Atmos. Chem. Phys. Discuss., 11, 19291–19355, 2011
www.atmos-chem-phys-discuss.net/11/19291/2011/
doi:10.5194/acpd-11-19291-2011

Atmospheric Chemistry and Physics Discussions

Observing requirements for geostationary satellites to enable ozone air quality prediction

P. D. Hamer[1], **K. W. Bowman**[1], **and D. K. Henze**[2]

[1]Jet Propulsion Laboratory, Pasadena, CA, USA
[2]Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA

Received: 2 April 2011 – Accepted: 4 June 2011 – Published: 6 July 2011

Correspondence to: P. Hamer (paul.d.hamer@jpl.nasa.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We conduct a variety of analyses to support mission planning for geostationary satellite measurements of atmospheric composition. We carry out a simplified observing system simulation experiment (OSSE) using a photochemical box model and its adjoint integrated with a Lagrangian 4-D-variational data assimilation system. Using this framework in conjunction with pseudo observational constraints we estimate surface emissions and assess the improvement in ozone air quality forecasting and prediction. We use an analytical model as our principle method of conducting uncertainty analyses, which is the primary focus of this work. We investigate the impacts of changing the observed species (e.g., ozone, carbon monoxide (CO), nitrogen dioxide (NO_2), and formaldehyde (HCHO)), observation frequency and quality upon the ability to predict the magnitude of summertime peak ozone events, characterize the uncertainties of those predictions, and the performance of the assimilation system. We use three observed species scenarios: CO and NO_2; ozone, CO, and NO_2; and HCHO, CO and NO_2. These scenarios are designed to test the effects of adding observations of either ozone or HCHO to an existing CO and NO_2 observing system. The studies were conducted using the photochemical model setup to simulate a range of summertime polluted environments spanning NO_x limited to volatile organic compound (VOC) limited conditions. As the photochemical regime changes the relative importance of trace gas observations to constrain emission estimates and subsequent ozone forecasts varies. For example, adding ozone observations to an NO_2 and CO observing system is found to decrease ozone prediction error under NO_x and VOC limited regimes, and complimenting the NO_2 and CO system with HCHO observations would improve ozone prediction in the transitional regime and under VOC limited conditions.

1 Introduction

Ozone is a hazard to human health, plants and animals and a greenhouse gas (Mustafa, 1990; Pryor, 1992; Murphy et al., 1999; Fumagalli et al., 2001; Nali et al., 2002; IPCC, 2007; Van Dingenen et al., 2009). The spatial distribution and temporal variability of air pollutants play an important role in controlling ground-level ozone. Knowledge of the processes that control the variability of ozone precursors is vital for understanding and predicting ozone air quality. Prediction of ozone air quality on local and regional scales is key for providing prior warning of impending ozone exceedances (Dabberdt et al., 2004, 2006). Currently, a wide variety of techniques are used to predict ozone concentrations ranging from statistically based models (Gardner and Dorling, 2000), neural networks (Yi and Prybutok, 1996), to prognostic models of atmospheric processes that include data assimilation (Grell et al., 2005; Otte et al., 2005; Zhang et al., 2008; Kang et al., 2010). For prognostic models, uncertainties result from meteorology, the limitations of the photochemical mechanisms, wet and dry deposition, uncertainties in the emissions of ozone precursors, and, for data assimilation, observation uncertainty (Dabberdt et al., 2004, 2006). Current predictive statistical and data assimilation forecasting techniques rely primarily on surface observing networks.

The US national surface air quality observing network typically observes a wide range of chemical species, but has sparse and inhomogeneous spatial coverage closely related to population density. For instance, surface monitoring sites within California (http://www.arb.ca.gov/adam/) have instruments that can measure in-situ ozone, CO, NO_2, nitrogen oxide, particulate matter 2.5 µm and 10 µm, sulphur dioxide, methane, total hydrocarbons, and hydrogen sulphide on hourly timescales. Most surface monitoring networks lack chemical species vertical profile information in the absence of towers, LIDAR, or regular sonde launches. Vertical profile information can improve estimates of ozone transport into the boundary layer (Parrington et al., 2009; Parrish et al., 2010). Due to the spatial limitations of the surface air quality monitoring network, space-borne remote sensing observations, which typically have greater spa-

tial sampling, are able to support air quality research. As a consequence investigations of air quality using remote sensing data of atmospheric composition have been a major research focus. Studies range from quantification of air quality indices (Martin, 2008; Duncan et al., 2010), to inverse emission estimates of ozone precursors (Jones et al., 2009; Bowman et al., 2009; Kurokawa et al., 2009; Konovalov et al., 2006; Millet et al., 2008; Kopacz et al., 2010; Arellano et al., 2006), direct observation of ozone and its impacts on health and agriculture (Dufour et al., 2010; Fishman et al., 2010), and use of observations of air pollutants within air quality models via data assimilation (Sandu et al., 2003a; Chai et al., 2007; Pierce et al., 2007; Zhang et al., 2008; Parrington et al., 2009). There are factors to be considered when comparing the characteristics of remote sensing observations to existing surface observing network. Surface monitoring involves a point measurement whereas the spatial footprint and vertical sampling of space-based remote sensing is generally coaser. For example, spatial footprints range from 5×8 km for the Tropospheric Emission Spectrometer (TES) (Beer, 2006) up to the relatively large 40×40 km for the Global Ozone Monitoring Experiment (GOME) (Burrows et al., 1999), and the vertical resolution ranges from a total ozone column to a vertical profile depending on the wavelengths measured and whether the instrument is passive or active. Satellite remote sensing observations of air quality relevant trace gases within the lower troposphere are limited to fewer chemical species than surface monitoring such as ozone, CO, NO_2, and HCHO. In addition, satellite observations of trace gases typically have degraded precision relative to in-situ observations and the error characterization is more complicated.

In the assimilation studies of Pierce et al. (2007), Zhang et al. (2008), and Parrington et al. (2009), and in the case of the examples mentioned earlier (Konovalov et al., 2006; Martin, 2008; Millet et al., 2008; Jones et al., 2009; Bowman et al., 2009; Kurokawa et al., 2009; Dufour et al., 2010; Duncan et al., 2010; Fishman et al., 2010), the satellites used are in low Earth orbit (LEO) and are limited to observing single locations once during the day in either the early afternoon (Aura and Aqua) or morning (Infrared Atmospheric Sounding Interferometer, IASI (Clerbaux et al., 2009)) and at most two obser-

vations per day, but do achieve global coverage. As such, the information provided aids understanding of continental and global processes affecting air quality. Geostationary (GEO) remote sensing offers denser temporal and spatial sampling and therefore have the potential to better serve air quality research and forecasting observational needs on local and regional scales (Edwards et al., 2009b; Campbell and Fishman, 2010). Much of the rational for moving to a monitoring system with denser temporal and spatial sampling is that there are a variety of physical processes controlling air quality that are occurring on far shorter timescales and smaller spatial scales compared to the LEO spatio-temporal sampling. For example, the National Research Council decadal survey (http://science.nasa.gov/earth-science/decadal-surveys/) outlines a mission titled Geostationary Coastal and Air Pollution Events (GEO-CAPE), which will be a geostationary satellite over North America designed to observe atmospheric composition with scientific aims to aid and improve air quality forecasting (http://geo-cape.larc.nasa.gov/) along with coastal ocean biophysical indices. For the purposes of our study we focus solely on the atmospheric composition component of this mission. The European Space Agency (ESA) and the Korean National Institute of Environmental Research are individually planning geostationary atmospheric composition monitoring instruments named Sentinel 4 (http://www.esa.int/esaLP/SEM3ZT4KXMF_LPgmes_0.html) and the Geostationary Environment Spectrometer (GEMS) (Lee et al., 2009), respectively. Aside from denser observing rates, the decadal survey and Lee et al. (2009) indicate that GEO-CAPE and GEMS will observe the following trace gases: ozone, CO, NO_2, HCHO, and sulphur dioxide (SO_2). A range of instruments and spectral wavelengths have been used for retrieving these four trace gases: tropospheric ozone has been retrieved using UV and visible wavelengths in the case of TOMS (Total Ozone Mapping Spectrometer) (Gleason et al., 1993), SCIAMACHY (SCanning Image Absorption spectroMeter for Atmospheric Cartography) (Bovensmann et al., 1999), OMI (Ozone Monitoring Instrument) (Levelt et al., 2006), and GOME (Burrows et al., 1999), and using the thermal IR (TIR) by AIRS (Atmosperic InfraRed Sounder) (Aumann et al., 2003), TES (Beer, 2006), MIPAS (Michelson Interferometer for Passive Atmospheric Sound-

ing) (Fischer et al., 2008), and IASI (Clerbaux et al., 2009); CO has been retrieved using the near IR (NIR) and IR by MOPITT (Measurements Of Pollution In The aTmosphere) (Drummond and Mand, 1996), and by AIRS (Aumann et al., 2003), TES (Beer, 2006), and IASI (Clerbaux et al., 2009); NO_2 has been retrieved using visible wavelengths by OMI (Levelt et al., 2006), SCIAMACHY, and GOME (Burrows et al., 1999); and SO_2 has been retrieved by GOME (Burrows et al., 1999), OMI (Levelt et al., 2006) and SCIAMACHY (Bovensmann et al., 1999) using UV and visible wavelengths and by MODIS (MODerate resolution Imaging Spectroradiometer) (Watson et al., 2004) and TES (Clerbaux et al., 2008) using thermal infrared wavelengths. In the case of GEO-CAPE the decadal survey indicates that sensitivity to the boundary layer trace gas concentrations should be achieved, which is a higher level of tropospheric vertical resolution than has currently been achieved. Ozone retrievals with high vertical information and boundary layer sensitivity have been proposed by combining different radiance wavelength bands in the UV, visible, and IR (Landgraf and Hasekamp, 2007; Worden et al., 2007). Similar efforts to improve vertical resolution in retrievals of CO have been successfully demonstrated by combining wavelength bands in the TIR and NIR (Worden et al., 2010). Future geostationary satellites focusing on supporting air quality observational forecasting needs have a variety of choices available regarding instrument design and spectral wavelength bands based on the heritage of instruments deployed in LEO orbit, e.g., OMI, TES, SCIAMACHY, AIRS, and MOPITT. Given the available mission design choices we seek to provide support for such missions by indicating how choices regarding observed species and observation precisions can affect air quality forecast prediction error. We have therefore designed a simplified OSSE to explore observing requirements for future geostationary atmospheric compostion instruments. OSSEs have been used to evaluate and design observing systems for use by the numerical weather prediction community since the 1950s (Arnold and Dey, 1986), and, more recently, observing systems for trace gases (Rayner et al., 1996; Jones et al., 2003; Edwards et al., 2009a). The main component of the simplified OSSE is a prototype air quality forecasting framework consisting of an idealized prog-

nostic model coupled with a data assimilation system.

Combining satellite observations of ozone and its precursors with data assimilation models could potentially reduce the uncertainties in prognostic models and improve their predictive skill of ozone air quality. Data assimilation has been used successfully in the context of ozone air quality in conjunction with chemical transport models and satellite remote sensing data of trace gas observations (Pierce et al., 2007; Zhang et al., 2008; Parrington et al., 2009; Sofiev et al., 2009). Our framework consists of a photochemical box model with idealized meteorology, its adjoint, and 4-D-variational data assimilation setup to use pseudo observations to constrain ozone precursor emission uncertainties (for NO_x, CO and VOC emissions) within scenarios where the only actual source of prediction error results from the emission uncertainties. Our OSSE is simplified relative to other OSSEs (Rayner et al., 1996; Jones et al., 2003; Edwards et al., 2009a) due to the relatively small model domain and lack of detailed meteorological simulation, since we use a box model rather than a regional or global model, and because our observation simulation is relatively crude as we don't simulate the full retrieval of trace gas species. We conduct an uncertainty analysis using a linear estimation technique to support the work conducted with the 4-D-variational data assimilation. In addition, we carry out a variety of supporting sensitivity analyses to test the robustness of our methodology.

2 Methodology

2.1 Overview

The photochemical box model is run over 3 days to represent a worsening period of ozone air quality during a stagnation event. Meteorological stagnation events under hot, sunlit conditions over urban areas typically lead to poor ozone air quality (Jacob et al., 1993; Valente et al., 1998). Under stagnation event conditions we can focus on the role played by chemistry in allowing different selections of observations of ozone

and its precursors to constrain emission uncertainties and reduce the air quality prediction error. Another advantage of selecting a photochemical box model is that we are able to generate a Jacobian describing the model response to emission parameter perturbations, which can be used within an analytical modeling framework to conduct uncertainty analysis. It would be very difficult to produce a Jacobian within a regional or global chemical transport models in a timely fashion given the size of the model state space. Further, the resulting Jacobian is sufficiently small to be used within our analytical framework and uncertainty analysis. Within the model framework days 1–2 represent the period over which observations are made and the final day represents the prediction and monitoring period.

We aim to demonstrate that the 4-D-variational data assimilation framework is capable of solving the non-linear optimal estimation problem in the context of constraining ozone precursor emission uncertainties (for NO_x, CO and VOC emissions) with observations. In addition, we aim to determine the variability of the prediction error of the 4-D-variational data assimilation prediction framework across a range of photochemical conditions for fixed observing precision, varying observed species scenario, and for a single realization of the observing noise.

4-D-variational data assimilation and the emission inversion problem are reflective of the state of the art in prognostic air quality forecast modeling development (e.g., in the case of the Community Multi Scale Air Quality Modeling System, Hakami et al. (2007) and the Sulfur Transport Eulerian Model, Zhang et al. (2008)) and thus our model framework is relevant to and is reflective of the future direction of air quality forecasting. In order to establish the utility of more complex air quality forecasting systems that might use 4-D-variational data assimilation, and the relevance of our presented results, our prototype forecasting system is demonstrated theoretically. Since the emission inversion problem that we explore only becomes more complex as the model state space increases and additional sources of uncertainty are introduced a failure to show sufficiently reduced prediction error in this setting would indicate that more complex systems are unlikely to fare better. Sufficient prediction model error within this framework is

therefore a necessary but not sufficient condition for more complex 4-D-variational data assimilation forecasting systems using space bourne observations to be successful.

