Evaluation of AQ models: what we miss with limited information

Gabriele Pfister (& FRAPPÉ and DISCOVER-AQ Science Teams)

National Center for Atmospheric Research, Atmospheric Chemistry Observations and Modeling Laboratory, Boulder CO, U.S.A.

Two major field campaigns - the NSF/NCAR and State of Colorado Front Range Air Pollution and Photochemistry Éxperiment (FRAPPÉ) and the NASA DISCOVER-AQ - took place jointly in summer 2014 to study the drivers of summertime ozone pollution in the Northern Colorado Front Range (NFR). A comprehensive suite of chemical and meteorological measurements was collected from four research aircraft, six heavily instrumented ground sites with in-situ and remote sensing instruments, additional surface ozone monitoring sites, six mobile labs as well as tethered balloons and ozone sondes to provide a 3D picture of the chemical and meteorological characteristics of the area. This contrast the about dozen of operational surface ozone monitoring sites in the NFR and the even fewer surface sites with CO or NO_x measurements and the infrequent canister samples at two locations that are typically available for evaluating air quality models. Using WRF-Chem, we demonstrate how the additional information from the field campaign might change the conclusions drawn about model performance compared to findings based on evaluation with operationally available data alone. We will not only demonstrate the importance of available information above the surface but also the additional benefit from information on solar radiation and boundary layer heights.

Motivation

- Accurate Air Quality Predictions are needed for more than next day's forecast, but data assimilation and other methods can only get so far.
- Need models that accurately represent relevant atmospheric processes

… and gain confidence in models and predictions through careful evaluations

FRAPPÉ and DISCOVER-AQ provide a compressive dataset for detailed evaluation of chemistry and meteorology.

