Atmospheric transport simulations in support of the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE)*

J. M. Henderson[1], J. Eluszkiewicz[1], M. E. Mountain[1], T. Nehrkorn[1], R. Y.-W. Chang[2], A. Karion[3], J. B. Miller[3], C. Sweeney[3], N. Steiner[4], S. C. Wofsy[2], and C. E. Miller[5]

[1]{Atmospheric and Environmental Research, Lexington, MA, USA}

[2]{Harvard University, Cambridge, MA, USA}

[3]{NOAA Earth Systems Research Laboratory, Global Monitoring Division, Boulder, CO, USA}

[4]{The City College of New York, New York, NY, USA}

[5]{Jet Propulsion Laboratory, Pasadena, CA, USA}

Correspondence to: J. M. Henderson (jhenders@aer.com)

*The authors wish to dedicate this article to the memory of their co-author Dr. Janusz Eluszkiewicz, who died 27 May 2014.

Abstract

This paper describes the atmospheric modeling that underlies the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE) science analysis, including its meteorological and atmospheric transport components [Polar variant of the Weather Research and Forecasting (WRF) and Stochastic Time Inverted Lagrangian Transport (STILT) models], and provides WRF validation for May-October 2012 and March-November 2013 - the first two years of the aircraft field campaign. A triply nested computational domain for WRF was chosen so that the innermost domain with 3.3-km grid spacing encompasses the entire mainland of Alaska and enables the substantial orography of the state to be represented by the underlying high-resolution topographic input field. Summary statistics of the WRF model performance on the 3.3-km grid indicate good overall agreement with quality-controlled surface and radiosonde observations. Two-meter temperatures are generally too cold by approximately 1.4 K in 2012 and 1.1 K in 2013, while 2-m dewpoint temperatures are too low (dry) by 0.2 K in 2012 and too high (moist) by 0.6 K in 2013. Wind speeds are biased too low by 0.2 m s^{-1} in 2012 and 0.3 m s^{-1} in 2013. Model representation of upper level variables is very good. These measures are comparable to model performance metrics of similar model configurations found in the literature. The high quality of these fine-resolution WRF meteorological fields inspires confidence in their use to drive STILT for the purpose of computing surface-influences ("footprints") at commensurably increased resolution. Indeed, footprints generated on a 0.1-degree grid show increased spatial detail compared with those on the more common 0.5-degree grid, lending itself better for convolution with flux models for carbon dioxide and methane across the heterogeneous Alaskan landscape. Ozone deposition rates computed using STILT footprints indicate good agreement with observations and exhibit realistic seasonal variability, further indicating that WRF-STILT footprints are of high quality and will support accurate estimates of CO_2 and CH_4 surface-atmosphere fluxes using CARVE observations.

1 Introduction

Increased concentrations of greenhouse gases (GHGs), including carbon dioxide (CO_2) and methane (CH_4), are warming the atmosphere (IPCC, 2013). The Arctic exhibits amplified signs of this warming, with unprecedented changes over the past two decades, including warmer surface temperatures and a moistening of the Arctic boundary layer (Cohen et al., 2012), as well as a dramatic decline in sea ice extent in late summer and early fall (Stroeve et al., 2012). The interplay between these processes creates feedbacks that further amplify environmental change, e.g., a decrease in sea ice reduces the surface albedo and increases latent and sensible heat fluxes into the atmosphere, resulting in warmer surface temperatures relative to mid-latitudes (the so-called Polar Amplification, Manabe and Stouffer, 1980; Holland and Bitz, 2003; ACIA, 2004; Serreze and Barry, 2011). These changes in the Arctic also affect hemispheric flow patterns (e.g., Liu et al., 2012), with detection and attribution of these influences (e.g., Screen et al., 2014; Barnes, 2013; Coumou et al., 2014; Hines et al., 2014) gaining important attention, especially in light of the many extremes of 2012 (Peterson et al., 2013)

A warmer and moister Arctic is expected to cause carbon (CO_2 and CH_4) releases from the vast shallow continental and marine megapool reservoirs (ACIA, 2004), but conclusive evidence for these releases is yet lacking (Bergamaschi et al., 2013). This creates the need for continuous monitoring and modeling of concentrations and surface emissions of GHGs, leading to a comprehensive quantification and process-based understanding of the Arctic carbon budget that ultimately will enable prescient discussion and action in this vulnerable and strategically important region. Responding to this need, the objective of the NASA Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE) is to "quantify correlations between atmospheric and surface state variables for the Alaskan terrestrial ecosystems through intensive seasonal aircraft campaigns, ground-based observations, and analysis sustained over a 5-year mission" (Miller and Dinardo, 2012). The aircraft campaigns obtain measurements during the spring through autumn months, while instrumented towers provide year-round observations. As described in the companion papers by Miller et al. (in preparation) and Karion et al. (in preparation), the overall approach to achieving the science objectives of CARVE is two-pronged: 1) Direct fieldwork using aircraft and ground measurements of atmospheric GHG concentrations, and 2) "Top-down" (inverse) estimates of surface fluxes using concentration measurements as inputs. The basic components of the

inverse method include measurements of atmospheric CH_4 and CO_2 (dry air mole fractions), an "invertible" transport model computing the sensitivity of the measurements to fluxes in the upwind source regions ("footprints"), and a priori flux models to be optimized by minimizing the model-data mismatch. CARVE also measures a suite of environmental variables to evaluate and improve the transport model, develop empirical flux models, and enable a physically based flux aggregation across spatial scales. This paper presents the atmospheric transport model that underpins the CARVE GHG flux inversion work, with a focus on two aspects: 1) Validation against observations of the meteorological fields driving the transport model, and 2) Demonstration of the benefits of high-resolution transport modeling for simulating concentrations of GHGs measured by CARVE and the collaborating tower sites.

The organization of this paper is as follows: Section 2 describes the STILT model that is used for atmospheric transport. Section 3 describes a customized version of the WRF numerical weather prediction (NWP) model that provides meteorological input fields to drive STILT. Section 4 describes WRF model verification for both surface and upper-air variables during the 2012 and 2013 CARVE campaigns. Examples of the impact of model resolution on the representation of meteorology in the WRF model and on STILT footprint calculations are described in Section 5. Conclusions are drawn in Section 6.

2 Atmospheric Transport Model

A central goal of CARVE is to infer surface-atmosphere fluxes of CH_4 and CO_2 from space- and time-varying atmospheric concentration measurements. This problem may be solved using inverse approaches ranging in complexity from simple scaling (Press et al., 1992) to formal Bayesian or geostatistical inversions (e.g., Matross et al., 2006; Gourdji et al., 2012; Miller et al., 2013). At the heart of flux inversions is the availability of an accurate transport model that is capable of computing the sensitivity of atmospheric concentrations to surface fluxes upwind. In an Eulerian (gridded) approach, the inverse method involves the calculation or approximation of model adjoints, a complex and demanding approach that must be performed for each version of the non-linear model, or the computationally intensive calculation of "basis functions" or ensembles of forward model runs for a set of predefined aggregated fluxes. Because of numerical diffusion (e.g., Eluszkiewicz et al., 2000), the Eulerian approach has inherent limitations in dealing with localized (subgrid-scale) sources, such as lakes, and in situ measurements, such as made from aircraft or towers. These

shortcomings are addressed by applying a Lagrangian Particle Dispersion Model (LPDM), which is computationally more flexible when applied to a shared computational resource. In an LPDM, atmospheric dispersion is simulated by advecting tracer particles (500 per receptor in these transport simulations) by the three-dimensional gridded wind field from an NWP model, plus a turbulent velocity component represented as a stochastic process (Markov chain). The inclusion of both the mean and stochastic wind components (Uliasz, 1994) sets LPDMs apart from trajectory models that only employ mean winds and thus cannot simulate dispersion or surface interactions (Stohl, 1998; Fuelberg et al., 2010).

When applied backward in time from a measurement location (receptor), the LPDM creates the adjoint of the transport model ("footprint"). The footprint, with units of mixing ratio/(micromole m^{-2} s^{-1}), quantifies the influence of upwind surface fluxes on concentrations measured at the receptor and is computed by counting the number of particles in a surface-influenced volume (defined as the lower half of the PBL) and the time spent in that volume (Lin et al., 2003). When multiplied by an a priori flux field (units of micromole m^{-2} s^{-1}), the footprint gives the associated contribution to the mixing ratio (units of ppm) measured at the receptor. Lagrangian methods minimize numerical diffusion, and through coupling to a mesoscale weather model, meteorological realism and mass conservation are achieved. These aspects enable the Lagrangian approach to compute realistic surface fluxes and their uncertainties for measurements from a variety of platforms including towers, aircraft, and satellites. Furthermore, the collection of footprints form a library of sensitivity functions that can be applied to a variety of trace gases without the need to rerun.

The flux inversion work for CARVE relies on the Stochastic Time-Inverted Lagrangian Transport (STILT) model (Lin et al., 2003, www.stilt-model.org), an LPDM rigorously tested by Hegarty et al. (2013) and widely used in regional GHG flux inversions (e.g., Zhao et al., 2009; Kort et al., 2008; Kort et al., 2010; Gourdji et al., 2012; Miller et al., 2012; Miller et al., 2013; McKain et al., 2012). STILT is an enhanced version of the NOAA Air Resources Laboratory's HYSPLIT model (Draxler and Hess, 1998) that combines the powerful features of HYSPLIT, such as the ability to make optimal use of highly-nested, high-resolution meteorological input fields from a large number of data sources, with enhancements aimed at mass conservation, a critical consideration for inversion work. These enhancements include a reflection/transmission scheme and the use of customized meteorological fields (time-

averaged mass fluxes and convective mass fluxes) in the dispersion calculations (Lin et al., 2003; Nehrkorn et al., 2010). Studies by Brioude et al. (2012) and Hegarty et al. (2013) have confirmed the benefits of using the customized meteorological output for LPDM computations, particularly in complex terrain. A sample comparison between STILT footprints and HYSPLIT trajectories (Fig. 1) shows general agreement between footprints and trajectories, but, in general, the convolution of footprints with surface-atmosphere flux models reveals details in flux distributions not captured by mean trajectories. Furthermore, the footprints are flux- and species-independent and can be efficiently applied to different flux models and species and incorporated into formal inversion frameworks. In all these aspects, our transport modeling for CARVE goes beyond previous airborne campaign meteorological and transport modeling that was predominantly focused on studying the origin of air masses through forward computation of mean trajectories or particle dispersion for selected species (e.g., Fuelberg et al., 2010 for the Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission).

3 WRF Mesoscale Model

3.1 WRF v3.4.1 baseline simulations for 2012 campaign

The STILT runs for CARVE carbon flux inversions are driven by customized meteorological fields from the Advanced Research version of the Weather Research and Forecasting (WRF) model (ARW; Skamarock et al., 2008). The atmospheric model was configured to generate high-quality, high-resolution meteorological fields over Arctic and boreal Alaska. This is a region of challenging atmosphere-ocean-cryosphere interactions that is subjected to numerous unique physical processes (e.g., Vihma et al., 2013). The WRF fields form the starting point of many lines of research by the CARVE team and were made available in a timely manner through use of NASA supercomputer resources. Initially (see Section 3.2 for additional runs), for the 2012 campaign, we used Polar version 3.4.1 of the WRF model (for brevity, hereafter, simply WRF), with features, primarily related to the Noah land surface model (Chen and Dudhia, 2001), optimized for polar applications (Wilson et al., 2011). The fields generated by WRF v3.4.1 were used by the CARVE team in their subsequent 2012 analyses (see Section 5.3), and thus our analysis of model performance in this paper will focus on output from this configuration.

Placement of the modeling domains, and specification of STILT receptor locations, was dictated mainly by the locations of the 2012 CARVE aircraft flight tracks (Fig. 2a) over mainland Alaska. The 2013 aircraft flight tracks (Fig. 2b) show a similar geographical extent. A triply nested computational domain on a polar stereographic grid (Fig. 2c) was chosen so that the innermost domain with 3.3-km grid spacing encompasses the entire mainland of Alaska and enables the substantial orography of the state to be represented by the underlying topographic input field. The edges of this domain are positioned distant from the location of the CARVE aircraft flights to avoid deleterious model domain edge effects. Domain 2 is similarly positioned so that the domain edges are a considerable distance from any tower locations, while domain 1 is sufficiently large for effects from the lateral boundary conditions to be minimized.

Table 1 summarizes the WRF v3.4.1 model configuration and physics options used for the baseline 2012 simulations. Our choice of major physics options follows the WRF configuration selected for the Arctic System Reanalysis (ASR; Bromwich et al., 2012), except that we used the ensemble Grell-Devenyi cumulus parameterization that is coupled with STILT. Clouds were parameterized in all three domains. Soil moisture and a binary (i.e., open water versus complete ice cover) sea ice field from the NCAR/NCEP Reanalysis product (Kalnay et al., 1996) were applied in our simulations. The default sea ice thickness of 3 m and snow depth on sea ice of 0.05 m were used in the Noah land surface model.

CARVE 2012 science flights took place between 23 May and 1 October (Alaska time). WRF runs were generated for 10 May through 2 October, 2012 (a total of 146 days), thus accommodating back trajectory simulations for measurements starting on 23 May. Meteorological fields for the entire period were formed from daily 30-h WRF runs that were initiated at 0000 UTC every 24 h with the first 6 h of each simulation dropped to avoid spin-up errors. That is, the WRF fields of forecast lengths 6-30 h from individual model simulations are retained for each 0600 to 0600 UTC period. The 0000 to 0600 UTC period is thus represented by the 24- to 30-h forecast fields from the prior day's simulation. This approach was found to enhance STILT accuracy and computational efficiency (Nehrkorn et al., 2010). The practice of using multiple short-length runs mitigates model drift and accumulation of errors versus long-term continuous runs (Lo et al., 2008). The use of 30-h simulations also minimizes the magnitude of the seams between model runs compared with longer runs. The boundary and initial conditions are derived from the NASA Modern Era

Retrospective-analysis for Research and Applications (MERRA, Rienecker et al., 2011). Grid nudging above the boundary layer every 3 h in domain 1 further prevents model drift. These initial WRF v3.4.1 results for 2012 represent the baseline against which later efforts aimed at optimizing the WRF performance through numerical and physics options are evaluated.

3.2 WRF v3.5.1 simulations for 2012 and 2013 campaigns

The preliminary meteorological fields from WRF v3.4.1 described above were ingested by STILT. An example of CARVE chemical analysis facilitated by these data is presented in Section 5.3. Subsequently, to improve the scientific fidelity of the CARVE modeling efforts, WRF simulations were repeated for the 2012 campaign, and then extended to the 2013 CARVE campaign, using the Polar variant of WRF v3.5.1. The modeling period for 2013 was expanded in time to accommodate the CARVE 2013 science flights that took place between 2 April and 28 October (Alaska time), and a subset of observations from the CARVE tower in Fox, AK, and other existing towers. WRF runs for 2013 were generated for 1 March to 30 November (a total of 275 days).

Using v3.5.1 of WRF, more rigorous implementation of cryospheric fields was enabled. The fractional MERRA snow cover field was used over land. Over water bodies in our modeling domains, daily PIOMAS v2.1 ice thickness and depth of snow cover over ice (Zhang and Rothrock, 2003) supplemental datasets were implemented as part of the model preprocessing following Hines et al. (2014). Sea ice thickness in the Noah land surface model was restricted within 0.1 to 10 m, while snow depth over ice was restricted within 0.001 to 1 m. Sea ice albedo was prescribed following Hines et al. (2014) to vary by Julian Day and latitude. The specified v3.5.1 model configuration, otherwise, was unchanged from that of the baseline v3.4.1 and software refinements and corrections inherent to any model update are not anticipated to have significant effect.