Uncertainty analysis using 4-D-variational data assimilation is difficult. Therefore we use an analytic model (derived from the photochemical box model) that is simplified relative to the full assimilation framework that is capable of conducting a full uncertainty analysis characterizing the performance of the 4-D-variational data assimilation system. This technique is a linear estimation technique based upon Rodgers (2000). We calculate the following diagnostics across a wide range of observational uncertainty, varying choices regarding targetted observing species, and different observing frequencies: posteriori ozone prediction error covariance, posteriori emission parameter error covariance, the averaging kernel, and the degrees of freedom of signal. Given the breadth of mission profiles explored the results contained herein will therefore be relevant to GEO and LEO satellite instrument planning for instruments designed to observe air quality indices, and to emission inversion studies using an adjoint model coupled with data assimilation.

The 4-D-variational data assimilation and uncertainty analysis using the linear estimation are therefore complementary methods. In both cases we use a forward photochemical model setup using a range of NO_x emission scenarios for polluted regions and meteorologically idealized conditions, which is then coupled with each technique to guide the observation requirements (observing frequency, noise, and observed species choice) for future GEO atmospheric composition missions, e.g., GEO-CAPE, Sentinel-4, and GEMS. We conduct a series of supporting analyses: since we assume a fixed diurnal variability of ozone precursor emissions, we study the impact on our conclusions of varying the assumptions regarding the diurnal variability of emissions; we carry out a sensitivity analysis to assess the relative performance of ozone and HCHO observations at reducing prediction error and to test the assumptions regarding HCHO observing noise; and because when conducting the VOC emission inversion we solve ethene emission uncertainties (rather than a more adverse range of VOCs) we test that assumption in a sensitivity analysis by assuming VOC emission errors for ethane

insead of ethene.

2.2 Photochemical box model

A 1-Dimensional photochemical box model was built using the Kinetic Pre-Processor (KPP) (Damian et al., 2002; Daescu et al., 2003; Sandu et al., 2003b). The Rosenbrock solver is used to integrate the KPP generated ordinary differential equations required to calculate trace gas concentrations (Eller et al., 2009). The photochemical mechanism consists of 171 gas phase species and 524 chemical reactions simulating the degredation of hydrocarbons from C_1–C_5 including isoprene and is based upon the Master Chemical Mechanism v3.1 (Jenkin et al., 1997) (http://mcm.leeds.ac.uk/MCM/). In addition, the model includes dry deposition for all relevant chemical species, contains a 2-parameter photolysis scheme, and simulates the emission of ozone precursors including NO_x, CO, and volatile organic compounds (VOCs). The diurnal emission variability of anthropogenic compounds is prescribed according to the National Atmospheric Emissions Inventory (NAEI) (http://www.naei.org.uk/emissions/) for an urbanised area (see Fig. 1), and the isoprene emission variability is parameterized to correlate to solar zenith angle offset by 2 h to consider both temperature and photon flux effects. The isoprene emissions have an average daily emission of 1.7×10^{10} molecules $m^2\,s^{-1}$ and an afternoon peak of 4.6×10^{10} molecules $m^2\,s^{-1}$, which yields modeled isoprene mixing ratios less than 10 pptv. The VOC speciation is defined according to NAEI and the total peak emission of carbon via VOCs (excluding isoprene) is 2.3×10^{12} carbon atoms $m^{-2}\,s^{-1}$ and the average emission is 1.2×10^{12} carbon atoms $m^{-2}\,s^{-1}$. Boundary layer dynamics are described with a prescribed variability in mixing height ranging from 500–1500 m and mixing between the boundary layer and free troposphere equivalent to a constant 10 % mass exchange per hour. Background free tropospheric concentrations of long lived species are assumed to remain constant, and are defined in Table 1. The box model is situated at 33° North and is run from 30 June to 2 July and has a humidity of 1.62 %, equivalent to the Southern Californian coastal region.

The model is run under a range of photochemical conditions. This is achieved by varying the NO emissions across 9 different scenarios that span the full range of modeled ozone response with respect to changing NO_x concentration. For the purposes of the emission inversion we define our ozone precursor emissions in a simplified form (excluding emitted species not considered in the inversion) as

$$\phi_i(t) = \boldsymbol{x}_i E_i(t),\ i = \mathrm{NO, CO, VOC} \quad (1)$$

where $\boldsymbol{x}_i$ are the time independent emission scaling factors for the emitted species, i, and $E_i(t)$ are the emissions with a prescribed and repeating diurnal cycle for each emitted species. The emission inversion solves for, $\boldsymbol{x}_i$, the time independent emission scaling factors, which can be represented as a vector, $\boldsymbol{x}$, for the emitted species, i, as shown by

$$[\boldsymbol{x}]_i = \boldsymbol{x}_i\ ,\ i = \mathrm{NO, CO, VOC} \quad (2)$$

The variability of $E_{\mathrm{NO}}(t)$ is shown in Fig. 2 and this variability is represented by

$$E_i(t) = \epsilon_i k(t) \quad (3)$$

where $k(t)$ is the temporal variability emission factor for all of the emitted species and ϵ_i is the time independent emission for each species. Note then that all of the species emissions, $E_i(t)$, share the same temporal variability. The variability of $k(t)$ is shown in Fig. 1. In our model simulations ϵ_{NO} is 4.8×10^{10} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$, ϵ_{CO} is 2.6×10^{12} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$, and ϵ_{VOC} is 4.3×10^{10} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ where in the emission inversion calculations we represent VOC emissions via ethene emissions. We define a range of different $k(t)$ scenarios in order to probe the emission solution sensitivity to diurnal emission variability and this variability is shown in Fig. 1. The different scenarios that we use are described in Sect. 2.5.

In the emission inversion calculations we represent VOC emissions via ethene emissions. $\overline{k(t)}$ is 1.89, and thus while ϵ_i is an important model parameter it is hard to relate to the total daily emission burden. The average emissions are therefore a factor of 1.89 larger than ϵ_i. In the case of NO, $\overline{E(t)}_{\mathrm{NO}}$ is 9×10^{10} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$. The scalings used x_{NO} = [0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, and 2.5] lead to a range in $\overline{E(t)}_{\mathrm{NO}}$ between 4.5×10^{10} and 2.3×10^{11} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$, and to modeled peak NO_x concentrations ranging between 4.0 ppbv and 24.0 ppbv (peak concentrations from 1 to 11.3 ppbv for NO and 3 to 16.9 ppbv for NO_2). These NO emission scalings are chosen to represent a wide range of photochemical conditions and given the VOC burden in the model scalings 0.5, 0.75 and 1.0 represent NO_x limited conditions, 1.25, 1.5 and 1.75 represent the transitional conditions, and 2.0, 2.25, and 2.5 represent VOC limited conditions. For CO and the reactive VOC (ethene) $\overline{E(t)}_{\mathrm{CO}}$ is 5×10^{12} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ and $\overline{E(t)}_{\mathrm{VOC}}$ 8.2×10^{10} molecules $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$. Given the latitude, humidity, dominance of the VOC burden from anthropogenic VOCs, and range of modeled NO_x concentrations these model runs can be viewed as somewhat analogous to a range of environments spanning the wider urbanized Southern Californian region. The emissions of CO and VOCs lead to modeled peak concentrations of CO and HCHO ranging between 590 and 820 ppbv and 6.5 and 8.1 ppbv, respectively.

2.3 Forecasting framework and 4-D-variational data assimilation

Several NO_x emissions scenarios are simulated to cover a wide range of photochemical conditions (x_{NO} = 0.5–2.5). Each emission scenario is represented mathematically as a forward model, $\mathbf{F}(\boldsymbol{x}, t)$, which are the concentrations as a function of time evaluated at emissions $\boldsymbol{x}$. Three observing scenarios were studied: CO and NO_2 (CN scenario); O_3, CO, and NO_2 (OCN scenario); and HCHO, CO, and NO_2 (HCN scenario). The model true state is sampled at 3 hourly intervals in the standard scenarios (used as default unless specified) and at intervals between 1 and 18 h in scenarios characterizing the impact of observing frequency on prediction error. The sampled species concentrations

are combined with an additive noise model to generate the pseudo observations, $\boldsymbol{y}$, represented by

$$\boldsymbol{y} = \mathbf{F}(\boldsymbol{x},t) + n \tag{4}$$

where $\boldsymbol{n}$ is the noise

$$\boldsymbol{n} = \overline{\mathbf{F}(\boldsymbol{x})} \times \beta \times r.n. \tag{5}$$

and where $\overline{\mathbf{F}(\boldsymbol{x})}$ is the average species concentration (values shown in Table 2), β is the noise scaling factor, and $r.n.$ is a random number with a gaussian distribution, a standard deviation of 1, and a mean of zero. Figure 3 shows a schematic of how we generate the pseudo observations, $\boldsymbol{y}$, from the forward model, $\mathbf{F}(\boldsymbol{x})$, and how the pseudo observations of ozone relate directly to the true ozone variability. The modeled concentrations for all species and times resulting from $\mathbf{F}(\boldsymbol{x})$ can be represented as a vector, $\boldsymbol{q}$,

$$\boldsymbol{q} = \mathbf{F}(\boldsymbol{x},t) \tag{6}$$

or for specific species, $\boldsymbol{z}$, at time, t, as $q_z(\boldsymbol{x},t)$,

$$q_z(\boldsymbol{x},t) = [\mathbf{F}(\boldsymbol{x},t)]_z \tag{7}$$

The contents of the vector, z, vary depending on the observing scenario. Note that in the case of Fig. 3 the pseudo observations are generated from $q_{\mathrm{O_3}}(\boldsymbol{x},t)$. We define a priori emission scaling factors, $\boldsymbol{x}_a$, with specified errors (Table 3 provides a summary of the values of $\boldsymbol{x}$ used to generate both $\boldsymbol{x}$ and $\boldsymbol{x}_a$), which are combined with the model to yield the a priori model state, $\mathbf{F}(\boldsymbol{x}_a)$. Note that within our framework the a priori is also the initial guess.

The assimilation is started at the first iteration with the forward model using the initial guess and is thus described as $\mathbf{F}(\boldsymbol{x}_a)$ after one iteration. A cost function, which is a

scalar, $J(\boldsymbol{x})$, is then evaluated

$$J(\boldsymbol{x}) = \frac{1}{2}((\boldsymbol{y} - \mathbf{F}(\boldsymbol{x}))^T \mathbf{S}_n^{-1}(\boldsymbol{y} - \mathbf{F}(\boldsymbol{x})) + (\boldsymbol{x} - \boldsymbol{x}_a)^T \mathbf{S}_a^{-1}(\boldsymbol{x} - \boldsymbol{x}_a)) \tag{8}$$

where $\mathbf{S}_a$ is the a priori constraint matrix and $\mathbf{S}_n$ is the observation error covariance. The 4-D-variational data assimilation method seeks the solution for $\boldsymbol{x}$, $\hat{\boldsymbol{x}}$, that minimizes $J(\boldsymbol{x})$

$$\hat{\boldsymbol{x}} = \min_{\boldsymbol{x}} J(\boldsymbol{x}) \tag{9}$$

such that the gradient of the cost function with respect to $\boldsymbol{x}$ is zero if the solution $\hat{\boldsymbol{x}}$ is equal to the true state (though this is never achieved in practical circumstances)

$$\nabla_{\boldsymbol{x}} J = \mathbf{K}^T \mathbf{S}_n(\boldsymbol{y} - \mathbf{F}(\hat{\boldsymbol{x}})) - \mathbf{S}_a^{-1}(\hat{\boldsymbol{x}} - \boldsymbol{x}_a) = 0 \tag{10}$$

where $\mathbf{K}$ is the Jacobian matrix (see Eq. 15) describing the forward model response to perturbations to the emission parameters, and $\nabla_{\boldsymbol{x}^t} J$ is the adjoint sensitivity (Daescu et al., 2003; Sandu et al., 2003b), calculated by the Rosenbrock solver (Eller et al., 2009), which indicates the sensitivity of the cost function to the emission parameters. The cost function and its adjoint senitivities are passed to the quasi-Newton L-BFGS algorithm (Zhu et al., 1997). The L-BFGS algorithm iteratively determines the optimal state of $\boldsymbol{x}$, $\hat{\boldsymbol{x}}$, that minimizes the difference between the model and observations subject to the a priori constraints.