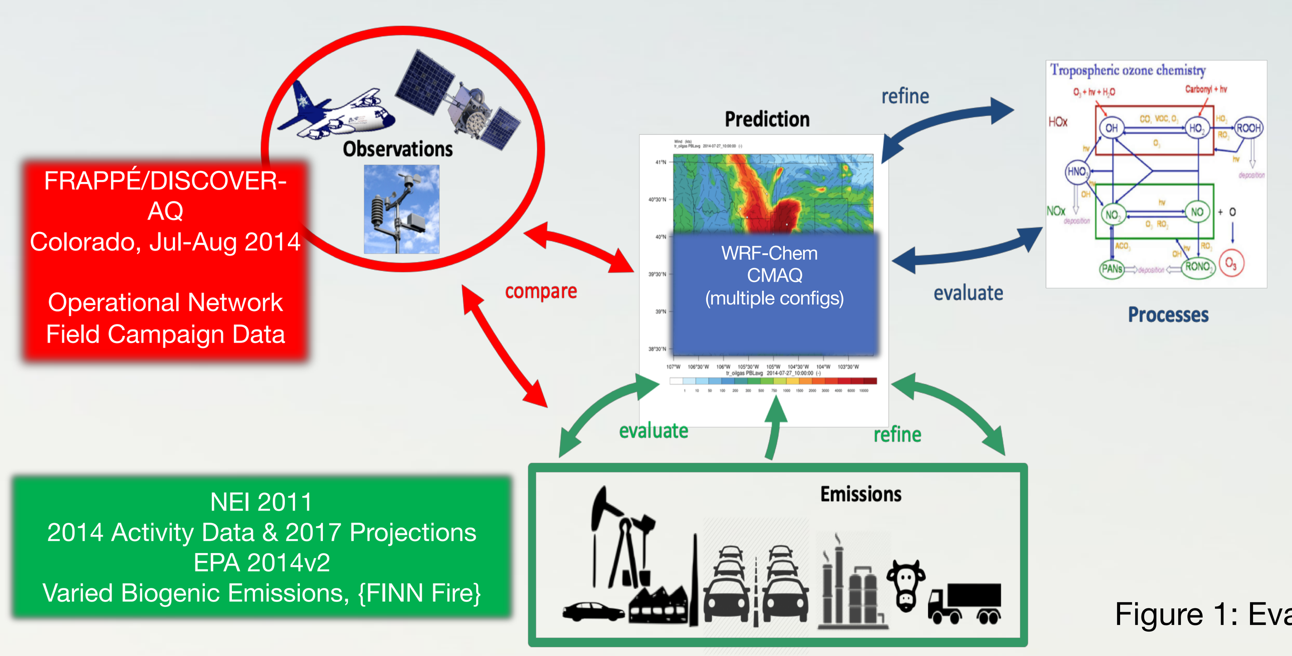

Figure 1: Evaluation framework

Example – Surface Ozone at one selected Surface Station

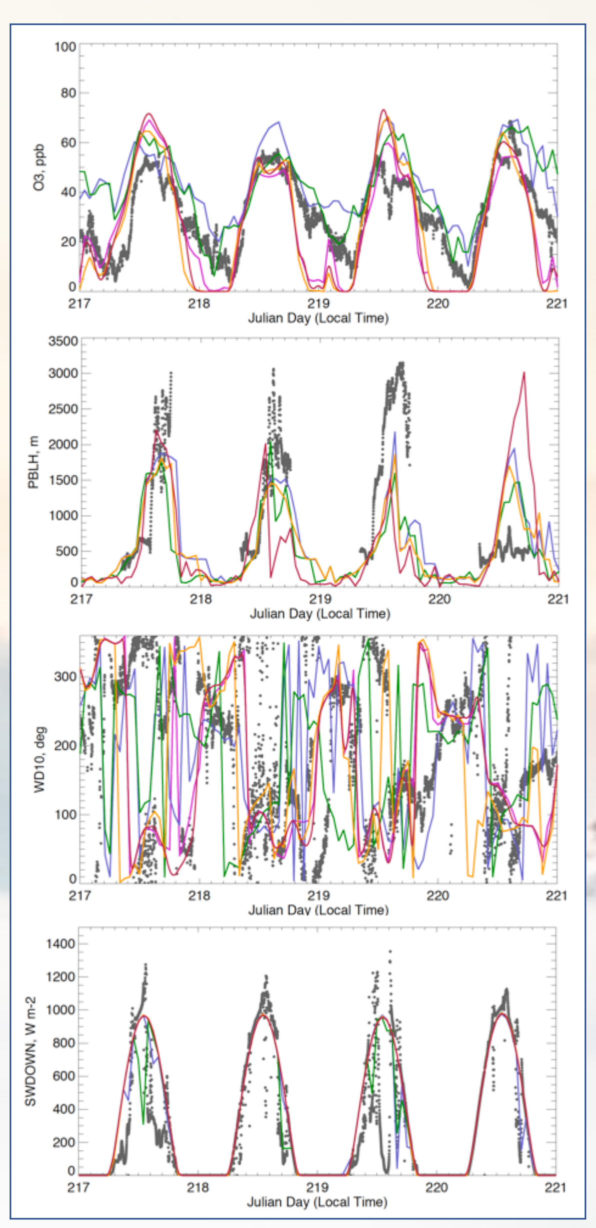

Figure 1: Evaluation of (from top to bottom) surface ozone, PBLH (derived from ceilometer data), wind direction and down-welling shortwave radiation for Platteville for July 5-8, 2014.

Large variability in performance for surface ozone amongst different model runs, but all tend to be too high during the day.
Reasons? Typically we blame: (1) Emissions, and (2) Model resolution

Observations of **boundary layer height** (PBLH) are essential since the PBL height and evolution determine the mixing of trace gases and aerosols and also the chemistry that takes place. PBLH can vary significantly among different model configurations.

Accurate representation of **winds** is crucial since transport determines not only how pollutants move around but also which and how different emission sources mix. Similar to PBLH, simulated winds can vary widely among different model configurations.

Solar radiation & clouds drive photochemistry but are often overlooked in assessing model performance. During the campaign, all models underestimate clouds (overestimate downwelling solar radiation) and overpredict ozone.

Modeled peak ozone too high independent of emission scenario – multiple factors (meteorological and chemical) can explain model disagreement.

