Evaluating the Impact of Storage-and-Release on Aircraft-based Mass-Balance Methodology Using a Regional Air Quality Model

Sepehr Fathi[1,2], Mark Gordon[3], Paul A. Makar[1], Ayodeji Akingunola[1], Andrea Darlington[1], John Liggio[1], Katherine Hayden[1], and Shao-Meng Li[4]

[1]Air Quality Research Division, Environment and Climate Change Canada, Toronto, Canada
[2]Physics and Astronomy, York University, Toronto, Canada
[3]Earth and Space Science, York University, Toronto, Canada
[4]College of Environmental Science and Engineering, Peking University, Beijing, China

Correspondence: Sepehr Fathi (sfathi@yorku.ca, sepehr.fathi@canada.ca) and Paul A. Makar (paul.makar@canada.ca)

Abstract. We investigate the potential for aircraft-based top-down emission rate retrieval over- and under-estimation using a regional chemical transport model, the Global Environmental Multiscale-Modeling Air-Quality and CHemistry (GEM-MACH). In our investigations we consider the application of the mass-balance approach in the Top-down Emission Rate Retrieval Algorithm (TERRA). Aircraft-based mass-balance retrieval methodologies such as TERRA require relatively constant meteorological conditions and source emission rates to reliably estimate emission rates from aircraft observations. Avoiding cases where meteorology and emission rates change significantly is one means of reducing emissions retrieval uncertainty, and quantitative metrics that may be used for retrieval accuracy estimation are therefore desirable. Using these metrics has the potential to greatly improve emission rate retrieval accuracy. Here, we investigate the impact of meteorological variability on mass-balance emission rate retrieval accuracy by using model simulated fields as a proxy for real world chemical and meteorological fields, in which virtual aircraft sampling of the GEM-MACH output was used for top-down mass balance estimates. We also explore the impact of upwind emissions from nearby sources on the accuracy of the retrieved emission rates. This approach allows the state of the atmosphere used for top-down estimates to be characterized in time and 3D space; the input meteorology and emissions are "known", and thus potential means for improving emission rate retrievals and determining the factors affecting retrieval accuracy may be investigated. We found that emissions retrieval accuracy is correlated with three key quantitative criteria, evaluated *a priori* from forecasts and/or from observations during the sampling period: (1) changes to the atmospheric stability (described as the change in gradient Richardson number), (2) variations in the direction of transport, as a result of plume vertical motion and in the presence of vertical wind shear, and (3) the combined effect of the upwind to downwind concentration ratio and the upwind to downwind concentration standard deviations. We show here that cases where these criteria indicate high temporal variability and/or high upwind emissions can result in "Storage-and-Release" events within the sampled region (control volume), which decrease emission rate retrieval accuracy. Storage-and-release events may contribute the bulk of mass-balance emission rate retrieval under- and over-estimates, ranging in the tests carried out here from -25% to 24% of the known (input) emissions, with a median of -2%. Our analysis also includes two cases with unsuitable meteorological conditions and/or significant upwind emissions to demonstrate conditions which may result in severe storage, which in turn cause emission rate underestimates by the mass-balance approach. We also introduce a flight strategy whereby the emission rate re-

1

25 trieval under- and over-estimates associated with storage-and-release are greatly reduced (to -14% to +5% respectively, relative to the magnitude of the known emissions). We recommend repeat flights over a given facility in order to measure and account for the factors associated with storage-and-release events – to estimate the temporal trends in the evolution of the system during the flight/sampling time, and partially correct for the effects of meteorological variability and upwind emissions. Our analysis allows the increased costs associated with longer flight times to be weighed against quantitatively reduced uncertainty in the

30 emission rate estimates.

1 Introduction

Aircraft-based measurements of air pollutants provide an effective top-down approach for estimating total integrated emissions of sources ranging from individual stacks to large industrial complexes and cities. These top-down estimates are commonly employed for the evaluation of bottom-up emissions reported to inventories and can provide insights on air quality and health

35 impacts of anthropogenic emissions. Top-down estimates can also be used to provide highly resolved emissions data for input into air quality models. The mass-balance technique and Gauss's divergence theorem can be employed to infer point and area source emission rates from aircraft measured data. Several studies have utilized aircraft-based mass balance emission rate retrievals in the past (Kalthoff et al., 2002; Mays et al., 2009; Turnbull et al., 2011; Karion et al., 2013, 2015; Cambaliza et al., 2014; Gordon et al., 2015; Nathan et al., 2015; Tadić et al., 2017; Conley et al., 2017; Ryoo et al., 2019). Flight patterns

40 from these studies include single-height transects, single-screen flights and box flights. Box flight refers to flight patterns with enclosed shapes (polygonal, cylindrical) and is accomplished by flying in closed paths (loops) at multiple altitudes around the emission source while making measurements of meteorological fields and pollutant concentrations. Kalthoff et al. (2002) used aircraft measured data to estimate CO and NOx emissions of the city of Augsburg in southern Germany, using data from two research aircrafts at several altitudes (polygonal box) around the boundaries of the city during the EVA project (Slemr et al.,

45 2002). Emission rate estimations were made using a mass-balance approach applied to aircraft measurements, and estimates were used in improving emissions input data for an air quality model (KAMM/DRAIS). Through model simulations, Panitz et al. (2002) analyzed the validity of the assumptions made by Kalthoff et al. (2002) and also studied the contributions of all relevant processes (advective fluxes, deposition, turbulent transports, and chemical transformations) to the mass budget of chemical species (e.g. CO and NOx). They concluded that, in an aircraft-based mass-balance approach, 85% and 95% of source

50 emissions can be determined from the advective fluxes. Gordon et al. (2015) conducted top-down aircraft-based emission rate retrievals by flying polygonal (e.g. rectangular) box flights around individual oil sands facilities in Alberta, Canada. This study also involved development of a methodology for extrapolation of aircraft measured data below the lowest flight level based on various assumptions and the analysis of the associated uncertainties. Gordon et al. (2015) estimated uncertainty in aircraft-based estimates as 2%-30%, with the source of this uncertainty attributed mainly to extrapolation of aircraft measurements

55 below the lowest flight track. Conley et al. (2017) and Ryoo et al. (2019) made aircraft-based emission estimations by flying cylindrical box flights around different point/area emission sources. The study by Conley et al. (2017) also involved evaluation

and optimization of the sampling strategy (e.g. optimized circling radius) in order to minimize the uncertainties due to data extrapolation below the lowest flight level, through tracer controlled release and large eddy simulations (LES).

These and other similar studies, employed either mass transfer (e.g. single-height transects) or a mass-balance approach
60 (box flights) to infer point and area source emission rates from aircraft measured data. Aircraft-based estimates are generally (and where possible) compared to bottom-up inventory reported emission data (e.g., Gordon et al., 2015; Li et al., 2017; Baray et al., 2018; Liggio et al., 2019; Cheng et al., 2020), and may also be evaluated using Continuous Emissions Monitoring System (CEMS) data (when these data are available). These comparisons and their conclusions are dependent in part on the magnitude of the uncertainty levels associated with the estimates. The algorithms for emission rate retrieval require steady-
65 state meteorological and emissions data. Cases where these conditions are not met have been avoided through the application of *a priori* criteria in flight planning, and post-flight retrieval decision making. For example, pre-flight decision criteria such as relatively invariant forecast wind speed, wind direction, and planetary boundary layer height, were used to determine "fly/no fly" decisions in past studies (e.g., Li et al., 2017; Liggio et al., 2019). We refine these concepts here to create quantitative metrics for assessing retrieval success which may also be used in forecasting flight conditions or during flights. We also
70 investigate the potential causes of emissions retrieval uncertainty, using a variety of cases, including unsuitable cases which would usually not be used for retrievals, as tests of our criteria. Therefore, it is important to further refine the methodology used to predict uncertainties in retrievals, as well as suggest revised methodologies which may reduce emissions retrieval uncertainty. Uncertainties in aircraft-based top-down retrievals may arise from various sources (e.g. measurements, data interpolation and extrapolation), most of which have been studied in the past (e.g., Cambaliza et al., 2014; Gordon et al., 2015; Angevine
75 et al., 2020). Here we focus on investigating the underlying assumption of time-invariant meteorological conditions (during observation time), which is shared among most (if not all) top-down retrieval methodologies, in particular the impact of localized variations in meteorology (e.g. change in atmospheric stability or wind direction) on the application of the mass-balance technique in box flight emission rate retrievals.

Under time-invariant and spatially uniform conditions, a mass-balance approach based on the Gauss's divergence theorem
80 equates the rate of emissions from sources within a control volume (box flight) to the net mass flux through the boundaries (box walls) by assuming the net rate of accumulation of mass within the volume to be negligible. When conditions (e.g. meteorology, emissions) deviate from steady state and/or localized inhomogeneity is developed in tracer concentration and meteorological fields, the rate of mass accumulation/release within the control volume may become significant enough to affect the accuracy of the estimates based on the mass balance approach. Herein, we refer to these circumstances as "**Storage-and-Release**" events.

85 Numerical dispersion modelling has been proven useful in assessing the uncertainties in top-down retrievals in the past (e.g., Panitz et al., 2002) and more recently (e.g., Karion et al., 2019; Angevine et al., 2020). Here, the numerical modelling system (a regional chemical transport model) we use is Environment and Climate Change Canada's (ECCC) air-quality model, Global Environmental Multiscale-Modeling Air-Quality and CHemistry (GEM-MACH). Through the use of GEM-MACH model 4D fields (known meteorology and input emissions), we have eliminated all other sources of uncertainty in the input information
90 used to retrieve emissions (e.g. measurement error, interpolation) and have extensively explored the impact of storage-and-release events on aircraft-based top-down emission rate retrievals. We investigate the potential magnitude and the causes

Atmospheric
Chemistry
and Physics
Discussions

of emissions over- and under-estimates, using GEM-MACH model's input emissions and output concentrations as a proxy for aircraft observations. This approach allows us to more precisely examine the root causes of potential emissions retrieval over/underestimates, through making use of ancillary information provided by the model, which may not be available from past

95 aircraft studies. Throughout this paper, as an example method, we refer to (and adopt schemes from) the aircraft-based retrieval methodology developed by ECCC and described in Gordon et al. (2015), the Top-down Emission Rate Retrieval Algorithm (TERRA). The concepts explored here are not specific to TERRA and are applicable to all top-down retrieval methodologies that use the mass-balance approach.

2 Methods

100 During an aircraft campaign conducted in August and September of 2013 as part of the Joint Canada Alberta implementation plan for Oil Sands Monitoring (JOSM), the National Research Council of Canada Flight Research Laboratory (NRC-FRL) Convair 580 research aircraft was used to make air-quality measurements over the Athabasca oil sands region (Alberta, Canada). Air quality forecasts created using ECCC's air-quality model GEM-MACH, provided advice on flight planning for the JOSM 2013 campaign (Makar et al., 2016). For this purpose, and for the post-study simulations conducted here, the GEM-MACH

105 system was configured to create high-resolution nested air-quality forecasts with a fine resolution domain (2.5km) over the Canadian provinces of Alberta and Saskatchewan, including the Athabasca oil sands region (Figure 1a). From 10 August to 10 September 2013 a total of 22 flights were conducted around and downwind of oil sands facilities; 13 of which were emission estimation (box) flights where the aircraft was flown around individual facilities at multiple altitudes, approximating a rectangular/polygonal shaped box (control volume), while making meteorological and air-quality measurements (JOSM, 2011). We

110 chose our model input test cases from the same times and locations as JOSM 2013 box flights, to sample conditions (meteorology and emissions) similar to the range existing within the real atmosphere which might be considered for retrievals. Nine box cases over three oil sands facilities with significant SO_2 emissions (Syncrude, CNRL and Suncor; see Fig. 1b) were chosen because: (1) SO_2 is emitted from stack sources, but has low background levels (Gordon et al., 2015), reducing uncertainties in emission rate retrievals based on the mass-balance approach and (2) hourly SO_2 emissions and stack parameter data are

115 available from the CEMS reports (Zhang et al., 2018) thereby ensuring the GEM-MACH inputs are realistic: CEMS-derived input emissions in GEM-MACH resulted in improved model plume rise and transport representation (model-simulated plume heights were evaluated in Akingunola et al., 2018). Output data from GEM-MACH model simulations for the period of the nine cases were analyzed and mass-balance estimates of model SO_2 input emissions were made. Model configuration and set-up, and simulations scenarios for this work are described in section 2.1. Section 2.2 describes the methodologies that were

120 developed here for direct and virtual aircraft-based retrievals of model input emissions using the GEM-MACH model output fields, as well as suggested improvements for flight planning.
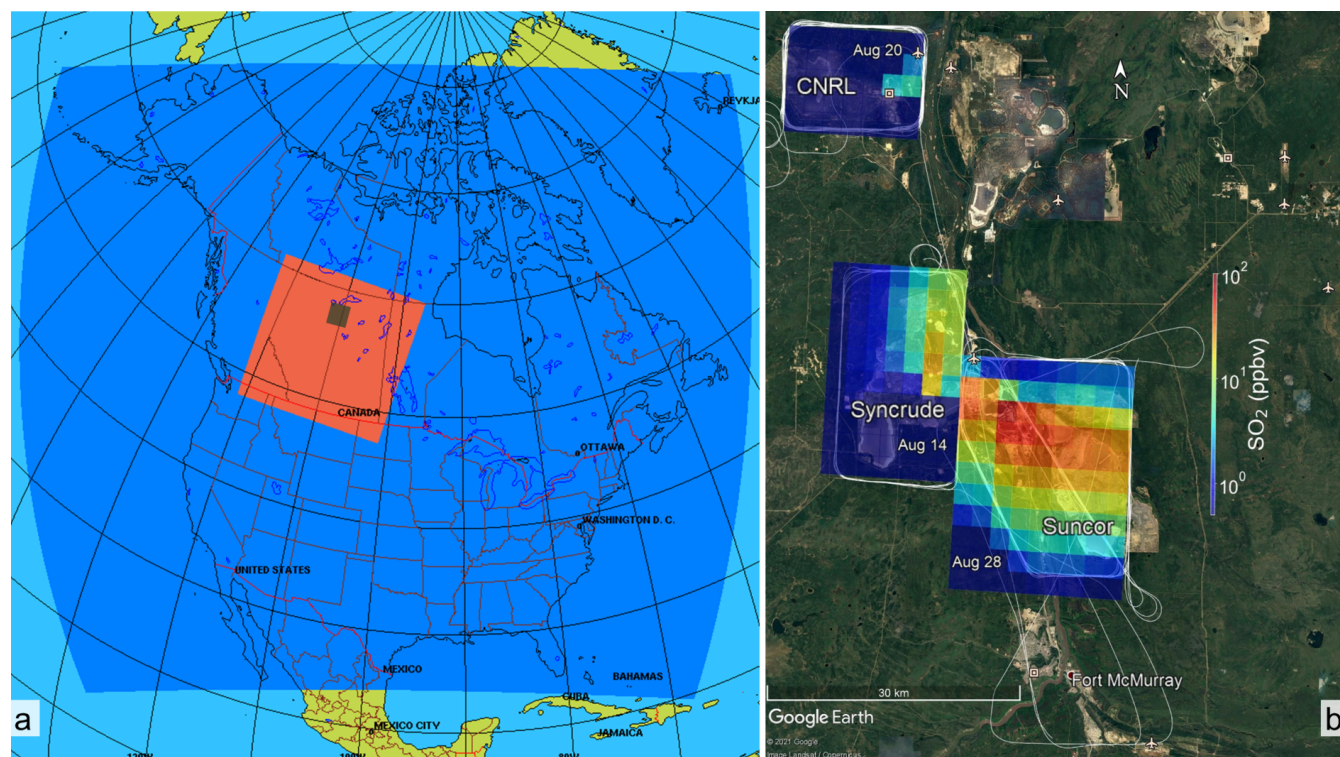
Figure 1. (a) GEM-MACH model horizontal grid structure with nested domains: the dark blue region shows the parent domain over North America at 10km horizontal grid resolution. Output from GEM-MACH 10km simulations provided initial and boundary conditions for the nested high-resolution GEM-MACH 2.5km domain (red). The high-resolution domain was configured over the Canadian provinces of Alberta and Saskatchewan including the Alberta Athabasca oil sands region, shown as a dark green shaded area. (b) © Google Earth map of the region with three oil sands facilities labeled. JOSM 2013 flight tracks on August 14, 20 and 28 are drawn with white curves. Model 4D fields were extracted for nine JOSM 2013 box flight cases over these three facilities, examples are shown as overlay SO$_2$ concentration maps.

2.1 Model Description and Simulations

ECCC's air quality model, GEM-MACH, is a fully coupled online air-quality and chemical transport model. The GEM-MACH modelling system resides within the Global Environmental Multiscale (GEM) numerical weather prediction model (Côté et al., 1998b,a; Girard et al., 2014) and includes an atmospheric chemistry module (Moran et al., 2010) with gas and particle process representation. The details of the GEM-MACH configuration used here appear in Table A1, Appendix A.

For this work, a series of retrospective high-resolution nested air-quality simulations were made for the period of the JOSM summer 2013 aircraft intensive campaign (JOSM, 2011). The GEM-MACH model grid configuration consisted of a coarse parent domain with a horizontal resolution of 10km covering North America, and a nested 2.5km high resolution domain including Canadian provinces of Alberta and Saskatchewan (see Fig. 1a). To prevent the model high-resolution meteorology from drifting chaotically from observed meteorology, the GEM-MACH 10km simulations were updated every 24 hours at 6

UT (local mid-night). Most of the JOSM 2013 cases examined here take place during the late mornings and early afternoons in local time (LT); by choosing to update the driving meteorology in the model at 6 UT sudden shifts in the model meteorological fields were avoided during the local daytime, and more than sufficient spin-up time (6h time-frame) was provided for model
135   meteorological fields (in particular the cloud microphysics fields) to reach equilibrium before local noon.

Flux box data, including meteorological fields and SO_2 concentrations, were extracted from the model output database for the period of nine JOSM 2013 flights over three different facilities (Fig. 1b). The mass-balance approach was employed to retrieve model input emissions using the extracted flux box data. Earlier work (Fathi, 2017) suggests that the use of regional air-quality models as proxies for observations must be taken with care, with three main considerations in their use for this
140   purpose:

    1. Air-quality models usually require the use of some form of "mass conservation" correction for their advection algorithms (de Grandpré et al., 2016), which can otherwise create spurious negative concentrations. However, mass conservation algorithms which conserve mass over the entire model domain, but do not conserve mass locally, effectively include spurious mass transport – in turn affecting emissions retrievals using model output. A local mass correction algorithm (3
145   shell ILMC approach, Sørensen et al., 2013) was therefore used in the GEM-MACH simulations carried out here. Model resolution also impacts mass conservation and needs to be sufficiently high that the sources may be resolved.

    2. Regional air-quality models provide instantaneous output at every grid cell on a fixed (2 min) time-step – however, if these are interpolated to a finer time resolution (e.g. along an aircraft flight track), errors in interpolation may affect the use of the model output in retrieval algorithms. Values of model variables at 3-D grid-points, in the vertical columns
150   making up the emissions "box", were therefore used in the analysis which follows to avoid these (temporal and spatial) interpolation errors in the use of model output as a proxy for observations.

    3. In aircraft-based sampling (both in real world aircraft measurements and in virtual sampling with model predicted data), temporal lag between data points collected along flight tracks may represent an additional source of error. This may introduce additional errors if the "static atmosphere" assumption is not met. Here, instantaneous model output at each
155   time-step (2min) was used for mass-balance estimates, to avoid the potential for interpolation from the model resolution (2.5km) and time-step to influence the retrievals generated here.