Using the estimated emissions, $\hat{\boldsymbol{x}}$, the forward model, $\mathbf{F}(\hat{\boldsymbol{x}})$, provides the air quality prediction of the ozone concentration, $q_{\mathrm{O_3}}(\boldsymbol{x},t)$, on the afternoon of the 3rd day of the simulation during the prediction and monitoring period. Figure 2 shows how the a priori emissions, $\boldsymbol{x}_a$, relate to the true emissions $\boldsymbol{x}$, and the posteriori emissions, $\hat{\boldsymbol{x}}$, after the 4-D-variational data assimilation, as well as the a priori, the true and the posteriori ozone levels (i.e., $q_{\mathrm{O_3}}(\boldsymbol{x}_a,t)$, $q_{\mathrm{O_3}}(\boldsymbol{x},t)$, and $q_{\mathrm{O_3}}(\hat{\boldsymbol{x}},t)$, respectively). Figure 2 therefore demonstrates the mechanism by which the forecasting framework improves

the forward model ozone predictions, i.e., by an optimization of the ozone precursor emissions. The left panel of Fig. 2 shows the a priori emission error for NO emissions and the right panel shows the posteriori NO emission error. The posteriori emission parameter error can be defined more generally as a vector $\tilde{\boldsymbol{x}}$.

$$\tilde{\boldsymbol{x}} = \hat{\boldsymbol{x}} - \boldsymbol{x} \tag{11}$$

Figure 4 provides an example representation of the ozone prediction, $q_{O_3}(\hat{\boldsymbol{x}},t)$, relative to the true state, $q_{O_3}(\boldsymbol{x},t)$, during the prediction and monitoring period on the third day. Figure 5 shows a full representation of the 4-D-variational assimilation results, the pseudo observations, the a priori and initial guess ozone, the ozone true state and the posteriori ozone prediction. In Fig. 5 E represents the posteriori ozone prediction error at time, t^{μ} (t^{μ} is 15:00 LT on day 3 during the prediction and monitoring period), defined by

$$E = q_{O_3}(\hat{\boldsymbol{x}},t^{\mu}) - q_{O_3}(\boldsymbol{x},t^{\mu}) \tag{12}$$

and in both Figs. 4 and 5 G represents the a priori ozone prediction error defined by

$$G = q_{O_3}(\boldsymbol{x}_a,t^{\mu}) - q_{O_3}(\boldsymbol{x},t^{\mu}) \tag{13}$$

Table 4 shows the values of E and G calculated using the 4-D-variational data assimilation. The air quality prediction error over the entire prediction and monitoring period for each of the species, z, can be defined as a vector, $\tilde{\boldsymbol{q}}$

$$\left[\tilde{\boldsymbol{q}}_z\right]_j = q_z(\hat{\boldsymbol{x}},t_j) - q_z(\boldsymbol{x},t_j)\,, j = 3,6.....21,24 \tag{14}$$

where j is the hour of day on the 3rd day during the prediction and monitoring period.

The Jacobian matrix is redundant within 4-D-variational data assimilation, but it can help characterize the uncertainties on $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{q}}$. Therefore it is advantageous to determine **K**. For large systems such as global and regional chemical transport models

it is impractical to calculate the Jacobian matrix. However, it is feasible to determine the Jacobian for a small system such as a photochemical box model with a limited state space. Within our framework we define each element of **K** as the forward model response, $\partial q_z(\boldsymbol{x},t)/\partial \boldsymbol{x}_i$, at time, t, and for observed species, z, to perturbations in emissions of species, i, in the case of the OCN scenario it is defined by

$$\mathbf{K} = \begin{bmatrix} \partial q_{O_3}(\boldsymbol{x},t_1)/\partial x_{NO} & \partial q_{O_3}(\boldsymbol{x},t_1)/\partial x_{CO} & \partial q_{O_3}(\boldsymbol{x},t_1)/\partial \boldsymbol{x}_{VOC} \\ \partial q_{O_3}(\boldsymbol{x},t_2)/\partial \boldsymbol{x}_{NO} & \partial q_{O_3}(\boldsymbol{x},t_2)/\partial \boldsymbol{x}_{CO} & \partial q_{O_3}(\boldsymbol{x},t_2)/\partial \boldsymbol{x}_{VOC} \\ \vdots & \vdots & \vdots \\ \partial q_{O_3}(\boldsymbol{x},t_{N_t})/dx_{NO} & \partial q_{O_3}(\boldsymbol{x},t_{N_t})/\partial x_{CO} & \partial q_{O_3}(\boldsymbol{x},t_{N_t})/\partial \boldsymbol{x}_{VOC} \\ \partial q_{CO}(\boldsymbol{x},t_1)/dx_{NO} & \partial q_{CO}(\boldsymbol{x},t_1)/\partial \boldsymbol{x}_{CO} & \partial q_{CO}(\boldsymbol{x},t_1)/\partial \boldsymbol{x}_{VOC} \\ \partial q_{CO}(\boldsymbol{x},t_2)/dx_{NO} & \partial q_{CO}(\boldsymbol{x},t_2)/\partial x_{CO} & \partial q_{CO}(\boldsymbol{x},t_2)/\partial x_{VOC} \\ \vdots & \vdots & \vdots \\ \partial q_{CO}(\boldsymbol{x},t_{N_t})/dx_{NO} & \partial q_{CO}(\boldsymbol{x},t_{N_t})/\partial x_{CO} & \partial q_{CO}(\boldsymbol{x},t_{N_t})/\partial x_{VOC} \\ \partial q_{NO_2}(\boldsymbol{x},t_1)/dx_{NO} & \partial q_{NO_2}(\boldsymbol{x},t_1)/\partial x_{CO} & \partial q_{NO_2}(\boldsymbol{x},t_1)/\partial x_{VOC} \\ \partial q_{NO_2}(\boldsymbol{x},t_2)/dx_{NO} & \partial q_{NO_2}(\boldsymbol{x},t_2)/\partial x_{CO} & \partial q_{NO_2}(\boldsymbol{x},t_2)/\partial x_{VOC} \\ \vdots & \vdots & \vdots \\ \partial q_{NO_2}(\boldsymbol{x},t_{N_t})/dx_{NO} & \partial q_{NO_2}(\boldsymbol{x},t_{N_t})/\partial x_{CO} & \partial q_{NO_2}(\boldsymbol{x},t_{N_t})/\partial x_{VOC} \end{bmatrix} = \frac{\partial \mathbf{F}(\boldsymbol{x},t)}{\partial \boldsymbol{x}} \tag{15}$$

where **K** has dimensions $N_i \times N$. N_i is the number of species in the emission factor state vector, $\boldsymbol{x}$ and is thus always three. We define N as the total number of observations for

all species

$$N = N_t \times N_y \tag{16}$$

where N_t is the number of points in time the model perturbations are sampled and N_y is the number of species whose perturbations are used in the Jacobian. In the case of Eq. (15) $z = O_3$, CO and NO_2 therefore $N_y = 3$. z includes HCHO in the HCN scenario.

Figure 6 plots columns of the Jacobian and it shows that ozone is more sensitive to changes in emissions during the afternoon, and that CO and NO_2 respond to changes in emissions during the rush hour periods.

2.4 Uncertainty analysis

2.4.1 Overview

We conduct an uncertainty analysis in order to characterize the performance of the assimilation system under a range of different observing scenarios and photochemical conditions. The uncertainty analysis has two separate foci: the estimation performance of the emissions and the posteriori ozone prediction error. Note that there is a direct synergy between these two analyses since uncertainties in the emissions estimate directly impact upon ozone prediction uncertainty. The diagnostics that we calculate in the analysis of the emissions uncertainties include the posteriori emission parameter error, the averaging kernel matrix, and the degrees of freedom of signal.

The Jacobian matrix, **K**, calculated using the forward model is central to the calculation of the various statistics and diagnostics discussed in the following sections. The key assumption in using the Jacobian is that changes in the emissions can be described approximately by (Rodgers, 2000)

$$\mathbf{F}(\boldsymbol{x}) - \mathbf{F}(\boldsymbol{x} + \delta\boldsymbol{x}) \approx \mathbf{K}\delta\boldsymbol{x} \tag{17}$$

this assumption has been validated using finite differencing to compare to solutions derived from the right side of Eq. (17).

2.4.2 Emission error characterization

We calculate various statistics to determine the emission estimation performance. First, we determine the posteriori emission parameter error covariance, which is defined by (Rodgers, 2000)

$$\mathrm{E}\left[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\right] = (\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_n^{-1}\mathbf{K})^{-1} \tag{18}$$

Next, we calculate the averaging kernel defined by

$$\mathbf{A} = (\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_n^{-1}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{S}_n^{-1}\mathbf{K} \tag{19}$$

and the degrees of freedom of signal that is calculated via

$$d.o.f. = \det(\mathbf{A}) \tag{20}$$

where both of these diagnostics provide information on the resolution of the emission retrieval, i.e., the ability of the estimate to uniquely distinguish between the emissions of individual species. For example, low values in the diagonal of **A** for species, i, and hence low values of $d.o.f.$, combined with a large width in the row of the averaging kernel for species, i, would indicate a low retrieval resolution and a diminished ability to individually resolve the emission parameter of species, i. The columns of **A** represent the sensitivity of the retrieved emission parameter to perturbations in the real emission parameters. The diagonals also represent this sensitivity for the specific case of $\hat{\boldsymbol{x}}_i$ to $\boldsymbol{x}_i$.

2.4.3 Ozone prediction error characterization

Using the posteriori emission error we can determine the posteriori ozone prediction error during the prediction time. In order to determine the posteriori ozone prediction

error beyond the observation period in the prediction and monitoring period on day 3 of the forward model run we need to define a new Jacobian matrix, $\mathbf{K}'$. $\mathbf{K}'$ describes the forward photochemical response to perturbations in the emissions at time, t, during the prediction and monitoring period on the 3rd day. Note that $\mathbf{K}$ and $\mathbf{K}'$ differ in that $\mathbf{K}$ only describes the model response in the observation period as opposed to the prediction and monitoring period. Each element of $\mathbf{K}'$ is $\partial q_z(\boldsymbol{x},t_j)/\partial \boldsymbol{x}_j$ where j is the index of time denoting when the model is sampled in the prediction and monitoring period on the 3rd day. The posteriori prediction error covariance on the 3rd day for all species in the vector, z, can be determined by

$$\mathrm{E}\left[\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}^T\right] = \mathbf{K}'\,\mathrm{E}\left[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T\right]\mathbf{K}'^T \tag{21}$$

2.5 Model applications

2.5.1 Overview

We conduct uncertainty analysis of the assimilation performance under different observed species scenarios, observing frequencies, observing noise, and varying photochemical environment using the diagnostics of both emission parameter and ozone posteriori prediction error, averaging kernel, and degrees of freedom of signal. In addition, we support the uncertainty analysis with a series of supporting sensitivity analyses used to test the robustness of our conclusions to methodological choices.

2.5.2 Uncertainty analysis

We calculate the diagnostics of both emission parameter and ozone posteriori prediction error, averaging kernel, and degrees of freedom of signal over the full range of observing scenarios (CN, OCN, and HCN), the full range of NO emission scenarios (x_{NO} = 0.5–2.5 with increments of 0.25), and eight different levels of observing error: 1 %, 5 %, 10 %, 25 %, 50 %, 100 %, 250 %, 500 % (β = 0.1–5). The observing errors

are absolute errors represented here as a percentage of the average species concentration over all of the photochemical scenarios. The final analysis involves investigating the sensitivity of posteriori ozone prediction error to observing frequency and to observation removal.

2.5.3 Supporting sensitivity analysis

We demonstrate the usage and performance of the 4-D-variational data assimilation. This model solves the non-linear estimation problem whereby it optimizes the ozone precursor emissions across the full range of photochemical conditions (x_{NO}=0.5–2.5) for the CN, OCN and HCN scenarios whilst assuming low levels of observational error ($\beta = 0.1$) represented in the observation error covariance matrix. Second, we probe the sensitivity of the success of emissions estimate, and hence the predictive model, to a range of assumed emission diurnal profiles. We assume the following profiles selected arbitrarily to test the model sensitivity: constant, sine wave, square wave, and offsets of the existing profile by 1 and 2 h shifts both forward and backward in time. Figure 1 shows the alternative emission variability profiles compared to the standard emission profile over the full diurnal period. These alternate emission profiles are taken to represent the new true state (using $x_{\mathrm{NO}} = 0.75$) and are used to generate the pseudo observations (using $\beta = 0.1$). We then attempt the assimilation using the pseudo observations generated from the alternative emission scenarios whilst assuming that the emissions temporal variability is the standard variability. The alternate emission profiles test the robustness of the 4-D-variational data assimilation method to diurnal uncertainty in the emissions. Finally, we test the sensitivity of the emission estimate and the ozone forecast error to the ethene emission assumptions. Specifically, we conduct a sensitivity test whereby we represent VOC emission uncertainties with uncertainties in the emission of ethane, which is a less reactive VOC compared to ethene.

3 Results

3.1 Uncertainty analyses

3.1.1 Emission error characterization and ozone prediction error

We present results from the uncertainty analysis over scenarios CN, OCN, and HCN, all nine NO emission scenarios, and the eight levels of observation error. These results includes the posteriori ozone prediction error (calculated by Eq. 21) and the posteriori emission parameter error (calculated by Eq. 18). We will further characterize the emission estimate using the averaging kernel and degrees of freedom of signal diagnostics. Figure 7 presents the posteriori ozone prediction errors across the complete range of parameter space and, in each panel, the results from the three observing scenarios CN, OCN, and HCN. The three scenarios CN, OCN, and HCN all exhibit similar general behavior in the derived posteriori ozone prediction errors: a first maximum in ozone prediction uncertainty in the NO_x limited scenarios (x_{NO} = 0.5–0.75), with a consistent minimum in ozone prediction error in the transition region that is both NO_x and VOC limited (x_{NO} = 1.0–1.75), and a second larger maximum in ozone prediction uncertainty in the VOC limited regime (x_{NO} = 2–2.5). Scenario CN (observing only CO and NO_2) yields the highest posteriori ozone prediction uncertainties of the three scenarios across the range of NO emission scenarios. In both scenarios OCN and HCN (observations of ozone, CO, and NO_2 and HCHO, CO, and NO_2) the inclusion of ozone and HCHO observations, respectively, reduces the posteriori ozone prediction uncertainties. Scenarios OCN and HCN each exhibit subtly different perturbations to the scenario CN (observations of CO and NO_2) prediction errors; both scenario OCN and HCN show significant improvement in the VOC limited emission scenarios (x_{NO} = 2.0–2.5) compared to scenario CN with each outperforming scenario CN by up to 2.4 ppbv. Despite scenario HCN performing comparably to scenario OCN within the most VOC limited regimes (x_{NO} >2.0), using the lowest of the VOC limited NO emission factor (x_{NO} = 2.0) scenario OCN outperforms scenario HCN by up to 1.4 ppbv. The largest

improvements in ozone prediction errors of any scenario comparison (2.6 ppbv) occur in scenario OCN compared with CN in the strongly NO_x limited emission scenarios (x_{NO} = 0.5–1.0). In addition, scenario OCN also has lower posteriori ozone prediction error compared to the HCN scenario under the NO_x limited conditions by up to 1.9 ppbv.

