Errors in mixed layer heights over North America: a multi-model comparison

by

Myung G. Kim

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Earth Sciences

Waterloo, Ontario, Canada, 2011

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Vertical mixing is an important process that relates surface fluxes to concentrations of pollutants and other chemical species in the atmosphere. Errors in vertical mixing have been identified as a major source of uncertainties in various atmospheric modeling efforts including tracer transport, weather forecasting, and regional climate simulation. This thesis aims to quantify uncertainties in model-derived mixed layer heights (z_i) over North America through direct comparisons between radiosonde observations and four models at different months of the year 2004 through the bulk Richardson number method. Results of this study suggest that considerable errors in z_i exist throughout the region with the spatial and temporal variations of the errors differ significantly among the selected models. Over all, errors in z_i were larger in global models than in the limited area mesoscale models, and the magnitude of the random error was two times larger than the bias. Notably, spatial regions of with extremely large positive biases correspond to those with especially large random errors. The biases and random errors, however, were not correlated linearly nor can be easily used to predict each other. Uncertainties in model-derived z_i were attributed, through errors in the bulk Richardson number, to temperature and horizontal winds. Errors in both horizontal winds and temperatures were found contributing more or less the same to uncertainties in z_i, with relative errors in both variables being the greatest in the lowest part of the troposphere. Lastly, independent observations from the cooperative profiler network suggest that data assimilation did not add qualitative advantages for the comparisons presented in this study. The mixed layer height uncertainties demonstrated in this study may provide a guide for selecting a model to simulate regional scale atmospheric transport and for interpreting flux estimation and inversions studies.

Acknowledgements

This research has been an invaluable experience to me, yielding a deeper understanding and appreciation of scientific work. Without the expertise, assistant, and support of many people, this work would not have been possible.

Especially, I would like to thank my supervisor Dr. Lin for this learning opportunity and also for his encouragement and patience from the beginning of this process. I would also like to thank fellow members of my research group, WatAIR, for aiding me to fulfill the research project.

Foremost, I thank my family for their love and dedication.

Dedication

To Eunjung for all she has done for me.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

Small errors in atmospheric transport modeling can result in significant uncertainties in linking fluxes of trace gases with the resulting concentration changes. Atmospheric carbon dioxide (CO_2), for instance, is taken up through photosynthesis with the surface biosphere and emitted through fossil fuel combustion during anthropogenic activities, with the resulting concentration changes in the lower troposphere dependent on both the surface flux as well as the depth of the mixing layer. Therefore, atmospheric transport models need to be able to relate concentration changes with atmospheric mixing in order to retrieve the surface fluxes [*Zhang and Rao*, 1999; *Gerbig et al*, 2003; *Lin and Gerbig*, 2005; *Ishizawa et al*, 2006; *Stephens*, 2007].

1.1 Background

Interactions between climate and land ecosystem processes are the central elements of potential climate impacts associated with habitats and components of the earth system such as human health, agriculture, hydrologic cycle, biogeochemical cycle, and biodiversity. Currently about half of anthropogenic carbon that has been emitted since the beginning of industrial age remains in the atmosphere to drive the climate change; the remainder is removed in about equal amounts by the land biosphere and the oceans. However, several researches using climate-carbon models indicate that the land carbon sink may be unstable under future climate change, leading to a potentially large positive feedback. Understanding what controls the cycle of the important gas between the atmosphere and land biosphere means there are hope to plan better for and, perhaps, mitigate climate change. If we learn that the boreal forests are absorbing CO_2, for example, then we can hope to better manage these forests to take up more CO_2.

To date, there have been efforts to consolidate regional details on or insights into processes driving fluxes to address the global carbon budgets. Employing transport models to link spatially distributed flux fields with atmospheric observations is one of the most promising techniques to achieve the goals. Hence, the needs to accurately account the uncertainties in the transport fields among models should be obvious: the transport, deposition, dispersion, or changes concentration depends on it.