For this purpose, GEM-MACH model four-dimensional data (3D space, 1D time) were extracted at every 2 min model chemistry time-step for the flux boxes over individual oil sands facilities for our nine studied cases. Aircraft-based mass-balance retrievals were simulated by employing the divergence theorem in analyzing the model extracted data. The estimated emis-
160   sion rate time-series and flight-time (1.5-2.5 hours) averages were then compared to model input emissions (MIE), to evaluate top-down mass balance retrievals.

2.2   Divergence Theorem, the Mass-balance technique and Emission Rate Retrievals

During an actual aircraft campaign, measurements are made along a flight trajectory on the lateral walls of a flux box (control volume); measured data on a single 2D screen (closed surface) comprised of box lateral walls is all that is available for emission

165 rate retrievals based on mass flux and mass balance calculations. Here we use instantaneous model values along lateral box walls as our 2D screen ("TR" approach, section 2.2.2).

The temporal evolution of the system can be studied by increasing the number of flights around the same facility. Here, this approach is explored with GEM-MACH simulated fields by analyzing multiple 2D screens at consecutive model time-steps (3D fields) and generating mass balance estimates ("TR*" approach, section 2.2.3).

170 A truly comprehensive analysis of the tracer mass budget within the flux box is only possible with 3D-volumetric data over the entire extent of the box and over time (4D fields). However, collection of 4D (volumetric time-series) data is not practically feasible during aircraft measurements as the aircraft cannot sample the entire volume of the box even with increased flights during a time window of only a few (e.g. 2-3) hours. One advantage of the use of a model such as GEM-MACH as a proxy for observations is that 4D forecast fields can be generated, to provide complementary information for mass balance retrieval

175 investigation. Here we use GEM-MACH model 4D chemical and meteorological fields to conduct a comprehensive analysis of the transport of the emitted mass from sources within the control volume and to further investigate uncertainties associated with storage-and-release events in top-down mass balance retrievals ("DR" approach, section 2.2.1).

The Gauss's divergence theorem states that the total integrated divergence of a vector field (e.g. fluid flow) F in a region of space (control volume) V equals to the total flux of F through the enclosing boundary (surface) S,

180
$$\iiint_V \nabla \cdot F dV = \iint_S F \cdot dS \tag{1}$$

A mass-balance approach (conservation of mass) based on the divergence theorem can be employed to estimate chemical (pollutant) emission rates by sources (e.g. stack emissions) within a control volume V where atmospheric dispersion and transport takes place. In order for mass to be conserved, the inflow of mass of species C into V ($E_{C,in}$) plus the rate of generation/emission of mass (E_C) within V, must equal the rate of mass removal from the control volume ($E_{C,out}$) by the

185 outgoing flux and processes such as surface deposition and the rate at which mass is accumulated/stored within the volume of the box ($E_{C,S}$); written in a format linked to the divergence theorem (Eq. 1),

$$E_C = (E_{C,out} - E_{C,in}) + E_{C,S} \tag{2}$$

Note that the left hand side and the terms in brackets on the right hand side of Eq. (2), comprise the implementation of the divergence theorem in the mass-balance equation. In steady state conditions, i.e., under the assumption of time-invariant

190 meteorology and input emissions, the net accumulation (rate) of mass can be assumed zero and therefore the divergence theorem can be utilized to equate E_C to the net flux through the boundaries of the box ($E_{C,out} - E_{C,in}$). The divergence theorem is commonly utilized in aircraft-based studies with flight paths approximating flux boxes of polygonal (e.g., Gordon et al., 2015; Liggio et al., 2019) and/or cylindrical (e.g., Ryoo et al., 2019) shapes, where aircraft observations along flight tracks are projected (and interpolated) onto a 2D screen comprised of box lateral walls surrounding the emitting facility. A

195 mass-balance technique is then employed to estimate point/area source (e.g. stack, tailing ponds) emission rates by calculating the net mass flux through the screen (e.g. TERRA).

We start by a comprehensive analysis of GEM-MACH 4D data for direct retrieval of model input emissions, where we explore all the processes (including storage-and-release) contributing to change in tracer mass within the flux box and over time (Sect. 2.2.1). We then simulate aircraft-based retrievals by limiting our analysis to 2D data on box lateral walls (Sect.

200 2.2.2). Finally, we combine the information from time-consecutive screens (3D) to explore the temporal trends in chemical and meteorological fields and introduce potential means for improving top-down estimates (Sect. 2.2.3).

2.2.1 Direct Retrieval (DR); Using GEM-MACH 4D Fields

Here we apply the divergence theorem to the GEM-MACH model four-dimensional (volumetric time-series) output data to estimate model input emissions. Data from volumes corresponding to the nine JOSM 2013 cases were extracted from the

205 GEM-MACH model output at every 2min model time-step (no interpolation), creating nine sets of 4D metadata with variables (including wind speed and direction, air density, temperature, species mixing ratio, ground surface deposition). Model-extracted 4D flux box metadata, spanning model vertical levels from the surface up to $\sim$ 2000 m above ground (average JOSM 2013 box flight top height), were analyzed using the mass balance approach to estimate the model input emissions. We note that this approach incorporates information not available for aircraft observation-based retrievals (i.e. volumetric time-series data), but

210 may be used to analyze the importance of storage-release, and as a standard against which the other methods which follow may be evaluated and compared.

The total mass of compound C within the control volume V ($B_C(t,V)$) is calculated at each model time-step (Δt) by integrating over the entire volume of the box. The rate of change in mass can then be calculated by taking the time derivative of $B_C(t,V)$ and following the chain rule,

$$\frac{\partial}{\partial t}B_C(t,V) = \frac{\partial}{\partial t}\left[M_R \iiint \chi_C(t,V)\rho_{air}(t,V)dV\right] = M_R \iiint \left[\frac{\partial \chi_C}{\partial t}\rho_{air} + \chi_C \frac{\partial \rho_{air}}{\partial t}\right]dV \tag{3}$$

where M_R is the ratio of the compound molar mass (e.g. 64.07 g/mol for SO_2) to the molar mass of air (28.97 g/mol), $\chi_C(t,V)$ and $\rho_{air}(t,V)$ are compound mixing ratio and air density at each model grid point and time-step, respectively; with the 3D integrand volume element $dV = dxdydz$. The first integral on the right hand side of Eq. (3) represents the rate of change in compound mass within V due to variations in compound mixing ratio $\chi_C(t,V)$ independent of the changing air density,

$$E_{C,S}(t) = M_R \iiint \frac{\partial \chi_C}{\partial t}\rho_{air}dV \tag{4}$$

The second integral represents the apparent change in compound mass (which is calculated as a product of air density and compound mixing ratio) due to temporal variations in air density $\rho_{air}(t,V)$, caused by meteorological processes, and does not represent an actual change in compound mass within the control volume,

$$E_{C,M}(t) = -M_R \iiint \chi_C \frac{\partial \rho_{air}}{\partial t}dV \tag{5}$$

225 The minus sign signifies that this is an apparent change in compound mass, which is accounted for in the mass balance equation. In applying the mass balance technique, we set the net rate of change of Eq. (3) as equal to the sum of all flux and chemical terms changing the compound mass (emissions, horizontal and vertical advection, vertical diffusion and chemical production).

Atmospheric
Chemistry
and Physics
Discussions

Solving for $E_{C,S}$, the processes contributing to the net change in compound mass within the control volume V can be described with the following expression:

$$E_{C,S} = E_C - E_{C,H} - E_{C,V} - E_{C,VD} + E_{C,M} + E_{C,X} \tag{6}$$

where E_C represents the emission rate of compound C from sources within the control volume. $E_{C,H}$ and $E_{C,V}$ represent the net horizontal and vertical fluxes through the lateral and top walls of the flux box, respectively (derivations appear in Appendix B, Eqs. B1-B3). $E_{C,VD}$ represents the surface deposition rate, which is available as an output variable from GEM-MACH. Change in compound mixing ratio due to variations in air density can be accounted for by calculating $E_{C,M}$ from Eq. (5). $E_{C,X}$ represents the rate of change (i.e. the formation rate of C) due to chemical processes, and can be assumed negligible for a tracer compound such as SO_2 relative to the emission rates for these facilities since it has a lifetime on the order of days. In the context of emissions estimation using observed concentrations and winds along an aircraft trajectory, Eq. (4) can't be estimated directly, since it requires a volume integral over the box enclosing the emission source; data that are not available along the flight path. However, when a regional air-quality model is used as a proxy for observations, and/or a strategy of repeat flights are used (see Sect. 2.2.3), $E_{C,M}$ and $E_{C,S}$ can be estimated directly. When model values are used these terms may be generated from equations (4) and (5). For example $E_{C,S}$ is estimated using the mixing ratio rate of change ($\partial\chi_C/\partial t$) multiplied by the air density (ρ_{air}) at each model grid point and integrating over the volume of the emissions enclosure box, with a similar approach used for $E_{C,M}$. $E_{C,S}$ represents the net rate of increase/decrease of tracer mass within the box; a residual rate of change associated with the difference given in Eq. (6). We therefore refer to it as a "storage" term for tracer mass. Observation-based emission rate estimate studies (e.g. TERRA) assume $E_{C,S} = 0$, while in reality a non-zero $E_{C,S}$ term is incorporated in the E_C term; here we examine the potential magnitude of $E_{C,S}$ using the GEM-MACH model's concentration predictions as a proxy for aircraft observations, and use the combination of mass-balance approach and GEM-MACH to fully investigate the potential impacts of storage of tracer mass on emission rates retrieval over/underestimates. We also use this combination of retrieval algorithm and regional air-quality model to generate meteorological forecast model variables which predict conditions in which the storage term may be of concern, reducing its influence on future emissions analysis (Sect. 4), and to demonstrate flight procedures which may be adopted *a priori* to reduce its influence on emissions retrievals (Sect. 2.2.3). Rearranging Eq. (6) and neglecting $E_{C,X}$ as being insignificant over the short distances (5-20km) between source and emissions box wall, compound emission rates can be estimated based on the other terms as,

$$E_C = E_{C,H} + E_{C,V} + E_{C,VD} - E_{C,M} + E_{C,S} \tag{7}$$

Eq. (7) equates the total integrated emission rate, originating from sources (point and/or area) within the box (control volume), to the flux leaving the box through the top and lateral walls $E_{C,V}$ and $E_{C,H}$, the surface deposition rate $E_{C,VD}$, and the volumetric storage rate $E_{C,S}$, while accounting for the effects of changing air density ($E_{C,M}$). Note that in our GEM-MACH driven analysis, $E_{C,H}$, $E_{C,V}$ and $E_{C,VD}$ are derived from GEM-MACH 2D fields (screens, box-top, ground surface deposition) at each time-step, while $E_{C,M}$ and $E_{C,S}$ are temporal gradient terms estimated from model 4D fields (volumetric time-series) and employing a second order central Finite Difference (FD) scheme (i.e. $\partial\chi_C/\partial t \cong (\chi_C(t+\Delta t) - \chi_C(t-\Delta t))/2\Delta t$); the

choice of FD scheme was made based on optimized performance and accuracy in our computations and it was determined that higher order schemes did not provide further advantages in terms of accuracy. A similar set of equations, lacking sources, sinks and a separate storage term, can be written for the air mass changes within the emissions box (see Appendix B), which lead to the mass balance expression for air

$$E_{air,M} = -E_{air,H} - E_{air,V} \tag{8}$$

Eqs. (4), (5) and (7) (with other terms in Appendix B) were utilized in analyzing GEM-MACH model 4D flux box data for the nine cases, and DR method estimates were compared to the model input emissions (MIE). Results are discussed in Sect. 3.1.

2.2.2 TERRA Retrieval (TR) approximation; Using GEM-MACH 2D Fields

Aircraft-based retrieval methods such as TERRA rely on measured data along the lateral walls of the box. In observation-based retrievals, the horizontal flux $E_{C,H}$ can be calculated (directly) from aircraft collected data and the other terms need to be estimated based on several assumptions about the atmospheric conditions and source properties (Gordon et al., 2015). Here, we explore the use of regional (GEM-MACH) model output in a similar fashion as aircraft observations, restricting our analysis to model output sampled along the lateral walls of a box enclosing an emitting facility within the model domain (GEM-MACH 2D fields). Figure 2a shows model simulated SO_2 concentrations (normalized to the maximum concentration) on the lateral walls of the flux box surrounding the CNRL facility (case 4). Flight tracks on 20 August 2013 (JOSM) are also plotted as dark dots around the facility, for reference. For emissions retrieval applications using aircraft observations, the sampled observations must be interpolated (e.g. Kriging) between successive circuits around the box and extrapolated to the ground surface in order for emissions to be estimated. Surface deposition rate $E_{C,VD}$ is also estimated using the extrapolated species mixing ratio data at the ground level. Gordon et al. (2015) estimated the uncertainty due to extrapolation of aircraft data to the ground surface as 0.4 and 12% in TERRA estimates for SO_2 plumes. Here we avoid the uncertainties arising from such assumptions by extracting model data along box lateral walls from the ground surface up to the box top, and by using GEM-MACH model surface deposition ($E_{C,VD}$) values in our calculations. For this portion of our analysis, in which we limit the inputs to mass-balance estimations to the information available on the box lateral walls in analogy to aircraft observations, instantaneous 2D screens of relevant fields (e.g. SO_2 concentrations) covering the entire area comprised of box lateral walls (as shown in Fig. 2) were extracted/generated from the GEM-MACH model output at every model time-step. Such screens were also generated from GEM-MACH output for air flux, based on the extracted air density ρ_{air}, the calculated normal wind $U_\perp$ (positive outwards), and SO_2 mass flux, based on concentration and air flux screens (Figures 2b, c and d). The integrated horizontal flux through the screen $E_{C,H}$ was then calculated using Eq. (B1). In observation-based applications, TERRA estimates the changes in mass due to variations in air density over flight time, $E_{air,M}$ and $E_{C,M}$, based on observed meteorological trends in air density ($\frac{\partial}{\partial t}\rho_{air}$) from nearby weather stations (towers) and average species mixing ratios (χ_C) around the box. Here, we substituted the weather data with a single column (GEM-MACH predicted) air density temporal rate of change $\frac{\partial}{\partial t}\rho_{air}(t, z)$ at point (x_o, y_o) located at the horizontal center of the box for each of the nine studied cases. Here, single column $\frac{\partial}{\partial t}\rho_{air}(t, x_o, y_o, z)$ is assumed to represent the spatial average rate of change in air density over the region of interest through the assumption of horizontal
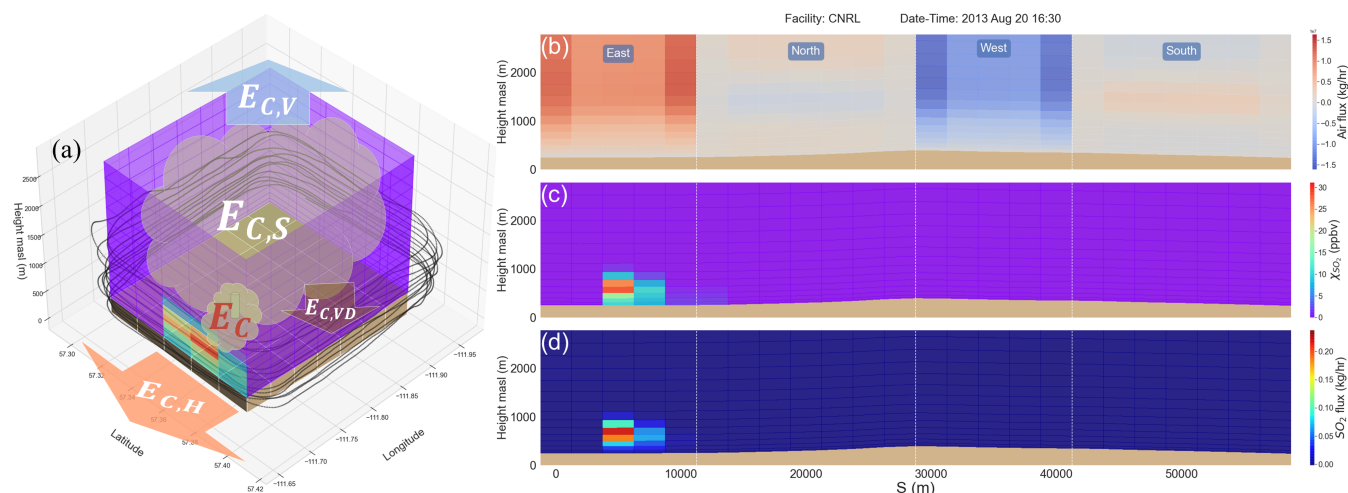
Figure 2. SO_2 mixing ratio screen around CNRL extracted from the GEM-MACH model output at the start of case 4 on 20 August 2013 at 16:30 UT (10:30 LT). (a) mixing ratio data (2D) on the lateral walls of the flux box, along with the flight tracks plotted as dark dots. The additional graphics (arrows and plumes) on (a) represent processes contributing to change in tracer mass within the box volume: source emission rate (E_C), horizontal flux ($E_{C,H}$), vertical flux ($E_{C,V}$), surface deposition ($E_{C,VD}$) and storage rate ($E_{C,S}$). The extracted fields around the box are unwrapped onto a 2D screen comprised of box lateral walls (East, North, West and South) and shown in (b) the normal air flux (positive outwards), (c) SO_2 mixing ratio and (d) SO_2 mass flux through the screen. Note that the net flux on box corners is the vector sum of the fluxes in west-east (U) and south-north (V) directions.

homogeneity, as in observational TERRA applications. Our comparisons between the GEM-MACH box average and the single
295   column rate of change in air density indicate < 5% difference between the two trends for the time periods of our studied cases.
This gives

$$E_{air,M}(t) = A \int \frac{\partial}{\partial t} \rho_{air}(t, x_o, y_o, z) \, dz \tag{9}$$

$$E_{C,M}(t) = -A \, M_R \int \overline{\chi_C}(t, z) \frac{\partial}{\partial t} \rho_{air}(t, x_o, y_o, z) \, dz \tag{10}$$

where A is the base area of the box. $\overline{\chi_C}(t, z)$ is the average GEM-MACH predicted species mixing ratio around the box along
300   the flight path (s) at height z, which is the same as in aircraft-based applications of TERRA.

Furthermore, due to lack of measurements at the box top, the vertical flux $E_{C,V}$ in TERRA in observation-based applications
is estimated in two steps: (a) the vertical air mass flux ($E_{C,V}$) is estimated based on the other terms by rearranging the mass
balance expression for air (Eq. 8), and (b) the species mixing ratio (e.g. SO_2) at the box top is assumed to be uniform and
equal to the average (along flight path s) screen top mixing ratio ($\overline{\chi_{C,top}}$) (Gordon et al., 2015). Here, the same procedure was
305   employed in deriving the vertical flux through box top using GEM-MACH output fields,

$$E_{C,V}(t) = A \, M_R \, \overline{\chi_{C,top}}(t) \, E_{air,V}(t) \tag{11}$$

As there are no volumetric data available in aircraft measurement applications of TERRA, the storage rate within the box cannot be inferred directly. Emission rate estimations in aircraft-based methods such as TERRA are made with the underlying assumption of steady state conditions and therefore $E_{C,S}$ is assumed to be negligible, effectively incorporating it into E_C. Thus, for aircraft-observation TERRA applications, mimicked here, Eq. (7) is reduced to,

$$E_C = E_{C,H} + E_{C,V} + E_{C,VD} - E_{C,M} \qquad (12)$$

Estimations of emission rates were made for each model time-step (using instantaneous 2D screens) during the time periods of the nine studied JOSM cases by using GEM-MACH output to approximate the input data for TERRA retrievals (TR) as described above. Note that the TR method, as described in this section, makes use of GEM-MACH 2D data at each time t to generate an estimate (one estimate at each time-step t during the entire sampling period) while making approximate estimations of the temporal gradient terms (Eqs. 9 and 10) at time t by following procedures similar to aircraft-based applications of TERRA. Flight-time averaged estimates are then compared to the flight-time averaged model input emissions (MIE), Sect. 3.2.