4 WRF Verification

In this section we evaluate domain 3, the innermost WRF domain with 3.3-km grid spacing, against surface and upper air observations to quantify the accuracy of the WRF meteorological fields. This high-resolution domain is expected to generate more realistic meteorological features than conventional meso- and global-scale atmospheric NWP models

that utilize horizontal grids at least an order of magnitude larger. The two-way grid configuration that we adopted permits feedback from domain 3 to the coarser resolution domains. While this results in more realistic simulations, it does not allow an independent evaluation of model performance with respect to grid resolution in the outer domains; they are consequently excluded from the current verification.

The meteorological evaluation of our WRF runs was performed using the Model Evaluation Toolkit (MET, version 4.1). MET is the official WRF validation software maintained by the NCAR Developmental Testbed Center (DTC, http://www.dtcenter.org/met/users/) and is tailored to ingest a variety of observation types available from the NCAR Research Data Archive dataset 337.0; http://rda.ucar.edu/datasets/ds337.0/) that subsequently is used in the NCEP Global Data Assimilation System. MET interpolates meteorological fields to observation locations to form matched pairs for a range of variables, including temperature, dewpoint, relative humidity, specific humidity, horizontal wind speed, and horizontal wind components. Extensive quality control was performed to identify unphysical observations and remove stations that do not meet data availability thresholds. The latter includes disuse of platforms, mesonets, and locations with fewer than one third of the expected hourly availability and those with unknown quality (e.g., due to poor exposure of the instruments). With these procedures in place, the remaining observations form the basis of a quantitative analysis aimed at ensuring that our WRF fields are adequate to drive STILT.

4.1 Surface variables

4.1.1 Monthly bias and RMS errors

WRF model performance is quantified against surface observations through two summary statistics, bias and root-mean-square (RMS) error, computed for three different modeling periods and configurations: a) v3.4.1 for 2012, b) v3.5.1 for 2012, and c) v3.5.1 for 2013. Hereafter, we denote the '2012 Campaign' as being the period 10 May – 2 October 2012 and the '2013 Campaign' as being the period 1 March – 30 November 2013. Statistics are also compiled separately for each month. (The two days in October 2012 are included only in the full 2012 Campaign statistics). Detailed statistics based on the 2012 WRF v3.4.1 simulations are provided below. Results for v3.5.1 WRF simulations are, in general, similar and tabular statistics for the v3.5.1 2012 and 2013 Campaigns are available in the Supplement.

The two calendar years of 2012 and 2013 had substantially different growing seasons: 2012 was cooler and wetter than the 1980-2010 mean, while 2013 was warmer and drier. In 2013 a very late thaw (~20 May in Fairbanks) transitioned rapidly into summer, while the autumn refreeze occurred very late (~1 Nov). Statistics provided here are computed using approximately 120 land sites located in domain 3, predominantly in mainland Alaska, with some on offshore Alaskan islands and in the Yukon and Northwest Territories of Canada. The following model fields are interpolated every hour to the location of the observations and included in the evaluation: 2-m temperature, 2-m dewpoint temperature and 10-m wind speed. No vertical correction is applied to account for mismatch between model topographic heights and the true station elevation. Such an adjustment is needed less for high-resolution grids at 3.3-km grid spacing than for coarser-resolution grids. Furthermore, the choice of lapse rate used in these adjustments is often not representative of the environmental lapse rate, and, for our purposes, obscures evaluation of the WRF fields as provided to the STILT model.

Table 2 shows that WRF v3.4.1 exhibits a bias of -1.40 K in 2-m temperature for the 2012 Campaign, and negative temperature biases are also present in each month. The model temperatures are relatively warmer and thus closer to observations during August and September, but remain too low overall. Dewpoint temperature has a negative bias of -0.16 K across the 2012 Campaign. The bias changes from positive (moist) early in the campaign in May and June to negative (dry) from July onwards. This evolving model performance with a change in sign may indicate challenges related to inaccurate representation of the underlying soil conditions and the melting of snow and ice cover. More evidence is provided later in this section. Two-meter wind speed has a small bias of -0.17 m s^{-1}, which is encouraging given the primary importance of the wind field as an input to STILT. A prolonged negative bias in wind speed decreases in magnitude from -0.56 m s^{-1} during the spring and summer and becomes positive by September. We retain model-observation pairs when the observed wind speeds are greater than 3 kt (~1.5 m s^{-1}). Below this value, mechanical and logistical influences, such as wind sensor starting thresholds, rounding, and administrative limits (Bellinger, 2011) complicate scientific interpretation of the errors. Indeed, standard Automated Surface Observing System (ASOS) sensors report variable direction under certain meteorological conditions. The count of observed wind speeds by wind speed bin for the entire 2012 Campaign demonstrates an unphysical distribution (Table 3). The imposition of a wind speed threshold by definition focuses the statistics on higher wind speeds that are more important to transport errors. Fox (2013) summarized the effects of ASOS implementation on wind speed

observations, including a note that there was a factor of 2.4 increase in the number of pre- versus post-ASOS incidences of calm observations. Indeed, inclusion in the statistics of an artificially large number of calm observations compared to large numbers of non-zero model values introduces an apparent positive bias in the model wind speeds. As well, the ambiguous wind direction coding in the observation database for calm, and light and variable, winds also results in inflated RMS wind direction errors. For all months of the 2012 Campaign, wind direction biases are small (less than or equal to 5 degrees), with July exhibiting the smallest values. Transport uncertainties related to the small wind speed and direction biases are similarly expected to be small.

RMS errors for WRF 2012 v3.4.1 reported in Table 4 are 2.97 K, 2.52 K and 2.19 m s^{-1} for temperature, dewpoint temperature and wind speed, respectively. Temperature and moisture RMS values decrease month-to-month from May-June (~3.0 to 3.5 K) to September (~2.0 K). Wind speed errors, however, remain relatively constant throughout the campaign, with an increase in September perhaps related to the passage of strong extratropical cyclones during this month. Such storms, while typical for the region as autumn commences, pose a significant modeling challenge, especially for the wind field. As detailed in Section 5.1, model representation of these events is fraught with timing and position errors associated with strong horizontal thermal and moisture advection from the south, mixing of strong low-level winds aloft to the surface, and the possibility of downslope wind events to the lee of mountain ranges. Strong cyclones affected Alaska around 5, 16 and 26 of September 2012. Model performance during a downslope wind event on 16-17 September follows in Section 5.1. Despite the varying amounts of cyclonic activity by month, wind direction RMS errors decrease slightly from May through September with a 2012 campaign average of 52.1 degrees.

4.1.2 Validation of Arctic modeling in literature: surface variables

As noted earlier, the purpose of this manuscript is to assess the suitability of CARVE WRF simulations for use in transport modeling. While there are many recent studies that have illustrated both subjective and objective characteristics of Arctic modeling (e.g., Cassano et al., 2011; Jakobson et al., 2012; Tilinina et al., 2014; Jung and Matsueda, 2014; Simmonds and Rudeva, 2012; Glisan and Gutowski Jr., 2014), we put our model performance in perspective by noting the most relevant recent modeling studies that share many components

similar to our study. We include comparisons against recent reanalysis products that are the result of intensive efforts to reduce errors in simulated fields by coupling the forecast model to a data assimilation system, in contrast to a free-running forecast. When making these comparisons, the evolving nature of WRF and polar modifications, differing choices of observational datasets, grid resolution, modeling domain, and simulation periods must be kept in mind. It should be noted that the different lengths of the CARVE 2012 and 2013 modeling periods strongly influence how seasonal and monthly differences in errors affect the Campaign averages. This should also be kept in mind when comparing model performance against those from the literature listed below where periods of performance vary greatly.

Hines et al. (2011) used Polar WRF version 3.0.1.1 in a 25-km Western Arctic domain during 2006-2007 and compared model performance to ten observing sites over Alaska. Model biases of 2-m temperature averaged over all sites for the months of May, June and July (i.e., the months that overlap with the 2012 CARVE campaign) were 1.7, 2.5 and 1.4 K, respectively. RMS errors for the same months were 3.4, 3.9 and 3.5 K, respectively. The biases (RMS errors) in 10-m wind speed for Barrow for May, June and July were -0.4 (1.5), 0.3 (1.8) and -0.3 m s^{-1} (1.4 m s^{-1}), respectively. Wilson et al. (2011), for an ASR-like domain with 60-km grid spacing during December 2006 to November 2007, reported an Arctic-wide bias of -1.3 K and an RMS error of 4.4 K for temperature. For dewpoint temperature, values were -0.4 K and 4.4 K, respectively. Individual sites experienced substantial biases ranging from -7.0 K to 5.9 K for temperature and -7.2 K to 6.5 K for dewpoint temperature. Wind speed bias was 0.5 m s^{-1}, with an RMS error of 2.7 m s^{-1}.

More recently, Hines et al. (2014) used Polar WRF v3.5 on a 20-km grid to perform a number of sensitivity experiments related to varying sea ice treatment. In the current work, we closely follow their baseline implementation of Polar WRF and use of supplemental sea ice thickness, snow depth over sea ice, and sea ice albedo. Using observations from the Surface Heat Budget of the Arctic Ocean (SHEBA; Uttal et al., 2002) drifting ice station for January 1998, they obtained surface temperature and wind speed bias (RMS error) values of -1.2 K (3.2 K) and 0.3 m s^{-1} (1.3 m s^{-1}), respectively. Bromwich et al. (2014) reported the performance of the ERA-Interim global reanalysis (ERA) and Arctic System Reanalysis (ASR) version 1 based on Polar WRF v3.3.1 against a large number of standard observations during the period December 2006-November 2007. They reported annual biases (RMS errors) for 2-m temperature for the ASR and ERA of 0.10 K (1.33 K) and 0.29 K (1.99 K), respectively. For

2-m dewpoint temperature, they reported annual biases (RMS errors) for the ASR and ERA of -0.02 K (1.72 K) and 0.32 K (2.04 K), respectively. For 10-m wind speed, they reported annual biases (RMS errors) for the ASR and ERA of -0.24 m s^{-1} (1.78 m s^{-1}) and 0.41 m s^{-1} (2.13 m s^{-1}), respectively. Errors were larger regionally where topography is complex and were compounded by use of a coarse grid spacing of 30 km underlying the reanalysis product. Wesslén et al. (2014) evaluated the performance of the global ERA-Interim reanalysis (~80 km grid spacing) and two versions of the developmental Arctic System Reanalysis (ASR1 and ASR2) against ship-based observations obtained by the Arctic Summer Cloud-Ocean Study (ASCOS) in the mostly ice covered Arctic Ocean in August and September 2008. For 2-m temperature, they obtained biases for the ASR1, ASR2 and ERA-I of -0.8, -1.3 and 1.3 K, respectively. Corresponding RMS errors were 2.3, 2.5 and 1.9 K, respectively. For wind speed, biases were -1.4, -1.6 and -0.4 m s^{-1}, respectively, and RMS errors were 2.2, 2.3 and 1.6 m s^{-1}. The above details strongly suggest that the current CARVE modeling effort has generated near-surface meteorological fields that are of comparable quality to those in the recent literature. It should be kept in mind that, while realism is improved, skill scores from traditional verification techniques are often degraded due to imperfect timing and placement of small-scale features (Mass et al., 2002).

4.1.3 Diurnal cycle of surface variables

We now investigate model bias of the WRF v3.4.1 simulations for 2012 as a function of time of day (all times are UTC unless otherwise indicated; subtract 8 h for AKDT; e.g., noon AKDT=2000 UTC) for the 2012 campaign and also each month of the 2012 runs. All subsequent figures and discussion, unless otherwise noted, refer to the WRF v3.4.1 model simulations for 2012, since these fields were used by CARVE in preliminary analyses and in the example that follows in Section 5.3. When assessing model representation of the true diurnal cycle, it should be kept in mind that the first 6 h of each daily 30-h simulation is discarded to avoid model spin up errors, with the 0000-0600 UTC period formed from the 24 to 30-h simulation period from the previous day's run. This splicing together of short model simulations results in a seam between adjacent model runs. Fields from forecast hour zero (not shown) exhibit very small bias and RMS error values. Since these hour-zero WRF fields are simply a spatial interpolation of the MERRA reanalysis to the WRF grid, the minimal error demonstrates the accuracy of the MERRA fields when used for initial conditions.

Conversely, deficiencies that are inherent to any numerical weather prediction model are responsible for the subsequent growth of errors with respect to the observations over the 30-h simulation periods. These are responsible for any discontinuities seen between 0600 and 0700 UTC. It should also be noted that the diurnal cycle is less well defined above the Arctic circle around the summer solstice when solar insolation persists 24 h a day, with resulting effects on the PBL moisture and temperature structure that can differ from those in the midlatitudes (Tjernström, 2007). In general, the summer Arctic PBL is strongly modulated by surface processes and the presence, and change of phase, of ice and snow. This is in contrast to the properties of the free troposphere, which retains characteristics of air masses advected from lower latitudes (e.g, Tjernström, 2004)

Performance of the WRF v3.4.1 simulations as a function of time of day for the entire 2012 campaign is similar to patterns seen for individual months, with exceptions noted below. A rather pronounced diurnal cycle in model performance for temperature (Fig. 3a.) results in increasingly negative biases starting around ~1500-1600 UTC, as the solar zenith angle decreases. This indicates a lack of sufficient sensible heating at the surface through the time of peak heating across all months. Of note is a pronounced increase in daytime positive bias (too moist) for dewpoint temperature for May (Fig. 3b), in contrast to other months that exhibit only small biases and are relatively insensitive to the time of day. This may indicate misrepresentation in the model of surface processes, such as changes of state, when (relatively) large diurnal cycles of solar radiation interact with extensive snow and ice cover that is starting to melt. Modeling of the timing and duration of freeze/thaw cycles during the spring thaw is challenging and small errors in timing can result in large errors in both temperature and dewpoint temperature. Wind speed bias is generally negative across all hours for all months, with several months showing a reduction in the negative bias starting around 1400-1500 UTC with increasing solar insolation. Of all months during the 2012 campaign, September exhibits the smallest cold bias in temperature and is unique with a positive wind speed bias across all hours (Fig. 3c). An increase in RMS error values of temperature (Fig. 4a) occurs in the afternoon as the bias becomes more negative. RMS errors for dewpoint temperature show (Fig. 4b) a marked discontinuity between 0600 and 0700 UTC reflecting the seam between model runs. The discontinuity is most pronounced in May and steadily decreases through the summer. Again this suggests that the quality of the model representation of surface dewpoint temperature degrades substantially with forecast length due to model deficiencies, including those related to changes of state in the spring months.

This increase in error accelerates around 1800 UTC (noon Alaska time) and remains large through the end of each 30-h simulation. A relatively marked seam of approximately 0.3 m s^{-1}) between 0600 and 0700 UTC in wind speed RMS error for September (Fig. 4c) is larger than for all other months, though still modest in magnitude.

