister G.G., Barth M.C., Bruyere C.L., Done J., Fast J., Holland G., Lamarque J.-F., Walters S., and Wong J., Projections of future summertime ozone over the U.S., Journal of Geophysical Research-Atmospheres, 119, 5559-5582, (2014) doi:10.1002/2013JD020932.

Thouret, V., Grouhel C., Logan J.A., Marenco A., and Nédélec P., Comparisons of ozone measurements from the MOZAIC airborne program and the ozone sounding network at eight locations, J. Geophys. Res., 103, 25,695-25,720, (1998A).

Thouret, V., Grouhel C., Marenco A., and Nédélec P., Ozone climatologies at 9-12 km altitude as seen by the MOZAIC airborne program between September 1994 and August 1996, J. Geophys. Res., 103, 25,653-25,679, (1998B).

Tilmes S., Conley A., Emmons L. K., Johnson B., Lamarque J.-F., Oltmans S.J., Saunois M., Schultz M.G., Tarasick D.W., Thompson A.M., and Thouret V., Technical Note: Ozonesonde climatology between 1995 and 2011: description, evaluation and applications. Atmos. Chem. Phys. 12, 7475-7497, (2012).

Zbinden R.M., Cammas J.-P., Carminati F., Nédélec P., Ricaud P., and Thouret V., Climatology of pure Tropospheric profiles and column contents of ozone and carbon monoxide using MOZAIC in the mid-northern latitudes (24° N to 50° N) from 1994 to 2009, Atmos. Chem. Phys., 13, 12363-12388, (2013).