2.2.3 Revised TERRA Retrieval (TR*); Using GEM-MACH 3D Fields

Utilizing time-consecutive instantaneous screens around an emission source, the temporal evolution of the system and its impact on mass-balance estimates can be studied in more detail, as discussed at the beginning of this section (Sect. 2.2). By using air density screens at every 2 min model time-step (2D space, 1D time), as opposed to using a single column (vertical) profile in the TR method, estimations of $E_{air,M}$ and $E_{C,M}$ can be improved (by 1-3%, determined through comparisons with GEM-MACH box averaged temporal trends):

$$E_{air,M}^*(t) = A \int \frac{\partial \overline{\rho_{air}}(t,z)}{\partial t} \, dz \qquad (13)$$

$$E_{C,M}^*(t) = -A \, M_R \int \overline{\chi_C}(t,z) \frac{\partial \overline{\rho_{air}}(t,z)}{\partial t} \, dz \qquad (14)$$

where $\overline{\rho_{air}}(t,z)$ and $\overline{\chi_C}(t,z)$ are the average GEM-MACH predicted air density and species mixing ratio around the box (along s) at height z. A similar procedure (using time-consecutive screens) can be used for estimating the storage term $E_{C,S}$. The main difference between DR (Eq. 7) and TR (Eq. 12) retrievals is that the former explicitly accounts for and calculates the storage term $E_{C,S}$. As discussed later in Sect. 3.3, the results from the Direct Retrieval (DR) method demonstrate the potentially significant contribution of the storage term $E_{C,S}$ to the total integrated estimated emission rates E_C (Eq. 7). Ignoring this term, as in Eq. (12), could result in substantial over/underestimates, if retrievals are attempted for conditions where the underlying assumption of time-invariant meteorology is not met and/or when significant upwind emissions (from other sources) enter the flux box. As noted in the Introduction, *a priori* metrics may be used to avoid these conditions. The key information for calculating $E_{C,S}$, is the estimate of the temporal gradient in species mixing ratio ($\partial \chi_C / \partial t$) as in Eq. (4). This is achieved in the DR method by utilizing GEM-MACH model 4D data (volumetric time-series) and integration over the entire volume of the box at each model time-step. These 4D data are not available for aircraft applications of retrieval algorithms, since instantaneous datasets cannot be created from flight information, and collection of volumetric data may not be operationally feasible in

aircraft campaigns, where flight durations may be limited to a few hours. However, data from time-consecutive 2D screens
may be used to approximate $E_{C,S}$, leading to improved estimates of emission rates – by providing ancillary information for
340   the study of temporal trends in chemical and meteorological fields. As we show later in Sect. 3.3, estimates of $E_{C,S}$ from time-
consecutive 2D screens are strongly correlated ($P_r = 0.9$) with the corresponding DR estimates of $E_{C,S}$ using volumetric (3D)
time-series. This high correlation suggests that $E_{C,S}$ can be approximated using $\partial \chi_C / \partial t$ generated from time-consecutive 2D
screens (3D fields) as representative of the rate of change in mass within the volume of the box. The approximation of $E_{C,S}$
using screen data alone is given by,

$$E_{C,S}^*(t) = A \, M_R \int \overline{\rho_{air}}(t,z) \frac{\partial \overline{\chi_C}(t,z)}{\partial t} \, dz \tag{15}$$

where, same as in Eq. (14), $\overline{\rho_{air}}(t,z)$ and $\overline{\chi_C}(t,z)$ are the average GEM-MACH predicted air density and species mixing ratio
along s (around the box) at height z. Note that the key difference between Eq. (4) and Eq. (15) is that the former integrates over
the volume, using consecutive time steps of chemical transport model output, while the latter integrates over average values
along s at each height level z on the screen walls using consecutive time steps, and multiplies the result by the box horizontal
350   area. However, Eq. (15) may also be generated from aircraft observations, something which is not possible for Eq. (4). For the
case of aircraft observations, a repeat flight would be required: rather than just a single flight path starting near the surface and
working upwards along the box boundary, the aircraft would execute (at least) two consecutive box flights around the same
facility. A better approach would be the use of two (multiple) aircrafts, or UAVs as they are becoming more popular for air
quality measurements (e.g., Elston et al., 2015; Mölders et al., 2015; Barchyn et al., 2018; Li et al., 2018), flying in tandem
355   following the same flight path in order to reduce time lag between samplings (in turn reducing the chance for meteorological
variability to be a concern). Differencing between the fields obtained from the two consecutive flights would be used to estimate
the value of $E_{C,S}^*$. The correlation between $E_{C,S}$ and $E_{C,S}^*$ also indicates that the latter may be used to estimate, from repeat
flight aircraft observations, the extent to which within-box mass storage may influence emission rate retrievals. We note that
the correlation between $E_{C,S}$ and $E_{C,S}^*$ was deduced here through GEM-MACH model simulations with model grid point
360   resolution (in the horizontal) of 2.5km; this correlation (and the corresponding time/length scales) can be further investigated
through model simulations at higher resolutions (e.g. 50m-250m) and/or field observations (e.g. controlled tracer release).

Eqs. (14) and (15) along with the other terms in Eq. (12), give the revised TERRA Retrieval (TR*) method,

$$E_C^* = E_{C,H} + E_{C,V} + E_{C,VD} - E_{C,M}^* + E_{C,S}^* \tag{16}$$

Using the consecutive screens at every model time-step (extracted from GEM-MACH model output fields), estimations
365   of $E_{C,S}^*$ were made for the period of the nine JOSM 2013 cases around Syncrude, CNRL and Suncor facilities. Retrieved
emissions by the TR* method were compared to the Model Input Emissions (MIE), as well as to the estimations made with the
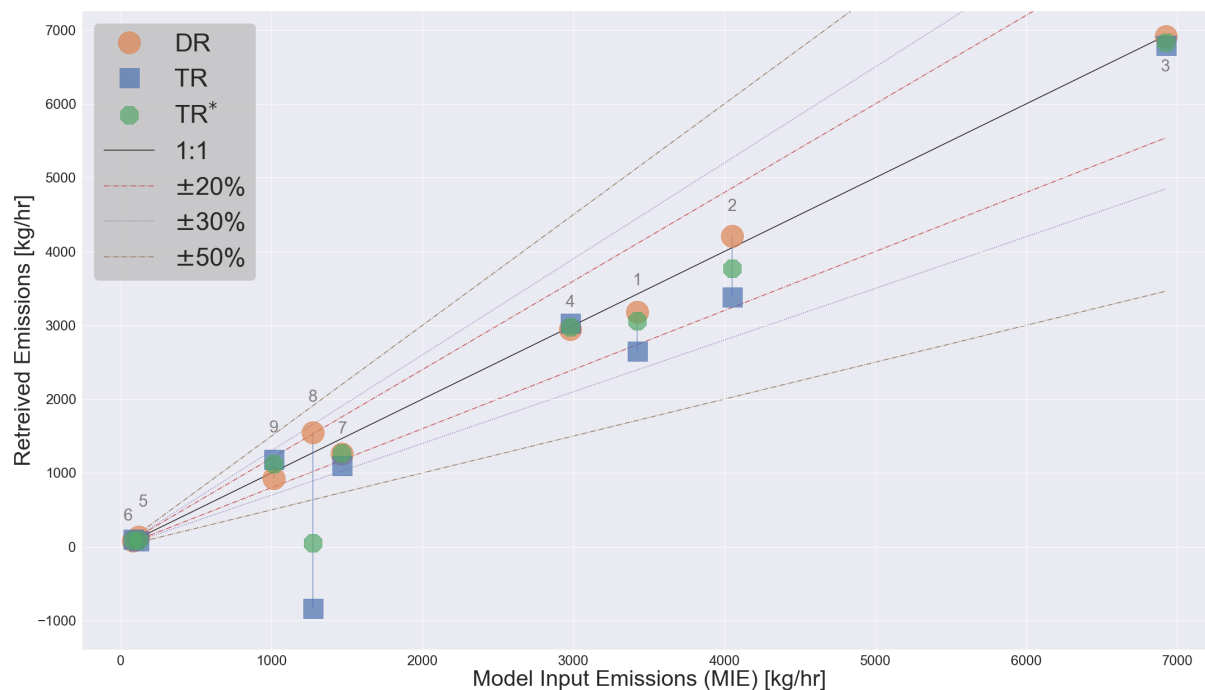DR and TR methods. Results and discussions follow.

Figure 3. Emission rates determined by the three methods (DR, TR and TR*) are plotted against model input emissions (MIE). The dark solid curve is the 1:1 MIE line, estimates below and above this line represent under and overestimation by the three methods. Case IDs are noted for each set. The vertical lines represent the error associated with storage-release events ($E_{C,S}$) in TR estimates relative to MIE. Also shown on the figure are the $\pm50\%$, 30% and 20% error limits (dashed lines) for reference.

3 Results and Discussions

Figure 3 summarizes the retrieved emissions by the three methods (DR, TR and TR*) and their performance against the model input emissions (MIE); vertical lines represent the over/underestimates by TR method due to the exclusion of the storage term ($E_{C,S}$) in Eq. (12), for each case. Also shown in Fig. 3 for reference are the $\pm50\%$, 30% and 20% error limits, with the $\pm30\%$ being the maximum estimated uncertainty for TERRA applications in Gordon et al. (2015). Estimates for most of the flight cases studied here using GEM-MACH output fell within $\pm30\%$ of MIE, exceptions being cases on August 26th and 28th (cases 5 and 8). The latter two flights were not analyzed in past work due to unfavorable conditions for reliable application of TERRA, namely, light variable winds (case 5) and high upwind emissions (case 8). However, they have been included here using GEM-MACH predictions in mass-balance estimates, as extreme cases in which storage-and-release events have a significant influence on resulting retrievals. We herein refer to cases 5 and 8 as the rejected cases. Using the DR method, we were able to determine the relative contribution of the different terms to the net estimate of E_C (Fig. 4).
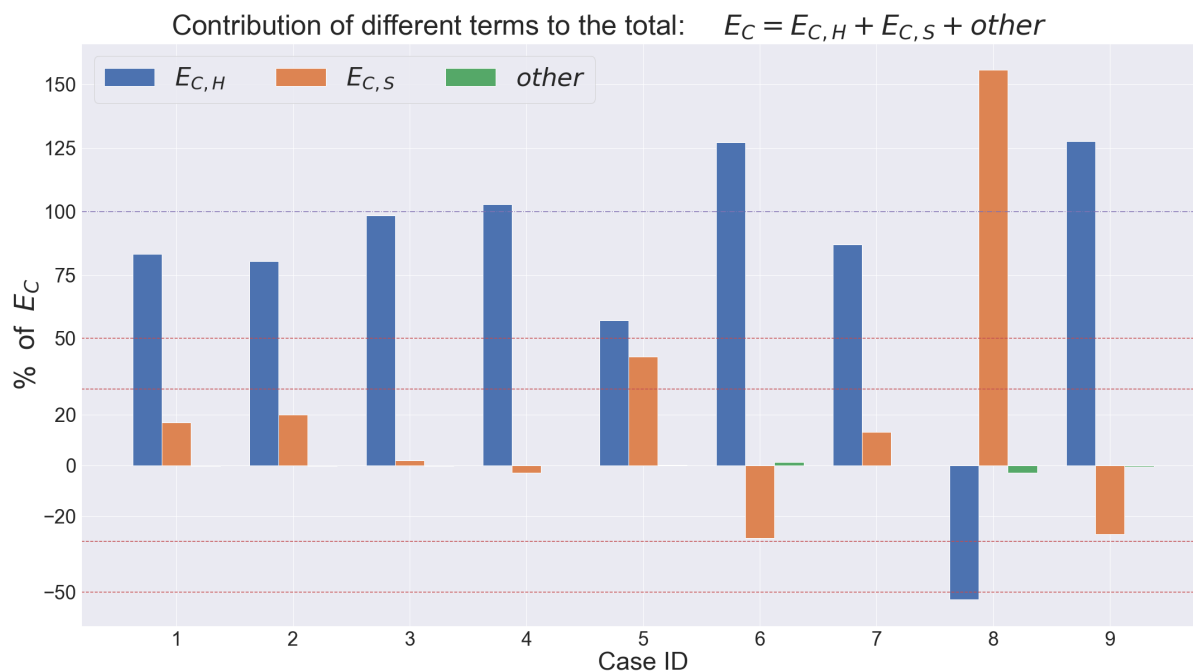
Figure 4. Contribution (%) of different terms in Eq. (7) to the total retrieved emissions (E_C), for the nine JOSM 2013 cases. The 100% contribution level is marked with a (purple) dot-dashed line, $\pm 30\%$ and 50% levels are also marked with (red) dashed lines on the figure. The main contributing terms were $E_{C,H}$ and $E_{C,S}$ ($> 97\%$). The other terms contributed $< 3\%$, combined.

3.1 Direct Retrieval (DR)

380 DR estimations were made by analyzing model 4D data for flux boxes approximating the nine cases. The numerical integration expressions for calculating each term, using model data, are described in Table B1. Estimates of E_C for SO_2 emission rates for the three oil sands facilities (Syncrude, CNRL and Suncor) were made by substituting the calculated terms from Table B1 in Eq. (7). Figure 4 demonstrates the contribution of each of the terms in Eq. (7) to the total integrated emission rate E_C by direct retrieval (DR) from the GEM-MACH model 4D data for the nine flight cases. As shown in Fig. 4 the main contribution came

385 from two terms, the net integrated horizontal flux $E_{C,H}$ and the storage term $E_{C,S}$, together contributing $> 97\%$ of the net change in mass. For 7 out of 9 studied cases, $E_{C,H}$ accounted for $> 80\%$ of the total emissions (E_C). This is in agreement with another model based study by Panitz et al. (2002), where the contribution of advective fluxes ($E_{C,H}$ and $E_{C,V}$) were estimated as 85% and 95% with no estimation of the storage term. The combined contribution of the other terms ($E_{C,V}$, $E_{C,VD}$ and $E_{C,M}$) were $< 1\%$ for all cases except for cases 6 and 8 (August 26th and 28th) were it was 1.2% and 2.9%, respectively.

390 For 2 out of 9 studied cases, the contribution of the storage term ($E_{C,S}$) was $> 30\%$, the maximum estimated uncertainty for cases analyzed in Gordon et al. (2015). As discussed earlier, these flights were not analyzed in past work – but they have been

included here as good examples of cases in which time-variant meteorology and/or high upwind emissions result in significant over/underestimates due to storage-and-release events. We also show later in Sect. 3.3 how these estimates may be improved.

395 DR method estimates were compared to the model input emissions (MIE), as shown in Figs. 3 and 5 and Table 2. Estimates were within 1-23% of MIE for the nine studied cases. We note that in some cases (4, 6 and 9), negative $E_{C,S}$ values reduced the net E_C estimate. Rejected cases 5 and 8 were the two cases with the largest storage term ($E_{C,S}$) contributions of 43% and 156% (averaged over flight time) of the total estimated emissions (E_C), respectively. Once corrected for storage (i.e. DR method), estimates for these two cases were within 12% and 23% of MIE. Cases 3 and 4 had the lowest storage rates (2% and -3%) where storage-corrected estimates were within 1% and 2% of MIE, respectively.

400 For the majority of our studied cases, our results indicate that the storage term (if not accounted for) contributes the bulk of emission rate over/underestimates for the variables investigated here, hence methods to predict and/or reduce its influence are desirable for aircraft retrieval algorithms. We note that the impact of storage on emission retrieval estimates will depend on the relative magnitude of the estimated emissions themselves in comparison to other sources of uncertainty. That is, an emissions estimate double that of previous estimates, which has a ±30% over or underestimate associated with storage, may still be

405 concluded as being significantly higher than the previous estimates (i.e. the emissions change is greater than the uncertainty associated with the retrieval itself). However, the storage term as defined here has been shown to be the main contributor to emissions over/underestimates, and will therefore be examined in more detail in the analysis which follows. We also note that the impact of $E_{C,S}$ is sometimes negative; a net release of mass from the box. Note that net releases of stored mass from the box would result in emissions overestimates; net storage of mass within the box would result in emissions underestimates.

410 Storage-and-release events occur when mass is accumulated within the volume of the box and released at a later time through box lateral walls, see the detailed discussions in Sect. 4. Such events can be detected on the box walls by comparing the temporal variation in the net horizontal flux, $E_{C,H}$ with $E_{C,S}$. Figures 7c and 7d (Sect. 4.1) show time series of the horizontal flux $E_{C,H}$ and the storage rate $E_{C,S}$ for case 4 on 20 August 2013 (CNRL). Storage-and-release events are represented by peaks and troughs in the $E_{C,S}$ curve, which coincide with troughs and peaks in the $E_{C,H}$ curve. $E_{C,H}$ and $E_{C,S}$ are anti-correlated

415 (with Pearson correlation coefficient of $P_r = -0.95$) for case 4. P_r values between $E_{C,H}$ and $E_{C,S}$ were in the range -0.68 to -0.98 for the nine cases, indicating strong anti-correlations for all the studied cases. Positive values of $E_{C,S}$ correspond to storage, which result in decreased net flux exiting through the box walls; negative values correspond to release, which results in increased net flux exiting through the box walls. Therefore, temporal variations in $E_{C,H}$ over the period of a flight around an emission source with relatively constant emission rates, can be attributed to mass storage/release. This provides a

420 potential practical means for detecting storage-release events in aircraft-based retrieval methods as the horizontal flux ($E_{C,H}$) can be calculated directly using airborne measurements along the box lateral walls, if ancillary information is available, such as from repeated flights. Changes to the calculated $E_{C,H}$ can indicate the potential presence of storage-and-release events within emissions observation data collection time periods.

Atmospheric
Chemistry
and Physics
Discussions

3.2 TERRA Retrieval (TR) approximations

425 By limiting the analysis to the extracted model data along box lateral walls (Fig. 2), TR estimates of model input emissions (MIE) were made. Terms in Eq. (12) were solved numerically using model data, with $E_{C,H}$ calculated as described in Table B1; numerical expressions for calculating the rest of the terms are described in Table B2. Model surface deposition rates for the base area of the box (A) were used in TR estimations, similar to the DR method, the main difference between these two methods is that, unlike Eq. (7), Eq. (12) does not account for the storage term. TR estimates of SO_2 emission rates were within

430 4-39% of MIE for 8 out of 9 studied cases, see Table 2. For cases 3 and 4, the two cases with the lowest net storage, results were within 4-6% of MIE. For rejected case 8 with the extreme storage event, the TR method underestimated the model input emissions by -166%; during the sampling period of this case, on average ∼1.5 times more SO_2 mass entered the flux box through the upwind wall than left through the downwind wall.