Regional scale weather forecasting and reanalysis models may be useful by providing high resolution meteorology fields to transport models. Uncertainties, however, in vertical mixing within

the atmospheric boundary layer simulated by meteorological models have been significant issue identified by various regional atmospheric modeling efforts [*Pielke and Uliasz*, 1998; *Kim and Stockwell*, 2008; *Gerbig et al*, 2008]. Lately, *Aulagnier et al* [2010] suggested that recent build-up of CO_2 in lower atmosphere over Europe is transport-induced (i.e., because of changes in boundary layer heights) rather than due to changes in underlying fluxes. Therefore, if our goal is to constrain uncertainties in concentration of greenhouse gases, aerosol precursors, and other trace gases essential to air quality, then reliable estimates of mixed layer heights are needed.

1.2 Goals and Objectives

In this study, by comparing modeled meteorology from multiple atmospheric models having various spatial resolutions to radiosonde observation, multi-model comparisons of transport uncertainties over North America with focus on model derived mixed layer heights ($\boldsymbol{z_i}$) is demonstrated.

1.2.1 Goals

The following two goals were designed to address challenges identified in Section 1.1.

1) to quantify biases and random errors in z_i from selected models
2) to attribute source of errors to relevant meteorological variables

1.2.2 Objectives

In order to realize goals in Section 1.2.1, meteorological fields from four atmospheric models with different grid spacing are directly compared to radiosonde measurements over North America. Several objectives that are needed to be met include:

1) determine the mixed layer height, which works both for the observation and for models;
2) have a measure of errors in z_i with which the model uncertainties are reliably quantified;
3) present the errors in a way that reveals spatial and temporal distribution; and
4) propagate the error through responsible sources.

The thesis is organized in the following manner. Chapter 2 describes the models and observations employed in the analysis, followed by the methods for error assessment. Chapter 3 begins by laying out results of the study and investigates the general characteristics and causes of the errors. Finally, the last chapter concludes with summary and discussion of the results. Recommendation for future research is presented at the end.

Chapter 2 Methodology

2.1 Models

The four models examined in this study include the ones that are commonly used for modeling of tracer transport. These include the Weather Research and Forecasting (WRF) mesoscale model and three models run at the National Centers of Environmental Prediction (NCEP): the Global Data Assimilation System (GDAS), Eta Data Assimilation System (EDAS), and North American Regional Reanalysis (NARR). Table 1 shows grid spacing of the models used in current study. The WRF output used in this study is archived hourly at 36 km resolution [*Nehrkorn et al*, 2010]. GDAS covers the globe and is archived 6-hourly at 190.5 km grid spacing [*Derber et al*, 1991]. The EDAS dataset is centered on continental U.S. with a 3-hourly, 40 km resolution [*Matross et al*, 2006]. NARR is a 3-hourly archived reanalysis with 32 km grid spacing covering the North American continent [*Mesinger et al*, 2006].

NCEP models have been a popular choice for predicting air quality transport and dispersion [*Saha et al*, 2010]. WRF is considered as the state-of-the-art mesoscale model that has been adopted by forecasting and various research communities ever since its initial development [*Skamarock et al*, 2005; *Skamarock*, 2006]. Considering its fast growth and extensive use for applications in atmospheric transport, it is important to evaluate z_i uncertainties in WRF, along with other regional scale models that have equivalent grid spacing. Also, this selection of models allows the analysis to compare and contrast potential outcomes of employing meteorological models that were run at different grid resolutions for global and regional transport modeling purposes.

Table 1 Characteristics of models used.

	Horizontal Grid (km)	Archive	RAOB*	Reference
GDAS	190.5	6 hourly	Assimilated	*Derber et al*, [1991]
EDAS	40	3 hourly	Assimilated	*Matross et al*, [2006]
NARR	32	3 hourly	Assimilated	*Mesinger et al*, [2006]
WRF	36	1 hourly	Not Assimilated	*Nehrkorn et al*, [2010]

*RAOB: radiosonde observation

Most NCEP models including the ones used here were nudged towards the radiosonde observations (RAOB) measurements. However, none of the RAOB data used for this study was assimilated by WRF that was run with initial and boundary conditions from NARR. This difference in data assimilation could complicate the uncertainty assessment although nudging coefficients are usually known to be small [*Stauffer et al*, 1991; *Seaman et al*, 1995; *Hanna and Yang*, 2001]. We will use profiler observations as an independent dataset to examine this point (see Chapter 2.2).