4.1.4 Aggregate wind rose plots

Wind rose plots showing modeled and observed winds are shown for all stations during the 2012 Campaign in Fig. 5. The plots contain the same thirty-six 10-degree wind direction bins used for encoding by the ASOS instrument to avoid the wind direction bias noted by Droppo and Napier (2008). As noted earlier, model-observation wind pairs are retained and plotted only if the observed wind speed is at least 3 knots. The wind roses demonstrate that the model well represents the wind direction frequency distribution, with subtle favored wind directions from the southwest and east in both the model (Fig. 5a) and observations (Fig. 5b). Of note is the higher frequency in the observations of the highest wind speeds (light green and orange).

4.1.5 Spatial distribution of bias and RMS errors

The spatial performance of WRF v3.41 shows expansive areas of negative bias across domain 3 for the 2012 Campaign (Fig. 6a). The 2-m model temperatures are too cold across all of Alaska, with sites on the North Slope adjacent to the Beaufort Sea and in the Yukon and Northwest Territories of Canada showing the largest negative biases. The persistence during June, July and August (monthly plots for the 2012 Campaign are provided as a Supplement) of substantial negative biases on the North Slope and over other coastal sites results in these locations exhibiting the largest Campaign-averaged cold bias. Similar negative biases are present in interior Alaska in May, but diminish in the summer months substantially following snow melt.

Dewpoint temperature (Fig. 6b) shows a small positive bias for most areas, except on the North Slope where a small negative bias exists. Similar to temperature, a pronounced seasonal change occurs over the 2012 Campaign. Monthly plots for the 2012 Campaign are provided as a Supplement. During May a large positive bias of 2-4 K exists across most interior stations. As summer progresses, the magnitude decreases everywhere, with a bias of -3 to -5 K developing and persisting across North Slope sites. The bias is minimal by September at

most sites. The coastal stations at Unalakleet (PAUN) and Point Hope (PAPO), however, have substantial positive biases of moisture, despite generally good model performance for other variables at these sites. Wind speed (Fig. 6c) is biased slightly high at many inland sites, but shows small negative bias for coastal regions and sites on the North Slope. There is no substantial seasonal signal, except that many sites in southern Alaska in September exhibited a noticeable positive wind speed bias of approximately 1-2 m s^{-1}.

RMS errors in temperature for the overall 2012 Campaign (Fig. 7a) are substantial (up to 6 K) and extensive along the coast of the North Slope, with smaller values elsewhere in Alaska. For these coastal locations, it is conceivable that errors in resolving the coastline on the 3.3-km grid, with related challenges in representing land and ocean surface characteristics, may profoundly affect model performance. Similar to the northward movement of the region of largest values for temperature bias, RMS values of temperature are maximized during July on North Slope stations, with some values of approximately 10 K. These indicate model simulations that are likely challenged by the arrival from the south of warm and moist air masses during the boreal summer over land surfaces still influenced by snow, ice or meltwater. RMS errors decrease noticeably across all sites during August and remain small (approximately 2 K) in September. RMS values for dewpoint temperature (Fig. 7b) do not exhibit substantial geographical variation when averaged over the entire 2012 Campaign, despite monthly values following a pattern similar to those for temperature, though with less extreme and more transient maximum errors on the North Slope in July of approximately 7 K. Wind speed RMS errors (Fig. 7c) are more uniform across all regions, with the most pronounced feature being a substantial regional increase in September in southeast Alaska. The increase in RMS error may be related to poor timing of high wind speeds associated with the strong cyclones that occurred in this mountainous region of Alaska during the month.

Of note are persistent and large errors in dewpoint temperature for stations PAUN and PAPO. Evidently erroneous low dewpoint temperatures (not shown) observed at PAPO are the primary cause for the moisture errors (both RMS and bias) at that station. These unphysical values are not present in the 2013 observations at these sites. The source of the large errors at PAUN, however, is not obvious and may be related to the fact that the coastline is not sufficiently resolved in the model.

4.1.6 Model performance at representative stations

We now show WRF v3.4.1 performance for three sites in Alaska often visited by the CARVE aircraft: McGrath (PAMC), Deadhorse (PASC) and Barrow (PABR). These regions are low-lying wetlands with abundant seasonal CO_2 and CH_4 fluxes and substantial evapotranspiration. We expect model performance at these stations to be representative of neighboring flight locations. Indeed, there is relatively small spatial variation (see Fig. 7) of errors in the vicinity of these three sites.

McGrath (62.95 N, 155.58 W) is located along the upper Kuskokwim River in Interior Alaska. At McGrath, a negative bias (Fig. 8a) exists in the model temperatures, and possibly a related positive (moist) bias (Fig. 8b) exists in model dewpoint temperature, from May through late June. Extensive snow cover ended in the vicinity of McGrath around 10 May – the date of the first model simulation. Wind speed (Fig. 8c) in the model is too high, in part because of frequent calm observations. The effect on the magnitude of wind speed errors will be small due to the low wind speeds involved here. Wind direction is well represented in the model, with prevailing winds from the south-southwest in both the model (Fig. 9a) and observations (Fig. 9b). Secondary peaks in the wind direction distribution from the west-northwest, north-northwest and east are also well modeled. The varied directions (compare to subsequent analyses for Deadhorse and Barrow) in part suggest complex source regions of air masses in the interior of Alaska. The frequency of lower wind speeds (blue colors) is higher in the observations.

Deadhorse (70.21 N, 148.51 W) is located on the North Slope adjacent to the Beaufort Sea amid low-lying tundra. While experiencing a very cold and dry climate, it is susceptible to marine influences during periods of open water. Local variability in the warmer months of the year is enhanced by its location at approximately 70^{o} N, as incursions of warm, moist air from the south result in brief and infrequent but pronounced departures from a cold annual state. Timing of such frontal passages and also local cooling effects of the nearby ocean are challenging for NWP models to reproduce. Temperatures (Fig. 10a) and dewpoint temperatures (Fig. 10b) in WRF generally exhibit a strong negative bias between the disappearance of extensive snow cover around 10 June 2012 and retreat of sea ice by the middle of July, with forecast errors especially large during and following a mid-June warm spell. The warm spell was associated with temperature and moisture advection from the south in the absence of low clouds and snow cover. Degradation in model performance, however,

appears strongly related to thawing of the top layer of soil. The appearance of a pronounced negative temperature and dewpoint temperature bias between 12-15 June 2012 coincides with an abrupt increase in near-surface soil moisture from snowmelt as soil temperatures rise above the melting point at a nearby Natural Resources Conservation Service Snow Telemetry (SNOTEL) instrument at Prudhoe Bay, AK; http://www.wcc.nrcs.usda.gov/nwcc/site?sitenum=1177; "near-surface" measurements are obtained at a depth of 5 cm). A similar bias appears in the 2013 time series between 9-11 June coincides with thawing of the soil. A potential source of error is that the Noah land surface model in WRF may not be completely spun-up, and therefore not reach equilibrium with the diurnal cycle at small spatial scales (e.g., Chen et al., 2007). Daily use of coarse-resolution soil temperature and moisture inputs from MERRA and NNRP presumably contributes to the degraded surface temperature simulations near the coastline in the presence of moisture from snow melt and thawing soil layers. This is despite the fact that the initial soil and temperature fields from MERRA and NNRP are likely in equilibrium for their respective much coarser-resolution analysis systems, though biases may still be present. Indeed, to minimize the effect of spin-up in the land surface model, Hines et al. (2011) performed an offline 10-year cycling of soil moisture and temperature fields. Wilson et al. (2011) and Hines et al. (2011) also adopted a longer spin-up of 24 h (compared to our 6 h) in an attempt to improve representation of the model atmosphere, specifically the surface interface and the boundary layer. However, their results, given earlier in Section 4.1.2, show bias and RMS errors comparable to results from CARVE. It is to be noted that the sign of the dewpoint temperature bias following melting of snow cover is different here than in McGrath. Model temperatures improve noticeably around the middle of July, likely related to the retreat of coastal sea ice, and errors remain small in August and September. The large negative bias in dewpoint temperature also decreases around the middle of July, but frequent episodes of negative bias persist into September. A time series plot of wind speed (Fig. 10c) shows a negative bias, especially during May, with peak speeds during the mid-June warm spell also too low. Wind roses show excellent agreement between model (Fig. 11a) and observations (Fig. 11b), with the predominant direction being east-northeast. A secondary peak is seen for winds from the southwest. The strongest observed winds (light green and orange in Fig. 11b) occurred just prior to the June warm spell but are absent in the model (Fig. 11a).

Barrow (71.29 N, 156.77 W) is located on a peninsula extending into the Beaufort Sea. It is the northernmost point in Alaska and is surrounded to the south by wetlands and low-lying

tundra. Results for Barrow are comparable to those for Deadhorse. Of note is the large negative temperature bias during the summer months starting in the middle of June (Fig. 12a), while there also are occurrences of excessive nighttime cooling in May and early June when snow cover is present. Extensive snow cover ended by the middle of June in Barrow and extensive sea ice in close proximity to the shoreline receded by the middle of August. In addition to the possibility of inadequate representation in the model of saturated soil conditions, poor representation of the extent and frequency of low clouds and coastal fog, which are modulated by the retreat of coastal sea ice starting in July, can affect modeled temperatures (e.g., Dong and Mace, 2003) and so may play an important role in the negative bias along the North Slope including Deadhorse and Barrow. Interestingly, simulation of Barrow summer temperatures by Hines et al. (2011) exhibited a positive bias (contrast to the negative bias reported here) of over 10 K on many occasions following snow melt in June. Coastal marine influences also appeared to be poorly modeled in their study, with much better agreement farther inland at, for example, Atqasuk, which lies approximately 110 km to the south of Barrow and experiences a more continental climate (McFarlane et al., 2009). Unfortunately, there is a dearth of non-coastal sites in our observational database with which to investigate further. Very low dewpoint temperatures in our simulations (Fig. 12b) in May were modeled on nights with excessive radiational cooling. A pronounced low (dry) bias from late June through late August was also seen. By September, low-level moisture in the model showed much improvement compared to observations. Wind speed (Fig. 12c) in the model is slightly too low, but does not exhibit any egregiously bad time periods. Similar to the wind roses for Deadhorse, those for Barrow indicate good agreement between model (Fig. 13a) and observations (Fig. 13b), with predominantly east-northeast winds and a secondary maximum from the southwest. The highest wind speeds (light green color) are more frequently seen in the observations (Fig. 13b).

4.2 Upper air variables

4.2.1 Campaign bias and RMS errors

To quantify model performance aloft, model bias and RMS error statistics are computed for all 0000 and 1200 UC upper air observations in the 2012 and 2013 Campaigns at 850, 700, 500, 300 and 200 hPa. Model values were interpolated to the location of 11 balloon-launch sounding sites (ten in Alaska and one in the Northwest Territories of Canada). For WRF

v3.4.1 in Table 5, model representation of upper-level variables is very good, with acceptably small bias and RMS errors. The 200 hPa level, which is frequently above the tropopause in the Arctic, often has a bias of a different sign than in the troposphere, but the magnitude remain modest. Temperature bias values are small (magnitude less than 0.2 K) and negative at lower levels. The largest bias is 0.39 K at 300 hPa. Geopotential height bias errors range from -2.5 to -6.4 m up to 300 hPa and are small and positive at 200 hPa. Model relative humidity is too high at all levels, ranging from +4.3 to +10.8 %, except at 200 hPa, where the bias is -7.6 %. Low-level wind speeds have a small positive bias of up to +0.17 m s^{-1}, with the bias remaining small but negative at 300 hPa and 500 hPa. A positive bias of +0.36 m s^{-1} exists at 200 hPa. RMS errors are modest – +0.98 to +1.80 K for temperature, 11.4 to 18.1 m for gepotential height, 14.1 to 22.7 % for relative humidity, and 3.00 to 4.88 m s^{-1} for wind speed -- and reasonably uniform at all levels.

We note that near-surface temperature errors are larger than those from each level of the free atmosphere. This may be a result of deficiencies in boundary layer and surface energy processes or the representation of clouds above the PBL that affects surface radiation budgets. Each of these potential sources may contribute to the episodes of large temperature errors at Deadhorse (Fig. 10a) and Barrow (Fig. 12a).

4.2.2 Validation of Arctic modeling in literature: upper-air variables

Bromwich et al. (2014) evaluated version 1 of the ASR and ERA-Interim against upper-level observations for December 2006 to November 2007. For temperature (values here refer to levels at 200 hPa and below to aid in direct comparison), they report mostly small negative biases of magnitude less than 0.2 K at all levels, with RMS errors ranging from 0.69 to 1.94 K. For geopotential height, bias values are approximately 2 m or less in magnitude, with RMS errors between 7 and 22 m. For relative humidity, biases are of mixed sign but generally less than 2 %, while RMS errors range from approximately 9-23 %. For wind speed, biases are almost exclusively negative with magnitudes less than 0.5 m s^{-1}; RMS errors are largest at high altitudes at approximately 3 m s^{-1}. While these values generally are in good agreement with the CARVE WRF v3.4.1 values reported in Table 5, the CARVE biases for relative humidity and geopotential height are larger.

CARVE errors are similar to those reported by Wilson et al. (2011). They reported the magnitude of the temperature bias through the entire column of less than 1 K, with RMS

errors of 1-2 K. Bias for geopotential height ranged from -11 m to +40 m and RMS errors ranged from 20 to 50 m. For relative humidity, they note that obtaining accurate measurements in cold conditions, such as often found in the Arctic, is challenging, thus making this field difficult to verify. Truncated to those (warmer) levels below 500 hPa, they report biases of less than 5 %, with RMS errors of 15-20 %. Biases in wind speed range from -1.1 m s^{-1} to 1.9 m s^{-1}.

5 Case Studies

5.1 Representation of small-scale features

Recent years have seen the emergence of reanalysis datasets with increasingly high resolution. These products typically have grid spacing of approximately 30 km and, as described above, generally compare favorably to observations. Bromwich et al. (2014) do note, however, that regions of complex topography still pose a modeling challenge, in part due to poor model representation of terrain and local wind effects. Here we demonstrate the ability of the WRF model on a 3.3-km grid to reasonably represent a damaging windstorm that is tied closely to mountain wave activity, while the downslope windstorm is absent on a coarse 30-km grid with resolution comparable to current reanalyses. It is our intent to illustrate the increased realism possible with higher-resolution modeling and suggest that transport studies will, in general, benefit from the increased detail in flow fields.

September 2012 featured the passage of several strong extratropical cyclones through Alaska. During the evening of September 16, damaging downslope winds (e.g., 62 kt at 1353 AKST at Delta Junction and an unofficial report of 99 kt around 2300 AKST at Dry Creek) were reported across extensive regions to the north of the Alaska Range (NOAA, 2012). The winds were associated with a deep low-pressure system of central pressure 975 hPa that moved along the west coast of Alaska. For 5 h starting around 2300 AKST (0700 UTC 17 September), the village of Tanacross observed severe wind gusts from a generally southerly direction. This event was noteworthy because of the presence of strong winds both at high elevations and also in the valleys where the village of Tanacross is located. It is hypothesized that decreased (relative to winter) static stability during the autumn months in the absence of strong surface-based temperature inversions may have contributed to the occurrence of the damaging wind event at Tanacross, which is located farther east than the favored high

elevation locations of wind storms in the region (personal communication, Rich Thoman, NWS Fairbanks). Here we demonstrate the improved realism of the downslope wind event that is afforded by use of high-resolution modeling. Mass et al. (2002) discussed similar benefits related to flow over orography that are associated with decreasing WRF grid spacing from 36 to 12 to 4 km. They also note that, while realism is improved, skill scores from traditional verification techniques are often degraded due to imperfect timing and placement of small-scale features.