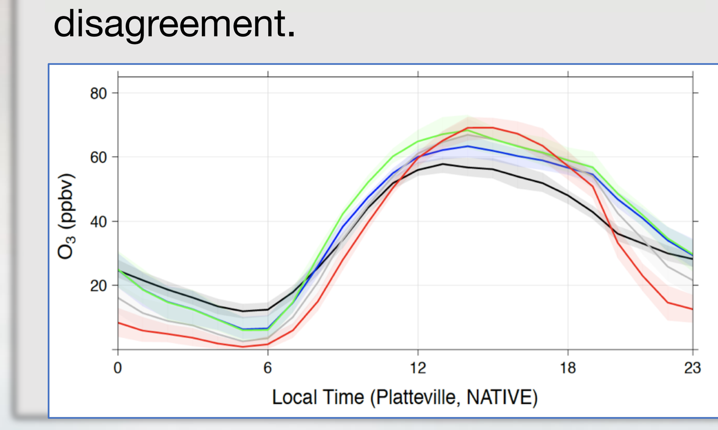

Figure 3: Average measured (black) and modeled (colors) diurnal surface ozone at Platteville for campaign period for different emission scenarios (results for WRF-CMAQ).

It's not always just emissions or model resolution!

Spatial Variability and Representativeness

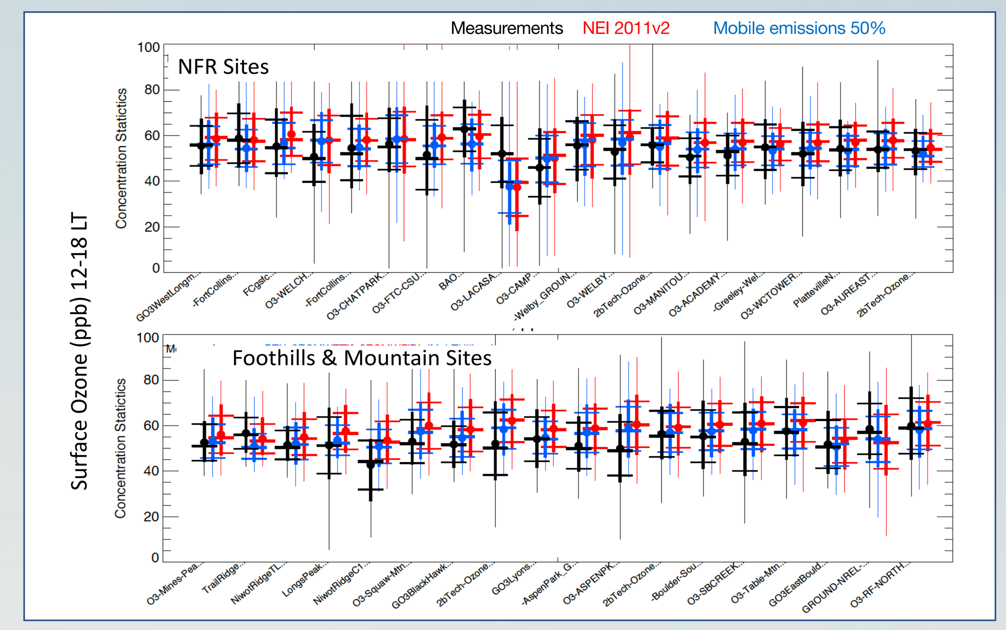

Figure 4: Statistics over the campaign period for measured and modeled surface daytime ozone for sites in the NFR (top) and the nearby mountains (bottom). Results are shown for WRF-Chem with original NEI 2011v2 emissions (red) and with a 50% reduction of mobile emissions (blue).