2.2 Observations

Model uncertainties were determined from direct comparison with RAOB downloaded from NOAA [*NOAA*, 2009]. The NOAA RAOB dataset has been carefully QA/QC'ed; see *Schwartz and Govett* [1992] for more information. Since most radiosondes are launched twice daily at 00UTC and 12UTC, these two hours are chosen to probe diurnal variations of z_i in the comparison. The comparisons concentrate on Canada and U.S.A., during four months spanning the different seasons of 2004: January, April, July, and October. The number of RAOB sites from the NOAA archive used for this study was 128, and locations are shown in Figure 1.

In addition to RAOB, an independent observational dataset is introduced which has not been assimilated by any of the models compared in this study. Cooperative Agency Profilers (CAP) is a cooperative measurement network between NOAA [*Benjamin et al*, 2004] and multiple participating agencies around the world. CAP primarily measures vertical profiles of horizontal winds using Doppler radars in near real-time. With the optional addition of Radio Acoustic Sounding System (RASS), the profilers can also measure virtual temperatures up to about 1 km above ground level. The CAP data is available for download from the Meteorological Assimilation Data Ingest System (MADIS, [*Miller, Barth, and Benjamin*, 2005]). Observations obtained with the profilers and RASS have comparable vertical resolutions to RAOB's, and models-to-CAP comparisons are focused at 00UTC and 12UTC in the four months of the year 2004. The locations of the profilers and RASS used in this study are shown in Figure 2.