To quantify the impact of model resolution on the ability of WRF to reproduce the mountain waves and subsequent strong surface winds associated with this episode, we have performed additional one-way nested runs of the non-polar WRF v3.5.1 for the period 11-17 September 2012. (Use of WRF v3.5.1 ensured that we took advantage of any software refinements and corrections inherent to any new model update, but these changes are not anticipated to have a significant effect.) The placement of 61 vertical levels follows the most recent configuration (available from http://www.mmm.ucar.edu/rt/amps/) of the Antarctic Mesoscale Prediction System (AMPS; Powers et al., 2012). The use of one-way nesting enables a direct comparison of the effects of grid spacing, whereas the production runs reported on earlier employed two-way nesting aimed at optimizing model performance. Recent studies (e.g., Moore, 2013) have used the Interim Arctic System Reanalysis (ASRI) product (that has the same 30-km grid spacing as domain 1 in this study) to investigate the climatology in polar WRF of high-speed wind events. Here, however, we demonstrate that use of coarse resolution at 30-km grid spacing versus 3.3-km grid spacing leads to a much more diffuse representation of model orography, mountain waves and low-level flow fields.

Figure 14 shows the model representation of a downslope windstorm along the Alaska Range. The 10-m wind field in the 3.3-km grid indicates extensive strong winds over and near the Alaska Range (Fig. 14c), including some locations over 70 kt, while the 30-km wind field offers a much more muted representation with maximum surface winds of about 30 kt (Fig. 14d). The spatial variation in the wind fields is considerably greater and more realistic in the higher-resolution domain. The cross-section for domain 3 (Fig. 14a) clearly suggests the presence of a mountain wave in both the potential temperature and wind fields, while mountain waves are absent in domain 1 (Fig. 14b). The unrealistically smooth cross-barrier terrain profile with insufficient vertical extent is a consequence of spatial averaging on a horizontal grid of insufficient resolution to fully resolve mountain waves (e.g., Chow et al.,

2012), despite the presence aloft of strong antecedent cross-barrier flow during the evening of 16 September (not shown). Indeed, a minimum of the Froude number (e.g., Vosper, 2004) below 1.5 occurs only in domain 3 and in the near-surface upstream flow. This supports the possibility of the flow transitioning from subcritical to supercritical in the vicinity of the mountain barrier and the occurrence of the hydraulic jump that is strongly suggested in the high-resolution fields in Fig 14a. A similar insufficient model representation of the Novaya Zemlya archipelago in the relatively coarse model simulations reported by Moore (2013) may have contributed to the underrepresentation of high-speed winds in their simulations. While damaging winds were observed at Tanacross at 1100 UTC, the high-resolution model generates strong downslope winds nearby that remain just uphill of the village and do not extend far into the lower elevations. Overall, strong surface winds in the model are limited to higher elevations, particularly to the west of Tanacross in the vicinity of the highest orography. It is encouraging then that the model, though several hours later than observed, was eventually able to reproduce the strong observed winds at the low-elevation village of Dry Creek (Fig. 15), with peak speeds in the vicinity of the village of approximately 100 kt at 1100 UTC 17 September due to breaking of a mature mountain wave. The ability of the high-resolution WRF configuration, with its more realistic representation of the Alaska Range, to simulate an extreme meteorological event is an indication that problems of lower resolution simulations under a variety of synoptic conditions can be avoided, along with concurrent improved transport in STILT.

Another example of the benefit of high-resolution modeling involves preservation of large horizontal gradients: At 0100 UTC 24 August 2012 the CARVE aircraft encountered a pronounced low-level atmospheric boundary on the North Slope during a flight southeast of Barrow. A wind direction change of about 80 degrees and a rapid decrease in atmospheric water vapor content occurred while crossing the boundary in a southward direction over a distance of approximately 500 m. Observed concentrations of CH_4 and CO_2 also decreased sharply, while CO increased across the front. While the model simulations placed the front too far south in this case (not shown), we anticipate the ability of high-resolution grids to "contain" the concentrations in a well-defined boundary whose sharpness could not be contained at coarser resolution.

5.2 Examples of STILT footprint calculations

To illustrate the effect of WRF resolution on STILT output, we have computed footprints using the WRF configuration in Section 5.1 (v3.5.1 with one-way nesting) for receptors placed at the CARVE tower located to the northeast of Fairbanks (64.986^{o} N, 147.600^{o} W) every hour from 0000 UTC 13 September to 2300 UTC 17 September 2012. Strong tropospheric-deep wind fields associated with migratory cyclones during September likely maximized the influence of orography on the near-surface receptors. The receptors in STILT are positioned at an altitude of 301 m above ground level (AGL), as well as at the actual altitude of 32 m AGL, to account for the 269-m discrepancy in model height (even as represented in domain 3) versus the true height of the base of the CARVE tower. STILT footprints for both heights are computed at 0.1-degree horizontal resolution using two sets of WRF fields: first a) from all domains (i.e., 30-, 10- and 3.3-km grid spacing) and then b) from just the outermost domain 1 with 30-km grid spacing. Resulting near-field footprints are summed every hour over each 24-h trajectory period. The corresponding footprints are substantially different, particularly for the 301-m receptors (Figs. 16a,b). For the receptors at 32 m (Figs. 16c,d), the footprint field is confined closer to the tower location as might be expected, making the differences appear smaller. The footprint fields generated using only the domain 1 wind field (Figs. 16b,d) are more diffuse because of the coarse grid spacing of the input wind field. While there is, obviously, no "ground truth" for establishing which footprints are more accurate, footprints derived from high-resolution WRF fields intrinsically can contain more horizontal spatial detail, and benefit from more realistic orography and presumably boundary layer processes, than is possible when using more traditional coarser-scale flow fields. At a minimum, these differences contribute to the transport uncertainties entering the measurement error budget for inverse flux estimates.

Another aspect of resolution is the impact of grid size employed in the STILT footprint calculations. For this purpose, we have utilized winds from all three domains of the v3.4.1 two-way nested runs and computed the footprints at 0.1- and 0.5-degree resolution for a subset of 79 receptors during the same five-day period as above. For a direct comparison, the 0.5-deg footprints have been scaled by a factor of 1/25, resulting in an effective resolution of 0.1-degrees. The two "native" STILT resolutions (0.5-deg coarse grid in Figs. 17a,c; 0.1-deg fine grid in Figs. 17b,d) result in noticeable differences in both magnitude and spatial patterns of the footprints. The 0.1-degree footprints (at 301 m in Fig. 16b and at 32 m in Fig. 16d)

exhibit substantially more detail, making them more suitable for applications in the heterogeneous Alaskan landscape. Note the color scale has been standardized across all four panels to highlight the 'washed out' look to the coarse-grid footprints.

5.3 Impact on CARVE chemical simulations

Chang et al. (2014) used the WRF-STILT footprints (based on WRF v3.4.1, described in Section 3.1 above) with vertical profiles of the CARVE aircraft methane mixing ratios to determine methane fluxes for Alaska for 2012. This set of vertical profiles comprised receptors located from near ground level to over 5000 m above sea level (ASL). The vertical profiles of six chemical and dynamic tracers measured by the CARVE aircraft (CH_4, CO_2, CO, O_3, water vapor and potential temperature) were used to identify the depth of the atmospheric column enhancement, which is defined as the well-mixed surface-influenced air from the ground to the bottom of the free troposphere. An independent estimation of the depth of the column enhancement is also provided by the height (ASL) at which the WRF-STILT surface-influence function (i.e., the footprint) becomes vanishingly small (< 0.1 ppm/μmol m^{-2} s^{-1}). The 10-day long footprint multiplied by a land mask for each of the receptors within a profile was summed to determine the total surface influence from land for that profile. Typically each profile contains between 200 and 400 individual receptors. For each of the 30 vertical profiles used by Chang et al. (2014), the WRF-STILT transport framework identifies the top of the column enhancement to within 500 m of the value identified by the CARVE aircraft in 67% of the profiles.

In Fig. 18a, we show a sample aircraft-observed vertical profile of CH_4 over interior Alaska near Fairbanks with the top of the atmospheric column enhancement at approximately 2400 m ASL. We see a well-mixed surface-influenced layer, with free tropospheric methane mixing ratios above the top of the column enhancement. The vertical profile of the WRF-STILT influence for receptors within this flight segment (Fig. 18b) is in good agreement and demonstrates that WRF-STILT is able to capture the shape of the CH_4 enhancement throughout the column, as well as the approximate depth of the column enhancement. This also ensures that the estimation of the volume of air that is affected by surface emissions is well estimated, which ultimately is an important aspect of the simulation of GHG concentrations.

We also compare modeled ozone loss with measured concentrations. Ozone can be used as a chemical tracer for land influence when dry deposition is the major loss process and photochemical sources and sinks are negligible. This is the case in the lowest 1.5 km of the Arctic atmosphere, where photochemistry is approximately 10 times slower than dry deposition (Jacob et al., 1992b; Walker et al., 2012). It is also more likely in the spring and fall, when incoming solar radiation is lower, resulting in less photochemistry and lower vegetative emissions of volatile organic compounds, both of which lead to lower ozone production. By studying the lowest layer of the atmosphere in which ozone flux can be assumed constant, the ozone loss can be calculated using WRF-STILT footprints and compared to measured ozone concentrations.

Model ozone loss was calculated at each receptor location by summing the portion of the footprint that had been in contact with land and multiplying it by an initial estimate of the deposition velocity (-0.3 cm s^{-1}) and the surface ozone concentration (determined for each flight by taking the mean ozone concentration below 500 m). Monthly results are shown in Fig. 19 for the lowest 1.5 km above ground level after the mean for each flight was subtracted for both the model and the observations. The best-fit lines are determined using standard major axis regression and the slope can be used to calculate the dry deposition velocity required to match the model with the observations. These computed values range from 0.1 to 0.5 cm s^{-1}, with a mean over the five months of 0.3±0.1 cm s^{-1}, where the uncertainty is the 95% confidence interval (2 × standard error). This range of deposition velocities is consistent with those measured over tundra, sub-Arctic fen, and sub-Arctic Norwegian spruce and Scots pine forests (0.18–0.60 cm s^{-1}) (Jacob et al., 1992a; Mikkelsen et al., 2004; Tuovinen et al., 1998, 2004). Furthermore, our simulated deposition rates exhibit a realistic seasonal cycle, with lower velocities in May (Fig. 19a) when the surface is snow covered and leaves are still unexposed, and higher velocities in the warmer months (e.g., July and August, respectively, in Figs. 19c and d). This analysis provides confidence in WRF-STILT as implemented for CARVE and lends credence to applying the WRF-STILT footprints in the science analysis and flux inversions.

6 Conclusions

We have presented a detailed description and validation of the atmospheric transport model used to estimate surface-atmosphere CO_2 and CH_4 fluxes from CARVE airborne and tower

observations. Polar WRF was run on a 3.3-km grid centered over Alaska to generate high-resolution atmospheric fields for input to the STILT transport model. Aircraft and tower-based receptor locations from 2012 and 2013 formed the starting points of backward trajectory computations by the STILT transport model. Model upgrades are ongoing. Related papers provide more details about research enabled by observations and modeling from the CARVE campaigns (Chang et al., 2014; Miller et al., in preparation; Karion et al., in preparation).

While the bulk statistics computed here for 2012 and 2013 CARVE model fields cannot be compared directly with published values that use different time periods and observations, the error magnitudes are in general agreement with others in the recent literature. The grid spacing of 3.3-km used for CARVE represents an order of magnitude increase in model resolution compared to standard reanalysis products. The high resolution permits more realistic depiction of flow, including the explicit modeling of downslope windstorms that are absent in coarser-scale model grids. The substantial influence of high-resolution model wind fields input to STILT on footprint fields is demonstrated, as is the increased detail when the footprints themselves are computed on a finer-scale grid. These approaches are likely to be beneficial in the complex orography and surface flux patterns of Alaska and the Arctic in general. These preliminary modeling results will be refined in future CARVE years to form a consistent modeling database of WRF simulations and STILT-based footprints that extends from the beginning to the end of the CARVE field campaigns.

The measurement-observation system developed for CARVE ultimately is applicable to other regions of the Arctic, such as the Mackenzie Delta in the Northwest Territories of Canada, Scandinavia and Siberia. The entire dataset (more than 3 TB of data) is publically available from NASA from the CARVE data portal (https://ilma.jpl.nasa.gov/portal/). The modeling framework is available to the general Arctic research community and the planned ABoVE NASA mission.

7 Author contributions

J.H. and M.M. performed model simulations. J.H, J.E. and R.C. prepared the manuscript. S.W. and C.M. designed the experiments. R.C., S.W., C.M., A.K., J.M. and C.S. collected observations. J.H., J.E., T.N., R.C., S.W. and N.S. conducted analysis.

8 Acknowledgements

Computing resources for this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center. The column enhancement analysis was provided in a timely manner by Roisin Commane and Jakob Lindaas at Harvard University. We appreciate the guidance provided by Jinlun Zhang at the University of Washington and Keith Hines at The Ohio State University. We gratefully acknowledge help provided by John Halley Gotway of the WRF-MET team at NCAR in setting up and adapting WRF-MET. The final version of this manuscript benefitted from the helpful comments of two anonymous reviewers. The research described in this paper was performed for the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), an Earth Venture (EV-1) investigation, under contract with the National Aeronautics and Space Administration.

9 References

ACIA: Impacts of a Warming Arctic: Arctic Climate Impact As- sessment, Tech. rep., Arctic Council, 140 pp. Available online at http://www.amap.no/documents/download/1058/Impacts-of-a-Warming-Arctic, 2004.

Barnes, E. A.: Revisiting the evidence linking Arctic amplification to extreme weather in midlatitudes, Geophys. Res. Lett., 40, doi:10.1002/grl.50880, 2013.

Bellinger, T. E.: Evaluating the wind data from the automated surface observing system in Oak Ridge, Tennessee Is KOQT the calmest site in the US?, in: Nuclear Utility Meteorological Data Users Group Meeting, Nuclear Utility Meteorological Data Users Group, Oak Brook, IL, available online at http://www.ornl.gov/ das/web/KOQTCalm.pdf, 2011.

Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheep- maker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric CH4 in the first decade of the 21st century: Inverse modeling analysis using SCIA- MACHY satellite retrievals and NOAA surface measurements, Journal of Geophysical Research: Atmospheres, 118, 7350 – 7369, doi: 10.1002/jgrd.50480, 2013.

Brioude, J., Angevine, W. M., McKeen, S. A., and Hsie, E.-Y.: Numerical uncertainty at mesoscale in a Lagrangian model in complex terrain, Geosci. Model Dev., 5, 1127–1136, doi:10.5194/gmd-5-1127-2012, 2012.

Bromwich, D. H., Bai, L., Hines, K. M., Wang, S.-H., Kuo, B., Liu, Z., Lin, H.- C., Barlage, M., Serreze, M. C., and Walsh, J. E.: Very high resolution Arctic System Reanalysis for 2000-2011, in: International Conference on Reanalyses, World Climate Research Programme, Silver Spring, MD, 2012.

Bromwich, D. H., Wilson, A. B., Bai, L.-S., Moore, G. W. K., and Bauer, P.: A comparison of the ability of the regional Arctic System Reanalysis and the global ERA-Interim Reanalysis to represent Arctic meteorology, Quart. J. Roy. Meteor. Soc., submitted, 2014.