We will now focus on explaining these differences in posteriori ozone prediction error highlighted above, i.e., the differences that occur under VOC limited conditions (x_{NO}2.0–2.5) between the OCN and HCN scenarios and the CN scenario of up to 2.4 ppbv, the difference of 1.4 ppbv between the OCN and HCN scenarios using the lowest of the VOC limited NO emission factor (x_{NO} = 2.0), and the difference of 2.6 ppbv and 1.9 ppbv between the OCN and CN scenarios, and OCN and HCN scenarios, respectively, under NO_x limited conditions. To gain further insight into this behavior Figs. 8 and 9 show the posteriori error for x_{NO} and x_{VOC}. Note that the posteriori error for x_{CO} is invariant with respect to the photochemical regime and is therefore unable to explain any of the observed variability of ozone prediction error over varying x_{NO}.

Figure 8 shows that scenario HCN is able to reduce x_{VOC} posteriori errors over the largest range of NO emission scenarios, followed by scenario OCN, and scenario CN. This reduction in VOC emission uncertainty in scenario HCN explains why it shows reduced posteriori ozone prediction error (by up to 2.4 ppbv) compared to the CN scenario under VOC limited conditions. Despite HCHO observations overall providing a better constraint on VOC emission uncertainties under all conditions this improved constraint only leads to lower posteriori ozone prediction error compared to the OCN scenario in the transition region between the NO_x and VOC limited regimes (x_{NO}= 1.0–1.75), and under the most VOC limited conditions (x_{NO} >2.0). This inconsistency between the ozone prediction error and the posteriori error for x_{VOC}, when comparing the OCN and HCN scenarios at x_{NO} = 2.0, occurs because under VOC limited conditions the posteriori ozone prediction error is also sensitive to the posteriori NO emission uncertainties and ozone and HCHO have different abilities to constrain NO emission uncertainties. As shown by Fig. 9 ozone is better able to constrain NO emission uncertainties as compared with HCHO under both NO_x limited and VOC limited conditions.

Note that ozone is still sensitive to changes in the NO emissions within the VOC limited regime, but the sensitivity is negative, i.e., increases in NO emission lead to reduction in ozone concentrations. Therefore, the posteriori ozone prediction uncertainties in the VOC limited regime are still sensitive to posteriori NO emission uncertainties, which is why scenario OCN has a lower posteriori ozone prediction uncertainty in the VOC limited regime as compared to scenario HCN by 1.4 ppbv. This improved estimation of x_{NO} within the OCN scenario also explains why it outperforms the CN scenario (by up to 2.4 ppbv) under VOC limited conditions (x_{NO}2.0–2.5) despite showing higher VOC emission uncertainty compared to the HCN scenario.

Figure 9 shows that the OCN scenario exhibits the smallest posteriori NO emission parameter errors of any of the other observing scenarios. This is particularly pronounced under VOC limited conditions and under NO_x limited conditions. In the case of NO_x limited conditions, this improvement in NO emission parameter error in the OCN scenario compared to the CN scenario leads to the observed difference of 2.5 ppbv in the posteriori ozone prediction errors between these two cases. These differences can be attributed to the addition of ozone observations in the OCN scenario and the fact that ozone sensitivity to NO emissions under both NO_x limited and VOC limited conditions allows NO emissions to be better constrained. Additionally, improvements in posteriori NO emission parameter error under NO_x limited conditions lead to direct improvements in posteriori ozone prediction error due to the sensitivity of the ozone to changes in NO emissions. This same effect leads to the OCN scenario out performing the HCN scenario under NO_x by 1.9 ppbv since HCHO observations do not constrain NO emissions very well.

We now briefly explore the benefits of combing the observed species from scenarios OCN and HCN. Figure 10 shows that a scenario combining ozone and HCHO observations with CO and NO_2 observations can improve ozone prediction errors by up to 2.9 ppbv and 3.1 ppbv under NO_x and VOC limited conditions, respectively, compared to the CN scenario. Combining ozone and HCHO observations improves ozone prediction errors by up to 0.3 ppbv and 0.8 ppbv under NO_x and VOC limited conditions,

respectively, compared to the OCN scenario, which is a modest improvement. The differences between the ozone and HCHO combined scenario and the OCN scenario under VOC limited conditions further highlight the potential for HCHO observations to improve ozone prediction errors under the most VOC limited conditions.

It should be noted that scenario HCN uses an assumption regarding the relative observing errors of HCHO versus ozone, which based on the specifics of the instrumentation, magnitude of absorption cross-section, and interferences from other absorbing gases, and the relative trace gas concentrations, are expected to favour ozone having lower observation errors. Therefore, we utilize a scenario where β is scaled upwards independently for HCHO by 50 % relative to the other species. This assumption is explored in a sensitivity study and the results from this are shown in Fig. 11. Figure 11 shows that scenario HCN only has lower posteriori ozone prediction uncertainties over the full range of NO emission scenarios under the optimistic scenario of lower HCHO observation uncertainties, and that in the other scenarios, that we assume would be closer to reality, scenario HCN only out performs scenario OCN in the transition region between the NO_x and VOC sensitive regimes. Under the assumptions of lower ozone observing uncertainty OCN out performs scenario HCN in the NO_x and VOC limited regimes by up to 1.9 ppbv.

3.1.2 Averaging kernel and degrees of freedom of signal

Using Eq. (19) we can calculate the averaging kernel matrix. The diagonal represents the sensitivity of the retrieved parameter for a particular species, i, to changes in the real emission parameter for species, i. Figure 12 shows the respective diagonals of the averaging kernel (for x_{VOC} and x_{NO}) varying in a manner consistent with the posteriori parameter errors as shown in Figs. 8 and 9. A comparison of the lower panels indicates that the NO emission parameter estimate using the OCN observing scenario is more sensitive to the true state of the NO emission parameter under both NO_x limited and VOC limited conditions than any of the other observing scenarios. The top panels show that the VOC parameter estimate shows the highest sensitivity to the true state of the

VOC emission parameter using the HCN observing scenario.

The variability of the degrees of freedom of signal with both x_{NO} and β is shown in Fig. 13 for both the OCN and HCN scenarios. Figure 13 shows that the HCN scenario is better able to uniquely retrieve and resolve the 3 separate emission parameters compared to the OCN scenario. This is because HCHO provides a better constraint on VOC emissions over a wider range of x_{NO} and β. However, ozone in general constrains ozone precursor emissions across a wider a variety of emission parameters, specifically for x_{NO}, which allows ozone observations to yield better posteriori ozone prediction errors. The OCN scenario constrains the NO emission parameter better under both NO_x and VOC limited conditions. The OCN scenario has a decrease in the degrees of the freedom of signal under NO_x limited conditions due to the lack of sensitivity of the retrieval when using these observations to the VOC emission parameter.

3.1.3 Observing time and observing frequency

Using scenario OCN with $\beta = 0.25$, we conduct two separate sensitivity analyses to investigate the effects of observing time and observing frequency upon posteriori ozone prediction error. Figure 14 shows how removing the observations for all observed species (ozone, CO and NO_2) at specific times results in increases in the posteriori ozone prediction errors in that perturbed scenario. Figure 14 shows posteriori ozone prediction errors are most sensitive to the removal of observations during the day particularly during the high emission periods in the morning and afternoon rush hours and particularly so during the period of elevated ozone in the afternoon. The timing and magnitude of the sensitivity and its peak to observation removal varies according to the 9 NO emission scenarios as well. In the more NO_x limited scenarios, $x_{NO} = 0.5$–1.0, the sensitivity to observation removal is distributed relatively evenly over the entire day including both rush hours and peaks close to 15:00 LT. In the VOC limited regimes, $x_{NO} = 1.75$–2.5, the sensitivity to observation removal is more tightly distributed within the afternoon period and peaks between 15:00 LT and 18:00 LT even showing a broad maximum out to 20:00 LT under the most VOC limited conditions. The temporal vari-

ability of the maximum sensitivity to observation removal with changing photochemical regime is due to the timing of afternoon peak ozone concentrations. This is because across all of the photochemical regimes maxima in ozone sensitivity to perturbations in emissions coincide with the daytime peak ozone concentration (see Fig. 6). Observations made during these key periods are therefore better able to constrain the emissions uncertainties. Ozone concentrations peak later in the afternoon under more VOC limited conditions compared to the NO_x limited conditions thus explaining some of the variability in maximum sensitivity to observation removal with changing photochemical regime.

Figure 15 shows how posteriori ozone prediction errors vary with changing observing frequency. Increasing observing frequency causes the largest decreases in posteriori ozone prediction uncertainty in the VOC limited regime and to a lesser extent in the NO_x limited regime due to the sensitivity of ozone prediction error to unresolved emission parameter errors in those regimes.

3.2 Supporting sensitivity analyses

3.2.1 4-D-variational data assimilation

We run the full data assimilation forecasting across the full range of 9 x_{NO} scenarios (0.5–2.5) and each observing scenario (CN, OCN, and HCN) and the results from these experiments are shown in Table 4. The results shown in Table 4 indicate that scenarios OCN and HCN (observations of ozone, CO, and NO_2 and ozone, CO, NO_2) yield acceptable prediction error under these idealised conditions ($\beta = 0.1$) within this prototype framework for all photochemical conditions. The more limited success of scenario CN (observations of CO and NO_2) is due to the lower sensitivity of CO and NO_2 observations to the emissions of NO_2, and VOCs. The magnitude of the adjoint sensitivities guides the L-BFGS algorithm to the global minimum. In cases where the adjoint sensitivities are low (e.g., in VOC limited conditions using the CN scenario) the optimization routine may only be able to find a non-global minimum, which leads to

larger posteriori emission factor errors, $\hat{\boldsymbol{x}} - \boldsymbol{x}$.

Table 4 indicates that there is variability of posteriori peak ozone prediction error over changing photochemical regime and x_{NO} for each observing scenario CN, OCN, and HCN. This variability with x_{NO} is due in part to the variations in modeled ozone sensitivity to the different ozone precursor emission parameters, $\partial q_{O_3}(\boldsymbol{x}, t)/\partial x_i$, and the posteriori emission parameter errors (i.e., $\hat{\boldsymbol{x}} - \boldsymbol{x}$). Generally, large sensitivity of predicted ozone to the emissions of ozone precursors, $\partial q_{O_3}(\boldsymbol{x}, t)/\partial \boldsymbol{x}_i$, combined with unresolved ozone precursor emission parameter errors, $\hat{\boldsymbol{x}} - \boldsymbol{x}$, can lead to larger posteriori peak ozone prediction error. For instance, in the NO_x limited regimes ($\boldsymbol{x}_{NO} = 0.5$–$1.0$) large residual error in the element of $\hat{\boldsymbol{x}}$ corresponding to NO emissions would lead to large posteriori ozone errors.

One example of this phenomenon occurs in the case of photochemically VOC limited NO_x emission scenarios ($x_{NO} = 1.75$–2.5). Table 5 shows the variability of posteriori VOC emission errors with $\boldsymbol{x}_{NO}$ and observing scenario. For observing scenario CN there is large unresolved error in $\boldsymbol{x}_{VOC}$ (Table 5), and this leads to larger posteriori ozone prediction error as compared to scenarios OCN and HCN (see Table 4), which are better able to resolve errors in VOC emissions.

Thus, there are a rather complex set of factors interacting to cause these resulting posteriori prediction errors and the analysis of the results is limited to identifying relationships between the observing scenario, the photochemical regime, the adjoint sensitivities and the resulting ozone posteriori prediction error. This demonstrates the utility of the analytical model in allowing a far more in-depth analysis. Overall, the 4-D-variational data assimilation framework seems capable of resolving emission uncertainties and in turn reducing ozone prediction error. This successful demonstration of the framework is a necessary but not sufficient condition for systems based upon more complex photochemical models to have utility and ozone predictive skill.

3.2.2 Probing emission solution sensitivity to diurnal emission variability

We investigate the sensitivity of the forward photochemical model ozone mixing ratios, the 4-D-var ozone prediction and the 4-D-var emissions estimate to the underlying assumptions of the diurnal variability of emissions using 7 perturbed scenarios, and the results from this analysis are presented in Table 6. Both the forward model and the 4-D-var were conducted for $x_{NO} = 0.75$ and the 4-D-variational data assimilation results were generated using pseudo observations from the perturbed emission scenarios (defined by Fig. 1) with $\beta = 0.1$. Table 6 indicates that the forward model shows peak ozone mixing ratios diverging from the base case run (standard assumed emission variability with $\boldsymbol{x}_{NO} = 0.75$) by up to 10.6 ppbv and that the forward model ozone mixing ratios are sensitive to the assumption of the diurnal emission variability. In addition, Table 6 shows that the 4-D-variational data assimilation is able to achieve posteriori peak ozone prediction errors of up to 2.4 ppbv relative to the true state, as defined by the perturbed scenario, despite using the un-perturbed diurnal emission scenario as its emission variability. Despite the relative success of the posteriori peak ozone prediction (only a maximum ozone prediction error of 2.4 ppbv) under these more challenging conditions the assimilation performs poorly in terms of the posteriori emission factor error. Errors range up to 0.46 (18–92 %), 0.17 (17 %), and 7.0 (108 %) for $\boldsymbol{x}_{NO}$, $\boldsymbol{x}_{CO}$, and $\boldsymbol{x}_{VOC}$ (relative to true scaling factors of 0.5–5.0, 1.0, and 6.5, respectively) and thus emission inversion success is strongly affected by errors in the assumed diurnal variability of ozone precursor emissions. Thus, in summary, we demonstrate forward model ozone sensitivity to perturbations in the diurnal variability of ozone precursor emissions, relative insensitivity of the 4-D-variational data assimilation posteriori prediction error to mismatches in the assumed versus observed diurnal variability of ozone precursor emissions, and sensitivity of the emissions inversion success to mismatches in the assumed versus true emissions variability.