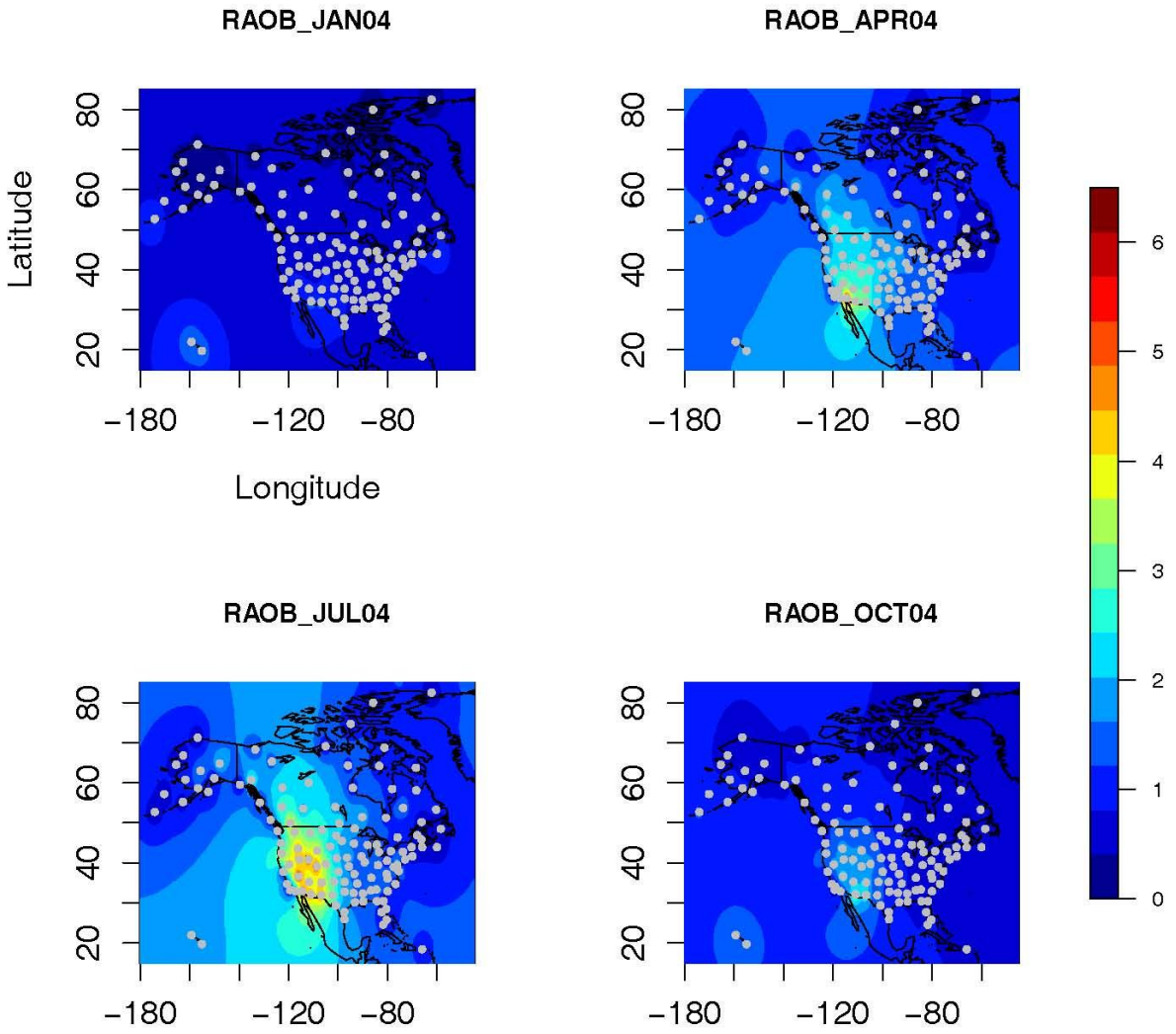

Figure 1 Monthly averaged mixing height [km] from RAOB, measured at 00UTC in January, April, July, and October of 2004. Grey dots indicate the RAOB locations.