Comparing the performance of the two methods (DR and TR) against model input emissions (MIE), reveals the significance

435 of the storage term ($E_{C,S}$) in emission rate retrievals. Figure 5 shows the time-series of retrieved emissions using the DR and TR methods for the nine cases, with model time-step temporal resolution ($\Delta t = 2$min). MIE time-series are also plotted for SO_2 emissions from all the sources within each facility. Note that in aircraft-based applications of TERRA, a single emission rate estimate (temporal average) is generated based on the observations during the entire flight-time (∼2 hours). Here, where we approximated the TERRA retrievals by the TR method (see Sect. 2.2.2), instantaneous 2D screen fields (extracted from

440 the GEM-MACH model output) were used to generate an emission rate estimate at each time-step using the TR method. The flight-time average estimates by DR and TR methods were compared to the flight-time average MIE; Normalized Root Mean Square (NRMS) error scores are noted for each method on Fig. 5.

By combining the information from Figures 4 and 5, TR estimates for the nine cases can be categorized in three groups: (a) net positive storage (resulting in emission rates underestimation), (b) net negative storage (i.e. release of stored mass from

445 the box during the time period studied, resulting in emission rates overestimation), (c) low storage amounts; best emission rate estimates. The first group (a) includes cases 1, 2, 5, 7 and 8. For these cases the net storage was positive, where 13-156% of the emitted SO_2 mass was accumulated within the box volume and was not released through box walls during the sampling time. The stored mass was accumulated within the box either before or during the sampling time, which in turn resulted in decreased flux through box walls and the consequent underestimation in TR estimations. Cases 6 and 9 fall into group (b) where the

450 previously stored mass was released (-29% and -27%) though the box walls along with the SO_2 mass from emission sources during the flight time, resulting in TR method overestimation. Group (c) include the ideal cases 3 and 4 where TR estimates were within 4% and 6% of MIE, respectively. These two cases captured consecutive storage-release pairs over the sampling period (see Sect. 4.1, Fig. 7c and d for case 4) that canceled each other out and therefor resulted in relatively low (sampling time average) storage rates ($E_{C,S}$) of 2% and -3%, respectively.
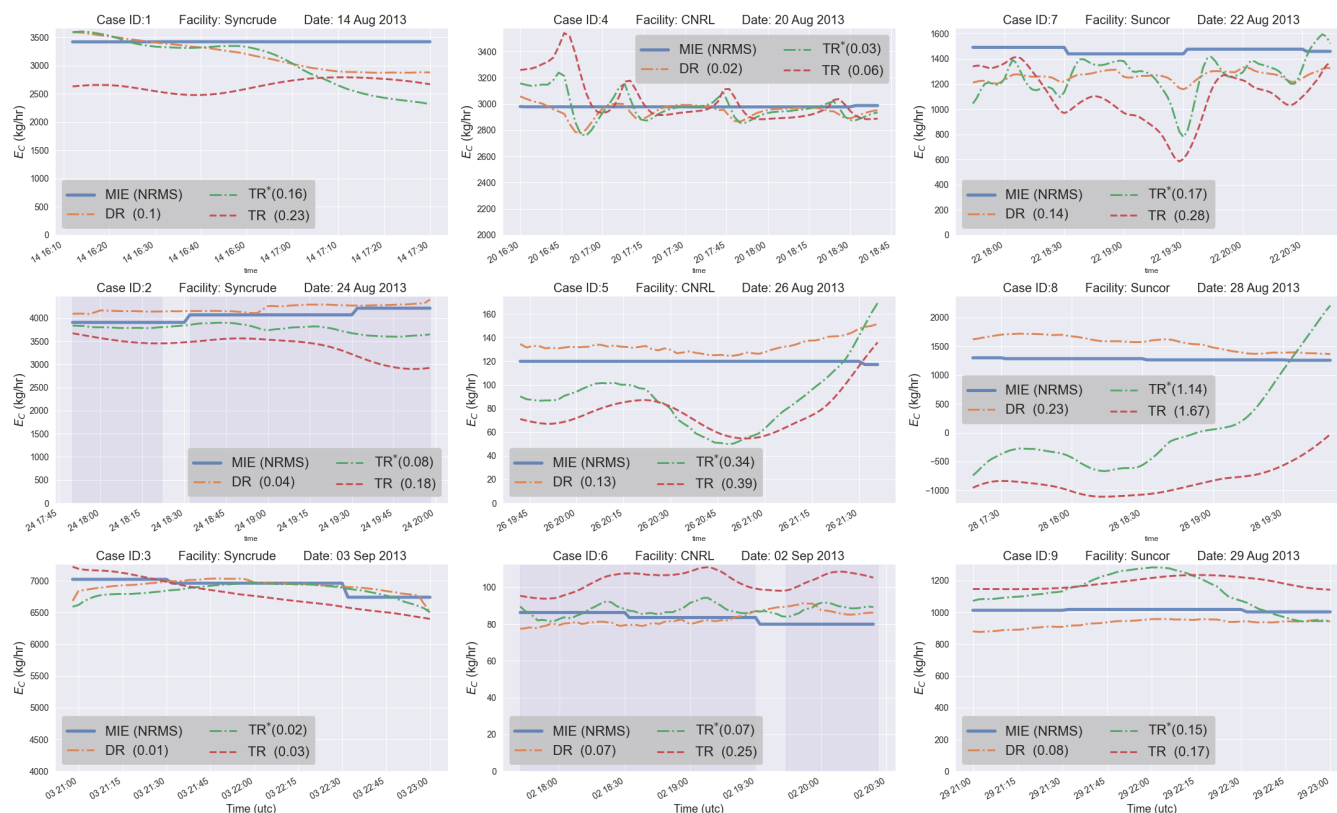
Figure 5. Time-series of the estimates by DR, TR and TR* methods, plotted along with model input emissions (MIE). Normalized-root-mean-square (NRMS) error for the three methods are noted on the plot for each of the nine JOSM 2013 cases. JOSM 2013 flights on August 24th (case 2) and September 2nd (case 6) included two consecutive box flights around the same facility, shown as shaded areas on their respective time-series plots. Note that the DR and TR* methods, unlike TR method, account for the storage rate ($E_{C,S}$ and $E_{C,S}^*$).

3.3 Revised TERRA (TR*) Retrieval

Here we show that change in tracer mass within the flux box can also be observed on box lateral walls over time (using time-consecutive screens), as described in Sect. 2.2.3. Figure 6 demonstrates the correlation between the 4D (volume) and 3D (screens) estimates of the temporal change in species mixing ratio ($\partial\chi_C/\partial t$) for all the data points (2 min resolution) corresponding to the period of the nine studied cases. The side plots show the range (and distribution) of the estimated $\partial\chi_C/\partial t$ by 3D (horizontal axis) and 4D (vertical axis) calculations. $\partial\chi_C/\partial t$ range is higher for 4D estimates (2.52 ppbv/hr) than for 3D estimates (0.98 ppbv/hr), but the two variables are correlated. The correlation coefficient $P_r = 0.7$ (with slope $m = 2.98$ and $b = 0.09$ intercept for the least-square fit line), indicates strong correlation between 4D and 3D estimates. The exception being rejected case 8 where the correlation is low due to relatively high upwind emissions during the time period of this case (on August 28th). Excluding case 8 from this analysis, increases the correlation between 3D and 4D estimates to $P_r = 0.9$.
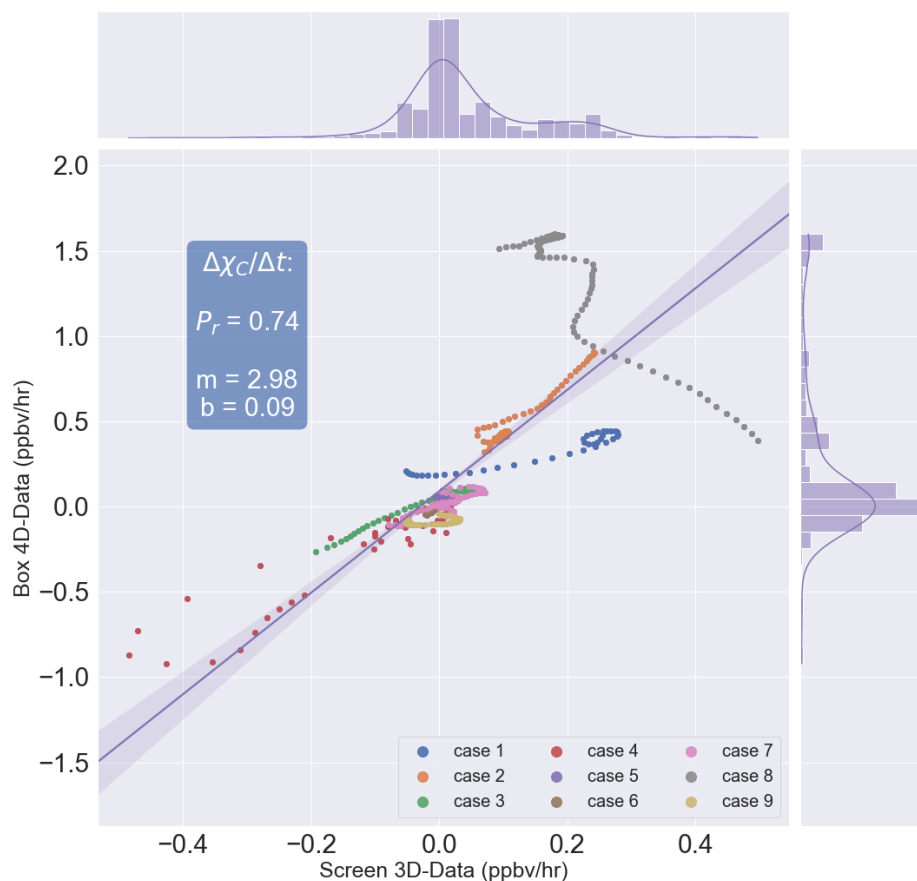
Figure 6. This three-panel graph demonstrates the correlation between estimates of $\partial\chi_C/\partial t$ by analyzing 4D (box) and 3D (consecutive screens) data sets. The vertical side plot shows the range and distribution of $\partial\chi_C/\partial t$ estimates for the period of nine studied flights by 4D analysis, the horizontal side plot shows the same for 3D analysis. The main panel demonstrates the correlation between the two methods of estimation. Pearson correlation coefficient ($P_r = 0.74$) shows the strength of this relation. The least-squares fit line has slope m=2.98 and intercept b=0.09. Note that case 8 estimates are in low correlation due to relatively high upwind emissions during the time of that particular flight (August 28th), rendering the 3D estimates less representative of the 4D estimates. Excluding case 8 from this analysis, results in $P_r = 0.88$. $E^*_{C,S}$ can be estimated from multiple consecutive 2D screens, by assuming 3D estimations of $\partial\chi_C/\partial t$ as representative of the rate of change in species mass within the flux box.

465 As discussed earlier, the lack of information about the storage rate was the primary source of under/overestimates in TR method estimates for the majority of our studied cases. Here we demonstrate the use of an estimate of the storage term ($E^*_{C,S}$, see Sect. 2.2.3) in a revised retrieval approach. We note that this term may be estimated from consecutive screens (Eq. 15) or tandem aircraft samples for ambient atmosphere applications. Data from time-consecutive screens (e.g. 2, 3) can be used to numerically solve Eqs. (14) and (15) (see Table B3 for discrete integral expressions). For GEM-MACH-driven retrievals as used here, these time-consecutive screens originate from consecutive model output times. Here, we show the effect of the

$E_{C,S}^*$ estimate on the resulting E_C estimates using GEM-MACH fields for top-down emission rate retrievals. Improved/revised TERRA Retrievals (TR*) can be made by plugging in $E_{C,M}^*$ and $E_{C,S}^*$ in Eq. (16). TR* method estimates are compared to the model input emissions (MIE) in Figs. 3 and 5 and Table 2. For 8 out of 9 cases, TR* estimates were within 2% to 34% of MIE. TR* estimates showed improvement over the TR method for all the cases (including the rejected case 8) by 2-52% of MIE,

475 see Table 2. TR* estimations rely on the availability of temporal data for estimating change in tracer concentrations ($\partial \chi_C / \partial t$) around the box for each emission rate retrieval case, which in our analysis with simulated fields were readily available with the GEM-MACH model data at each time-step. Again, for aircraft applications, a repeat box flight procedure can be used, where an aircraft carries out a second box flight immediately after the first flight, or two aircraft (or UAVs) follow the same flight path in tandem, one positioned a fixed distance behind the other.

480 ## 4 Three Quantitative Predictors of Storage and Release Events

The main processes contributing to the change in SO_2 mass within the volume of the box are illustrated in Fig. 2a (E_C, $E_{C,S}$, $E_{C,H}$, $E_{C,V}$ and $E_{C,VD}$). An SO_2 emission source continuously adds mass to the box at emission rate E_C. Most of the mass is advected out of the box through the top and lateral walls ($E_{C,V}$ and $E_{C,H}$), and some is deposited to the ground surface ($E_{C,VD}$). All the while, mass is continuously stored within the volume of the box and may later be released at net storage

485 rate $E_{C,S}$ (positive to negative). Under time-invariant conditions (e.g. wind field, atmospheric stability, emission rate), the rate of generation of mass (E_C) is at net (mass) balance with the other mass-removing processes ($E_{C,H}$, $E_{C,V}$ and $E_{C,VD}$) and the total accumulated mass within the box volume (B_C) remains relatively constant over time, and the storage rate negligible ($E_{C,S} \simeq 0$). The mass-balance approach based on the divergence theorem (Eq. 12) can therefore accurately equate the source emission rate to the net integrated flux leaving the box through the enclosing surfaces. However, localized (spatial) variations in

490 meteorological fields (e.g. atmospheric stability, wind) and chemical fields (e.g. temporal changes in source emission rate (E_C) and/or significant incoming upwind emissions) during the retrieval time, can shift the mass balance towards $E_{C,S} \neq 0$ (Eq. 7) - i.e. storage-and-release events. The scale and duration of the storage-release events contribute to over- and under-predictions in estimated emissions based on the mass-balance approach. Using GEM-MACH model 4D data we carried out extensive analysis of the storage-release events during our studied cases and devised three quantitative predictors for the severity of these events.

495 In the following subsections, using results from DR, TR and TR* estimates, we discuss these parameters and their correlation with over- and under-estimates in top-down mass balance emission rate retrievals due to storage-and-release events.

4.1 Gradient Richardson number (Ri) and changing atmospheric stability

Aircraft-based emission retrieval methods such as TERRA, utilize the mass-balance approach with the underlying assumption of mean steady-state conditions over the region of study ($\sim 100 \times 100$ km^2) during the flight time ($\sim$2-3 hours). However,

500 localized and transient variations in atmospheric stability at the vicinity of the stack/source location can impact the pollutant plume rise and transport within the box and result in storage-release events. Horizontal wind velocity usually increases logarithmically with height. The height the plumes reach in the atmosphere depend on the stack parameters (exit velocity, exit

Atmospheric
Chemistry
and Physics
Discussions

temperature, stack height) and on the local stability conditions (temperature and wind profile near the stack location). Thus plume height may change with varying stack parameters and stability conditions during the course of a retrieval. This in turn

505 would move the mass to a different flow regime – e.g. higher horizontal velocities if the plume height has increased, and lower horizontal velocities if the plume height has decreased. These changes in plume height, in response to changes in atmospheric stability, may thus lead to a change in the rate at which emitted mass is transported out of the box. Stability changes are thus a good predictor of the possibility for mass storage and release. A measure of the change in dynamic stability is therefore desirable as a predictor for conditions under which this aspect of the "static meteorology" condition, required for reliable top-down

510 emission rate retrievals, is not met.

Atmospheric dynamic stability can be described in terms of the gradient Richardson number (Ri), which is a dimensionless ratio, related to static stability and the shear induced turbulence (AMS, 2020),

$$Ri = \frac{\frac{g}{T_v}\frac{\partial \theta_v}{\partial z}}{\left(\frac{\partial U}{\partial z}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2} \tag{17}$$

where θ_v and T_v are virtual potential temperature and virtual absolute temperature, respectively, g is the gravitational constant

515 and U and V are the horizontal wind components. The sign of Ri is determined by the static instability term ($\partial \theta_v / \partial z$) in Eq. (17); the denominator is always positive. In the presence of a strong wind shear Ri approaches 0. The critical gradient Richardson number (Ri_c) is about 0.25, below this value the atmosphere is dynamically unstable. The atmospheric Physics module within the GEM model includes the estimation of atmospheric stability and calculations of the gradient Richardson number, Ri (Mailhot and Benoit, 1982), which is also used as a parameter (in addition to emission stack/source information

520 such as temperature, exit velocity) to calculate plume rise height (Akingunola et al., 2018). Changes in Ri with respect to time are thus an indicator of the potential for the plume to change height during an aircraft emission rates retrieval; Ri may be predicted using a weather forecast model as part of pre-flight planning.

For our analysis, 4D fields of Ri were extracted from the model output dataset for each of the nine studied cases. Figure 7 shows the case for the JOSM flight on 20 August (case 4) with vertical profile time-series at the location of the emitting stack,

525 and compares Ri in (a), with the SO_2 mixing ratio in (b), the net horizontal flux $E_{C,H}$ in (c) and the storage rate $E_{C,S}$ in (d). As shown in Fig. 7a, the region of unstable air (blue) was gradually increasing in height during the sampling time, resulting in an increase in the vertical extent of the mixing layer. Note that the color-scale is symmetric around the critical gradient Richardson number $Ri_C = 0.25$ in Fig. 7a. The increase in the vertical extent of the mixing layer during the period of case 4, resulted in the SO_2 plume mixing into heights with higher wind speeds (Fig. 7b). The incremental plume rise resulted in a

530 sequence of fast-release events, that is, the simulated plume was moved into a faster (higher level) horizontal flow, and SO_2 mass was carried to the boundaries of the box faster – the net rate of mass loss through the downwind walls thus increased during the sampling time period ($\frac{\partial}{\partial t}(E_{C,H}) > 0$, Fig. 7c; $\frac{\partial}{\partial t}(E_{C,S}) < 0$, Fig. 7d). Note the anti-correlation between $E_{C,S}$ and $E_{C,H}$ in Fig. 7c, d. We also note that although faster moving flow can result in more mixing and the subsequent plume dilution, the decrease in compound mixing ratio is not necessarily enough to compensate for the increased exiting flux. If emissions