Having demonstrated that the photochemical box forward model ozone mixing ratio is sensitive to changes in the diurnal variability of emissions we also explore what

the real-world variability is in terms of day-to-day emission magnitude and apparent emission profile for a specific case. This investigation is necessary because we assume that there is no day-to-day variation in either emission magnitude or the profile of the emissions. Observation data for ozone, CO and NO_2 collected by the South Coast Air Quality Monitoring District at Wilson Ave., Pasadena (see Fig. 16) show that this assumption is valid for a consecutive three day period consisting of Wednesday, Thursday, and Friday. Our assumption of no day-to-day variability in ozone precursor emissions is reasonable for this region. However, our assumptions regarding day-to-day emissions variability would be unreasonable for a consecutive three day period covering part or all of the weekend due to the variability in ozone precursor emissions between the working week and the weekend. Though our assumption of no day-to-day emission variability would provide an unrealistic description of the true emission variability for a simulation period covering the mid-week and weekend the previous analysis shows that the 4-D-variational data assimilation framework can still provide acceptable ozone forecasts in the presence of either incorrect emission magnitudes or emission variability profiles.

3.2.3 Emission inversion and ozone predictive skill sensitivity to VOC species selection

Figure 17 shows the results from the analysis probing emission solution sensitivity to choice of VOC used in the emission inversion. In this study we substitute ethene emission uncertainty for ethane emission uncertainty. Figure 17 shows that the VOC emission inversion is severely degraded by building the Jacobian by perturbing $\boldsymbol{x}_{\text{ethane}}$ as opposed to $\boldsymbol{x}_{\text{ethene}}$ across scenarios CN, OCN and HCN. The posteriori $\boldsymbol{x}_{\text{VOC}}$ parameter error relaxes to our chosen a priori of 1.5 to within 1 significant figure for most of the scenarios explored, but note that this does not affect ozone prediction error (Fig. 17) since the degraded VOC emission uncertainty is mitigated by the lower model sensitivity to that uncertainty because of the decreased reactivity and decreased ozone yield resulting from ethane oxidation relative to ethene.

4 Discussion

4.1 Observed species

The variability of ozone prediction error with both photochemical regime and observing species scenario (CN, OCN and HCN) is complex and no single observed species is ideal for all photochemical conditions. Certain pairs or groups of observations have greater impact given their complementary nature across a range of scenarios. Under NO_x limited conditions ozone prediction error is strongly controlled by the posteriori NO emission errors and therefore observations of NO_2 and ozone would be highly advantageous. Ozone provides a particularly good constraint upon NO emissions under very NO_x limited conditions. The value of NO_2 observations in constraining NO emissions improves as the NO_x lifetime increases under higher x_{NO} and the photochemical conditions become less NO_x limited. Much of the troposphere is in fact highly NO_x limited outside of the most polluted of urban areas (Duncan et al., 2010), e.g., in pristine, rural, and most suburban locations within the United States where there is plentiful vegetation and therefore significant biogenic VOC emissions. Under VOC limited conditions ozone prediction error is sensitive to both posteriori $\boldsymbol{x}_{\text{NO}}$ (due to the negative sensitivity of ozone to NO_x) and $\boldsymbol{x}_{\text{VOC}}$ errors and thus observations of ozone, HCHO and NO_2 allow significant improvements in ozone prediction error. Ozone allows constraints to be placed upon VOC and NO emission uncertainties, HCHO provides an excellent constraint upon reactive VOC emissions, which due to their reactivity are more relevant to air quality compared to less reactive VOCs, and NO_2 provides an excellent constraint upon NO emissions (more so than under NO_x limited conditions due to the longer NO_x lifetime). Despite the fact that large geographical portions of the US are NO_x limited a disproportionately large percentage of the populous live within or are exposed to ozone arising from VOC limited conditions due to the significant extent of urbanization within the US. Large urbanized areas of the South West that lack significant native vegetative biomass typically have a larger VOC limited regime that extends over the urban as well as sub-urban areas. In contrast, US cities in the East are located in regions with of-

ten times dense vegetative biomass (e.g., Atlanta) and thus the VOC limited region is far more geographically limited to the urban center itself. Therefore, improving ozone predictive skill within VOC limited conditions won't yield forecasting improvements over a wide geographical area but will yield improvements within certain regions with large populations where improved ozone predictive skill would be advantageous.

Our findings with respect to the utility of NO_2 and HCHO observations for constraining NO_x and VOC emissions, respectively, and in turn for improving ozone estimation are broadly consistent with the findings of Zhang et al. (2008), which used satellite observations of NO_2 and HCHO in conjunction with 4-D-variational data assimilation to solve for NO_2 and HCHO emissions and to improve the model's ozone estimation. This confirms that frameworks of this nature have the potential to function operationally and to improve air quality forecasting.

There are two further advantages to observations of ozone and HCHO made under VOC limited conditions. Often times plumes of NO_x polluted and VOC limited air can be exported from the VOC limited centralised urban region into sub-urban areas that are NO_x limited, and this can lead to significant variability in the photochemical regime in the regions surrounding an urban center. Therefore, observations of HCHO and ozone in addition to NO_2 observations could help to understand such events and in turn reduce ozone prediction errors. The second reason is that observations of ozone and HCHO can place constraints upon reactive VOC emissions, and in many cases that information regarding emissions can allow valid assumptions to be made regarding the geographical variability of VOC emissions considering land types and density of urbanization. However, often times this will extend the information regarding the geographical distribution of VOC emissions into NO_x limited regions. Although posteriori VOC emission uncertainties have the largest impact on ozone prediction error under VOC limited conditions, posteriori VOC emission uncertainties can still impact ozone prediction errors under NO_x limited conditions, so constraining VOC emission uncertainties will create benefits to ozone forecasting.

4.2 Temporal considerations

There is strong sensitivity of ozone prediction error to observation removal in the daytime, particularly in the afternoon, and therefore observations made during the day present greater returns in terms of improved forecasting ability. There is some temporal variability in the sensitivity of prediction error between the different x_{NO} scenarios. The NO_x limited regimes favour observations made throughout the day with increased observing density close to 15:00 LT. The VOC limited regimes favour a greater concentration of observations within the afternoon even up to 6pm in the most VOC limited cases.

4.3 Implications for emission inversion

Aside from the relevance of these results to GEO mission planning and air quality forecasting in general, we believe these results are also relevant for emission and flux estimation via inversion methologies. Our prototype framework is mechanically very similar to recent work using 4-D-variational data assimilation methologies (Henze et al., 2009; Stavrakou et al., 2009; Kopacz et al., 2010) using chemistry transport models that have focused on emission inversion. Since much of the emission inversion performance shown in this study is driven by the photochemistry, and the fundamental photochemical behaviour shown by our mechanism should be reproducable across all photochemical mechanisms it is reasonable to suppose that some of our conclusions are relevant to future work conducted using 4-D-variational data assimilation in emission inversion studies. From this premise, we recommend that emission inversion studies for NO_x utilize both observations of NO_2 and ozone observations since ozone observations add information to the x_{NO} estimation under both strongly positively and negatively NO_x limited conditions and NO_2 observations constrain emission parameter uncertainties the most under less NO_x limited conditions through to the negatively sensitive regime, so these observations complement each other. Likewise for emission inversions of VOCs we recommend observations of HCHO and ozone since HCHO

observations can constrain VOC emission uncertainties under a wide variety of photochemical conditions and ozone can constrain VOC emission uncertainties under VOC limited conditions.

It should be noted that the conclusions regarding VOC emission inversion are sensitive to our choice of representing VOC emission uncertainties with ethene. The success of the VOC emission inversion is significantly limited by solving for ethane emission uncertainties as opposed to ethene emission uncertainties. This inability to constrain unreactive VOC emissions is due to the lack of impact on secondary chemical species such as HCHO. The inability to place useful constraints on unreactive VOC emissions is one reason why previous emission inversion modeling studies have focused on constraining reactive VOCs like isoprene (Millet et al., 2006, 2008; Palmer et al., 2003, 2006).

In the supporting sensitivity analysis probing emission solution sensitivity to diurnal emission variability we demonstrate that emission inversions are potentially highly sensitive to the assumed variability of the emissions and that even perfect observations would lead to such errors. In our system such emission inversion errors would be hard to characterize in the absence of any information regarding the true state of the emissions variability. We recommend that such uncertainties should be considered and characterized in emissions inversion studies. Currently diurnal emission variabilities are determined in the process of building bottom-up emission inventories. Although our prototype assimilation system can only currently solve for time independent scaling factors it could be modified to solve for time dependent scaling factors and the diurnal emissions variability. Future assimilation forecasting systems should also possess this ability to solve for time dependent emission scaling factors. Increased temporal sampling density provided by GEO sounding satellites could place increased constraints upon temporally varying emissions as opposed to current LEO satellites.

5 Conclusions

We demonstrate the relevance of a prototype air quality forecasting box model to future air quality forecasting systems that might utilize state of the art assimilation and GEO orbiting satellite observations. Based on this framework, we show that ozone can be forecast within idealized conditions when using observing scenarios consisting of observations of ozone, CO, NO_2 and HCHO, CO, and NO_2, as a function of photochemical conditions. Uncertainties in ozone prediction are relatively insensitive to assumed diurnal emission variability even though the emission fluxes are sensitive to the diurnal profile.

Results from analytical calculations indicate that combined observations of ozone and NO_2 over regions within NO_x limited photochemical conditions will yield the largest decreases in ozone prediction error whereas observations of ozone, HCHO and NO_2 observations reduce ozone prediction errors the most under VOC limited conditions. Within the transition region between NO_x and VOC limited conditions, observations of HCHO and NO_2 reduce ozone prediction errors the most though posteriori ozone prediction errors are at their minimum within this photochemical region. Overall we find that no single species observation is capable of yielding acceptable ozone prediction error under all photochemical conditions. Further, observing only CO and NO_2 limits ozone predictive skill across both NO_x and VOC limited photochemical regimes. Our results indicating the potential for observations of NO_2 and HCHO to constrain NO_x and VOC emissions are in turn improve ozone prediction errors are consistent with previous work by Zhang et al. (2008)

If instead the objective and focus is upon optimizing ozone precursor emissions, then emissions of NO_x are best optimised using a combination of NO_2 and ozone observations, and VOC emissions are best optimized using both HCHO and ozone observations. Observations of HCHO provide a better constraint on VOC emission uncertainties compared to ozone under all photochemical conditions, but they only improve posteriori ozone prediction error relative to ozone observations in the transition

region. Investigations probing the temporal frequency of observations and the degree to which observations made at different times of day affect prediction error indicate that increasing the observing frequency always leads to decreases in prediction error and that observations made during the afternoon through to the early evening decrease prediction error the most.

Observations within scenarios OCN and HCN with uncertainties of up to an absolute error equivalent to 33 % of the average over polluted regions can achieve adequate ozone prediction errors (i.e., within 0–5 ppbv for most photochemical scenarios) within our prototype framework. This is a necessary but not sufficient condition in order for these observations to achieve acceptable ozone prediction error in more complex air quality forecasting systems.

Acknowledgements. PDH thanks ORAU and the NPP program for funding and support, D. Millet and D. Jacob for helpful discussion and comments, and A. M. Aghedo for help and formatting tips. DKH supported by NASA Applied Science Program.

References

Arellano, A., Kasibhatla, P., Giglio, L., van der Werf, G., Randerson, J., and Collatz, G.: Time-dependent inversion estimates of global biomass-burning CO emissions using Measurement of Pollution in the Troposphere (MOPITT) measurements, J. Geophys. Res.-Atmos., 111, D09303, doi:10.1029/2005JD006613, 2006. 19294

Arnold, C. and Dey, C.: Observing-systems simulation experiments: Past, present, and future, B. Am. Meteorol. Soc., 1986. 19296

Aumann, H., Chahine, M., Gautier, C., Goldberg, M., Kalnay, E., McMillin, L., Revercomb, H., Rosenkranz, P., Smith, W., Staelin, D., Strow, L., and Susskind, J.: AIRS/AMSU/HSB on the aqua mission: Design, science objectives, data products, and processing systems, IEEE Trans. Geosci. Remote Sens., 41, 253–264, doi:10.1109/TGRS.2002.808356, 2003. 19295, 19296

Beer, R.: TES on the Aura mission: Scientific objectives, measurements,

and analysis overview, IEEE Trans. Geosci. Remote Sens., 44, 1102–1105, doi:10.1109/TGRS.2005.863716, 2006. 19294, 19295, 19296

Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V., Chance, K., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, J. Atmos. Sci., 56, 127–150, 1999. 19295, 19296

Bowman, K. W., Jones, D. B. A., Logan, J. A., Worden, H., Boersma, F., Chang, R., Kulawik, S., Osterman, G., Hamer, P., and Worden, J.: The zonal structure of tropical O3 and CO as observed by the Tropospheric Emission Spectrometer in November 2004-Part 2: Impact of surface emissions on O-3 and its precursors, Atmos. Chem. Phys., 9, 3563–3582, doi:10.5194/acp-9-3563-2009, 2009. 19294

Burrows, J., Weber, M., Buchwitz, M., Rozanov, V., Ladstatter-Weissenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K., and Eisinger, M.: The global ozone monitoring experiment (GOME): Mission concept and first scientific results, J. Atmos. Sci., 56, 151–175, 1999. 19294, 19295, 19296

Campbell, J. and Fishman, J.: GEO-CAPE Mission NASA Workshop Report, Tech. rep., NASA Langley, USA, 2010. 19295