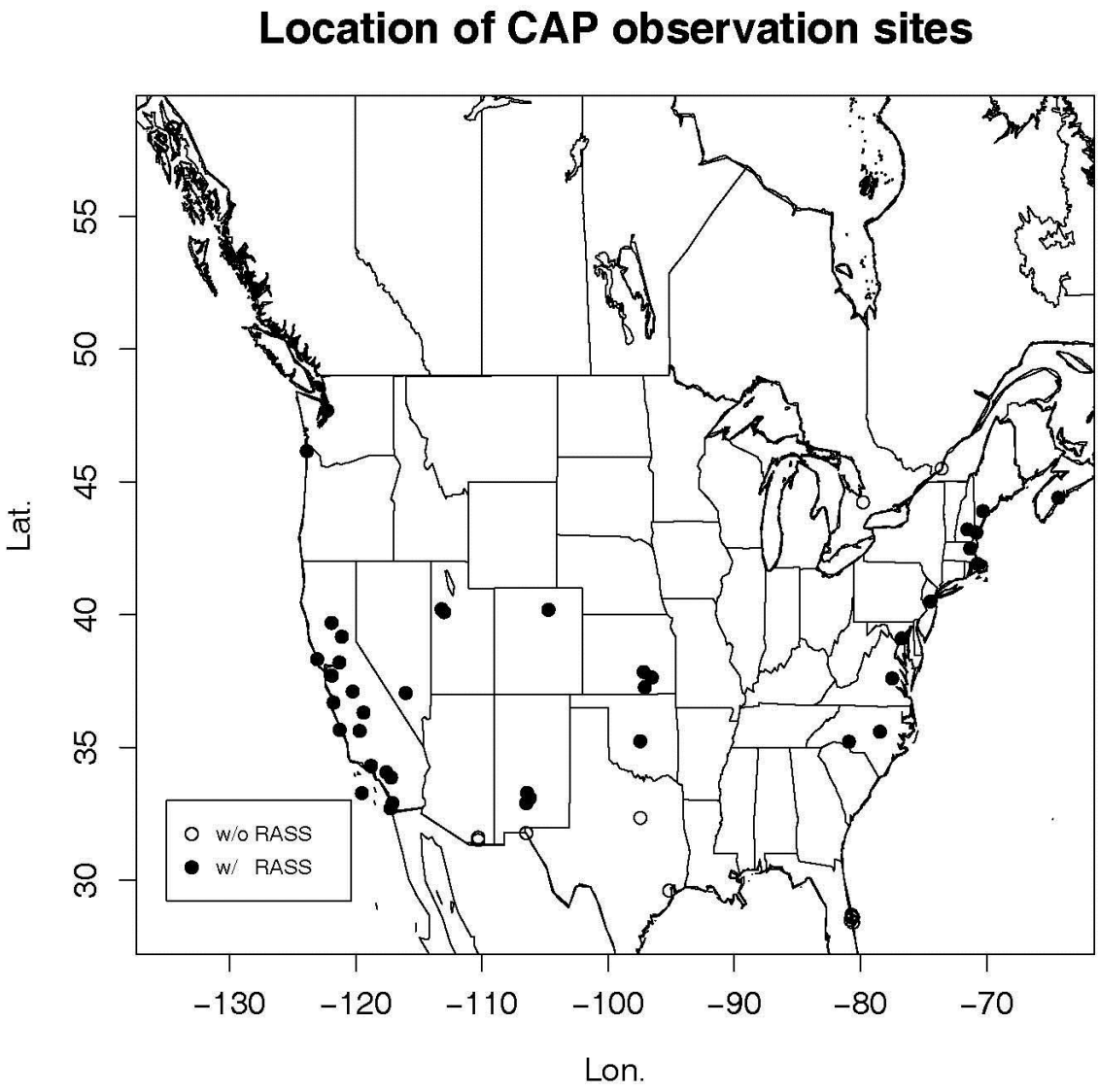

Figure 2 Locations of the Cooperative Agency Profilers used in this study for comparisons. Solid dots indicate wind profiler locations with an addition of temperature data available whereas circles indicate wind profilers only locations.

2.3 Analysis of mixed layer uncertainties

2.3.1 Determining mixed layer heights (z_i)

The Richardson number can be used to predict occurrence of turbulence in a fluid. The non-dimensional number is used as a dynamic stability measure to determine whether the turbulence exists and defined as the ratio between parameters related to buoyancy and shear. Thus it can be used both in convective and mechanical turbulence conditions [*Wang*, 1981; *Wallace and Hobbs*, 2006]. Ideally, to determine whether the flow is dynamically stable using the Richardson numbers, an infinitely shallow layer of air is required. Since RAOB provides profiles of meteorological variables at a series of discrete intervals, the bulk Richardson number (Ri_b) method was adopted to determine z_i.

$$Ri_b = \frac{g*(z-z_0)*[\theta(z)-\theta(z_0)]}{\theta(z)*[u(z)^2+v(z)^2]} \quad (1)$$

where g is gravitational acceleration; z is the measuring height; z_0 is the reference height; $\theta(z)$ is the potential temperature at the heights; and $u(z)$ and $v(z)$ are the zonal and meridional wind components at the heights, respectively. Certain variables in the RAOB data are measured only at mandatory and/or significant levels. When RAOB missed any of the necessary variables (e.g., temperature) at a measuring height, the value is filled in using a log-pressure linear interpolation such that RAOB measurements at all heights are utilized for computing Ri_b to determine z_i. By the interpolation, the layer of the atmosphere at which Ri_b is derived becomes sufficiently shallow so that the ideal condition of using the Richardson number is practically met. The interpolation also justifies the chosen critical number (Ri_c) at 0.25 [*Stull*, 1988] which serves as the chosen stability criterion of the Richardson number. When Ri_b exceeds Ri_c, the atmosphere is considered fully decoupled from the turbulence below.

The bulk Richardson number method is a popular approach among other objective methods available to determine z_i, in spite of its limitations [*Volgelezang and Holtslag*, 1996; *Siebert et al*, 2000]. Since z_i was determined in both the models and the observations (RAOB) using the same method, Ri_b at Ri_c = 0.25, potential methodological issues can be minimized. This study excludes investigation of effects from interpolation of the RAOB measurements between observational levels. However, a previous study in Europe demonstrated that interpolation in altitude is not the dominant cause of errors in ECMWF-to-radiosonde comparisons [*Gerbig et al*, 2008].

To reduce spatial discrepancies between RAOB and the models, variables are extracted at the same model locations where the RAOB profiles exist. This was accomplished by spatial interpolation that is carried out when the model fields are run through an off-line transport model, the Stochastic Time-Inverted Lagrangian Transport model (STILT, [*Lin et al*, 2003]). In other words, STILT is used to extract the temperature and wind vectors necessary to calculate Ri_b (Eq. 1) interpolated down to the RAOB locations. In this way the model errors are expected to be smaller than those from comparisons to gridded values in Eulerian models, since observations are compared to model values that are spatially interpolated to a measurement location instead of an area-averaged value representative for an Eulerian gridcell.