Cassano, J. J., Higgins, M. E., and Seefeldt, M. W.: Performance of the Weather Research and Forecasting Model for Month-Long Pan-Arctic Simulations, Monthly Weather Review, 139, 3469 – 3488, doi: 10.1175/MWR-D- 10-05065.1, 2011.

Chang, R. Y.-W., Miller, C. E., Dinardo, S. J., Karion, A., Sweeney, C. S., Daube, B. C., Henderson, J. M., Mountain, M. E., Eluszkiewicz, J., Miller, J. B., Bruhwiler, L. M. P., and Wofsy, S. C.: Methane emissions from Alaska in 2012 from CARVE airborne observations, Proceed. National Academy Sci., submitted, 2014.

Chen, F. and Dudhia, J.: Coupling an advanced land-surface hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Mon. Wea. Rev., 129, 569–586, 2001.

Chen, F., Manning, K. W., LeMone, M. A., Trier, S. B., Alfieri, J.G., Roberts, R., Tewari, M., Niyogi, D., Horst, T. W., Oncley, S. P., Basara, J., and Blanken, P. D.: Description of Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System During IHOP-02. J Appl. Meteorol. Climatol., 46, 694-713, 2007.

Chow, F. K., Snyder, B. J., and de Wekker, S. F. J.: Mountain weather research and forecasting: Recent progress and current challenges, first edn., 750 pp., 2012.

Cohen, J. L., Furtado, J. C., Barlow, M., Alexeev, V. A., and Cherry, J. E.: Asymmetric seasonal temperature trends, Geophys. Res. Lett., 39, L04 705, 7 pp. doi:10.1029/2011GL050582, 2012.

Coumou, D., Petoukhov, V., Rahmstorf, S., Petri, S., and Schellnhuber, H. J.: Quasi-resonant circulation regimes and hemispheric synchronization of extreme weather in boreal summer, Proceed. National Academy Sci., pp. 12331–12336, doi: 10.1073/pnas.1412797111, volume 111, number 34, 2014.

Dong, X. and Mace, G. G.: Arctic Stratus Cloud Properties and Radiative Forcing Derived from Ground-Based Data Collected at Bar- row, Alaska, Journal of Climate, 16, 445 – 461, doi: 10.1175/1520- 0442(2003)016<0445:ASCPAR>2.0.CO;2, 2003.

Draxler, R. R. and Hess, G. D.: An overview of the HYSPLIT 4 Modeling System for Trajectories, Dispersion, and Deposition, Australian Meteorological Magazine, 47, 295–308, 1998.

Droppo, J. G. and Napier, B. A.: Wind Direction Bias in Generating Wind Roses and Conducting Sector-Based Air-Dispersion Modeling, Journal of the Air and Waste Management Association, 58(7), 913–918, doi:10.3155/1047- 3289.58.7.913, 2008.

Eluszkiewicz, J., Hemler, R. S., Mahlman, J. D., Bruhwiler, L., and Takacs, L. L.: Sensitivity of Age-of-Air Calculations to the Choice of Advection Scheme, J. Atmos. Sci., 57, 3185–3201, 2000.

Fox, T.: Use of ASOS meteorological data in AERMOD dispersion modeling, EPA memorandum, EPA, Research Triangle Park, NC, 35 pp. Available online at http://www.epa.gov/scram001/guidance/clarification/20130308_Met_Data_Clarification.pdf , 2013.

Fuelberg, H. E., Harrigan, D. L., and Sessions, W.: A meteorological overview of the ARCTAS 2008 mission, Atmos. Chem. Phys., 10, 817–842, doi:10.5194/acp-10-817-2010, 2010.

Glisan, J. M. and Gutowski, W. J.: WRF summer extreme daily precipitation over the CORDEX Arctic, Journal of Geophysical Research: Atmospheres, 119, 2013JD020 697, doi: 10.1002/2013JD020697, 2014.

Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: Inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.

Hegarty, J. D., Draxler, R. R., Stein, A. F., Brioude, J., Mountain, M., Eluszkiewicz, J., Nehrkorn, T., Ngan, F., and Andrews, A. E.: Evaluation of Lagrangian Particle Dispersion Models from Controlled Tracer Releases, J. Appl. Meteor. Climatol., 52, 2623–2637, doi:10.1175/JAMC-D-13-0125.1, 2013.

Hines, K. M., Bromwich, D. H., Bai, L.-S., Barlage, M., and Slater, A. G.: Development and Testing of Polar WRF. Part III: Arctic Land, J. Climate, 24, 26–48, doi:10.1175/2010JCLI3460.1, 2011.

Hines, K. M., Bromwich, D. H., Bai, L., Bitz, C. M., Powers, J. G., and Manning, K. W.: Sea ice enhancements to Polar WRF, Mon. Wea. Rev., submitted, 2014.

Holland, M. and Bitz, C.: Polar amplification of climate change in coupled models, Climate Dynamics, 21, 221–232, doi: 10.1007/s00382-003-0332-6, 2003.

IPCC: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Groups I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Tech. rep., Intergovernmental Panel on Climate Change, Geneva, Switzerland, 29 pp. Editted by T.F. Stocker and D. Qin and G.-K. Plattner and M. Tignor and S.K. Allen and J. Boschung and A. Nauels and Y. Xia and V. Bex and P.M. Midgley, 2013.

Jacob, D. J., Fan, S.-M., Wofsy, S. C., Spiro, P. A., Bakwin, P. S., Ritter, J. A., Browell, E. V., Gregory, G. L., Fitzjarrald, D. R., and Moore, K. E.: Deposition of ozone to tundra, Journal of Geophysical Research: Atmospheres, 97, 16 473 – 16 479, doi: 10.1029/91JD02696, 1992a.

Jacob, D. J., Wofsy, S. C., Bakwin, P. S., Fan, S.-M., Harriss, R. C., Talbot, R. W., Bradshaw, J. D., Sandholm, S. T., Singh, H. B., Browell, E. V., Gregory, G. L., Sachse, G. W., Shipham, M. C., Blake, D. R., and Fitzjarrald, D. R.: Summertime photochemistry of the troposphere at high northern latitudes, Journal of Geophysical Research: Atmospheres, 97, 16421 – 16431, doi: 10.1029/91JD01968, 1992b.

Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., and Jaagus, J.: Validation of atmospheric reanalyses over the central Arctic Ocean, Geophysical Research Letters, 39, L10 802, doi: 10.1029/2012GL051591, 2012.

Jung, T. and Matsueda, M.: Verification of global numerical weather forecasting systems in polar regions using TIGGE data, Quart. J. Roy. Meteor. Soc., in press. doi:10.1002/qj.2437, 2014.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, B., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bull. Amer. Meteor. Soc., 77, 437–471, 1996.

Karion, A., Miller, J. B., Sweeney, C., Newberger, T., Wolter, S., Henderson, J., Lehman, S., Dinardo, S., and Miller, C.: In situ greenhouse gas measurements from Boreal Alaska, in preparation.

Kort, E. A., Eluszkiewicz, J., Stephens, B. B., Miller, J. B., Gerbig, C., Nehrkorn, T., Daube, B. C., Kaplan, J. O., Houweling, S., and Wofsy, S. C.: Emissions of CH_4 and N_2O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations, Geophys. Res. Lett., 35, L18 808, doi:10.1029/2008GL034031, 2008.

Kort, E. A., Andrews, A. E., Dlugokencky, E., Sweeney, C., Hirsch, A., Eluszkiewicz, J., Nehrkorn, T., Michalak, A., Stephens, B., Gerbig, C., Miller, J. B., Kaplan, J., Houweling, S., Daube, B. C., Tans, P., and Wofsy, S. C.: Atmospheric constraints on 2004 emissions of methane and nitrous oxide in North America from atmospheric measurements and a receptor-oriented modeling framework, Journal of Integrative Environmental Sciences, 7, 125–133, available online at http://www.informaworld.com/10.1080/19438151003767483, 2010.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport Model (STILT), J. Geophys. Res., 108, 4493, 10.1029/2002JD003 161, 2003.

Liu, J., Curry, J. A., Wang, H., Song, M., and Horton, R. M.: Impact of declining Arctic sea ice on winter snowfall, Proceed. National Academy Sci., 109, 4074–4079, doi:10.1073/pnas.1114910109, 2012.

Lo, J. C.-F., Yang, Z.-L., and Pielke, R. A.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, Journal of Geophysical Research: Atmospheres, 113, D09 112, doi: 10.1029/2007JD009216, 2008.

Manabe, S. and Stouffer, R. J.: Sensitivity of a global climate model to an increase of CO2 concentration in the atmosphere, Journal of Geophysical Research: Oceans, 85, 5529 – 5554, doi: 10.1029/JC085iC10p05529, 1980.

Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A.: Does Increas- ing Horizontal Resolution Produce More Skillful Forecasts?, Bulletin of the American Meteorological Society, 83, 407 – 430, doi: 10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2, 2002.

Matross, D. M., Andrews, A., Pathmathevan, M., Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Lee, J. T., Zhao, C., Bakwin, P. S., Munger, J. W., and Hollinger, D. Y.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, Tellus B, 58, 344–358, doi:10.1111/j.1600-0889.2006.00206.x, 2006.

McFarlane, S. A., Shi, Y., and Long, C. N.: A year of radiation measurements at the North Slope of Alaska, ARM and Climate Change Prediction Program Metric Report DOESC-ARMP-09-010, Norman, OK, available online at http://www.arm. gov/publications/programdocs/doe-sc-arm-p-09- 010.pdf, 2009.

Mckain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of Ground-based Atmospheric Observations for Verification of Greenhouse Gas Emissions from an Urban Region, Proceed. National Academy Sci., 109, 8423–8428, doi:10.1073/pnas.1116645109, 2012.

Mikkelsen, T., Ro-Poulsen, H., Hovmand, M., Jensen, N., Pilegaard, K., and Egelv, A.: Five-year measurements of ozone fluxes to a Danish Norway spruce canopy, Atmospheric Environment, 38, 2361–2371, doi: http://dx.doi.org/10.1016/j.atmosenv.2003.12.036, 2004.

Miller, C. E. and Dinardo, S. J.: CARVE: The Carbon in Arctic Reservoirs Vulnerability Experiment, in: Aerospace Conference, IEEE, Big Sky, MT, doi:10.1109/AERO.2012.6187026, 2012.

Miller, S. M., Kort, E. A., Hirsch, A. I., Dlugokencky, E. J., Andrews, A. E., Xu, X., Tian, H., Nehrkorn, T., Eluszkiewicz, J., Michalak, A. M., and Wofsy, S. C.: Regional sources of

nitrous oxide over the United States: Seasonal variation and spatial distribution, J. Geophys. Res., 117, D06 310, doi:10.1029/2011JD016951, 2012.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, Proceed. National Academy Sci., 110, 20018–20022, doi:10.1073/pnas.1314392110, 2013.

Miller, C. E., Dinardo, S. J., Bruhwiler, L., Fung, I., Koven, C., McDonald, K., Miller, J., Oechel, W., Podest, E., Randerson, J., Rayner, P., Rider, D., Sweeney, C., Wennberg, P., Wofsy, S., Henderson, J. M., Karion, A., Zona, D., and Fisher, J.: CARVE – The Carbon in Arctic Reservoirs Vulnerability Experiment: Overview and Preliminary Results, in preparation.

Moore, G. W. K.: The Novaya Zemlya Bora and its impact on Barents Sea air-sea interaction, Geophysical Research Letters, 40, 3462 – 3467, doi: 10.1002/grl.50641, 2013.

Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S.: Coupled weather research and forecasting- stochastic time-inverted Lagrangian transport (WRF-STILT) model, Meteor. Atmos. Phys., 107, 51–64, doi:10.1007/s00703-010-0068-x, 2010.

NOAA: Storm Reports, Tech. rep., NOAA, available from http://www.ncdc.noaa.gov/IPS/sd/sd.html, 2012a.

Peterson, T. C., Hoerling, M. P., Stott, P. A., and Herring, S.: Explaining Extreme Events of 2012 from a Climate Perspective, Bull. Amer. Meteor. Soc., 94(9), S1-S74.

Powers, J. G., Manning, K. W., Bromwich, D. H., Cassano, J. J., and Cayette, A. M.: A Decade of Antarctic Science Support Through Amps, Bulletin of the American Meteorological Society, 93, 1699 – 1712, doi: 10.1175/BAMS-D- 11-00186.1, 2012.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numeric Recipes in FORTRAN: The Art of Scientific Computing, Second Edition, Cambridge University Press, New York, 1982.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A.,

da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Screen, J., Deser, C., Simmonds, I., and Tomas, R.: Atmospheric impacts of Arctic sea-ice loss, 19792009: separating forced change from atmospheric internal variability, Climate Dynamics, 43, 333–344, doi: 10.1007/s00382-013-1830-9, 2014.

Serreze, M. C. and Barry, R. G.: Processes and impacts of Arctic amplification: A research synthesis, Global and Planetary Change, 77, 85 – 96, doi: http://dx.doi.org/10.1016/j.gloplacha.2011.03.004, 2011.

Simmonds, I. and Rudeva, I.: The great Arctic cyclone of August 2012, Geophys. Res. Lett., 39, L23 709, 6 pp. doi:10.1029/2012GL054259, 2012.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Wang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, Technical Note 475+STR, MMM Division, NCAR, Boulder, CO, 113 pp. Available online at http://www.mmm.ucar.edu/wrf/users/docs/arw v3.pdf, 2008.

Stohl, A.: Computation, accuracy and applications of trajectories – a review and bibliography, Atmos. Environ., 32, 947–966, 1998.

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, Geophysical Research Letters, 39, L16502, doi: 10.1029/2012GL052676, 2012.

Tilinina, N., Gulev, S. K., and Bromwich, D. H.: New view of Arctic cyclone activity from the Arctic system reanalysis, Geophysical Research Letters, 41, 2013GL058 924, doi: 10.1002/2013GL058924, 2014.

Tjernström, M., Leck, C., Persson, P. O. G., Jensen, M. L., Oncley, S. P., and Targino, A.: The Summertime Arctic Atmosphere: Meteorological Measurements During the Arctic Ocean Experiment 2001, Bull. Amer. Meteor. Soc., 85, 1305–1321, 2004.

Tjernström, M.: Is There a Diurnal Cycle in the Summer Cloud-Capped Arctic Boundary Layer?, Journal of the Atmospheric Sciences, 64, 3970 – 3986, doi: 10.1175/2007JAS2257.1, 2007.

Tuovinen, J.-P., Aurela, M., and Laurila, T.: Resistances to ozone deposition to a flark fen in the northern aapa mire zone, Journal of Geophysical Research: Atmospheres, 103, 16 953 – 16 966, doi: 10.1029/98JD01165, 1998.

Tuovinen, J.-P., Ashmore, M., Emberson, L., and Simpson, D.: Testing and improving the EMEP ozone deposition module, Atmospheric Environment, 38, 2373 – 2385, doi: http://dx.doi.org/10.1016/j.atmosenv.2004.01.026, 2004.

Uliasz, M.: Lagrangian particle modeling in mesoscale applications, in: Environmental Modeling II, edited by Zannetti, P., pp. 71–102, Computational Mechanics Publications, Berlin, Heidelberg, 1994.