Chai, T., Carmichael, G. R., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E. V., Avery, M. A., Nedelec, P., Merrill, J. T., Thompson, A. M., and Williams, E.: Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements, J. Geophys. Res.-Atmos., 112, D12S15, doi:10.1029/2006JD007763, 2007. 19294

Clerbaux, C., Coheur, P. F., Clarisse, L., Hadji-Lazaro, J., Hurtmans, D., Turquety, S., Bowman, K., Worden, H., and Carn, S. A.: Measurements of SO2 profiles in volcanic plumes from the NASA Tropospheric Emission Spectrometer (TES), Geophys. Res. Lett., 35, L22807, doi:10.1029/2008GL035566, 2008. 19296

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P. F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, Atmos. Chem. Phys., 9, 6041–6054, doi:10.5194/acp-9-6041-2009, 2009. 19294, 19296

Dabberdt, W., Carroll, M., Baumgardner, D., Carmichael, G., Cohen, R., Dye, T., Ellis, J., Grell, G., Grimmond, S., Hanna, S., Irwin, J., Lamb, B., Madronich, S., McQueen, J., Meagher, J., Odman, T., Pleim, J., Schmid, H., and Westphal, D.: Meteorological research needs for improved air quality forecasting - Report of the 11th prospectus development team of the US

Weather Research Program, B. Am. Meteorol. Soc., 85, 563–586, doi:10.1175/BAMS-85-4-563, 2004. 19293
Dabberdt, W. F., Carroll, M. A., Appleby, W., Baumgardner, D., Carmichael, G., Davidson, P., Doran, J. C., Dye, T. S., Grimmond, S., Middleton, P., Neff, W., and Zhang, Y.: USWRP Workshop on Air Quality Forecasting, B. Am. Meteorol. Soc., 87, 215–221, 2006. 19293
Daescu, D., Sandu, A., and Carmichael, G.: Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: II – Numerical validation and applications, Atmos. Environ., 37, 5097–5114, doi:10.1016/j.atmosenv.2003.08.020, 2003. 19300, 19304
Damian, V., Sandu, A., Damian, M., Potra, F., and Carmichael, G.: The kinetic preprocessor KPP – a software environment for solving chemical kinetics, Comp. Chem. Eng., 26, 1567–1579, 2002. 19300
Drummond, J. and Mand, G.: The measurements of pollution in the troposphere (MOPITT) instrument: Overall performance and calibration requirements, J. Atmos. Ocean. Technol., 13, 314–320, 1996. 19296
Dufour, G., Eremenko, M., Orphal, J., and Flaud, J. M.: IASI observations of seasonal and day-to-day variations of tropospheric ozone over three highly populated areas of China: Beijing, Shanghai, and Hong Kong, Atmos. Chem. Phys., 10, 3787–3801, doi:10.5194/acp-10-3787-2010, 2010. 19294
Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., Hu, Y., Pickering, K. E., Retscher, C., Allen, D. J., and Crawford, J. H.: Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation, Atmos. Environ., 44, 2213–2223, doi:10.1016/j.atmosenv.2010.03.010, 2010. 19294, 19320
Edwards, D., Arellano, A., and Deeter, M.: A satellite observation system simulation experiment for carbon monoxide in the lowermost troposphere, J. Geophys. Res.-Atmos., 2009a. 19296, 19297
Edwards, D. P., DeCola, P., Fishman, J., Jacob, D. J., Bhartia, P., Diner, D., Burrows, J. P., and M, G.: Community Input to the NRC Decadal Survey from the NCAR Workshop on Air Quality Remote Sensing From Space: Defining an Optimum Observing Strategy, http://geo-cape.larc.nasa.gov/docs/AirQualityfromSpace_Workshop_Rpt.pdf, 2009b. 19295
Eller, P., Singh, K., Sandu, A., Bowman, K., Henze, D., and M., L.: Implementation and evaluation of an array of chemical solvers in the Global Chemical Transport Model GEOS-Chem, Geosci. Model Dev., 2, 89–96, doi:10.5194/gmd-2-89-2009, 2009. 19300, 19304
Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia,

A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., Lopez-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151–2188, doi:10.5194/acp-8-2151-2008, 2008. 19296
Fishman, J., Creilson, J. K., Parker, P. A., Ainsworth, E. A., Vining, G. G., Szarka, J., Booker, F. L., and Xu, X.: An investigation of widespread ozone damage to the soybean crop in the upper Midwest determined from ground-based and satellite measurements, Atmos. Environ., 44, 2248–2256, doi:10.1016/j.atmosenv.2010.01.015, 2010. 19294
Fumagalli, I., Gimeno, B., Velissariou, D., De Temmerman, L., and Mills, G.: Evidence of ozone-induced adverse effects on crops in the Mediterranean region, Atmos. Environ., 35, 2583–2587, 2001. 19293
Gardner, M. and Dorling, S.: Statistical surface ozone models: an improved methodology to account for non-linear behaviour, Atmos. Environ., 34, 21–34, 2000. 19293
Gleason, J., Bhartia, P., Herman, J., McPeters, R., Newman, P., Stolarski, R., Flynn, L., Labow, G., Larko, D., Seftor, C., Wellemeyer, C., Komhyr, W., Miller, A., and Planet, W.: Record Low Global Ozone In 1992, Science, 260, 523–526, 1993. 19295
Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957–6975, doi:10.1016/j.atmosenv.2005.04.027, 2005. 19293
Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D., and Li, Q.: The adjoint of CMAQ, Enviro. Sci. Technol., 41, 7807–7817, 2007. 19298
Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic PM2.5 precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem. Phys., 9, 5877–5903, doi:10.5194/acp-9-5877-2009, 2009. 19322
IPCC: Inter-Governmental Panel on Climate Change Fourth Assessment Report Working Group 1 Report "The Physical Basis", United Nations, 2007. 19293
Jacob, D., Logan, J. A., Gardner, G., Yevich, R., Spivakosky, C., Wofsy, S., Sillman, S., and Prather, M.: Factors Regulating Ozone Over The United-States and its Export to the Global Atmosphere, Journal of Geophys. Res.-Atmos., 98, 14817–14826, 1993. 19297
Jenkin, M., Saunders, S., and Pilling, M.: The tropospheric degradation of volatile organic compounds: A protocol for mechanism development, Atmos. Environ., 31, 81–104, 1997. 19300
Jones, D., Bowman, K., Palmer, P., Worden, J., Jacob, D., Hoffman, R., and Bey, I.: Potential of

observations from the Tropospheric Emission Spectrometer to constrain continental sources of carbon monoxide, J. Geophys. Res.-Atmos., 108(D24), 4789 doi:10.1029/2003JD003702, 2003. 19296, 19297

Jones, D. B. A., Bowman, K. W., Logan, J. A., Heald, C. L., Liu, J., Luo, M., Worden, J., and Drummond, J.: The zonal structure of tropical O_3 and CO as observed by the Tropospheric Emission Spectrometer in November 2004 – Part 1: Inverse modeling of CO emissions, Atmos. Chem. Phys., 9, 3547–3562, doi:10.5194/acp-9-3547-2009, 2009. 19294

Kang, D., Mathur, R., and Rao, S. T.: Real-time bias-adjusted O3 and PM2.5 air quality index forecasts and their performance evaluations over the continental United States, Atmos. Environ., 44, 2203–2212, doi:10.1016/j.atmosenv.2010.03.017, 2010. 19293

Konovalov, I. B., Beekmann, M., Richter, A., and Burrows, J. P.: Inverse modelling of the spatial distribution of NOx emissions on a continental scale using satellite data, Atmos. Chem. Phys., 6, 1747–1770, doi:10.5194/acp-6-1747-2006, 2006. 19294

Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaia, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V., and Nedelec, P.: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), Atmos. Chem. Phys., 10, 855–876, doi:10.5194/acp-10-855-2010, 2010. 19294, 19322

Kurokawa, J.-i., Yumimoto, K., Uno, I., and Ohara, T.: Adjoint inverse modeling of NOx emissions over eastern China using satellite observations of NO2 vertical column densities, Atmos. Environ., 43, 1878–1887, doi:10.1016/j.atmosenv.2008.12.030, 2009. 19294

Landgraf, J. and Hasekamp, O. P.: Retrieval of tropospheric ozone: The synergistic use of thermal infrared emission and ultraviolet reflectivity measurements from space, J. Geophys. Res.-Atmos., 112, D08310, doi:10.1029/2006JD008097, 2007. 19296

Lee, S., Hong, Y., Song, C., Lee, M., Ryoo, S., Kim, J., Yong, S., Bhartia, P. K., Park, R., Woo, J., Kim, Y. J., Song, C. H., Kim, J. H., Lee, K., Ho, C., Park, S. K., Lee, Y., Lee, J., Kim, M., Eom, Y., and Hong, J.: Geostationary Environment Monitoring Spectrometer(GEMS) onboard MP-GEOSAT (Multi Purpose Geostationary Satellite) over Asia-Pacific region, AGU, 2009. 19295

Levelt, P., Van den Oord, G., Dobber, M., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J., and Saari, H.: The Ozone Monitoring Instrument, IEEE Trans. Geosci. Remote Sens., 44, 1093–1101, doi:10.1109/TGRS.2006.872333, 2006. 19295, 19296

Martin, R. V.: Satellite remote sensing of surface air quality, Atmos. Environ., 42, 7823–7843,

doi:10.1016/j.atmosenv.2008.07.018, 2008. 19294

Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B., Anderson, B. E., and Clarke, A. D.: Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission, J. Geophys. Res.-Atmos., 111, D24S02, doi:10.1029/2005JD006853, 2006. 19323

Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K., Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, J. Geophys. Res.-Atmos., 113, D02307, doi:10.1029/2007JD008950, 2008. 19294, 19323

Murphy, J., Delucchi, M., McCubbin, D., and Kim, H.: The cost of crop damage caused by ozone air pollution from motor vehicles, J. Environ. Manage., 55, 273–289, 1999. 19293

Mustafa, M.: Biochemical Basis of Ozone Toxicity, Free Radical Biology and Medicine, 9, 245–265, 1990. 19293

Nali, C., Pucciariello, C., and Lorenzini, G.: Ozone distribution in central Italy and its effect on crop productivity, Agr. Ecosyst. Environ., 90, 277–289, 2002. 19293

Otte, T., Pouliot, G., Pleim, J., Young, J., Schere, K., Wong, D., Lee, P., Tsidulko, M., McQueen, J., Davidson, P., Mathur, R., Chuang, H., DiMego, G., and Seaman, N.: Linking the Eta Model with the Community Multiscale Air Quality (CMAQ) modeling system to build a national air quality forecasting system, Weather Forecast., 20, 367–384, 2005. 19293

Palmer, P., Jacob, D., Fiore, A., Martin, R., Chance, K., and Kurosu, T.: Mapping isoprene emissions over North America using formaldehyde column observations from space, J. Geophys. Res.-Atmos., 108, 4180, doi:10.1029/2002JD002153, 2003. 19323

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer, C., Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B., and Sumner, A. L.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, J. Geophys. Res.-Atmos., 111, D12315, doi:10.1029/2005JD006689, 2006. 19323

Parrington, M., Jones, D. B. A., Bowman, K. W., Thompson, A. M., Tarasick, D. W., Merrill, J., Oltmans, S. J., Leblanc, T., Witte, J. C., and Millet, D. B.: Impact of the assimilation of ozone from the Tropospheric Emission Spectrometer on surface ozone across North America, Geophys. Res. Lett., 36, L04802, doi:10.1029/2008GL036935, 2009. 19293, 19294, 19297

Parrish, D. D., Aikin, K. C., Oltmans, S. J., Johnson, B. J., Ives, M., and Sweeny, C.: Impact of transported background ozone inflow on summertime air quality in a California ozone exceedance area, Atmos. Chem. Phys., 10, 10093–10109, doi:10.5194/acp-10-10093-2010, 2010. 19293

Pierce, R. B., Schaack, T., Al-Saadi, J. A., Fairlie, T. D., Kittaka, C., Lingenfelser, G., Natarajan, M., Olson, J., Soja, A., Zapotocny, T., Lenzen, A., Stobie, J., Johnson, D., Avery, M. A., Sachse, G. W., Thompson, A., Cohen, R., Dibb, J. E., Crawford, J., Rault, D., Martin, R., Szykman, J., and Fishman, J.: Chemical data assimilation estimates of continental US ozone and nitrogen budgets during the Intercontinental Chemical Transport Experiment-North America, J. Geophys. Res.-Atmos., 112, D12S21, doi:10.1029/2006JD007722, 2007. 19294, 19297

Pryor, W.: How Far Does Ozone Penetrate Into The Pulmonary Air Tissue Boundary Before It Reacts, Free Radical Biology And Medicine, 12, 83–88, 1992. 19293

Rayner, P., Enting, I., and Trudinger, C.: Optimizing the co2 observing network for constraining sources and sinks, Tellus B, 48, 433–444, 1996. 19296, 19297

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding – Theory and Practice, World Scientific, 65–79, 2000. 19299, 19307, 19308

Sandu, A., Daescu, D., and Carmichael, G.: Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I - theory and software tools, Atmos. Environ., 37, 5083–5096, doi:10.1016/j.atmosenv.2003.08.019, 2003a. 19294

Sandu, A., Daescu, D., and Carmichael, G.: Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I - theory and software tools, Atmos. Environ., 37, 5083–5096, doi:10.1016/j.atmosenv.2003.08.019, 2003b. 19300, 19304

Sofiev, M., Vankevich, R., Lotjonen, M., Prank, M., Petukhov, V., Ermakova, T., Koskinen, J., and Kukkonen, J.: An operational system for the assimilation of the satellite information on wild-land fires for the needs of air quality modelling and forecasting, Atmos. Chem. Phys., 9, 6833–6847, doi:10.5194/acp-9-6833-2009, 2009. 19297