535 retrievals are calculated under these conditions without accounting for the storage term (negative in this case), emissions estimates will be biased high (overestimation). These fast-release events temporarily disturb the required balance between the
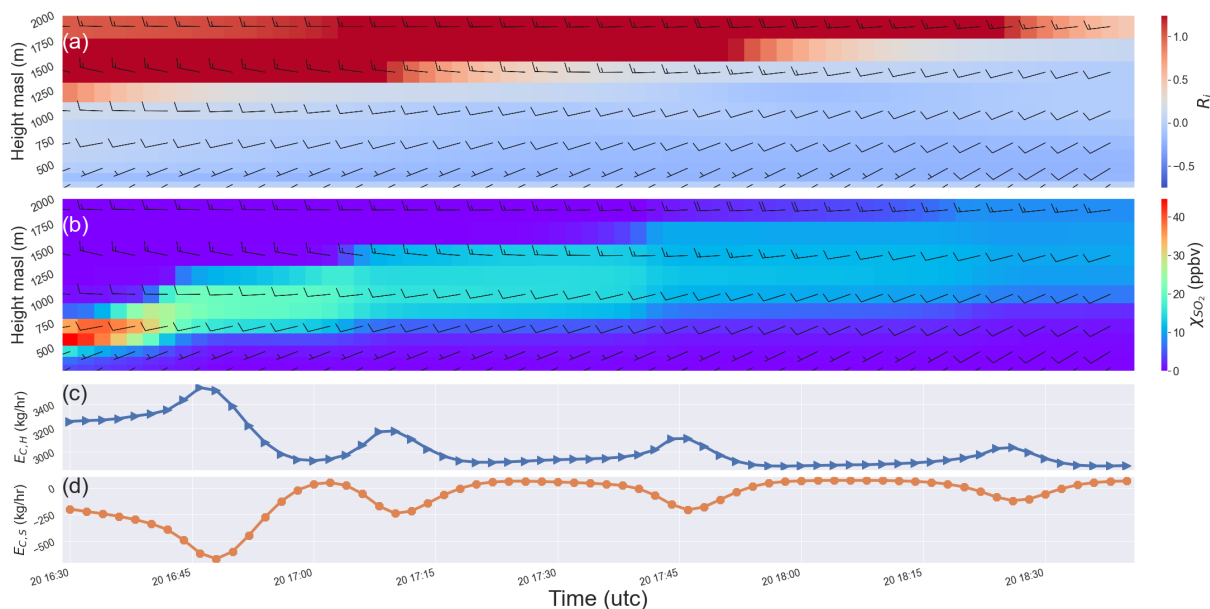
Figure 7. This schematic demonstrates (qualitatively) the correlation between change in atmospheric stability represented by the gradient Richardson number (Ri), plume vertical motion and storage-release events for the period of case 4. (a) Ri vertical profiles at stack location over the flight time, with the color-scale symmetric around the critical gradient Richardson number $Ri_C = 0.25$). (b) SO_2 mixing ratio vertical profiles at stack location. (c) Time-series of net horizontal flux ($E_{C,H}$) and (d) the storage rate ($E_{C,S}$), over sampling time. Note the level of anti-correlation between $E_{C,H}$ and $E_{C,S}$ ($P_r = -0.95$); oscillations in storage and release are occurring: changes in $E_{C,S}$ are matched by opposite changes in $E_{C,H}$. Wind speed and direction are shown as barb plots with units m/s (half =5, full=10). For every incremental rise in plume height, a corresponding release event occurs. As the plume ascends into regions with faster wind speeds, more mass is released ($E_{C,S} < 0$) through the lateral walls resulting in an increase in $E_{C,H}$ as can be seen in (c).

addition of mass to the box volume by source emissions (E_C) and removal of the mass through vertical and horizontal fluxes and surface deposition, causing a negative change in the total stored mass (B_C) within the volume of the box ($E_{C,S} < 0$). These release events can be seen as troughs (peaks) in the $E_{C,S}$ ($E_{C,H}$) time-series plot (Fig. 7c, d). While the plume remained at a

540   constant height, the balance was restored over time ($E_{C,S} \rightarrow 0$); this can be seen as peaks/plateaus in $E_{C,S}$ time-series in Fig. 7d. After these periods of changing stability, the plume started to rise again, and the balance was shifted, resulting in negative $E_{C,S}$ values. This process was repeated several times over the period of case 4, resulting in a sequence of release events. Note that the trough to peak duration (period) and scale (magnitude) of each event are inversely proportional to the wind speed at plume height, and they are reduced as the plume ascends into faster moving layers of air resulting in a damped oscillatory

545   storage and release sequence. In this particular example, this periodic storage-release during the sampling time, with subsequent equilibrium intervals, resulted in relatively small net storage of $E_{C,S} = -3\%$ over the sampling time and TR estimates within 6% of MIE (slight overestimation) for case 4. When consecutive storage-release pairs are captured during the flight/sampling time as in case 4, the average error due to storage term ($E_{C,S}$) is reduced. However, if the storage or release is of longer
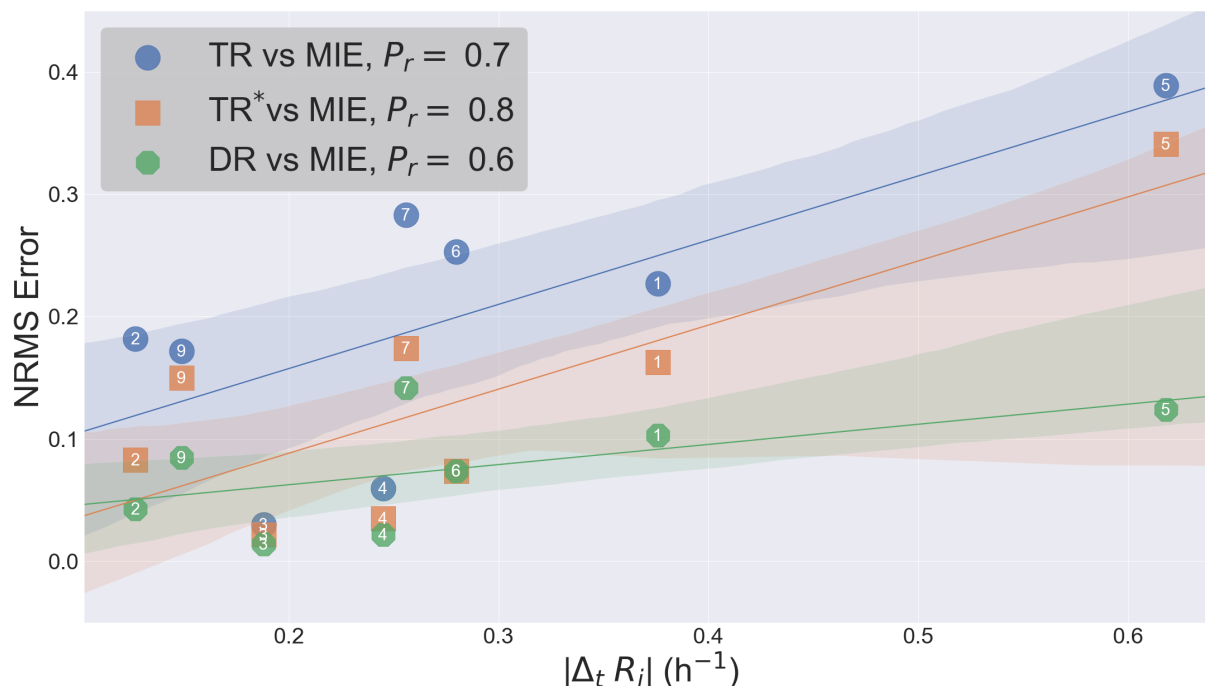
Atmospheric
Chemistry
and Physics
Discussions

Open Access

EGU

Figure 8. Correlation between the normalized root-mean-square (NRMS) error and the change in atmospheric stability represented by the absolute change in the gradient Richardson number $|\Delta_t R_i|$. The shaded areas indicate the 90% confidence region. Case ID is indicated for each data point on the graph along with the calculated Pearson correlation coefficients (P_r) for the three methods. Results indicate a strong positive correlation between the NRMS error and $|\Delta_t R_i|$.

duration, so that an unpaired event occurs during observations, the net error increases, as was the case for flight case 9 where

550   an unpaired release event (one storage-release pair and one additional release event) resulted in net storage of $E_{C,S} = -27\%$ and consequent TR overestimates. Monotonic storage or release events occur when mass is continuously stored (e.g. when a plume descends into slower moving air throughout a sampling interval; $\frac{\partial}{\partial t}(E_{C,S}) > 0$ and $\frac{\partial}{\partial t}(E_{C,H}) < 0$) or released (e.g. release of previously stored mass during a sampling interval; $\frac{\partial}{\partial t}(E_{C,S}) < 0$ and $\frac{\partial}{\partial t}(E_{C,H}) > 0$); such events result in TR under/overestimates as in case 2 where net storage of $E_{C,S} = 20\%$ (i.e. storage event: $E_{C,S} > 0$ and $\frac{\partial}{\partial t}(E_{C,H}) < 0$) resulted

555   in TR underestimations.

The average value $Ri = -0.14$ at plume height, indicates unstable conditions ($Ri < Ri_C$) during the period of case 4 – the average value was calculated by spatially averaging Ri at the vertical layer containing the plume center (maximum concentration in the vertical) over the horizontal extent of the box at each time-step and then averaging over the sampling time. Atmospheric stability was further decreased at an average rate of $\Delta_t Ri = -0.25 \text{ h}^{-1}$ over the period of this case. Change in

560   atmospheric stability (represented by $\Delta_t Ri$) is thus related to plume vertical motion and mixing, which in turn contributes to storage-release events and a corresponding over- or under-estimate in TR retrievals; $\Delta_t Ri$ is thus a useful forecast metric to use as an *a priori* indicator of the potential for changing atmospheric stability affecting emission rates retrieval accuracy, and may

also be calculated from retrieved aircraft meteorological data. Table 2 lists (flight time) average Ri and $\Delta_t Ri$ values for each of the nine studied cases. Our analysis indicates a strong positive correlation between the absolute value $|\Delta_t Ri|$ and the NRMS

565 error for DR, TR and TR* methods with the Pearson correlation coefficients (P_r) of 0.6, 0.7 and 0.8, respectively. Figure 8 shows the correlation between $|\Delta_t Ri|$ and NRMS error for each method. Note that the rejected case 8 with $\Delta_t Ri = -0.34$ h^{-1} (see Table 2) was excluded from correlation analysis as an outlier, due to the extreme impact of high upwind emissions (and not meteorological variability) on emission rate estimates (further discussed in Sect. 4.3). $\Delta_t Ri$ can be calculated from aircraft-measured data along the lateral walls of the box and the measured data from additional profiling spiral flights upwind

570 and downwind the source, such as those conducted during JOSM 2013 campaign. Analysis of $\Delta_t Ri$ can provide insights regarding the change in atmospheric stability and the probability of plume vertical motion and the resulting storage-release events during flight time; the uncertainty levels associated with the storage rate ($E_{C,S}$) in the retrieved emissions can therefore be quantified.

4.2 Plume vertical motion and variations in the direction of the transport

575 The direction of the wind experienced by the plume may change when the plume remains at one altitude, or may change as the plume rises or falls in the atmosphere (e.g. via the well-known "Ekman spiral" of wind direction changes with increasing height, AMS 2020). Changes in wind direction for a fixed plume height would result in the rejection (for top-down retrievals) of the flight – here we note that changes in wind direction due to plume vertical motion may also be a potential reason for flight rejection. Using GEM-MACH predicted wind fields for the nine studied cases (chosen *a priori* for relatively constant

580 wind speeds at each height), we here investigated the extent of variations in wind direction at the time-varying vertical level containing the plume center,

$$\Delta_t \alpha = \frac{\sigma_\alpha}{\Delta t^{SP}} \qquad (18)$$

where σ_α is the standard deviation in wind direction α_{wind} time-series (spatially averaged over the horizontal extent of the box including lateral walls at each time step) and Δt^{SP} is the *Sampling Period* time duration. In aircraft-based retrievals (e.g.

585 TERRA), $\Delta_t \alpha$ can be estimated from aircraft measurements around the box at plume center-point height. Our analysis of the wind fields for the nine studied cases indicates a positive correlation ($P_r = 0.6$) between the rate of variation in wind direction ($\Delta_t \alpha$) and the rate of change in atmospheric stability ($\Delta_t Ri$). Variations in the direction of the flow at plume height ($\Delta_t \alpha$) can result in the re-circulation of the plume, accumulation of the emitted mass within the volume of the box and its release at a later time (storage-release event). Such an event occurred during our GEM-MACH model simulations for the period of the

590 JOSM flight on 26 August around CNRL (case 5). Figure 9 shows horizontal cross-sections of the SO$_2$ mixing ratio data at two (consecutive) plume-maximum heights and over time for case 5. Starting from top left at 20:20UT and height $H_1 = 909$m (meters above sea level - masl) to the bottom right at 21:30UT and height $H_2 = 1079$m (masl), horizontal cross-sections are plotted at plume center-point (maximum concentration in the vertical) height at each time. Air flow direction and strength are shown as grey arrows on each panel. The model shows strong vertical wind shear between these two heights, with a $\sim 180°$shift

595 in the wind direction between the upper and lower row of panels which remains relatively constant with time. This is a subtle
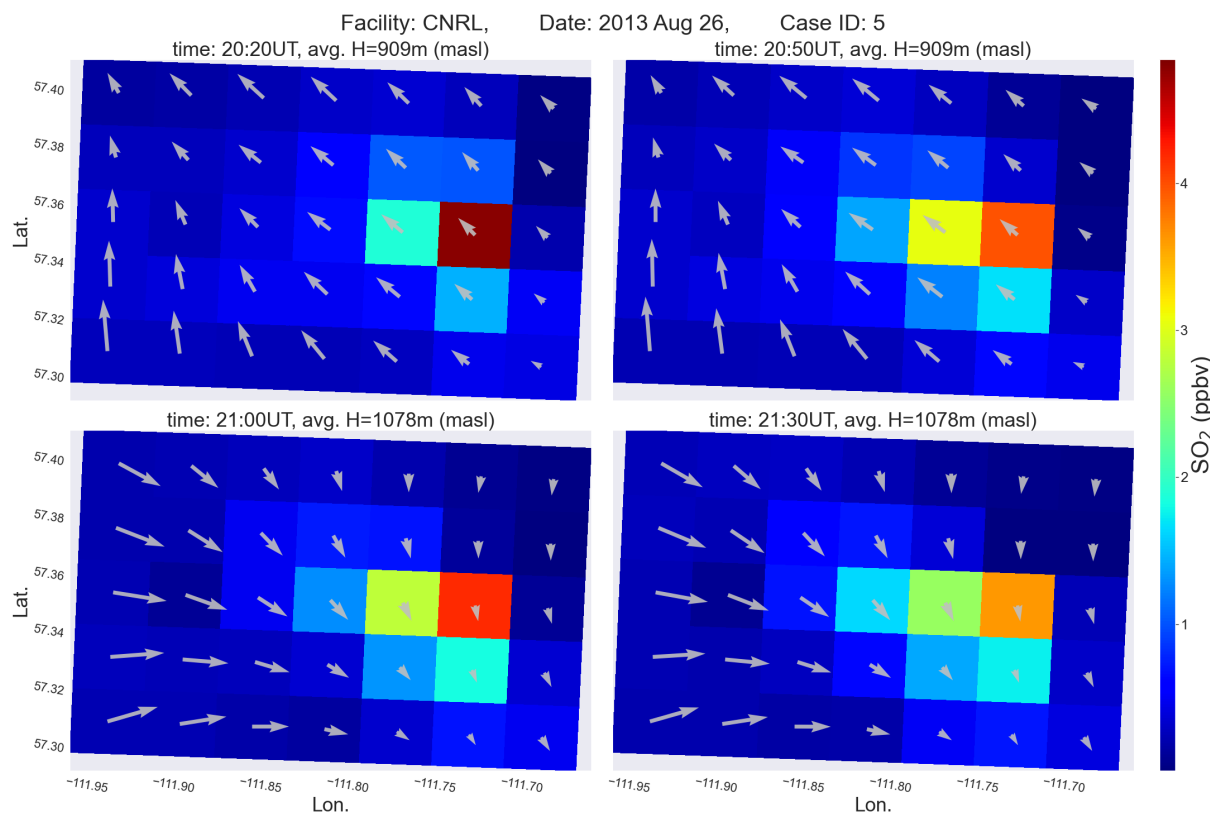
Figure 9. SO$_2$ mixing ratio horizontal cross-sections at two (consecutive) plume center (max. conc.) heights (meters above sea level - masl) and four time-steps, during the period of case 5 over CNRL. The strength and direction of the airflow at each cross-section is represented by grey arrows. Corresponding time and average height are noted for each panel. As the plume rises into layers with different wind direction over time, it re-circulates within the volume of the box which in turn results in a storage event.

example of how wind direction changes may interact with changes in stability: while the wind direction is relatively static at each height, changes in stability may cause the plume to rise to different heights, where the wind direction may change or even reverse relative to the level below. The changes in turn will alter (slow) the rate of mass release from the walls of the box ($\frac{\partial}{\partial t}(E_{C,H}) < 0$); a storage event, with $\frac{\partial}{\partial t}(E_{C,S}) > 0$. For this particular case (5), between 20:20UT and 20:50UT the plume

600   center was at H_1 and was advected in a weak south-easterly wind; between 20:50UT and 21:00UT, the plume mixed upwards and rose to H_2 where wind direction was north north-westerly and remained there until 21:30UT (which afterwards descended back to H_1). This change in wind direction at plume height ($\Delta_t \alpha$) as a consequence of plume vertical motion due to change in atmospheric stability ($\Delta_t Ri \neq 0$), resulted in a net storage of $E_{C,S} = 43\%$ during case 5 and consequently an underestimation in emission rate using the TR method.

605   Cases 6 and 9 (both release events) along with rejected case 5 (storage event) were the three cases most impacted by variations in wind direction at plume-center height. Flight time average $\Delta_t \alpha$ values for all nine cases are given in Table 2. Our
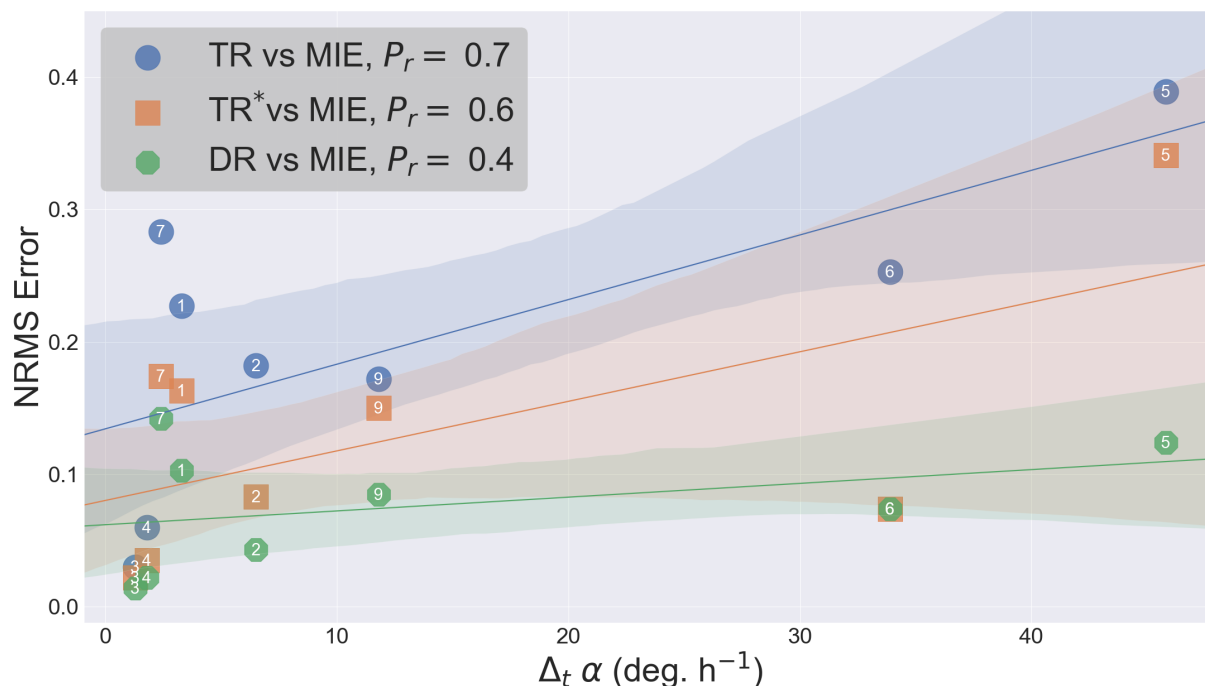
Figure 10. Correlation between the normalized root-mean-square (NRMS) error and the change in wind direction at plume height ($\Delta_t \alpha$). The shaded areas show the 90% confidence region. Case ID is noted for each data point on the graph along with the calculated Pearson correlation coefficients (P_r) for the three methods. Results indicate a positive correlation between the NRMS error and $\Delta_t \alpha$.

results indicate a positive correlation between $\Delta_t \alpha$ and NRMS error in DR, TR and TR* estimates, similar to $\Delta_t Ri$; Figure 10 demonstrates this correlation for the three methods, with P_r scores 0.4, 0.7 and 0.6, respectively. As mentioned earlier, there is also a positive correlation between $\Delta_t \alpha$ and $\Delta_t Ri$ ($P_r = 0.6$); the combined effect of change in the direction of the transport

610 and changing stability increases the uncertainty in the retrieved emission rates. In aircraft-based emission rate retrievals, wind shear strength in terms of $\Delta_t \alpha$ can be included in uncertainty analysis associated with plume vertical motion due to change in atmospheric stability ($\Delta_t Ri$). Aircraft-measured data along the box walls combined with measurements during additional spiral flights can be used to evaluate these metrics.