Uttal, T., Curry, J. A., Mcphee, M. G., Perovich, D. K., Moritz, R. E., Maslanik, J. A., Guest, P. S., Stern, H. L., Moore, J. A., Turenne, R., Heiberg, A., Ser- reze, M. C., Wylie, D. P., Persson, O. G., Paulson, C. A., Halle, C., Morison, J. H., Wheeler, P. A., Makshtas, A., Welch, H., Shupe, M. D., Intrieri, J. M., Stamnes, K., Lindsey, R. W., Pinkel, R., Pegau, W. S., Stanton, T. P., and Grenfeld, T. C.: Surface Heat Budget of the Arctic Ocean, Bulletin of the American Meteorological Society, 83, 255 – 275, doi: 10.1175/1520-0477(2002)083<0255:SHBOTA>2.3.CO;2, 2002.

Vihma, T., Pirazzini, R., Renfrew, I. A., Sedlar, J., Tjernström, M., Nygrd, T., Fer, I., Lpkes, C., Notz, D., Weiss, J., Marsan, D., Cheng, B., Birnbaum, G., Ger- land, S., Chechin, D., and Gascard, J. C.: Advances in understanding and parameterization of small-scale physical processes in the marine Arctic climate system: a review, Atmos. Chem. Phys. Discuss., 13, 32703 – 32816, doi: 10.5194/acpd-13-32703-2013, 2013.

Vosper, S. B: Inversion effects on mountain lee waves, Quart. J. Roy. Meteor., 130, 1723-1748, doi: 10.1256/qj.03.63.

Walker, T. W., Jones, D. B. A., Parrington, M., Henze, D. K., Murray, L. T., Bottenheim, J. W., Anlauf, K., Worden, J. R., Bowman, K. W., Shim, C., Singh, K., Kopacz, M., Tarasick, D. W., Davies, J., von der Gathen, P., Thompson, A. M., and Carouge, C. C.: Impacts of midlatitude precursor emissions and local photochemistry on ozone abundances in the Arctic, Journal of Geophysical Research: Atmospheres, 117, D01 305, doi: 10.1029/2011JD016370, 2012.

Wesslén, C., Tjernström, M., Bromwich, D. H., de Boer, G., Ekman, A. M. L., Bai, L.-S., and Wang, S.-H.: The Arctic summer atmosphere: an evaluation of reanalyses using ASCOS data, Atmos. Chem. Phys., 14, 2605 – 2624, doi: 10.5194/acp-14-2605-2014, 2014.

Wilson, A. B., Bromwich, D. H., and Hines, K. M.: Evaluation of Polar WRF forecasts on the Arctic System Reanalysis domain: Surface and upper air analysis, J. Geophys. Res., 116, D11 112, doi:10.1029/2010JD015013, 2011.

Yver, C. E., Graven, H. D., Lucas, D. D., Cameron-Smith, P. J., Keeling, R. F., and Weiss, R. F.: Evaluating transport in the WRF model along the California coast, Atmos. Chem. Phys., 13, 1837–1852, doi:10.5194/acp-13-1837-2013, 2013.

Zhang, J. and Rothrock, D. A.: Modeling Global Sea Ice with a Thickness and Enthalpy Distribution Model in Generalized Curvilinear Coordinates, Monthly Weather Review, 131, 845 – 861, doi: 10.1175/1520- 0493(2003)131<0845:MGSIWA>2.0.CO;2, 2003.

Zhao, C., Andrews, A. E., Bianco, L., Eluszkiewicz, J., Hirsch, A., MacDon- ald, C., Nehrkorn, T., and Fischer, M. L.: Atmospheric inverse estimates of methane emissions from Central California, J. Geophys. Res., 114, D16 302, doi:10.1029/2008JD011671, 2009.

List of Tables

Table 1. CARVE 2012 WRF v3.4.1 Configuration and Physics Options.

Table 2. 2012 WRF v3.4.1 Model Biases for Selected Surface Variables.

Table 3. Count of observed wind speeds (ws) by wind speed (in kt) for the 2012 Campaign.

Table 4. 2012 WRF v3.4.1 Model RMS Error for Selected Surface Variables. Wind direction errors computed using observations greater or equal to 3 kt (~1.5 m s^{-1}) in this table and all later tables. Positive direction is clockwise.

Table 5. 2012 Campaign Model Bias/RMS Error for Upper-Level Variables for WRF v3.4.1.

Figure Captions

Figure 1. STILT footprints (shaded, units are ppmv/(micromole m^{-2} s^{-1})) and HYSPLIT trajectories from receptors located at 100 m (solid), 500 m (dashed) and 1000 m (dotted black lines) at 1935 UTC 20 August 2012.

Figure 2. CARVE flight tracks from a) 2012 and b) 2013 superimposed on WRF innermost domain 3, and c) placement of nested WRF domains used for CARVE modeling, with model topography field (shaded, m) for outermost domain 1 (30-km grid spacing). Nested subdomains shown by green rectangles: domain 2 (d2) with 10-km grid spacing covers Eastern Russia and Western Canada, while innermost domain (domain 3, d3, 3.3-km grid spacing) covers mainland Alaska. Innermost domain model topography field (shaded, m) is shown in panels a) and b).

Figure 3. WRF v3.4.1 model bias by time of day (UTC; subtract 8 h for Alaska Daylight Savings Time, AKDT) for a) temperature (K), b) dewpoint temperature (K) and c) wind speed (m s^{-1}). Each curve and color indicates a different period of time (either monthly or the full 2012 Campaign) over which the errors are computed.

Figure 4. As in Fig. 3, but for RMS error.

Figure 5. Wind roses for a) model and b) observations for the entire 2012 Campaign. Model-observation pairs are retained only if the observed wind speed is >= 3 knots. The unit of wind speed in these plots is m s^{-1}.

Figure 6. Spatial distribution of WRF v3.4.1 bias on innermost domain (3.3-km grid spacing) for the entire 2012 Campaign for a) temperature (K), b) dewpoint temperature (K) and c) wind speed (m s^{-1}).

Figure 7. As in Fig. 6, but for RMS error.

Figure 8. Time series from WRF v3.4.1 of model (red) and observations (black) at McGrath, AK (station ID: PAMC) for the 2012 Campaign for a) temperature (K), b) dewpoint temperature (K) and c) wind speed (m s^{-1}). All wind observations are plotted.

Figure 9. As in Fig. 5, but for McGrath, AK, for a) model and b) observations.

Figure 10. As in Fig. 8, but for Deadhorse, AK (station ID: PASC).

Figure 11. As in Fig. 5, but for Deadhorse, AK.

Figure 12. As in Fig. 8, but for Barrow, AK (station ID: PABR).

Figure 13. As in Fig. 5, but for Barrow, AK.

Figure 14. Cross-sections of model potential temperature (panels a and b, contoured every 2 K) from domain 3 (panels a and c) and domain 1 (panels b and d) and locator maps (panels c and d, red lines near 143^{o} W) valid at 1100 UTC 17 September 2012. Cross-sections are along model grid column 439 of domain 3 (panel c) and grid column 220 of domain 1 (panel d). Wind speed (kt) on panels c and d is shaded (note scale change between top and bottom rows). Model surface pressure trace shown as joined blue circles on each cross-section plot (panels a and b). Locations of orange and light blue circles on surface pressure trace (panels a and b) are shown in panels c and d. Line near 144.5^{o}W on panel c) shows location of cross-section in Fig. 15. MSLP (contoured) and wind field, using standard representation of grid-point wind barbs (every second grid point for domain 3), shown in panels c and d. Locations of Tanacross (63.38^{o} N, 143.36 ^{o} W) and Dry Creek (63.68 ^{o} N, 144.60 ^{o} W) denoted by black-outlined circle and solid black circles, respectively.

Figure 15. As in Fig. 14a, but along column 418 of domain 3. Location of cross-section is denoted in Fig. 14c near 144.5^{o} W.

Figure 16. Aggregate of all STILT footprints (units are ppmv/(micromole m^{-2} s^{-1}) on 0.1-degree grid for 120 hourly receptors placed at the CARVE tower during the period 0000 UTC 13 September to 2300 UTC 17 September 2012. Panels a and b (c and d) are for receptors at 301 (32) m AGL. Winds from domains 1, 2 and 3 are used in panels a and c; winds only from domain 1 (the outermost domain) are used in panels b and d. The black circle in the center of the grid marks the tower location. Note the change in colour scale between the top and bottom rows.

Figure 17. As in Fig. 16, except for the aggregate of 79 CARVE tower footprints on a 0.5-degree grid (panels a and c) and 0.1-degree grid (panels b and d). The 0.5-degree footprints are scaled by a factor of 1/25.

Figure 18. Vertical profiles of methane mixing ratios near Fairbanks, AK, on 21 August 2012 based on a) chemical and dynamical tracers as observed by the CARVE aircraft (in black), and b) modeled profile based on an aggregation of 370 WRF-STILT footprint fields (in red). Dashed black line at approximately 2400 m in each panel represents the top of the

1 atmospheric column enhancement as determined using observations from panel a. Red dashed
2 line in panel b at approximately 2000 m denotes the top of the atmospheric column
3 enhancement as defined by WRF-STILT.

4 Figure 19. Deviation of modeled and measured ozone concentrations (ppb) from the means of
5 each flight for each month of the 2012 Campaign (panels a-e). Computed deposition velocity
6 (cm s^{-1}) is shown in the lower right corner of each panel.
7

Tables

Table 1. CARVE 2012 WRF v3.4.1 Configuration and Physics Options.

Physics	**Option Selected**
Microphysics	Morrison 2-moment
LW Radiation	RRTMG
SW Radiation	RRTMG
Surface Layer	MYNN
Land Surface Model	Noah LSM with polar modifications
Land Surface dataset	USGS
PBL Scheme	MYNN 2.5 TKE
Cumulus	G-D ensemble (d1+d2)
Fractional Sea Ice	Binary, source: NCAR/NCEP Reanalysis Project (NNRP)
Time step	90 s
Horizontal grid spacing	d1=30, d2=10 and d3=3.3 km
Grid dimensions	418x418, 799x649 and 550x550 grid points
Vertical resolution	41 eta levels
Feedback	Two-way nesting

Table 2. 2012 WRF v3.4.1 Model Biases for Selected Surface Variables. Wind direction errors are computed using observations greater or equal to 3 kt (~1.5 m s^{-1}) in this table and all later tables. Positive direction is clockwise.

Surface Variable	May	June	July	August	September	2012 Campaign
2-m Temperature (K)	-2.25	-1.74	-1.57	-1.06	-0.66	-1.40
2-m Dewpoint temperature (K)	1.01	0.07	-0.78	-0.67	-0.08	-0.16
10-m Wind Speed (m s^{-1})	-0.56	-0.34	-0.18	-0.21	0.36	-0.17
10-m Wind Direction (deg)	5.0	3.5	1.9	4.7	4.1	3.8

Table 3. Count of observed wind speeds (ws) by wind speed (in kt) for the 2012 Campaign.

0.0	**0<ws<=1**	**1<ws<=2**	**2<ws<=3**	**3<ws<4**	**4<ws<=5**	**5<ws<=6**
71371	2524	6933	33495	8259	38795	38774

Table 4. 2012 WRF v3.4.1 Model RMS Error for Selected Surface Variables. Wind direction errors computed using observations greater or equal to 3 kt (~1.5 m s^{-1}) in this table and all later tables. Positive direction is clockwise.

Surface Variable	May	June	July	August	September	2012 Campaign
2-m Temperature (K)	3.43	3.46	3.01	2.39	1.97	2.97
2-m Dewpoint temperature (K)	2.88	2.67	2.51	2.17	2.09	2.52
10-m Wind speed (m s^{-1})	2.12	1.99	2.00	2.11	2.48	2.19
10-m Wind Direction (deg)	58.0	55.5	53.2	49.0	46.7	52.1

Table 5. 2012 Campaign Model Bias/RMS Error for Upper-Level Variables for WRF v3.4.1.

Pressure Level (hPa)	Temperature (K)	Geopotential Height (m)	Relative Humidity (%)	Wind Speed ($m\ s^{-1}$)
200	-0.06/1.80	0.2/16.2	-7.6/14.1	0.36/3.11
300	0.39/1.38	-4.9/18.1	10.8/21.1	-0.20/4.88
500	-0.05/0.98	-4.1/13.7	5.6/22.7	-0.06/3.15
700	-0.08/1.05	-2.5/11.4	4.5/20.1	0.08/3.00
850	-0.17/1.16	-6.4/12.6	4.3/16.3	0.17/3.05

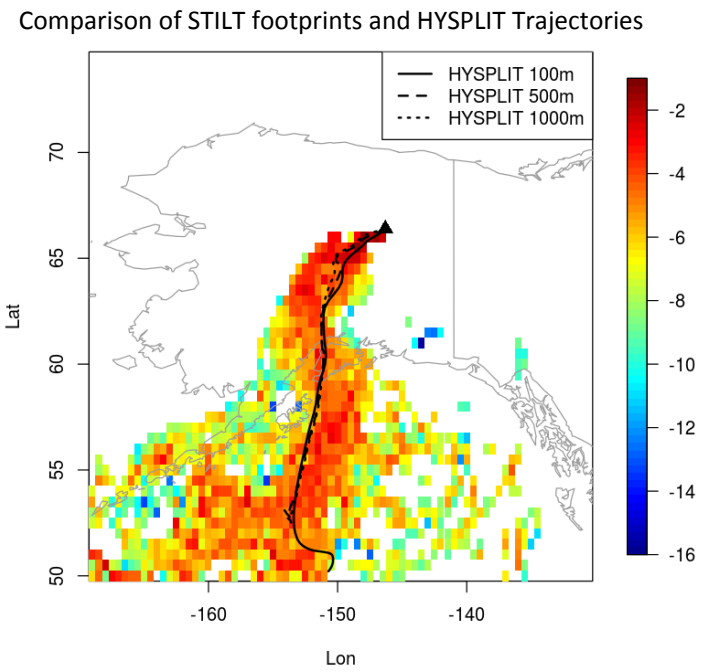

Figure 1. STILT footprints (shaded, units are ppmv/(micromole m^{-2} s^{-1})) and HYSPLIT trajectories from receptors located at 100 m (solid), 500 m (dashed) and 1000 m (dotted black lines) at 1935 UTC 20 August 2012.

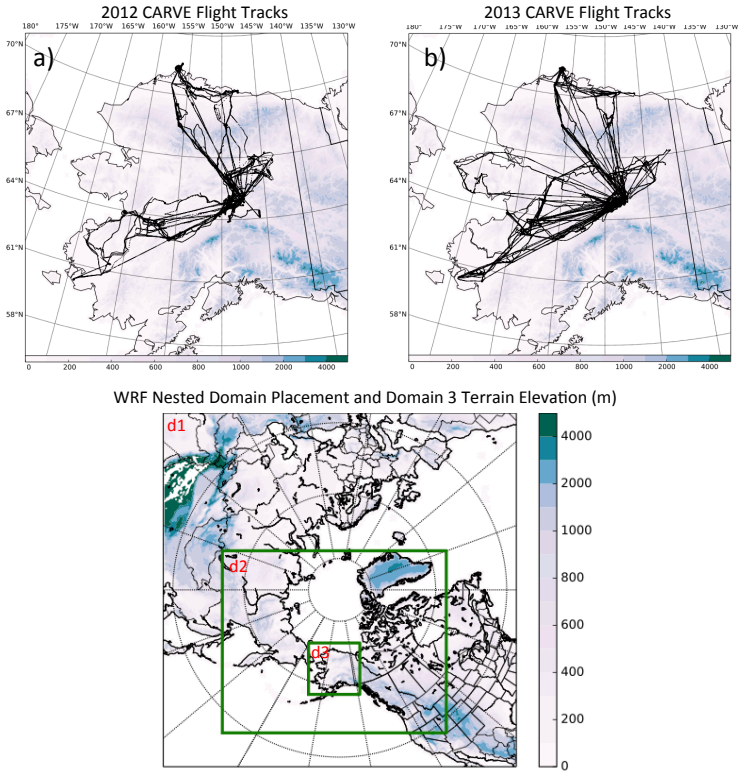

Figure 2. CARVE flight tracks from a) 2012 and b) 2013 superimposed on WRF innermost domain 3, and c) placement of nested WRF domains used for CARVE modeling, with model topography field (shaded, m) for outermost domain 1 (30-km grid spacing). Nested subdomains shown by green rectangles: domain 2 (d2) with 10-km grid spacing covers Eastern Russia and Western Canada, while innermost domain (domain 3, d3, 3.3-km grid spacing) covers mainland Alaska. Innermost domain model topography field (shaded, m) is shown in panels a) and b).