Stavrakou, T., Mueller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., and Guenther, A.: Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003-2006, Atmos. Chem. Phys., 9, 3663–3679, doi:10.5194/acp-9-3663-2009, 2009. 19322

Valente, R., Imhoff, R., Tanner, R., Meagher, J., Daum, P., Hardesty, R., Banta, R., Alvarez, R., McNider, R., and Gillani, N.: Ozone production during an urban air stagnation episode over

Nashville, Tennessee, J. Geophys. Res.-Atmos., 103, 22555–22568, 1998. 19297

Van Dingenen, R., Dentener, F. J., Raes, F., Krol, M. C., Emberson, L., and Cofala, J.: The global impact of ozone on agricultural crop yields under current and future air quality legislation, Atmos. Environ., 43, 604–618, doi:10.1016/j.atmosenv.2008.10.033, 2009. 19293

Watson, I., Realmuto, V., Rose, W., Prata, A., Bluth, G., Gu, Y., Bader, C., and Yu, T.: Thermal infrared remote sensing of volcanic emissions using the moderate resolution imaging spectroradiometer, J. Volcanol. Geotherm. Res., 135, 75–89, doi:10.1016/j.jvolgeores.2003.12.017, 2004. 19296

Worden, H. M., Deeter, M. N., Edwards, D. P., Gille, J. C., Drummond, J. R., and Nedelec, P.: Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals, J. Geophys. Res.-Atmos., 115, D18314, doi:10.1029/2010JD014242, 2010. 19296

Worden, J., Liu, X., Bowman, K., Chance, K., Beer, R., Eldering, A., Gunson, M., and Worden, H.: Improved tropospheric ozone profile retrievals using OMI and TES radiances, Geophys. Res. Lett., 34, L01809, doi:10.1029/2006GL027806, 2007. 19296

Yi, J. and Prybutok, V.: A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area, Envrion. Pollut., 92, 349–357, 1996. 19293

Zhang, L., Constantinescu, E. M., Sandu, A., Tang, Y., Chai, T., Carmichael, G. R., Byun, D., and Olaguer, E.: An adjoint sensitivity analysis and 4D-Var data assimilation study of Texas air quality, ATMOSPHERIC ENVIRONMENT, 42, 5787–5804, doi:10.1016/j.atmosenv.2008.03.048, 2008. 19293, 19294, 19297, 19298, 19321, 19324

Zhu, C., Byrd, R., Lu, P., and Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization, ACM Trans. Math. Softw., 23, 550–560, 1997. 19304

Table 1. Background free tropospheric concentrations of trace gases mixed into the boundary layer in the photochemical model.

Chemical Species	Background Mixing Ratio
Ozone	30 ppbv
NO	100 pptv
NO_2	50 pptv
CO	80 ppbv
CH_4	1.76 ppm
NMHCs	100–200 pptv each

Table 2. Values of $\overline{\mathbf{F}(\boldsymbol{x})}$ used to calculate $\boldsymbol{y}$.

$\mathbf{F}(\hat{\boldsymbol{x}})$	Mixing Ratio
Ozone	44.4 ppbv
CO	620 ppbv
NO_2	6.5 ppbv
HCHO	3.9 ppbv

Table 3. Values of $\boldsymbol{x}$ and $\boldsymbol{x}_a$ used in the 4-D-variational data assimilation model.

$\boldsymbol{x}$			$\boldsymbol{x}_a$		
NO	CO	VOC	NO	CO	VOC
0.5	1.0	6.5	0.475	0.95	0.1
0.75	–	–	0.7125	–	–
1.0	–	–	0.95	–	–
1.25	–	–	1.1875	–	–
1.5	–	–	1.425	–	–
1.75	–	–	1.8375	–	–
2.0	–	–	2.1	–	–
2.25	–	–	2.3625	–	–
2.5	–	–	2.625	–	–

Table 4. Initial peak ozone predictions, true state peak ozone, initial guess ozone prediction error, and prediction error across the full range of x_{NO} and the three observing scenarios CN, OCN and HCN. The ozone values and absolute differences in ozone mixing ratio are listed for 15:00 LT during the final day of the prediction model. Figures 4 and 5 show what E and G represent.

x_{NO} Scenario	$q_{O_3}(\boldsymbol{x}_a, t^u)$ (ppbv)	$q_{O_3}(\boldsymbol{x}, t^u)$ (ppbv)	G (ppbv)	E (ppbv) Scenario CN	E (ppbv) Scenario OCN	E (ppbv) Scenario HCN
0.5	72.7	79.3	−6.6	−6.3	−0.4	−1.0
0.75	81.3	89.7	−8.4	−8.3	−0.5	−0.7
1.0	85.2	96.3	−11.1	−4.5	−0.6	−0.5
1.25	85.5	100.3	−15.1	−3.3	−0.6	−0.3
1.5	79.7	101.5	−21.8	−4.2	−0.5	−0.1
1.75	66.1	98.7	−32.6	2.2	0.3	−0.2
2.0	52.8	89.0	−36.2	1.9	0.3	−0.2
2.25	43.6	73.0	−29.4	1.4	0.3	−0.2
2.5	37.1	58.8	−21.7	1.0	0.3	−0.2

Table 5. The posteriori x_{VOC} error resulting from the 4-D-variational data assimilation. The table shows the variablity of the posteriori VOC emission error both with observing scenario and NO emission factor. Errors are represented as absolute errors of x_{VOC}.

	$\hat{x}_{VOC}$ - x_{VOC}		
x_{NO}	Scenario CN	Scenario OCN	Scenario HCN
0.5	−6.4	0.40	8.5×10^{-2}
0.75	9.1	0.33	5.0×10^{-2}
1.0	−2.7	−0.01	3.3×10^{-2}
1.25	−1.6	9.87	-2.6×10^{-2}
1.5	−1.7	2.71	-3.6×10^{-2}
1.75	0.77	0.21	2.4×10^{-2}
2.0	0.54	0.20	3.3×10^{-2}
2.25	0.40	0.18	4.5×10^{-2}
2.5	0.35	0.18	4.8×10^{-2}

Table 6. Results from a study exploring the sensitivity of the 4-D-variational data assimilation forecast of peak ozone to varying assumptions regarding, $k(t)$, the diurnal variability of ozone precursor emissions. Note that in each scenario the cumulative daily emission burden remains constant for each scenario and thus each scenario has identical $\overline{E(t)}$. The table shows (in ppbv) the modeled ozone for each alternative $k(t)$ scenario, the differences in true state peak ozone between these alternative $k(t)$ scenarios and the standard $k(t)$ scenario, and the absolute posteriori ozone prediction errors of these alternative $k(t)$ scenarios relative to both the standard and alternative $k(t)$ scenario true states. All of the ozone mixing ratios are listed for 15:00 LT during the final day of the prediction and monitoring period.

Assumed $k(t)$ Scenario	Alternative Emission Scenario True State (ppbv)	Alternative Emission Scenario True State – Standard Emission Scenario True State (ppbv)	Alternative Posteriori Prediction – Standard True State (ppbv)	Alternative Posteriori Prediction – Alternative True State (ppbv)
Constant	92.5	2.8	3.5	0.7
Sine Wave	97.6	7.9	8.3	0.5
Saw-Tooth	100.3	10.6	9.3	−1.4
Offset −1	93.8	4.1	4.2	0.1
Offset −2	98.9	9.0	8.8	−0.2
Offset +1	86.2	−3.5	−4.9	−1.4
Offset +2	83.5	−6.2	−8.6	−2.4

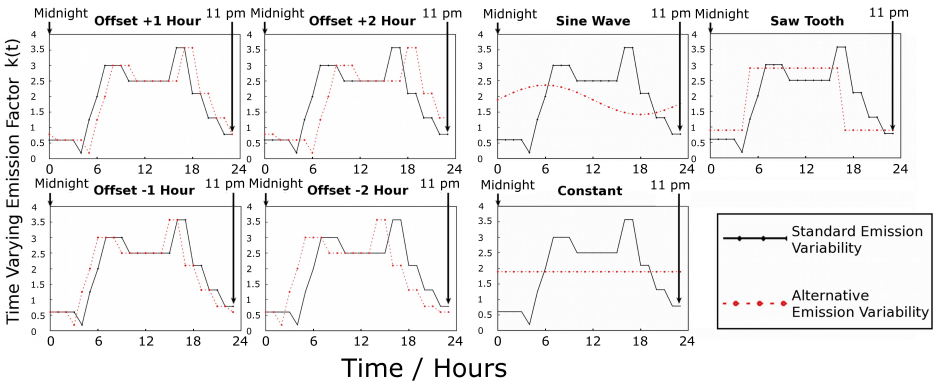

Fig. 1. The various different profiles of the temporal variability emission factor, $k(t)$, used in the analysis of the emission solution sensitivity to diurnal emission variability. The red dashed and the solid black lines indicate the alternative and standard emissions variabilities, respectively.

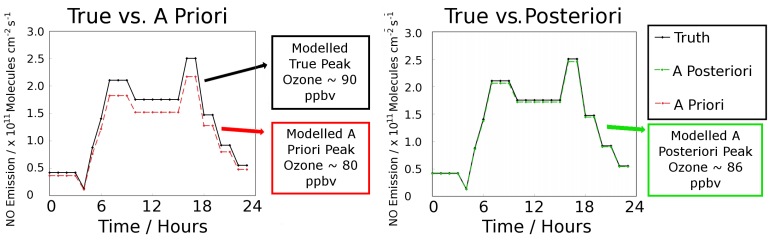

Fig. 2. A schematic showing how both the a priori and posteriori emissions relate to the true emissions of NO, and the modeled peak afternoon ozone that results from these emission variabilities. Note that the same emission variability is used for all of the anthropogenic chemical species emitted in the model. The a priori and posteriori emissions are scaled relative to the true emissions and these differences can be characterized as being due to different emission scaling factors (i.e., x_{NO}) for the a priori, posteriori and true emissions. The black solid, green dashed and red dashed lines show the truth, a posteriori, and a priori emissions, respectively.

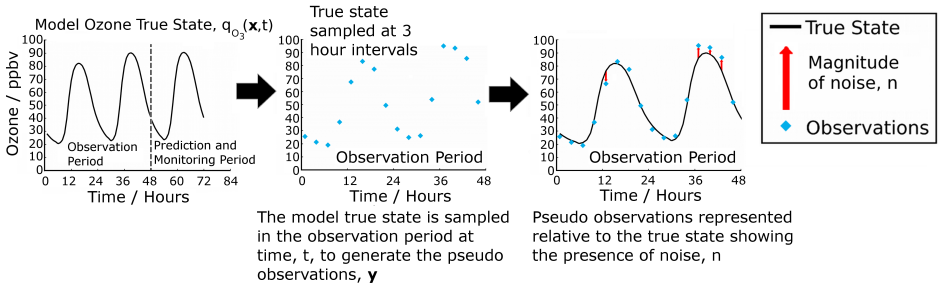

Fig. 3. A schematic showing an example of how the ozone true state ($q_{O_3}(\boldsymbol{x},t)$) is sampled and is combined with the noise to generate the pseudo observations $\boldsymbol{y}$. The black solid line indicates the truth and the blue diamonds indicate the pseudo observations.

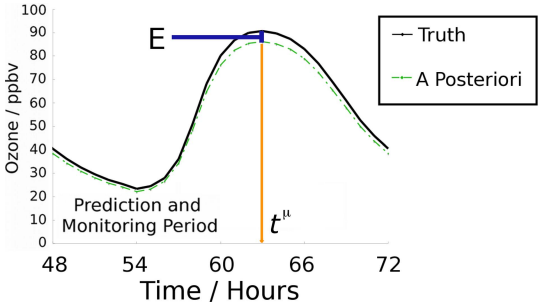

Fig. 4. True and posteriori ozone mixing ratios during the prediction and monitoring period resulting from the 4-D-variational data assimilation, i.e., $q_{O_3}(\hat{\boldsymbol{x}},\tau_3)$ and $q_{O_3}(\boldsymbol{x},\tau_3)$, respectively. The posteriori ozone prediction results directly from the optimization of emission scaling factors to yield $\hat{\boldsymbol{x}}$. E represents the posteriori ozone prediction error at 15:00 LT, i.e., $q_{O_3}(\hat{\boldsymbol{x}},t^{\mu}) - q_{O_3}(\boldsymbol{x},t^{\mu})$. The solid black and green dashed lines indicate the truth and a posteriori, respectively.

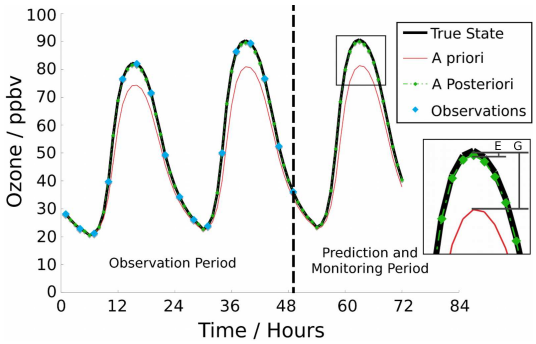

Fig. 5. A representation of the ozone prototype forecasting framework and the 4-D-variational data assimilation results for scenario OCN with $\beta = 0.1$. The observation period covers the first 48 h period of the assimilation during which time pseudo observations are made (at a frequency of every 3 h in this case) and are used within the assimilation. The observations are used to constrain the emissions of ozone precursors, which in turn allows the forecasting model to produce the posteriori ozone prediction. During the prediction and monitoring period the model true state now plays the monitoring role allowing comparisons to be made to the ozone forecast. The posteriori ozone prediction represents the forecast for ozone concentrations one day in the future. E represents the posteriori prediction model error and G represents the a priori and intitial guess prediction error. The black solid line, red solid line, green dahsed line, and blue diamonds represent the truth, a priori, a posteriori, and pseudo observations, respectively.