4.3 Weighted upwind to downwind concentration ratio (ϕ)

615 In applying the mass-balance technique, the mass flux associated with upwind emissions (and regional background levels) entering the box volume (inflow) is subtracted from the downwind outgoing flux containing the emissions within the box (outflow) to estimate the net emission rate from sources within the box volume. The mass inflow associated with regional background levels (e.g. for CH_4 and CO with relatively high background concentrations) is balanced out with the mass flux out of the box due to spatial (horizontal) homogeneity of background concentrations. However, emissions from upwind sources

620 carried through the sampling region (box volume) are spatially heterogeneous (plumes) and therefore their inflow and outflow

Atmospheric
Chemistry
and Physics
Discussions
Open Access
EGU

through a control volume is not necessarily equal and canceling. For instance, converging horizontal wind fields and/or plume vertical motion and mixing can result in temporary accumulation of the mass of these upwind plumes within the box (control volume). If the emissions from upwind sources are large relative to the within-box source(s), the emission rates retrieval may therefore become less accurate. That is to say, all the processes (including the ones described above in terms of $\Delta_t Ri$ and $\Delta_t \alpha$)

625   that contribute to storage-release of the mass emitted by the source(s) within the flux box, may also affect the mass emitted from upwind sources which are entering the box through the upwind wall(s) – leading to further increased uncertainty in the retrieved emission rates for sources within the box. The box upwind-to-downwind concentration ratio is in strong correlation with the relative magnitude of upwind emissions entering the box to the emissions from a given source within the box, with $P_r \geq 0.8$ as determined for our nine studied cases. To distinguish upwind emissions from regional background levels, this ratio

630   can be weighted by temporal standard deviations in upwind and downwind concentrations. Hence we define,

$$\phi = \left(\frac{\overline{\chi_{C,u}}}{\overline{\chi_{C,p}}} \right) \left(\frac{\sigma_{\chi_C,u}}{\sigma_{\chi_C,d}} \right) \tag{19}$$

where $\overline{\chi_{C,u}}$ is the screen average upwind concentration and $\overline{\chi_{C,p}}$ is the plume average concentration; here the plume is defined as downwind concentrations that are one standard deviation above the mean. $\sigma_{\chi_C,u}$ and $\sigma_{\chi_C,d}$ are upwind and downwind concentration standard deviations. Note that while the first ratio in Eq. (19) represents the magnitude of upwind to source emis-

635   sions, the second ratio distinguishes upwind plumes from background concentrations: in the absence of upwind plumes, $\sigma_{\chi_C,u}$ approaches zero due to the uniform nature of background concentrations. Parameter ϕ can therefore provide a quantitative indication of the level of confidence in the retrieved emissions based on the mass-balance approach, due to this factor. Table 2 lists ϕ values in % for the nine studied cases, determined from our GEM-MACH simulated fields.

For an isolated source with relatively low upwind emissions (which in turn depends on the mean wind direction and the

640   location of the nearby sources) ϕ would be small and the application of the mass-balance technique would be expected to yield more accurate results. This was true for cases 3 and 4 with $\phi < 0.2\%$ (see Fig. 2d for case 4). JOSM 2013 cases 3 and 4 were conducted over CNRL and Syncrude oil sands facilities on 20 August and 3 September 2013. For the period of these two flight cases, wind directions were also consistently from the south-west, a condition that was accurately simulated in our GEM-MACH model runs for the same period. CNRL and Syncrude are the two facilities located on the western side of the

645   Athabasca oil sands complex, with no facilities to their west and therefore no upwind sources of emissions when wind direction is from the west (Fig. 1b). These conditions resulted in low $E_{C,S}$ rates of 2% and -3% for cases 3 and 4. For these two cases, TR estimates were within 4-6% of the model input emissions (MIE). Our TR estimates of SO_2 emissions for CNRL were in strong agreement with TERRA SO_2 emission rate estimates by Gordon et al. (2015) using aircraft-measured data during JOSM 2013 campaign, where estimates were 4.8% higher than minute-by-minute emissions reported by CNRL. Our analysis

650   indicates that about 3% of this overestimation could be due to release of the stored mass during the flight on 20 August (case 4).

Small ϕ values represent favorable conditions for applying the mass-balance technique for emission rate retrievals. For 7 out of 9 studied cases, ϕ was less than 10% with NRMS errors between 4-28% for TR estimates. The opposite was true for rejected cases 5 and 8 with ϕ scores 11% and 13%, with NRMS errors 39% and 166% for TR estimates. This under performance
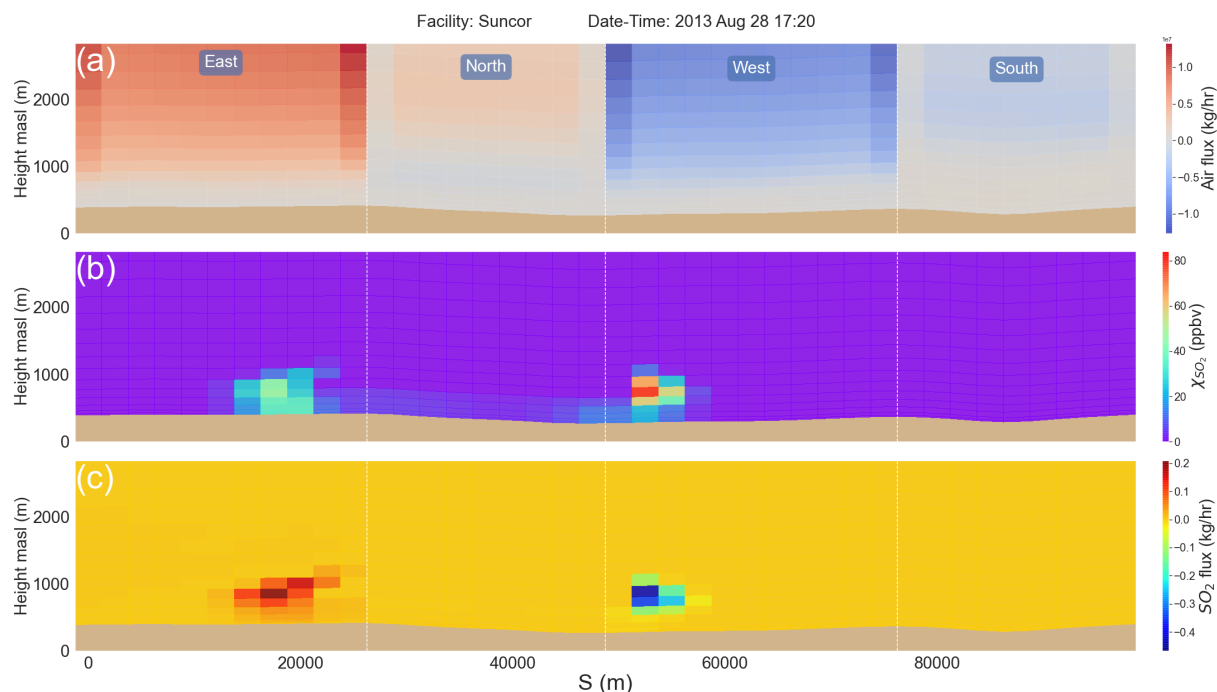
Figure 11. Case 8 screens around Suncor: (a) air flux screen (b) SO$_2$ mixing ratio screen and (c) SO$_2$ mass flux screen. Note the strong negative flux (inflow) on the west wall and the weaker positive flux (outflow) on the east wall resulting in (sampling time avg.) $\phi = 13\%$.

655   was a combined result of upwind emissions which were higher than the source emissions, consistent converging wind fields (slower wind speeds within the flux box relative to the upwind wall throughout the flight time), and plume vertical motion due to change in atmospheric stability ($\Delta_t Ri \neq 0$). For rejected case 5, there was an additional contribution from variations in the direction of plume transport ($\Delta_t \alpha = 46°$,   sampling time average). Rejected case 8 was conducted around the Suncor facility on 28 August 2013, during which the mean wind direction was from the west; as a result large upwind SO$_2$ plumes originating

660   from the Syncrude facility, located on the west side of Suncor (see Fig. 1b), continuously entered the flux box during the period of this case. SO$_2$ mass was accumulated within the volume of the box resulting in a flight/sampling time average storage rate of 156% of the total retrieved emission rates (see Fig. 4). Note that a large upwind source was a condition previously identified for rejecting this flight from emissions calculations (which is why emissions from this flight were not estimated in previous work). The ϕ predictor introduced here allows this effect to be quantified (from model forecasts used to plan aircraft flights,

665   and the analysis of aircraft measured pollutant data time-series). Figure 11 shows (a) air flux, (b) SO$_2$ mixing ratio (χ_{SO_2}) and (c) SO$_2$ mass flux screens around Suncor at the start of rejected case 8; note the strong negative (inwards) flux on the west wall (originating from upwind emissions from the Syncrude facility) and the weaker positive (outwards) flux on the east wall (Fig. 11). By accounting for the storage rate ($E^*_{C,S}$), using model 3D data from time-consecutive screens, NRMS error was reduced by 52% of MIE in TR* estimates for this case. Direct retrieval (DR) of the Suncor emission rates for the period of rejected case

670   8, where storage was estimated directly from the model 4D data (volumetric time-series), was within 23% of MIE. $\phi = 13\%$
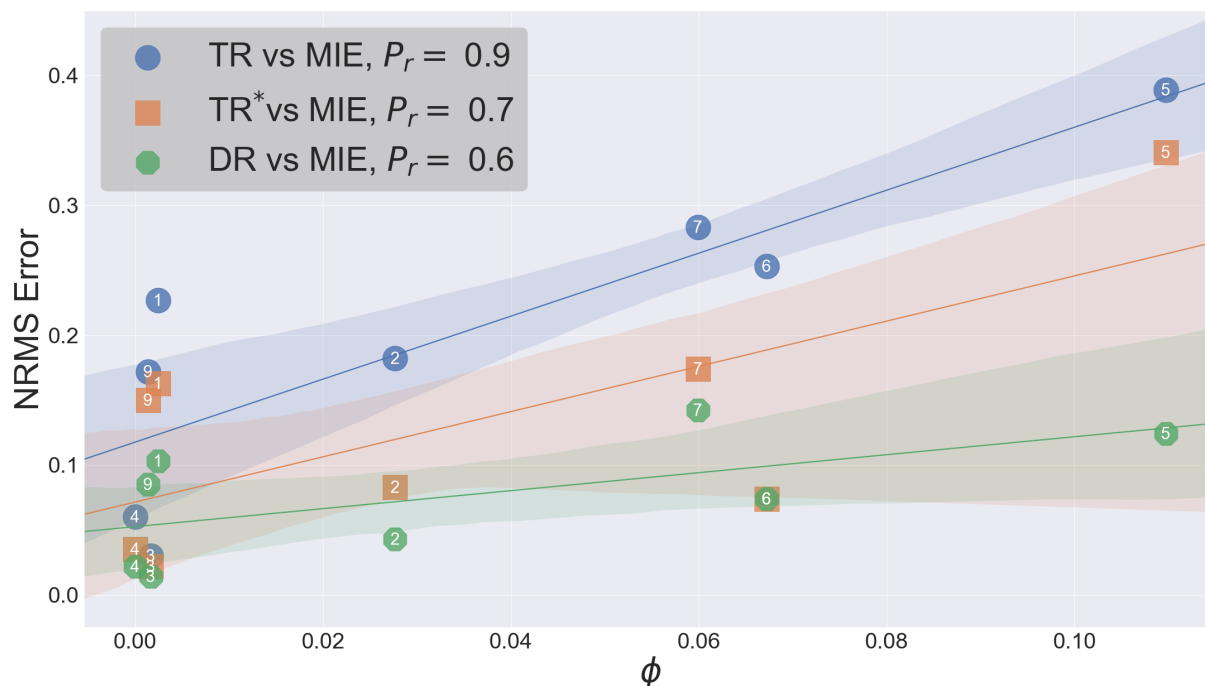
Figure 12. Correlations between the normalized root-mean-square (NRMS) error and the forecast parameter ϕ, for DR, TR and TR* estimates. The shaded areas indicate the 90% confidence region. Case ID is indicated for each data point. Pearson correlation coefficients (P_r) for the three methods are also noted on the graph. Results indicate a strong correlation between NRMS error and the ϕ parameter.

for rejected case 8 implies unsuitable conditions ($> 30\%$ uncertainty) for the application of the mass-balance technique, as the TR method failed to retrieve facility emissions ($E_C < 0$); the ϕ parameter allows a quantification of the conditions for rejection of flight data for emission rates retrieval purposes.

675   Our results indicate a strong positive correlation between ϕ and the NRMS error (strongest among the three forecast parameters), independent of the method. Figure 12 shows this correlation for the three retrieval methods, note that rejected case 8 was excluded from correlation analysis as an outlier (extreme storage event) since its inclusion would dominate the correlation and obscure the information from the other data points. Pearson correlation coefficients (P_r) for DR, TR and TR* were calculated as 0.6, 0.9 and 0.7, respectively. That is, a significant proportion of the variability in the emission retrieval NRMS error can be accounted for by each of the metrics proposed here. In aircraft-based retrievals, analysis of ϕ can provide insights on the

680   expected reliability of the retrieved emissions and the probability of storage-release events. Pre-flight model forecasts of ϕ can be used for the same purpose. For ϕ scores greater than 10%, the uncertainty in the retrieved emission rates due to storage-release is over 30%. For such conditions, it is advised that the storage rate ($E_{C,S}$) be estimated from time-consecutive screens of data. This can be achieved by conducting multiple box flights in order to estimate $E_{C,S}^*$ for improved top-down emission rate retrievals (TR*, see Sect. 2.2.3 and Sect. 3.3).

Table 1. Summary of correlations between NRMS error in TR estimates and the three forecast parameters for storage-release events, in terms of the Pearson correlation coefficient (P_r), and fit line (least squares) slopes (m) in Figures 8, 10 and 12. The last row indicates the amount of change in forecast parameters that would result in $>= 30\%$ error in estimates. $\phi < 10\%$ corresponds to NRMS error $< 30\%$.

Parameter	$|\Delta_t R_i|$	$\Delta_t \alpha_i$	ϕ
P_r correlation	0.7	0.7	0.9
$m = \frac{\Delta(NRMSe)}{\Delta(Param.)}$	0.525 [h]	0.005 [$^{\circ -1}$h]	2.4
$\frac{30\%}{m}$	0.571 [h^{-1}]	60°[h^{-1}]	0.125

4.4 Summary of predictive methods

The three forecast parameters $\Delta_t Ri$, $\Delta_t \alpha$ and ϕ as discussed above provide a new approach for analyzing the uncertainty levels associated with potential storage-release events during aircraft-based emission rate retrieval flight time, either using *a priori* meteorological forecasts to provide guidance for emission retrieval flight decision-making, or in post-processing of already completed flights. Table 1 summarizes the correlations between these parameters and the NRMS error in estimates by TR method in terms of P_r correlation coefficients and the slopes (m) of the best fit lines (least squares) in Figures 8, 10 and 12. The last row in Table 1, demonstrates the correlation between the changes in the two meteorological forecast parameters $\Delta_t Ri$ and $\Delta_t \alpha$ and change in NRMS error levels; accordingly, net changes of 0.57 h^{-1} in $|\Delta_t Ri|$ and 60° h^{-1} in $\Delta_t \alpha$ during the flight/sampling time may result in 30% (or more when both contribute) uncertainty in the estimates, the maximum estimated uncertainty (30%) for TERRA SO$_2$ emission rate retrievals in Gordon et al. (2015). The combined effects of $|\Delta_t Ri| = 0.62$ h^{-1} and $\Delta_t \alpha = 46°$ h^{-1} during the ~ 2 h flight time of case 5, contributed to a net error of 39% in TR estimates for this case. Further, for any forecasted/estimated rate of change in meteorology in terms of $\Delta_t Ri$ and $\Delta_t \alpha$, shorter sampling durations (e.g. two flights in tandem) may result in more confidence in the estimated emission rates; small values of $\Delta_t Ri$ and $\Delta_t \alpha$ correspond to steady state conditions, whereas higher values correspond to increased uncertainty in the estimates due to storage-release events. The NRMS errors in TR estimates for cases with $\phi < 10\%$ (7 out of 9), were less than 30%. $\phi > 10\%$ correspond to over 30% uncertainty in the estimates, as for rejected cases 5 and 8. Note that cases 5 and 8 were excluded (rejected) from TERRA retrievals in previous work due to unfavorable conditions for the application of TERRA. Here, using GEM-MACH predicted fields, we were able to quantify the impact of changing meteorology and/or high upwind emissions in mass-balance emission rate retrieval for these two cases in terms of the three forecast parameters $\Delta_t Ri$, $\Delta_t \alpha$ and ϕ as described above.

Table 2 provides a summary of all the meteorological conditions and emission scenarios in our GEM-MACH model simulations for the period of the nine studied JOSM 2013 cases. The flight/sampling time average $\Delta_t Ri$, $\Delta_t \alpha$ and ϕ scores are all provided in the table. Estimates by the three described methods (DR, TR and TR*) are compared to model input emissions (MIE), and the corresponding normalized mean and root-mean-square (NRMS) errors are listed for each of the studied cases.

Atmospheric
Chemistry
and Physics
Discussions

Open Access

Table 2. The table provides a summary of the meteorological conditions and source emission scenarios by the three oil sands facilities for the period of the nine JOSM 2013 box flight cases in our GEM-MACH model simulations. Flight date-time and duration is provided for each case, LT stands for local time (UT−6hr). The flight time average gradient Richardson number (R_i) and wind direction at plume height (α_{wind}) are provided. The three forecast parameters are also provided: the average rate of change in atmospheric stability represented by $\Delta_t R_i$, the average rate of change in wind direction at plume height represented by $\Delta_t \alpha$ and the weighted upwind-to-downwind concentration ratio parameter (ϕ). DR, TR and TR* method estimates are compared to model input emissions (MIE). Storage (S) and/or release (R) events, their order of occurrence (for the cases where both were present) and their relative magnitude (represented by font-size) are given. The flight time average storage rate as a fraction (%) of the total retrieved emission rates are also provided for each case. The performance of the three retrieval methods against MIE, are shown in terms of normalized (to MIE) mean and root-mean-square (RMS) errors.