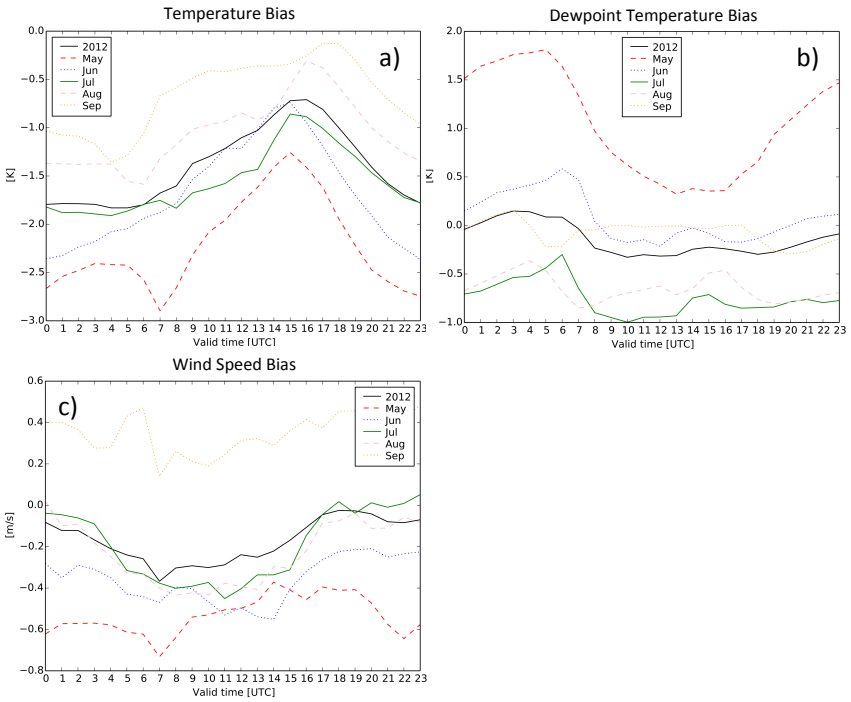

Figure 3. WRF v3.4.1 model bias by time of day (UTC; subtract 8 h for Alaska Daylight Savings Time, AKDT) for a) temperature (K), b) dewpoint temperature (K) and c) wind speed (m s^{-1}). Each curve and color indicates a different period of time (either monthly or the full 2012 Campaign) over which the errors are computed.

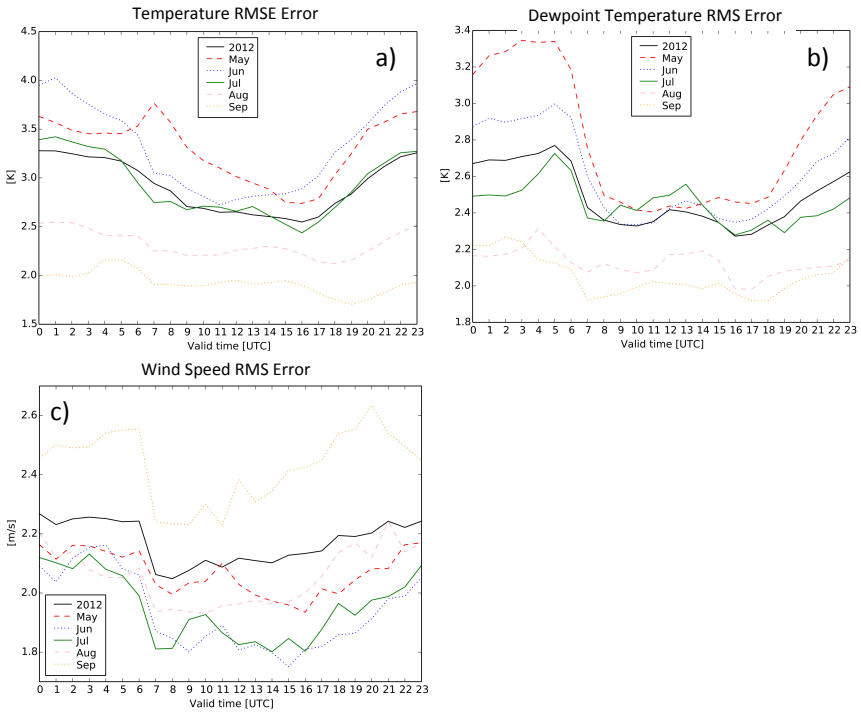

Figure 4. As in Fig. 3, but for RMS error.

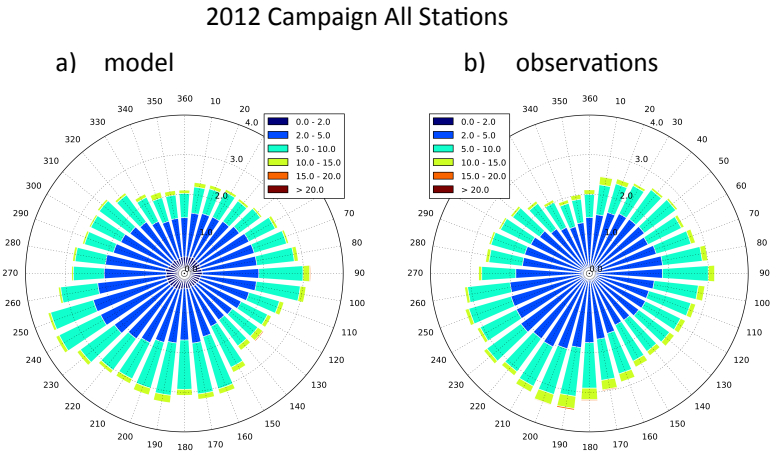

Figure 5. Wind roses for a) model and b) observations for the entire 2012 Campaign. Model-observation pairs are retained only if the observed wind speed is >= 3 knots. The unit of wind speed in these plots is m s^{-1}.

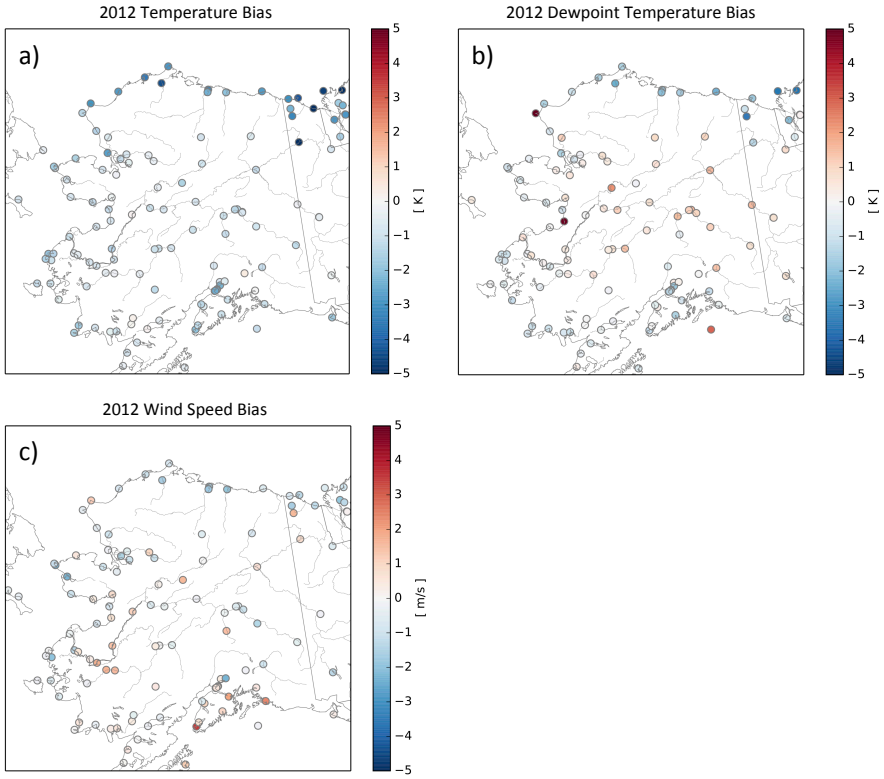

Figure 6. Spatial distribution of WRF v3.4.1 bias on innermost domain (3.3-km grid spacing) for the entire 2012 Campaign for a) temperature (K), b) dewpoint temperature (K) and c) wind speed (m s^{-1}).

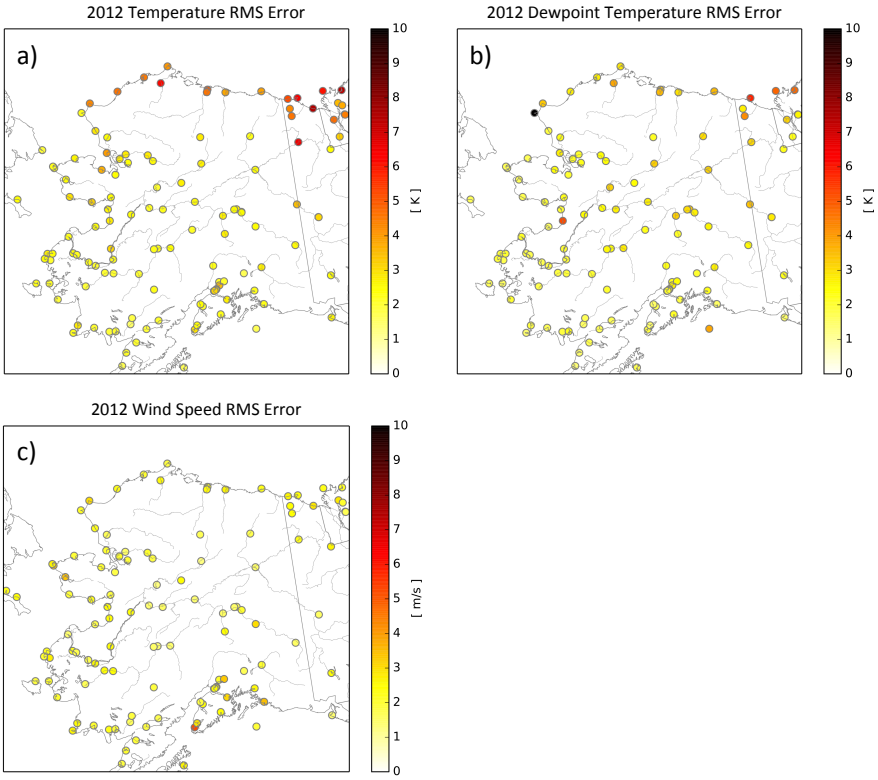

Figure 7. As in Fig. 6, but for RMS error.

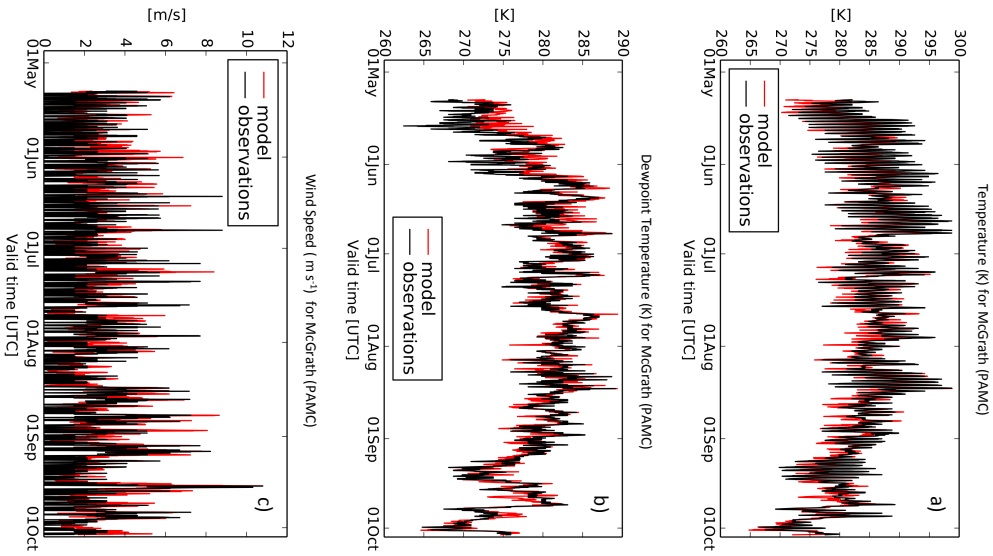

Figure 8. Time series from WRF v3.4.1 of model (red) and observations (black) at McGrath, AK (station ID: PAMC) for the 2012 Campaign for a) temperature (K), b) dewpoint temperature (K) and c) wind speed (m s^{-1}). All wind observations are plotted.

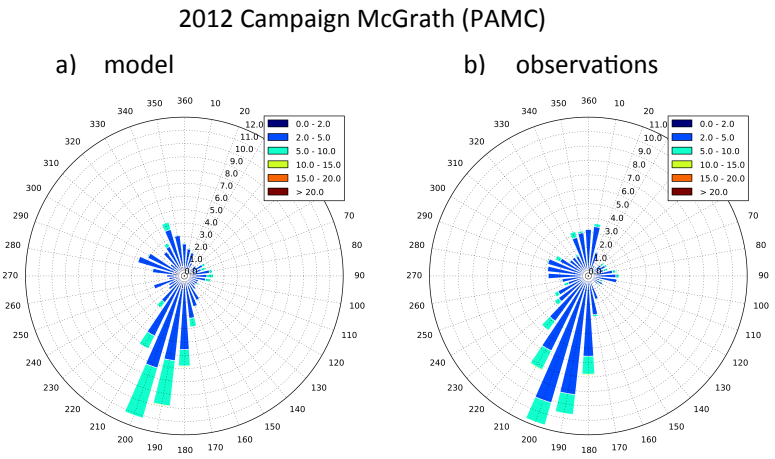

Figure 9. As in Fig. 5, but for McGrath, AK, for a) model and b) observations.