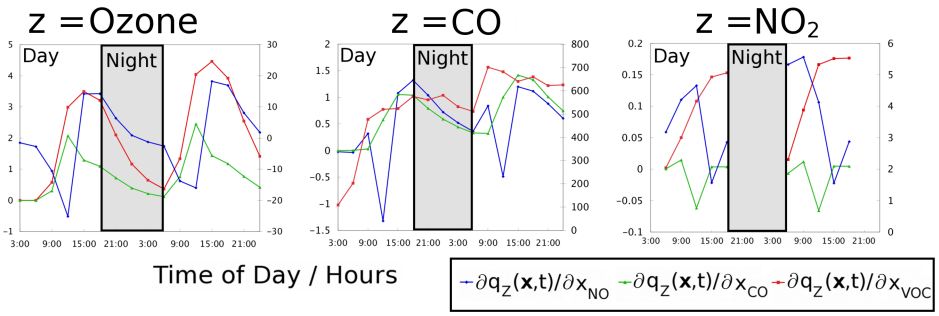

Fig. 6. These plots show the columns of the Jacobian matrix, **K**, that correspond to the perturbations of the three observed species in scenario OCN, i.e., ozone (left), CO (middle), and NO_2 (right). This Jacobian is for the NO emission scenario where $\boldsymbol{x}_{NO}$ is equal to 1.25. The shaded area represents observations made during the night. Note that since NO_2 observations made using visible remote sensing instruments that observations can only be made during the daytime, so there is no need to include a row in the Jacobian corresponding to night time NO_2 observations. The blue, green, and red solid lines represent $\partial q_Z(\boldsymbol{x},t)/\partial x_{NO}$, $\partial q_Z(\boldsymbol{x},t)/\partial x_{NO}$, and $q_Z(\boldsymbol{x},t)/\partial x_{NO}$, respectively.

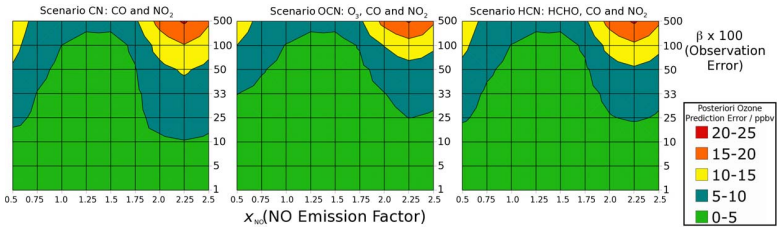

Fig. 7. Ozone posteriori prediction errors across the complete range of parameter space for x_{NO} (0.5–2.5) on the x-axis and β (0.1–5) along the y-axis with each panel presenting the results from the three observing scenarios CN, OCN and HCN. The coloured contours represent the posteriori prediction error in units of ppbv. The green and red colors indicate low and high levels of posteriori ozone prediction error, respectively.

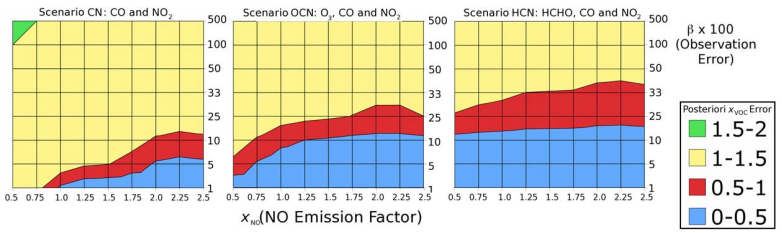

Fig. 8. x_{VOC} posteriori errors across the complete range of parameter space for x_{NO} (0.5–2.5) on the x-axis and β (0.1–5) along the y-axis with each panel presenting the results from the three observing scenarios A–C. The coloured contours represent the posteriori error. To allow comparison of the error in x_{VOC} to the true state we note that the true state is defined as $x_{VOC} = 6.5$. The light blue and green colors indicate low and high posteriori error on x_{VOC}, respectively.

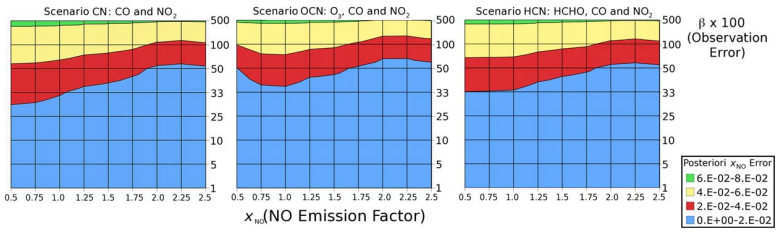

Fig. 9. x_{NO} posteriori errors across the complete range of parameter space for x_{NO} (0.5–2.5) on the x-axis and β (0.1–5) along the y-axis with each panel presenting the results from the three observing scenarios CN, OCN and HCN. The coloured contours represent the posteriori error. To allow comparison of the error in x_{NO} to the true state we note that the true state is defined as the x axis value. The light blue and green colors indicate low and high posteriori error on x_{NO}, respectively.

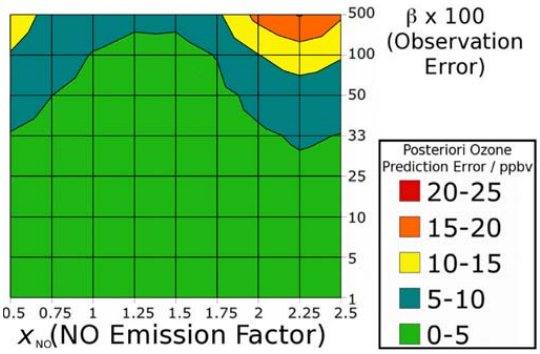

Fig. 10. The posteriori ozone prediction ozone prediction error for an observing scenario using HCHO, ozone, CO and NO_2 observations across the 9 x_{NO} emission scenarios (x_{NO} = 0.5–2.5) and full range of β. The green and red colors indicate low and high levels of posteriori ozone prediction error, respectively.

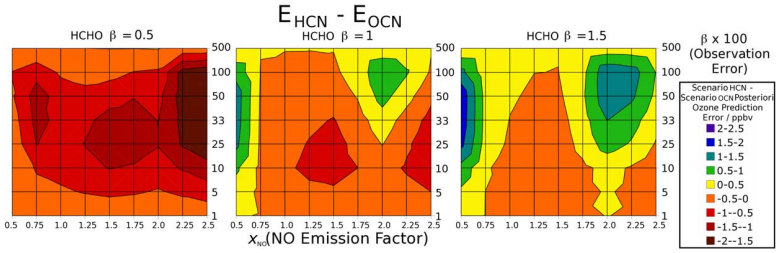

Fig. 11. The difference between the scenario HCN and OCN posteriori ozone prediction error for a range of assumed HCHO observing error scenarios. In all of the previous analyses and results β has been identical for all observed species, but in this sensitivity analysis we scale β for HCHO independently from the other observed species. From left to right HCHO observing errors are assumed to be 50 %, 100 %, and 150 % of the observing error for the other species. Thus the right hand panel indicates a scenario with HCHO observations to be of poorer quality relative to the other species, and represents the difference in ozone prediction error between the right and middle panels of Fig. 7, and the left panel indicates a rather optimistic case with assumed HCHO observation errors to be less than the other observed species errors. The brown and purple contour colors indicate the negative and positive differences between the scenario HCN and OCN posteriori ozone prediction error, respectively.

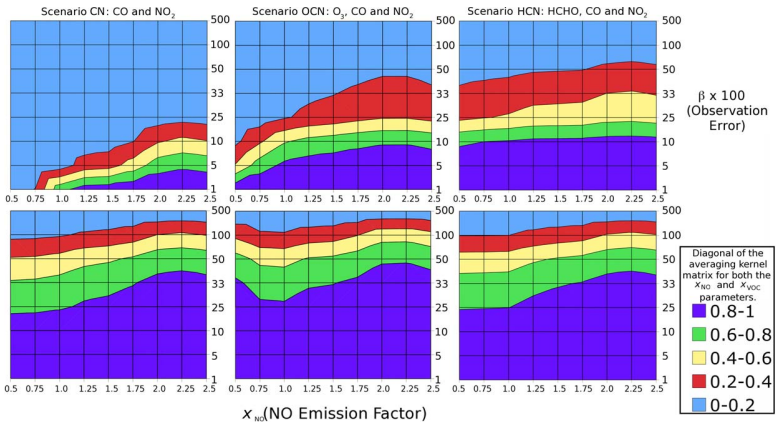

Fig. 12. The diagonal of the averaging kernel for x_{VOC} on the lower row and x_{NO} on the upper row. Each column represents a different observing scenario (CN, OCN, and HCN). The x-axis denotes the varying value of x_{NO} and the y-axis shows β (0.1–5). The contours represent the varying magnitude of the diagonal of the averaging kernel matrix from 0 to 1. The purple and light blue contour colors indicate high and low values of the diagonal of the averaging kernel matrix, respectively.

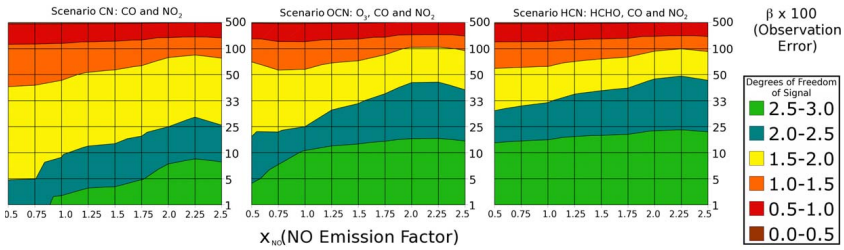

Fig. 13. The degrees of freedom of signal for the parameter retrieval from the thre observing scenarios (CN, OCN, and HCN). The x-axis indicates varying x_{NO} and the y-axis shows β (0.1–5.0). The contours represent the magnitude of the degrees of freedom of signal. The green and brown colors represent high and low values of the degrees of freedom of signal, respectively.

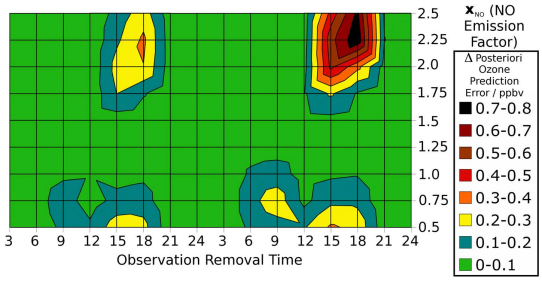

Fig. 14. The absolute increase in posteriori ozone prediction error between scenario OCN with $\beta = 0.25$ and the same scenario with observations removed form specific times over the course of 2 days (perturbed case), e.g., hour 15 on the second day indicates that no observations were included in the analytical model calculation of posteriori ozone prediction error for the perturbed case from 15:00 LT on the second day. The green and black colors indicate low and high values, respectively.

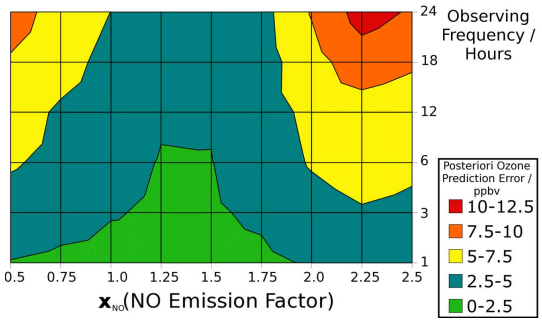

Fig. 15. The posteriori ozone prediction error for a variety of observation frequency scenarios ranging from an observing frequency of 1 h to once per day. These were calculated for scenario OCN with $\beta = 0.25$. The green and red colors indicate low and high levels of posteriori ozone prediction error, respectively.

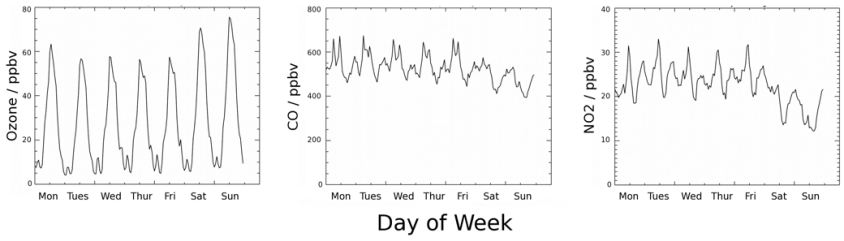

Fig. 16. Weekly averaged later summer and early fall ozone, CO, and NO_2 variability for the years 2005 through 2008. Data from the months July, August and September are included in the analysis. These results show persistent day to day variability for these trace gases related to the specific day of the week. The plots on the left, center, and right show the ozone, CO, and NO_2 mixing ratios, respectively.

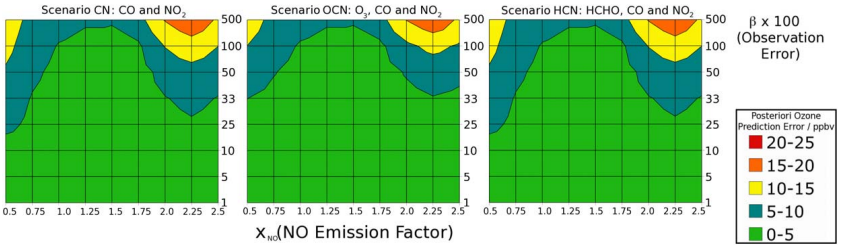

Fig. 17. The ozone posteriori ozone prediction error across the full range of 9 x_{NO} emission scenarios ($x_{NO} = 0.5$–2.5) and full range of β for the scenario where VOC emission uncertainties are represented by ethane emission uncertainties rather than ethene emission uncertainties. The green and red colors indicate low and high levels of posteriori ozone prediction error, respectively.