Facility	Syncrude			CNRL			Suncor		
Case ID	1	2	3	4	5	6	7	8	9
Flight Date &	Aug 14	Aug 24	Sep 3	Aug 20	Aug 26	Sep 2	Aug 22	Aug 28	Aug 29
Start Time (LT)	10:12	11:50	14:58	10:30	13:43	11:43	11:45	11:18	15:01
Duration (hh:mm)	1:18	2:06	2:07	2:10	1:52	1:45	2:58	2:30	2:13
Atmos. Conditions	stable	unstable	unstable	unstable	stable	unstable	unstable	unstable	unstable
$\overline{R_i}$	0.26	-0.03	-0.66	-0.14	1.01	-0.26	-0.19	-0.44	-0.58
$\overline{\alpha}_{wind}$	200°	264°	215°	259°	243°	100°	246°	297°	55°
$\Delta_t R_i$ (h^{-1})	-0.38	-0.13	0.19	-0.25	0.62	-0.28	-0.26	-0.34	0.15
$\Delta_t \alpha$ (h^{-1})	3.3°	6.5°	1.3°	1.8°	45.8°	33.9°	2.4°	3.3°	11.8°
ϕ (%)	0.244	2.766	0.167	0.004	10.97	6.725	5.992	13.04	0.137
avg. Model Input	3423	4050	6923	2979	120	83	1467	1273	1013
Emissions (MIE)	kg/hr	kg/hr	kg/hr	kg/hr	kg/hr	kg/hr	kg/hr	kg/hr	kg/hr
Storage and/or Release	S-R	S	R-S	R-S	S	R	R-S	S	R
Net Contribution	17 %	20 %	2 %	-3 %	43 %	-29 %	13 %	156 %	-27 %
Norm. Mean Error:									
DR vs. MIE	-0.07	0.04	-0.00	-0.01	0.11	-0.00	-0.14	0.22	-0.08
TR vs. MIE	-0.22	-0.16	-0.02	0.02	-0.36	0.24	-0.25	-1.66	0.17
TR*vs. MIE	-0.11	-0.07	-0.01	-0.00	-0.26	0.05	-0.14	-0.96	0.11
Norm. RMS Error:									
DR vs. MIE	0.10	0.04	0.01	0.02	0.12	0.07	0.14	0.23	0.08
TR vs. MIE	0.23	0.18	0.04	0.06	0.39	0.25	0.28	1.66	0.17
TR*vs. MIE	0.16	0.08	0.02	0.03	0.34	0.07	0.17	1.14	0.15

Atmospheric
Chemistry
and Physics
Discussions

5 Conclusions

We have carried out a series of retrospective air quality simulations employing Environment and Climate Change Canada's
(ECCC) air quality model, Global Environmental Multiscale - Modelling Air-quality and Chemistry (GEM-MACH), for the
period of the JOSM 2013 campaign over the Athabasca oil sands region, with the primary objective of evaluating aircraft-
based mass balance emission rate retrievals such as applied in TERRA (i.e. the Top-down Emission Rate Retrieval Algorithm).
We considered the simulations of nine JOSM 2013 emission estimation flight cases, over three oil sands facilities with high
SO_2 emissions (CNRL, Syncrude and Suncor), and used GEM-MACH model 4D output data as a surrogate for aircraft mea-
surements in emission rate estimations to evaluate the application of the mass-balance approach in top-down methodologies.
Although the emphasis was on SO_2 emissions, the explored concepts and the three forecast parameters developed here are also
applicable to species with high background levels (e.g. CH_4, CO). Emission retrievals were made by analyzing GEM-MACH
model output data through three different approaches: (a) Direct Retrieval (DR) of the model input emissions (MIE) by analyz-
ing GEM-MACH model 4D data (volumetric time-series) through a comprehensive analysis of all of the processes contributing
to the change in mass within the control volume (flux box) including vertical and horizontal fluxes through box top and lateral
walls, surface deposition and rate of storage of mass within the volume of the box, (b) TERRA Retrieval (TR) approximations
by limiting our analysis to the model data on 2D screens comprised of lateral walls of the flux box surrounding the emission
source, mimicking aircraft-based retrievals during an emission estimation (box) flight, and (c) improved TERRA Retrievals
(TR*) through estimation of the rate of change in the stored mass within the volume of the flux box by analyzing the data on
time-consecutive 2D screens at each model time-step (3D).

Results were compared to the model input emissions (MIE) and the performance of the three methods was analyzed in
terms of Normalized (to MIE) Root-Mean-Square (NRMS) error. DR method estimates were within 1-23% of MIE for all 9
studied cases. TR and TR* estimates for 7 out of 9 cases were within 4-28% and 2-17% of MIE, respectively. This range of
emissions retrieval error was similar to the range ($\pm30\%$) estimated from aircraft observations in previous work (Gordon et
al, 2015), with the remaining 2 flights (rejected cases 5 and 8) excluded in previous work due to unsuitable meteorological
conditions and knowledge of upwind emissions. Estimates for the JOSM case on 26 August 2013 around the CNRL facility
(rejected case 5) were impacted by change in atmospheric stability at rate $|\Delta_t Ri| = 0.62$ h^{-1} and shift in wind direction at
plume center-point (maximum concentration) height at rate $\Delta_t \alpha = 46°$ h^{-1} during the $\sim$2 h flight time; NRMS errors for this
case in estimates by TR and TR* methods were 39% and 34%, respectively. Estimates for the case on August 28th around
the Suncor facility (rejected case 8) were impacted by high emissions (relative to source emission rates) from the upwind
oil sands facility Syncrude (avg. $\phi = 13\%$) and change in atmospheric stability at rate $|\Delta_t Ri| = 0.34$ h^{-1} during the $\sim$2.5
h flight/sampling time; NRMS errors for this case were 166% and 114% by TR and TR* methods, respectively. The under-
performance in rejected cases 5 and 8 estimates were the consequence of several different factors and conditions during the
flight time; we refer to the collective impact of these conditions as **Storage-and-Release** events. The nine studied cases were
affected by storage-release events at various degrees; ranging from -29% to 156% with a median of 13%. This range for the 7
cases that are included in past work with TERRA applications is -29% to 20% with a 2% median.

Atmospheric
Chemistry
and Physics
Discussions
Open Access
EGU

Storage-release events occur when temporal and spatial variations in meteorological conditions and/or source emission rate result in a temporary imbalance between the addition of mass to the volume of the box through source emissions and removal of mass by vertical and horizontal fluxes through box top and lateral walls and the deposition to ground surface. The transient

745 storage of the emitted mass within box volume and its later release contribute to emission rate over- and under-estimations based on the mass-balance technique. We introduced a correction term, $E_{C,S}$, in the mass-balance equation to represent the rate of storage-release. Utilizing GEM-MACH model output data, we investigated the contribution of this term to the total integrated retrieved emission rates. The storage rate term ($E_{C,S}$) contributed to over 20% of the total retrieved emissions for 5 out of 9 cases, signifying the impact of storage-release events in emission rate retrievals based on the mass-balance approach.

750 The corresponding uncertainties in estimated emission rates were quantified for each of the nine studied cases, and conditions giving rise to storage-release events were extensively explored.

We have introduced an approach for estimating the storage term as the TR* method, based on estimations of the rate of change in the stored mass within the box volume from data on time-consecutive 2D screens around the same emission source (3D data-set). We have shown that by estimating the storage term ($E_{C,S}^*$) in this way, estimates can be improved

755 by 2-52% of MIE. Three forecast parameters were introduced for predicting/identifying the probability of storage-release events: variations in atmospheric dynamic stability represented by change in gradient Richardson number ($\Delta_t Ri$), shift in the direction of transport represented by variations in wind direction at plume height ($\Delta_t \alpha$), and the weighted upwind to downwind concentration ratio (ϕ). We have demonstrated the strong correlation between these three parameters and the NRMS error in mass-balance estimates and discussed the potential of utilizing the three forecast parameters in analysis of the uncertainty

760 levels in aircraft-based top-down emission rate retrieval methodologies. Model-based forecasts of these parameters can also provide practical advice for the planning of future airborne measurement campaigns.

The investigation carried out here has resulted in the creation of quantitative measures for the extent to which storage-and release-events may impact on aircraft emissions retrieval accuracy. Retrieval uncertainties, using regional air-quality model output, were shown to be of similar magnitude to previously published values, with the exception of cases where the underlying

765 assumption of time-invariant meteorological conditions was not valid and/or where significant upwind emissions impacted the estimates. We have also devised a methodology to reduce the impact of this form of emissions retrieval error by estimating the storage rate ($E_{C,S}^*$) through the use of repeat flights, around the same facility, either with a single or multiple aircraft(s).

Appendix A

Table A1. GEM-MACH Configuration Details

Model Component	Description	Reference
Driving numerical weather prediction model	Global Environmental Multiscale (GEM), v5.0	Côté et al. (1998a,b); Girard et al. (2014)
Driving air-quality model	Global Environmental Multiscale – Modelling Air-quality and Chemistry (GEM-MACH) v2	Moran et al. (2010); Makar et al. (2015a,b)
Model grid and nesting, time stepping	North American 10km resolution parent domain provides boundary conditions for 2.5km high resolution, 64 vertical staggered hybrid level (terrain-following coordinate) Alberta/Saskatchewan domain (Figure 1a). Time stepping: Parent domain: 5 min physics, 15 min chemistry. Alberta/Saskatchewan domain: 1 min physics and 2 min chemistry.	Girard et al. (2014); Makar et al. (2015a,b)
Weather/Aerosol feedbacks	Direct effect via binary water-dry aerosol mixtures with *Mie* algorithm optical property calculations. Indirect effect via aerosols providing cloud condensation nuclei via the Abdul-Razzak and Ghan scheme.	Abdul-Razzak and Ghan (2002); Gong et al. (2015); Makar et al. (2015a,b)
Gas-phase chemistry	Acid Deposition and Oxidant Mechanism, version 2 (ADOM-II) represents gas-phase chemistry for 42 gas species, integrated using a Young and Boris solver.	Stockwell and Lurmann (1989)
Particle microphysics	Sectional approach – 8 particle species (sulphate, nitrate, ammonium, primary organic carbon, secondary organic carbon, black carbon, sea salt, crustal material), and 12 particle bins	Gong et al. (2002, 2003)
Aqueous chemistry and gas and aerosols scavenging	Cloud scavenging of gases and aerosols, aqueous phase chemistry using a Young and Boris solver (combined time-resolved and steady-state chemistry).	Gong et al. (2015)
Deposition	Gas (Robichaud scheme) and particle dry deposition (Zhang scheme) as described in Makar et al. (2018)	Makar et al. (2018)
Particle inorganic thermodynamics	Sulphate-nitrate-ammonium non-ideal (high concentration) thermodynamic equilibrium system solved using a nested iterative approach.	Makar et al. (2003)
Anthropogenic plume rise	Modified Briggs (1984) equations for calculating plume rise. CEMS stack parameters data (e.g. stack height and diameter, exit temperature, exit velocity) were used for plume rise calculations in the model simulations.	Akingunola et al. (2018); Gordon et al. (2018)
Advection and mass conservation	Chemical transport in GEM-MACH is solved utilizing an implicit semi-Lagrangian (SL) advection space-time integration scheme. The SL scheme is not inherently mass conserving, and therefore requires the use of a post advection mass conservation step (3D ILMC approach used here).	Bermejo and Conde (2002); Sørensen et al. (2013); de Grandpré et al. (2016)
Emissions data	Emissions are processed based upon the Sparse Matrix Operator Kernel Emissions – SMOKE; emissions data from hybrid Oil Sands Database. Large stack data derived from Continuous Emissions Monitoring (CEMS).	Coats (1996); Zhang et al. (2018)

Appendix B

B1 Derivations for mass-balance terms

770   The net horizontal flux exiting through lateral walls of the box ($E_{C,H}$) is calculated by extracting compound mixing ratio (χ_C), air density (ρ_{air}) and the wind normal to the lateral walls of the box from the model output, and integrating over the surface (2D screen) comprised of box lateral walls at each time-step,

$$E_{C,H}(t) = M_R \iint \chi_C(t,s,z)\rho_{air}U_\perp \, dsdz \tag{B1}$$

775   where $s(x,y)$ is the screen path coordinate, which follows the horizontal projection of the box flight path around the facility in a counter clockwise direction.

The normal wind (positive outwards) along the screen is calculated as,

$$U_\perp(t,s,z) = \frac{U\frac{ds(x,y)}{dy} - V\frac{ds(x,y)}{dx}}{\sqrt{\left(\frac{ds(x,y)}{dx}\right)^2 + \left(\frac{ds(x,y)}{dy}\right)^2}} \tag{B2}$$

where U and V are wind components towards east (x) and north (y), respectively.

780   The vertical flux through the box top can be calculated as,

$$E_{C,V}(t) = M_R \iint \chi_{C,top} \, \rho_{air,top} \, W_{top} \, dxdy \tag{B3}$$

where $\chi_{C,top}(t,x,y)$ and $W_{top}(t,x,y)$ are compound mixing ratio and the vertical wind component (positive upwards) at the box top. In the observational application of the retrieval methodology, the top of the box corresponds to an aircraft measurement altitude which is above the plume.

785   The horizontal air flux through the screen (positive outwards) is calculated from extracted air density and wind fields along the screen at time t as,

$$E_{air,H}(t) = \iint \rho_{air}(t,s,z) \, U_\perp \, dsdz \tag{B4}$$

The rate of change in the air mass (accumulation) within the box due to changing air density (compressible fluid) can be estimated as,

790
$$E_{air,M}(t) = \iiint \frac{\partial \rho_{air}(t,V)}{\partial t} dV \tag{B5}$$

The vertical air flux leaving the box (positive upwards) is calculated similar to Eq. (B3),

$$E_{air,V}(t) = \iint \rho_{air,top}(t,x,y) \, W_{top}(t,x,y) \, dxdy \tag{B6}$$

B2 Discrete Integral Expressions

The numerical integration expressions for calculating the terms in Eqs. (7), (12) and (16), using model data, are described in

795   Tables B1, B2 and B3. Throughout, the 2nd order central finite difference scheme (Jacobson, 2005) was used to numerically

solve the time derivatives ($\Delta/\Delta t$) with the residual error of order $O(\Delta t^2)$, where Δt is equal to the 2 min (1/30 h) model time-step. Surface deposition rates ($E_{C,VD}^t$) were extracted directly from the model output for the area covered by the base of the flux box for each case, to eliminate deposition estimates as a potential source of error in retrievals. Note that in Table B2, the species vertical mass flux ($E_{C,V}^t$) is calculated using the estimated vertical air mass flux ($E_{air,V}^t$) from the mass-balance equation for air (Eq. 8) and the box top average mixing ratio ($\bar{\chi}_{C,top}^t$).

Table B1. Discrete integral expressions of Eqs. (4), (5), (B1) and (B3) to be substituted in Eq. (7). Constants Δx, Δy and Δs are equal to model grid resolution in the horizontal, 2.5 km. t is the time index; i, j and k are the 3D-space indices and s is the path index around the box.

Term	Numerical Integration
$E_{C,S}^t =$	$M_R \, \Delta x \Delta y \, \sum_{i,j,k}^{n_x,n_y,n_z} \frac{\Delta \chi_{C,ijk}^t}{\Delta t} \rho_{air,ijk}^t \Delta z_k$
$E_{C,M}^t =$	$-M_R \, \Delta x \Delta y \, \sum_{i,j,k}^{n_x,n_y,n_z} \chi_{C,ijk}^t \frac{\Delta \rho_{air,ijk}^t}{\Delta t} \Delta z_k$
$E_{C,H}^t =$	$M_R \, \Delta s \, \sum_{s,k}^{n_s,n_z} \chi_{C,sk}^t \rho_{air,sk}^t U_{\perp,sk}^t \Delta z_k$
$E_{C,V}^t =$	$M_R \, \Delta x \Delta y \, \sum_{i,j}^{n_x,n_y} \chi_{C,ij}^{t,top} \rho_{air,ij}^{t,top} W_{ij}^{t,top}$

800

Table B2. Discrete integral expressions for Eqs. (B4), (8), (9) and (10) to be substituted in Eq. (12) for TR method estimations. Terms and variables are listed in the order needed for calculating the dependent terms. Δs is constant and equal to model horizontal grid resolution, 2.5 km. t and s are time and path indices, respectively. A is the box base area. Point (i_o, j_o) is at box horizontal center.

Term	Numerical Integration
$E_{air,H}^{t} =$	$\Delta s \sum_{s,k}^{n_s,n_z} \rho_{air,sk}^{t} U_{\perp,sk}^{t} \Delta z_k$
$E_{air,M}^{t} =$	$A \sum_{k}^{n_z} \frac{\Delta}{\Delta t} \rho_{air,i_o j_o k}^{t} \, \Delta z_k$
$\overline{\chi}_{C,k}^{t} =$	$\frac{1}{n_s} \sum_{s}^{n_s} \chi_{C,sk}^{t}$
$E_{C,M}^{t} =$	$-A \, M_R \sum_{k}^{n_z} \overline{\chi}_{C,k}^{t} \, \frac{\Delta}{\Delta t} \rho_{air,i_o j_o k}^{t} \, \Delta z_k$
$E_{air,V}^{t} =$	$-E_{air,M}^{t} - E_{air,H}^{t}$
$\overline{\chi}_{C,top}^{t} =$	$\frac{1}{n_s} \sum_{s}^{n_s} \chi_{C,s \, top}^{t}$
$E_{C,V}^{t} =$	$A \, M_R \, \overline{\chi}_{C,top}^{t} \, E_{air,V}^{t}$

Table B3. Discrete integral expressions for Eqs. (13), (14) and (15) to be substituted in Eq. (16) for TR* method estimations. t is the time index; s and k are the horizontal and vertical space indices along the screens, respectively. A is the box base area.

Term	Numerical Integration
$\overline{\rho}^{\,t}_{air,k} =$	$\frac{1}{n_s} \sum_s^{n_s} \rho^t_{air,sk}$
$E^t_{air,M} =$	$A \sum_k^{n_z} \frac{\Delta \overline{\rho}^{\,t}_{air,k}}{\Delta t} \Delta z_k$
$E^{t}_{C,M}{}^* =$	$-A\, M_R \sum_{sk}^{n_s, n_z} \overline{\chi}^t_{C,sk} \frac{\Delta \overline{\rho}^{\,t}_{air,sk}}{\Delta t} \Delta z_k$
$E^{t}_{C,S}{}^* =$	$A\, M_R \sum_{sk}^{n_s, n_z} \overline{\rho}^{\,t}_{air,sk} \frac{\Delta \overline{\chi}^t_{C,sk}}{\Delta t} \Delta z_k$

Author contributions. SF wrote the paper with supervision and input from MG and PAM. SF ran the GEM-MACH simulations and implemented the mass-balance retrieval algorithm into GEM-MACH. PAM coordinated the GEM-MACH work. AA setup the GEM-MACH code and provided simulation support. AD, JL, KH, and SML provided advice during analysis and contributed to paper revisions.

Competing interests. The authors have the following competing interests: John Liggio is an editor for ACP.

Atmospheric
Chemistry
and Physics
Discussions

References

American Meteorological Society, cited 2020: The Gradient Richardson number, and The Ekman spiral. Glossary of Meteorology. Available online, http://glossary.ametsoc.org/wiki/Gradient_richardson_number, 2020.

Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation 3. Sectional representation, Journal of Geophysical Research: Atmospheres, 107, AAC 1–1–AAC 1–6, https://doi.org/https://doi.org/10.1029/2001JD000483, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD000483, 2002.

Akingunola, A., Makar, P. A., Zhang, J., Darlington, A., Li, S.-M., Gordon, M., Moran, M. D., and Zheng, Q.: A chemical transport model study of plume-rise and particle size distribution for the Athabasca oil sands, Atmospheric Chemistry and Physics, 18, 8667–8688, https://doi.org/10.5194/acp-18-8667-2018, https://acp.copernicus.org/articles/18/8667/2018/, 2018.

Angevine, W. M., Peischl, J., Crawford, A., Loughner, C. P., Pollack, I. B., and Thompson, C. R.: Errors in top-down estimates of emissions using a known source, Atmospheric Chemistry and Physics, 20, 11 855–11 868, https://doi.org/10.5194/acp-20-11855-2020, https://acp.copernicus.org/articles/20/11855/2020/, 2020.