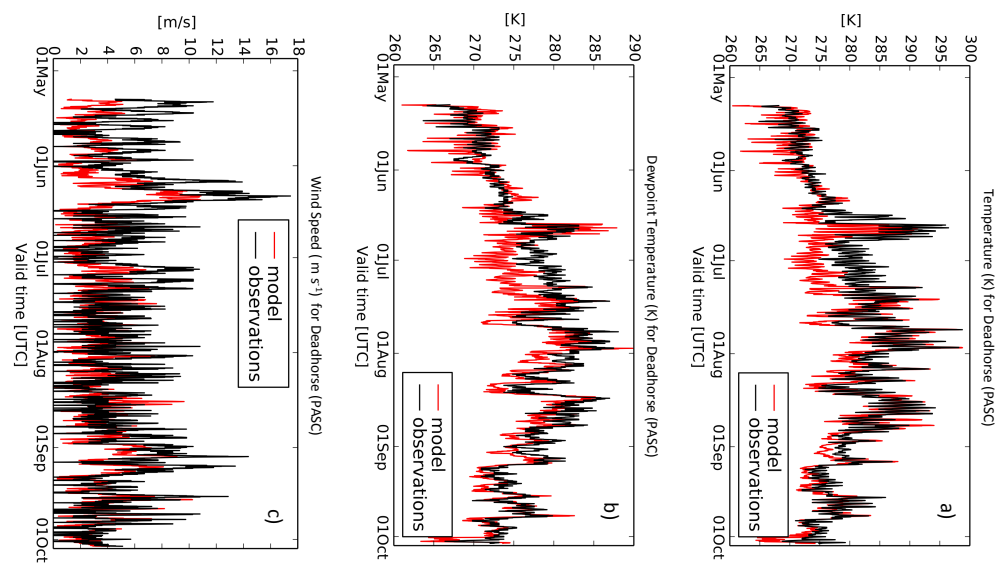

Figure 10. As in Fig. 8, but for Deadhorse, AK (station ID: PASC).

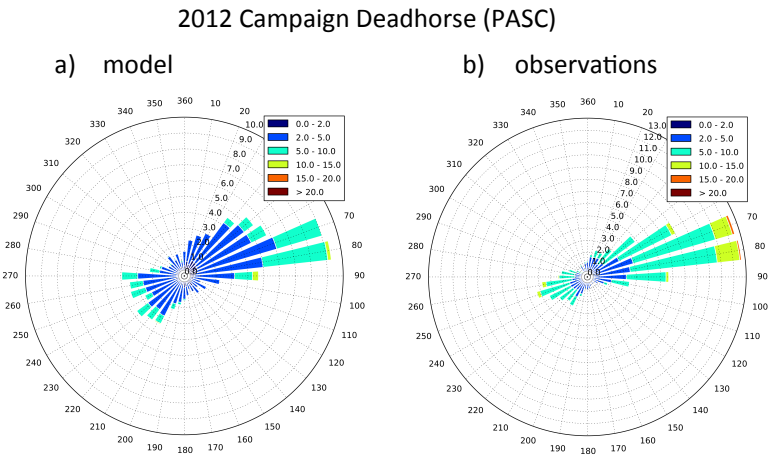

Figure 11. As in Fig. 5, but for Deadhorse, AK.

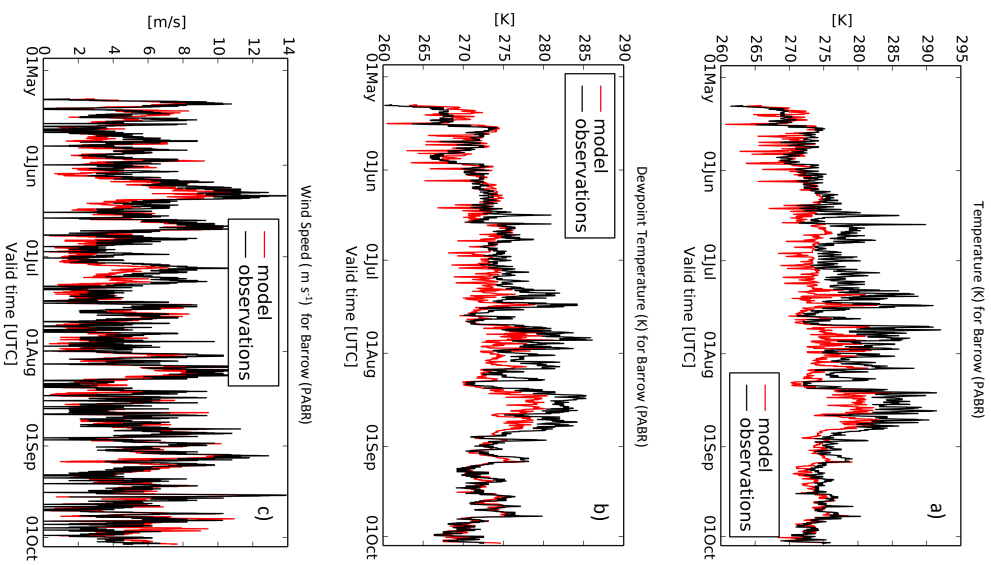

Figure 12. As in Fig. 8, but for Barrow, AK (station ID: PABR).

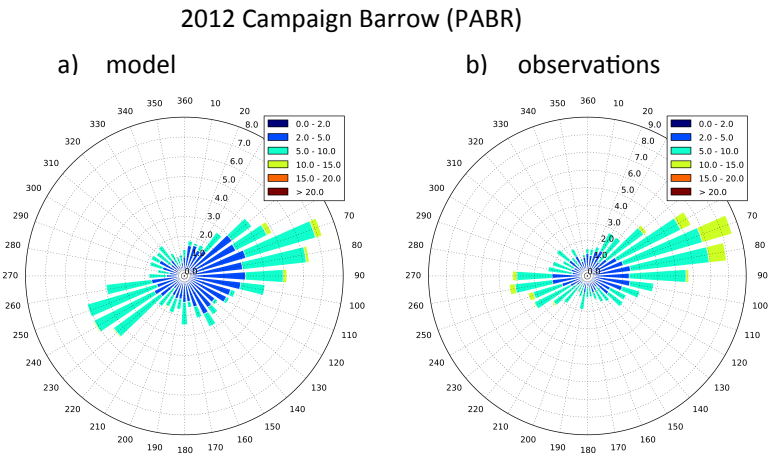

Figure 13. As in Fig. 5, but for Barrow, AK.

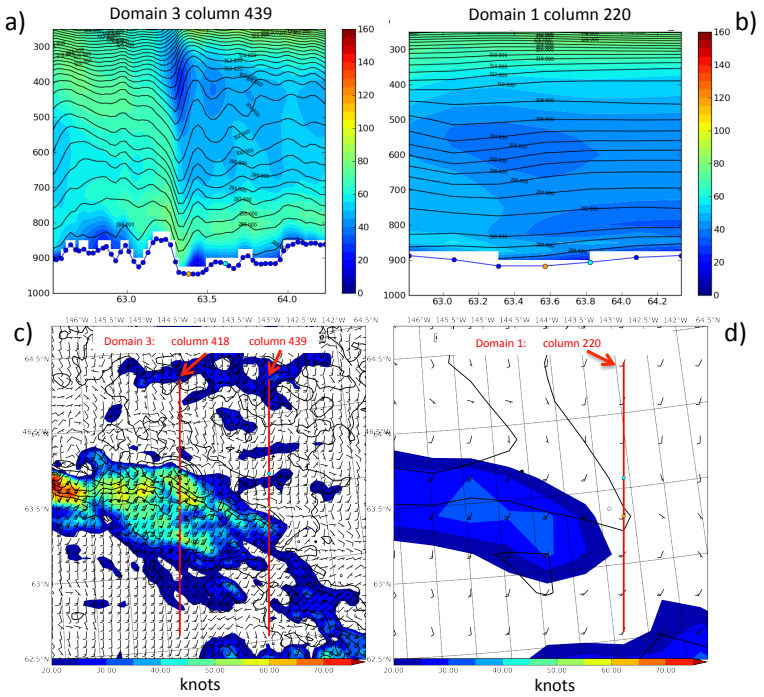

Figure 14. Cross-sections of model potential temperature (panels a and b, contoured every 2 K) from domain 3 (panels a and c) and domain 1 (panels b and d) and locator maps (panels c and d, red lines near 143° W) valid at 1100 UTC 17 September 2012. Cross-sections are along model grid column 439 of domain 3 (panel c) and grid column 220 of domain 1 (panel d). Wind speed (kt) on panels c and d is shaded (note scale change between top and bottom rows). Model surface pressure trace shown as joined blue circles on each cross-section plot (panels a and b). Locations of orange and light blue circles on surface pressure trace (panels a and b) are shown in panels c and d. Line near 144.5°W on panel c) shows location of cross-section in Fig. 15. MSLP (contoured) and wind field, using standard representation of grid-point wind barbs (every second grid point for domain 3), shown in panels c and d. Locations of Tanacross (63.38° N, 143.36 ° W) and Dry Creek (63.68 ° N, 144.60 ° W) denoted by black-outlined circle and solid black circles, respectively.

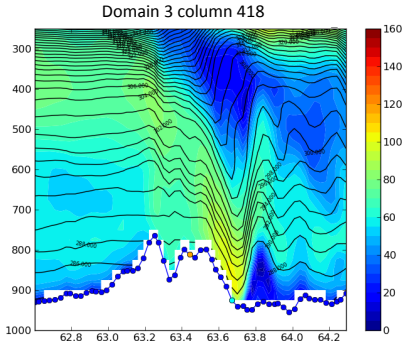

Figure 15. As in Fig. 14a, but along column 418 of domain 3. Location of cross-section is denoted in Fig. 14c near 144.5^{o} W.

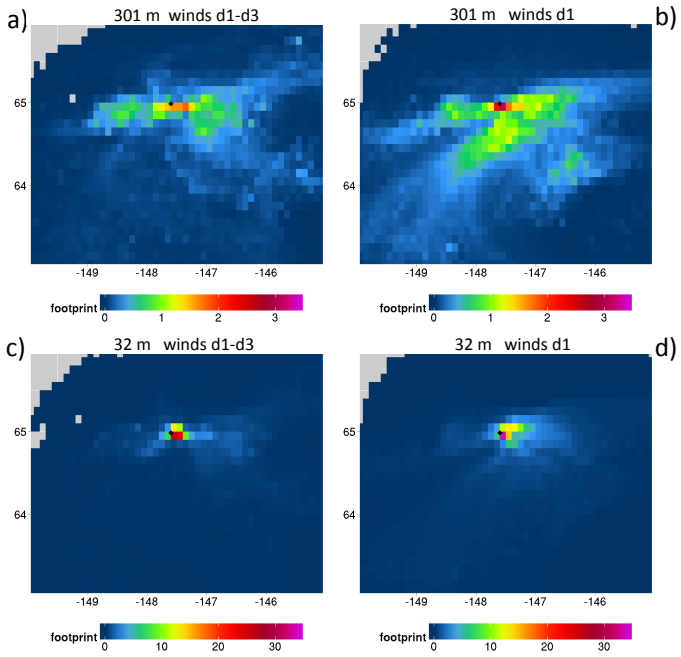

Figure 16. Aggregate of all STILT footprints (units are ppmv/(micromole m^{-2} s^{-1}) on 0.1-degree grid for 120 hourly receptors placed at the CARVE tower during the period 0000 UTC 13 September to 2300 UTC 17 September 2012. Panels a and b (c and d) are for receptors at 301 (32) m AGL. Winds from domains 1, 2 and 3 are used in panels a and c; winds only from domain 1 (the outermost domain) are used in panels b and d. The black circle in the center of the grid marks the tower location. Note the change in colour scale between the top and bottom rows.

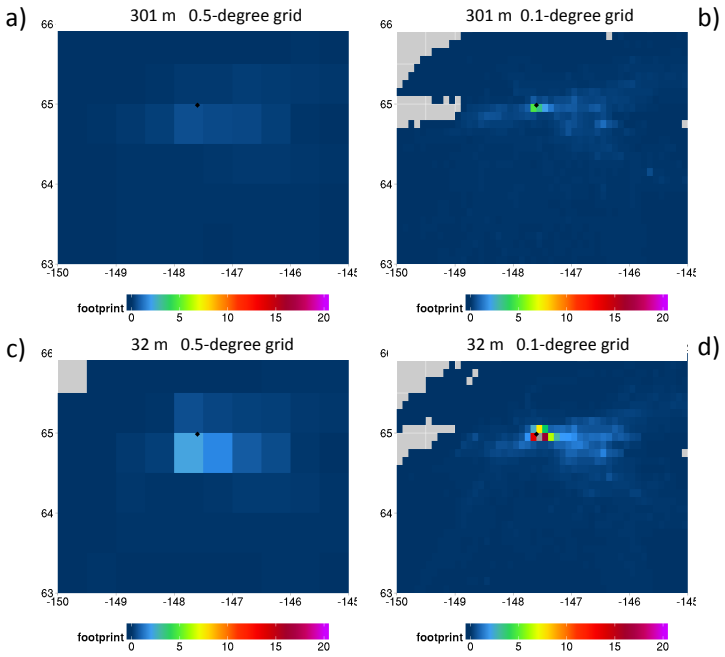

Figure 17. As in Fig. 16, except for the aggregate of 79 CARVE tower footprints on a 0.5-degree grid (panels a and c) and 0.1-degree grid (panels b and d). The 0.5-degree footprints are scaled by a factor of 1/25.

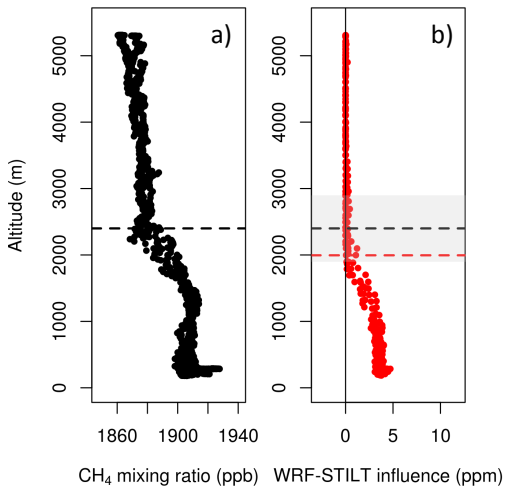

Figure 18. Vertical profiles of methane mixing ratios near Fairbanks, AK, on 21 August 2012 based on a) chemical and dynamical tracers as observed by the CARVE aircraft (in black), and b) modeled profile based on an aggregation of 370 WRF-STILT footprint fields (in red). Dashed black line at approximately 2400 m in each panel represents the top of the atmospheric column enhancement as determined using observations from panel a. Red dashed line in panel b at approximately 2000 m denotes the top of the atmospheric column enhancement as defined by WRF-STILT.

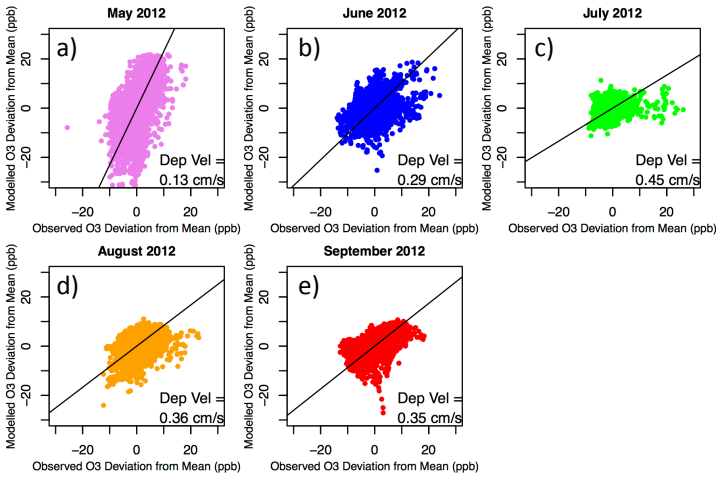

Figure 19. Deviation of modeled and measured ozone concentrations (ppb) from the means of each flight for each month of the 2012 Campaign (panels a-e). Computed deposition velocity (cm s^{-1}) is shown in the lower right corner of each panel.