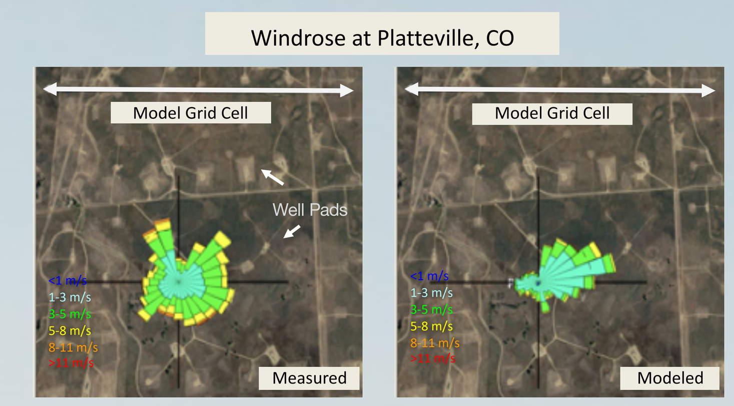

Figure 5: Modeled and measured windrose at Platteville overlaid over a Google map of the area. The 4 km model grid is indicated.

Localized features (topographic depression, nearby intense point sources, …) limits the representativeness of a surface measurement at a single site – ***not a "fair" model intercomparison.***

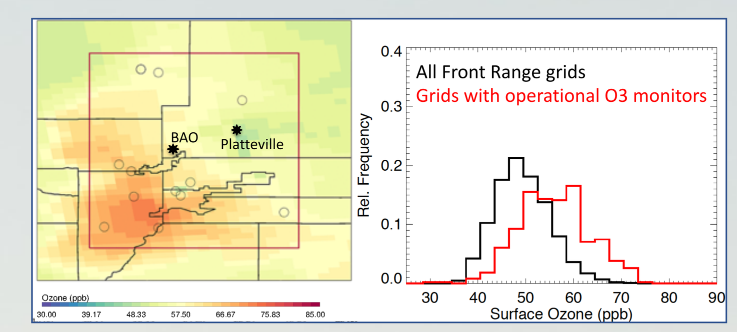

Figure 6 Left: Average daytime surface ozone; circles show location of operational monitoring sites. Right: Frequency distribution of model ozone at sites and for the region indicated by the red area in the map.

Varying performance across sites

Representativeness of measurements?

Averaging over a number of observations does not necessarily eliminate representativeness errors. Care need to be taken of the spectrum of conditions that can be evaluated with available observations. Often, measurements only provide limited insight into model performance –> **confidence is gained on some model aspects only!**

Vertical Information

The surface is of main interest for air quality, but it by no means is disconnected from the rest of the atmosphere!

Entrainment of free tropospheric ozone can significantly impact surface concentrations indicating the importance of correctly simulating atmospheric composition across all altitudes as well as the dynamic exchange processes.

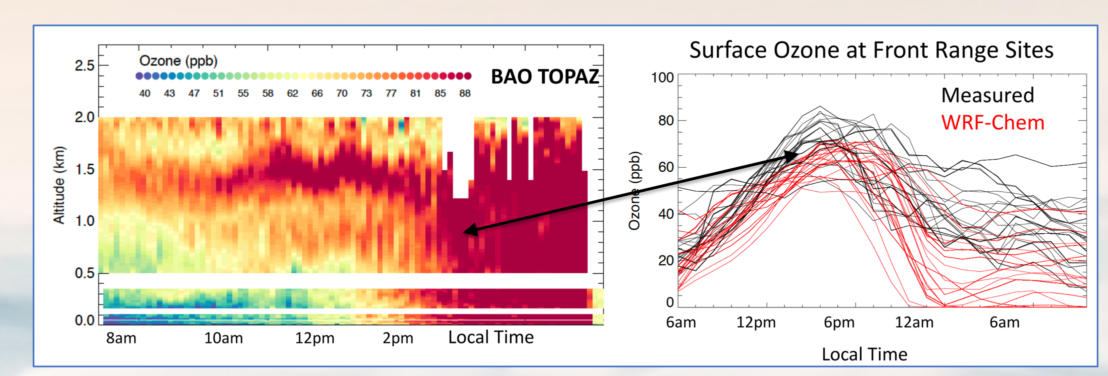

Figure 7: Left: Ozone Lidar data at BAO for 29 July 2014 (Data by C. Senff, NOAA). Right: Measured and modeled surface ozone at NFR sites.