Baray, S., Darlington, A., Gordon, M., Hayden, K. L., Leithead, A., Li, S.-M., Liu, P. S. K., Mittermeier, R. L., Moussa, S. G., O'Brien, J., Staebler, R., Wolde, M., Worthy, D., and McLaren, R.: Quantification of methane sources in the Athabasca Oil Sands Region of Alberta by aircraft mass balance, Atmospheric Chemistry and Physics, 18, 7361–7378, https://doi.org/10.5194/acp-18-7361-2018, https://acp.copernicus.org/articles/18/7361/2018/, 2018.

Barchyn, T. E., Hugenholtz, C. H., Myshak, S., and Bauer, J.: A UAV-based system for detecting natural gas leaks, Journal of Unmanned Vehicle Systems, 6, 18–30, https://doi.org/10.1139/juvs-2017-0018, https://doi.org/10.1139/juvs-2017-0018, 2018.

Bermejo, R. and Conde, J.: A Conservative Quasi-Monotone Semi-Lagrangian Scheme, Monthly Weather Review, 130, 423–430, https://doi.org/10.1175/1520-0493(2002)130<0423:ACQMSL>2.0.CO;2, https://doi.org/10.1175/1520-0493(2002)130<0423:ACQMSL>2.0.CO;2, 2002.

Cambaliza, M. O. L., Shepson, P. B., Caulton, D. R., Stirm, B., Samarov, D., Gurney, K. R., Turnbull, J., Davis, K. J., Possolo, A., Karion, A., Sweeney, C., Moser, B., Hendricks, A., Lauvaux, T., Mays, K., Whetstone, J., Huang, J., Razlivanov, I., Miles, N. L., and Richardson, S. J.: Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions, Atmospheric Chemistry and Physics, 14, 9029–9050, https://doi.org/10.5194/acp-14-9029-2014, https://acp.copernicus.org/articles/14/9029/2014/, 2014.

Cheng, Y., Li, S.-M., Liggio, J., Gordon, M., Darlington, A., Zheng, Q., Moran, M., Liu, P., and Wolde, M.: Top-Down Determination of Black Carbon Emissions from Oil Sand Facilities in Alberta, Canada Using Aircraft Measurements, Environmental Science & Technology, 54, 412–418, https://doi.org/10.1021/acs.est.9b05522, https://doi.org/10.1021/acs.est.9b05522, 2020.

Coats, Jr, C. J.: High performance algorithms in the Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system, 1996.

Conley, S., Faloona, I., Mehrotra, S., Suard, M., Lenschow, D. H., Sweeney, C., Herndon, S., Schwietzke, S., Pétron, G., Pifer, J., Kort, E. A., and Schnell, R.: Application of Gauss's theorem to quantify localized surface emissions from airborne measurements of wind and trace gases, Atmospheric Measurement Techniques, 10, 3345–3358, https://doi.org/10.5194/amt-10-3345-2017, https://amt.copernicus.org/articles/10/3345/2017/, 2017.

Côté, J., Desmarais, J.-G., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The Operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part II: Results, Monthly Weather Review, 126, 1397–1418, https://doi.org/10.1175/1520-0493(1998)126<1397:TOCMGE>2.0.CO;2, https://doi.org/10.1175/1520-0493(1998)126<1397:TOCMGE>2.0.CO;2, 1998a.

Atmospheric
Chemistry
and Physics
Discussions

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The Operational CMC–MRB Global Environmental Multiscale

845   (GEM) Model. Part I: Design Considerations and Formulation, Monthly Weather Review, 126, 1373–1395, https://doi.org/10.1175/1520-
0493(1998)126<1373:TOCMGE>2.0.CO;2, https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2, 1998b.

de Grandpré, J., Tanguay, M., Qaddouri, A., Zerroukat, M., and McLinden, C. A.: Semi-Lagrangian Advection of Stratospheric Ozone on a
Yin–Yang Grid System, Monthly Weather Review, 144, 1035–1050, https://doi.org/10.1175/MWR-D-15-0142.1, https://doi.org/10.1175/
MWR-D-15-0142.1, 2016.

850   Elston, J., Argrow, B., Stachura, M., Weibel, D., Lawrence, D., and Pope, D.: Overview of Small Fixed-Wing Unmanned Aircraft for
Meteorological Sampling, Journal of Atmospheric and Oceanic Technology, 32, 97–115, https://doi.org/10.1175/JTECH-D-13-00236.1,
https://doi.org/10.1175/JTECH-D-13-00236.1, 2015.

Fathi, S.: Evaluating the Top-down Emission Rate Retrieval Algorithm (TERRA) Using Virtual Aircraft-based Sampling Within the GEM-
MACH Model, Master's thesis, York University, http://hdl.handle.net/10315/34547, 2017.

855   Girard, C., Plante, A., Desgagné, M., McTaggart-Cowan, R., Côté, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch,
M., Spacek, L., Tanguay, M., Vaillancourt, P. A., and Zadra, A.: Staggered Vertical Discretization of the Canadian Environmen-
tal Multiscale (GEM) Model Using a Coordinate of the Log-Hydrostatic-Pressure Type, Monthly Weather Review, 142, 1183–1196,
https://doi.org/10.1175/MWR-D-13-00255.1, https://doi.org/10.1175/MWR-D-13-00255.1, 2014.

Gong, S. L., Barrie, L. A., and Lazare, M.: Canadian Aerosol Module (CAM): A size-segregated simulation of atmospheric aerosol pro-
860   cesses for climate and air quality models 2. Global sea-salt aerosol and its budgets, Journal of Geophysical Research: Atmospheres, 107,
AAC 13–1–AAC 13–14, https://doi.org/https://doi.org/10.1029/2001JD002004, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/
2001JD002004, 2002.

Gong, S. L., Barrie, L. A., Blanchet, J.-P., von Salzen, K., Lohmann, U., Lesins, G., Spacek, L., Zhang, L. M., Girard, E., Lin, H., Leaitch,
R., Leighton, H., Chylek, P., and Huang, P.: Canadian Aerosol Module: A size-segregated simulation of atmospheric aerosol processes
865   for climate and air quality models 1. Module development, Journal of Geophysical Research: Atmospheres, 108, AAC 3–1–AAC 3–16,
https://doi.org/https://doi.org/10.1029/2001JD002002, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD002002, 2003.

Gong, W., Makar, P., Zhang, J., Milbrandt, J., Gravel, S., Hayden, K., Macdonald, A., and Leaitch, W.: Modelling
aerosol–cloud–meteorology interaction: A case study with a fully coupled air quality model (GEM-MACH), Atmospheric Environ-
ment, 115, 695 – 715, https://doi.org/https://doi.org/10.1016/j.atmosenv.2015.05.062, http://www.sciencedirect.com/science/article/pii/
870   S1352231015301333, 2015.

Gordon, M., Li, S.-M., Staebler, R., Darlington, A., Hayden, K., O'Brien, J., and Wolde, M.: Determining air pollutant emission rates
based on mass balance using airborne measurement data over the Alberta oil sands operations, Atmospheric Measurement Techniques, 8,
3745–3765, https://doi.org/10.5194/amt-8-3745-2015, http://www.atmos-meas-tech.net/8/3745/2015/, 2015.

Gordon, M., Makar, P. A., Staebler, R. M., Zhang, J., Akingunola, A., Gong, W., and Li, S.-M.: A comparison of plume rise algorithms to stack
875   plume measurements in the Athabasca oil sands, Atmospheric Chemistry and Physics, 18, 14 695–14 714, https://doi.org/10.5194/acp-18-
14695-2018, https://acp.copernicus.org/articles/18/14695/2018/, 2018.

Jacobson, M. Z.: Fundamentals of Atmospheric Modeling, Cambridge University Press, 2 edn., https://doi.org/10.1017/CBO9781139165389,
2005.

JOSM: Joint Oil Sands Monitoring Plan, Integrated Monitoring Plan for the Oil Sands, Air Quality Component, p. 72, http://publications.gc.
880   ca/site/eng/394253/publication.html, 2011.

Kalthoff, N., Corsmeier, U., Schmidt, K., Kottmeier, C., Fiedler, F., Habram, M., and Slemr, F.: Emissions of the city of Augsburg determined using the mass balance method, Atmospheric Environment, 36, Supplement 1, 19 – 31, https://doi.org/http://dx.doi.org/10.1016/S1352-2310(02)00215-7, http://www.sciencedirect.com/science/article/pii/S1352231002002157, evaluation of Modeled Emission Inventories of Ozone Precursors. A Case Study for an Urban Area (Augsburg, Germany), 2002.

885 Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R., Tans, P., Trainer, M., Zamora, R., and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophysical Research Letters, 40, 4393–4397, https://doi.org/https://doi.org/10.1002/grl.50811, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50811, 2013.

Karion, A., Sweeney, C., Kort, E. A., Shepson, P. B., Brewer, A., Cambaliza, M., Conley, S. A., Davis, K., Deng, A., Hardesty, M., Herndon, 890 S. C., Lauvaux, T., Lavoie, T., Lyon, D., Newberger, T., Pétron, G., Rella, C., Smith, M., Wolter, S., Yacovitch, T. I., and Tans, P.: Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region, Environmental Science & Technology, 49, 8124–8131, https://doi.org/10.1021/acs.est.5b00217, https://doi.org/10.1021/acs.est.5b00217, pMID: 26148550, 2015.

Karion, A., Lauvaux, T., Lopez Coto, I., Sweeney, C., Mueller, K., Gourdji, S., Angevine, W., Barkley, Z., Deng, A., Andrews, A., Stein, A., and Whetstone, J.: Intercomparison of atmospheric trace gas dispersion models: Barnett Shale case study, Atmospheric Chemistry and 895 Physics, 19, 2561–2576, https://doi.org/10.5194/acp-19-2561-2019, https://acp.copernicus.org/articles/19/2561/2019/, 2019.

Li, S.-M., Leithead, A., Moussa, S. G., Liggio, J., Moran, M. D., Wang, D., Hayden, K., Darlington, A., Gordon, M., Staebler, R., Makar, P. A., Stroud, C. A., McLaren, R., Liu, P. S. K., O'Brien, J., Mittermeier, R. L., Zhang, J., Marson, G., Cober, S. G., Wolde, M., and Wentzell, J. J. B.: Differences between measured and reported volatile organic compound emissions from oil sands facilities in Alberta, Canada, Proceedings of the National Academy of Sciences, 114, E3756–E3765, https://doi.org/10.1073/pnas.1617862114, https://www. 900 pnas.org/content/114/19/E3756, 2017.

Li, X.-B., Wang, D., Lu, Q.-C., Peng, Z.-R., Fu, Q., Hu, X.-M., Huo, J., Xiu, G., Li, B., Li, C., Wang, D.-S., and Wang, H.: Three-dimensional analysis of ozone and PM2.5 distributions obtained by observations of tethered balloon and unmanned aerial vehicle in Shanghai, China, Stochastic Environmental Research and Risk Assessment, 32, 1189–1203, https://doi.org/10.1007/s00477-018-1524-2, https://doi.org/10.1007/s00477-018-1524-2, 2018.

905 Liggio, J., Li, S.-M., Staebler, R. M., Hayden, K., Darlington, A., Mittermeier, R. L., O'Brien, J., McLaren, R., Wolde, M., Worthy, D., and Vogel, F.: Measured Canadian oil sands CO2 emissions are higher than estimates made using internationally recommended methods, Nature Communications, 10, 1863, https://doi.org/10.1038/s41467-019-09714-9, https://doi.org/10.1038/s41467-019-09714-9, 2019.

Mailhot, J. and Benoit, R.: A Finite-Element Model of the Atmospheric Boundary Layer Suitable for Use with Numerical Weather Prediction Models, Journal of Atmospheric Sciences, 39, 2249 – 2266, https://doi.org/10.1175/1520-0469(1982)039<2249:AFEMOT>2.0.CO;2, 910 https://journals.ametsoc.org/view/journals/atsc/39/10/1520-0469_1982_039_2249_afemot_2_0_co_2.xml, 1982.

Makar, P., Bouchet, V., and Nenes, A.: Inorganic chemistry calculations using HETV—a vectorized solver for the SO42-–NO3-–NH4+ system based on the ISORROPIA algorithms, Atmospheric Environment, 37, 2279–2294, https://doi.org/https://doi.org/10.1016/S1352-2310(03)00074-8, https://www.sciencedirect.com/science/article/pii/S1352231003000748, 2003.

Makar, P., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baró, R., Bianconi, R., Cheung, P., 915 Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M., Pabla, B., Pérez, J., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, part 2: Effects on chemistry, Atmospheric Environment, 115, 499 – 526, https://doi.org/https://doi.org/10.1016/j.atmosenv.2014.10.021, http://www.sciencedirect.com/science/article/pii/S1352231014008103, 2015a.

Makar, P., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Im, U., Balzarini, A., Baró, R., Bianconi, R., Cheung, P.,
920   Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M., Pabla, B., Pérez, J., Pirovano, G., San
       José, R., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, Part 1: Effects on weather,
       Atmospheric Environment, 115, 442 – 469, https://doi.org/https://doi.org/10.1016/j.atmosenv.2014.12.003, http://www.sciencedirect.com/
       science/article/pii/S1352231014009510, 2015b.

Makar, P. A., Stroud, C., Zhang, J., Moran, M., Akingunola, A., Gong, W., Gravel, S., Pabla, B., Cheung, P., Zheng, Q., Marson, G., Li,
925   S. M., Brook, J., Hayden, K., Liggio, J., Staebler, R., and Darlington, A.: High Resolution Model Simulations of the Canadian Oil Sands
       with Comparisons to Field Study Observations, pp. 503–508, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-
       319-24478-5_80, http://dx.doi.org/10.1007/978-3-319-24478-5_80, 2016.

Makar, P. A., Akingunola, A., Aherne, J., Cole, A. S., Aklilu, Y.-A., Zhang, J., Wong, I., Hayden, K., Li, S.-M., Kirk, J., Scott, K.,
       Moran, M. D., Robichaud, A., Cathcart, H., Baratzedah, P., Pabla, B., Cheung, P., Zheng, Q., and Jeffries, D. S.: Estimates of ex-
930   ceedances of critical loads for acidifying deposition in Alberta and Saskatchewan, Atmospheric Chemistry and Physics, 18, 9897–9927,
       https://doi.org/10.5194/acp-18-9897-2018, https://acp.copernicus.org/articles/18/9897/2018/, 2018.

Mays, K. L., Shepson, P. B., Stirm, B. H., Karion, A., Sweeney, C., and Gurney, K. R.: Aircraft-Based Measurements of the Carbon Foot-
       print of Indianapolis, Environmental Science & Technology, 43, 7816–7823, https://doi.org/10.1021/es901326b, https://doi.org/10.1021/
       es901326b, pMID: 19921899, 2009.

935 Mölders, N., Butwin, M., Madden, J., Tran, H., Sassen, K., and Kramm, G.: Theoretical Investigations on Mapping Mean Distributions of
       Particulate Matter, Inert, Reactive, and Secondary Pollutants from Wildfires by Unmanned Air Vehicles (UAVs), Open Journal of Air
       Pollution, 04, 149—-174, https://doi.org/10.4236/ojap.2015.43014, 2015.

Moran, M., Ménard, S., Talbot, D., Huang, P., Makar, P., Gong, w., Landry, H., Gravel, S., Gong, S., Crevier, L.-P., Kallaur, A., and Sassi,
       M.: Particulate-matter forecasting with GEM-MACH15, a new Canadian air-quality forecast model, Air Pollution Modelling and Its
940   Application XX, pp. 289–292M, 2010.

Nathan, B. J., Golston, L. M., O'Brien, A. S., Ross, K., Harrison, W. A., Tao, L., Lary, D. J., Johnson, D. R., Covington, A. N., Clark, N. N.,
       and Zondlo, M. A.: Near-Field Characterization of Methane Emission Variability from a Compressor Station Using a Model Aircraft,
       Environmental Science & Technology, 49, 7896–7903, https://doi.org/10.1021/acs.est.5b00705, https://doi.org/10.1021/acs.est.5b00705,
       pMID: 26011292, 2015.

945 Panitz, H.-J., Nester, K., and Fiedler, F.: Mass budget simulation of NOx and CO for the evaluation of calculated emissions for the
       city of Augsburg (Germany), Atmospheric Environment, 36, Supplement 1, 33 – 51, https://doi.org/http://dx.doi.org/10.1016/S1352-
       2310(02)00216-9, http://www.sciencedirect.com/science/article/pii/S1352231002002169, evaluation of Modeled Emission Inventories of
       Ozone Precursors. A Case Study for an Urban Area (Augsburg, Germany), 2002.

Ryoo, J.-M., Iraci, L. T., Tanaka, T., Marrero, J. E., Yates, E. L., Fung, I., Michalak, A. M., Tadić, J., Gore, W., Bui, T. P., Dean-Day,
950   J. M., and Chang, C. S.: Quantification of CO_2 and CH_4 emissions over Sacramento, California, based on divergence theorem using
       aircraft measurements, Atmospheric Measurement Techniques, 12, 2949–2966, https://doi.org/10.5194/amt-12-2949-2019, https://www.
       atmos-meas-tech.net/12/2949/2019/, 2019.

Slemr, F., Friedrich, R., and Seiler, W.: The research project EVA - General objectives and main results, Atmospheric Environment, 36, 1–6,
       www.scopus.com, cited By :4, 2002.

955   Sørensen, B., Kaas, E., and Korsholm, U. S.: A mass-conserving and multi-tracer efficient transport scheme in the online integrated Enviro-HIRLAM model, Geoscientific Model Development, 6, 1029–1042, https://doi.org/10.5194/gmd-6-1029-2013, https://gmd.copernicus.org/articles/6/1029/2013/, 2013.

Stockwell, W. and Lurmann, F.: Intercomparison of the ADOM and RADM Gas-Phase Chemical Mechanisms, 1989.

Tadić, J. M., Michalak, A. M., Iraci, L., Ilić, V., Biraud, S. C., Feldman, D. R., Bui, T., Johnson, M. S., Loewenstein, M., Jeong, S.,
960   Fischer, M. L., Yates, E. L., and o, J.-M.: Elliptic Cylinder Airborne Sampling and Geostatistical Mass Balance Approach for Quantifying Local Greenhouse Gas Emissions, Environmental Science & Technology, 51, 10 012–10 021, https://doi.org/10.1021/acs.est.7b03100, https://doi.org/10.1021/acs.est.7b03100, pMID: 28727429, 2017.

Turnbull, J. C., Karion, A., Fischer, M. L., Faloona, I., Guilderson, T., Lehman, S. J., Miller, B. R., Miller, J. B., Montzka, S., Sherwood, T., Saripalli, S., Sweeney, C., and Tans, P. P.: Assessment of fossil fuel carbon dioxide and other anthropogenic trace gas
965   emissions from airborne measurements over Sacramento, California in spring 2009, Atmospheric Chemistry and Physics, 11, 705–721, https://doi.org/10.5194/acp-11-705-2011, https://acp.copernicus.org/articles/11/705/2011/, 2011.

Zhang, J., Moran, M. D., Zheng, Q., Makar, P. A., Baratzadeh, P., Marson, G., Liu, P., and Li, S.-M.: Emissions preparation and analysis for multiscale air quality modeling over the Athabasca Oil Sands Region of Alberta, Canada, Atmospheric Chemistry and Physics, 18, 10 459–10 481, https://doi.org/10.5194/acp-18-10459-2018, https://acp.copernicus.org/articles/18/10459/2018/, 2018.