Just as surface ozone, free tropospheric ozone varies in time. Evaluating the model in terms of averages in time (and space) not necessarily gives the right picture. Model evaluation across all altitudes is needed.
Free tropospheric composition in regional models is strongly influenced by lateral boundary conditions.
-> **Representative and Evaluated Boundary Conditions**

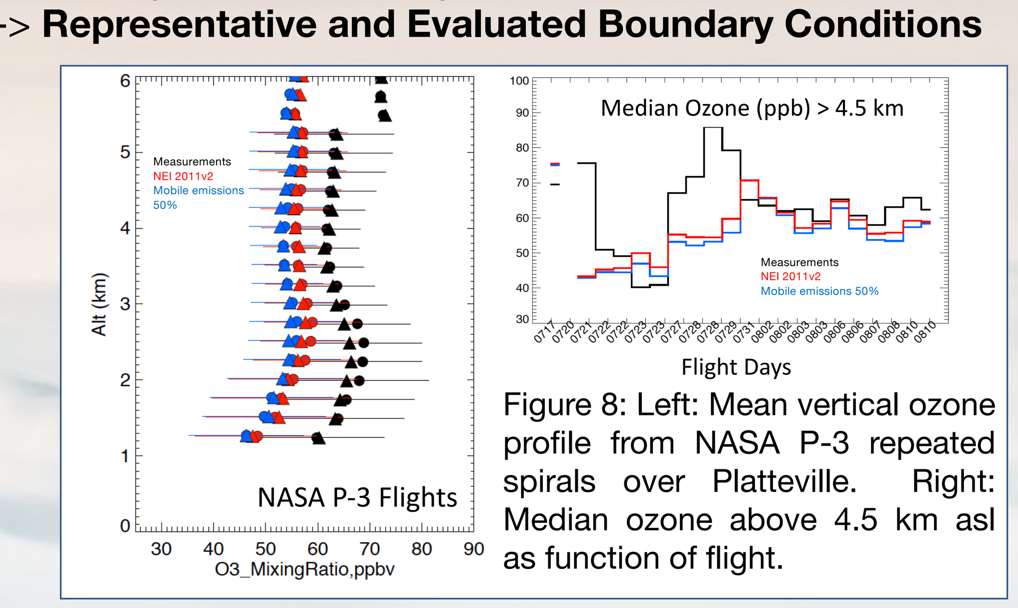

Figure 8: Left: Mean vertical ozone profile from NASA P-3 repeated spirals over Platteville. Right: Median ozone above 4.5 km asl as function of flight.

Conclusions

- Data assimilation and post-processing methods are valuable but get one only so far.
- Air quality predictions rely on both correctly simulating meteorology and chemistry. Evaluation of underlying synoptic-scale meteorology together with chemical processes needed for building confidence.
- Essential to evaluate models and inputs for a range of parameters and conditions – and for a range of parameters in addition to parameters of interest.
- Measurements need to provide actual and fair information for model performance (accuracy, representativeness, coverage, ….).
- Dense, high frequency, long-term, and reliable measurements necessary for evaluating model skill in representing frequency, intensity and duration of AQ episodes – comprehensive observations (even if snapshots) are essential in gaining confidence in a model.

-> Bringing together operational predictions and targeted field studies

Acknowledgements
We thank the FRAPPÉ and DISCOVER-AQ Science teams for observational data which are available from the NASA DISCOVER-AQ archive. We acknowledge Stu McKeen for help with NEI 2011 emissions and U.S. EPA for NEI 2014. Funding for FRAPPÉ has been provided by NSF and the State of Colorado. NASA funded DISCOVER-AQ. NCAR is sponsored by the National Science Foundation.

For more information about the model simulations:
FRAPPÉ final report (downloadable from https://www2.acom.ucar.edu/frappe)
Pfister et al., Chemical Characteristics and Ozone Production in the Northern Colorado Front Range, *J. Geophys. Res.* , under revision.

Contact: pfister@ucar.edu

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.