2.3.2 Error statistics

The normalized gross error (NGE) is a common diagnostic to quantify uncertainties in models [*Pielke*, 2002; *Jacobson*, 2005]. The NGE measures biases and is defined here as

$$NGE = \frac{1}{N}\sum_{i=0}^{N}\left(\frac{Model_i - RAOB_i}{RAOB_i}\right) \quad (2)$$

where $Model_i$ and $RAOB_i$ are individual modeled and RAOB-observed z_i out of N values. Relative random errors (RRE) are also examined in an approach adopted by *Gerbig et al* [2008] to quantify random errors:

$$RRE = \frac{stdev(Model_i - RAOB_i)}{mean(RAOB_i)} \quad (3)$$

RRE defined here measures the random errors as the relative spread in z_i discrepancies, as normalized to the "observed" z_i from RAOB. RRE can be directly compared against NGE, since both of them are normalized by observed z_i. Observations and predictions are paired in space and time. As a result, comparisons are made at the same location and time, and the model errors can translate to percentage of discrepancies relative to observation (when multiplied by 100).

The errors in model-derived z_i can be attributed to components of the Ri_b through an error propagation method [*Taylor*, 1997]:

$$Ri_b = \frac{A}{B}, where\ A \equiv \frac{g(z-z_0)[\theta(z)-\theta(z_o)]}{\theta(z)}\ and\ B \equiv u(z)^2 + v(z)^2 \qquad (4)$$

Since Ri_b is equal to the ratio between A and B, given the error of A and the errors of B are independent of one other, the relative uncertainty of Ri_b can be approximated by relative uncertainties of A and B:

$$\frac{\delta Ri_b}{Ri_b} = \frac{\delta A}{A} - \frac{\delta B}{B} + higher\ order\ terms\ ... \qquad (5)$$

It should be noted that both model-derived and RAOB-based mixed layer heights are determined using the bulk Richardson number method. Hence, the quantified errors can be directly comparable between different models, and the impacts of varied model physics such as convective schemes should be minimized.

Chapter 3 Results and Discussions

3.1 Observed z_i

Figure 1 shows maps for observed z_i at 00UTC for the four months in 2004 over North America. The visualization of spatial distributions is accomplished through a kriging technique with the inverse distance weighting scheme [*Bivand et al*, 2008]. The maps depict a west-east gradient in mixed layer heights reflecting the different time zones: 00UTC translates to late afternoon in the western part of North America and to early evening in the eastern part. Values of z_i derived from RAOB during July 2004 has average values of approximately 3.65 km above ground level over the western states of the U.S., where high z_i is expected over dry surfaces where sensible heat fluxes dominate over the latent fluxes. Also, the Rocky Mountains induce strong mechanical mixing over the area. z_i in the western U.S., while appearing high, fall within the range estimated by an independent study: the Support Center for Regulatory Atmospheric Modeling (SCRAM) at the U.S. Environmental Protection Agency offers data archives for z_i values across the U.S. for 1984-1991. SCRAM also made use of radiosonde soundings taken at 00 and 12UTC. The July average z_i in this study using RAOB is about 23% higher than the monthly mean mixed layer height in the SCRAM archive. SCRAM utilized the z_i detection method by *Holzworth* [1964], and natural variation also might have come into play to make such differences between RAOB and SCRAM. However, these sources of the differences should matter little because the focus of this study is to examine uncertainties in z_i from different atmospheric models by directly comparing the values to observation-based ones, not to determine which mixed layer detection scheme produces a more accurate result.

3.2 Biases and random errors in z_i

Monthly averages of NGE in model-derived z_i at 00 and 12UTC are displayed in the spatial maps in Figures 3 ~ 6 and 7 ~ 10, respectively. The spatial and temporal variations of NGE shown in the figures are complex. As expected, regional scale meteorological models with finer grid spacing (EDAS, WRF, and NARR) represent z_i better than the global product (GDAS). In particular, GDAS predominantly underestimates z_i while regional models reveal more varied spatial variability. For all models, however, large biases are found at certain sites.

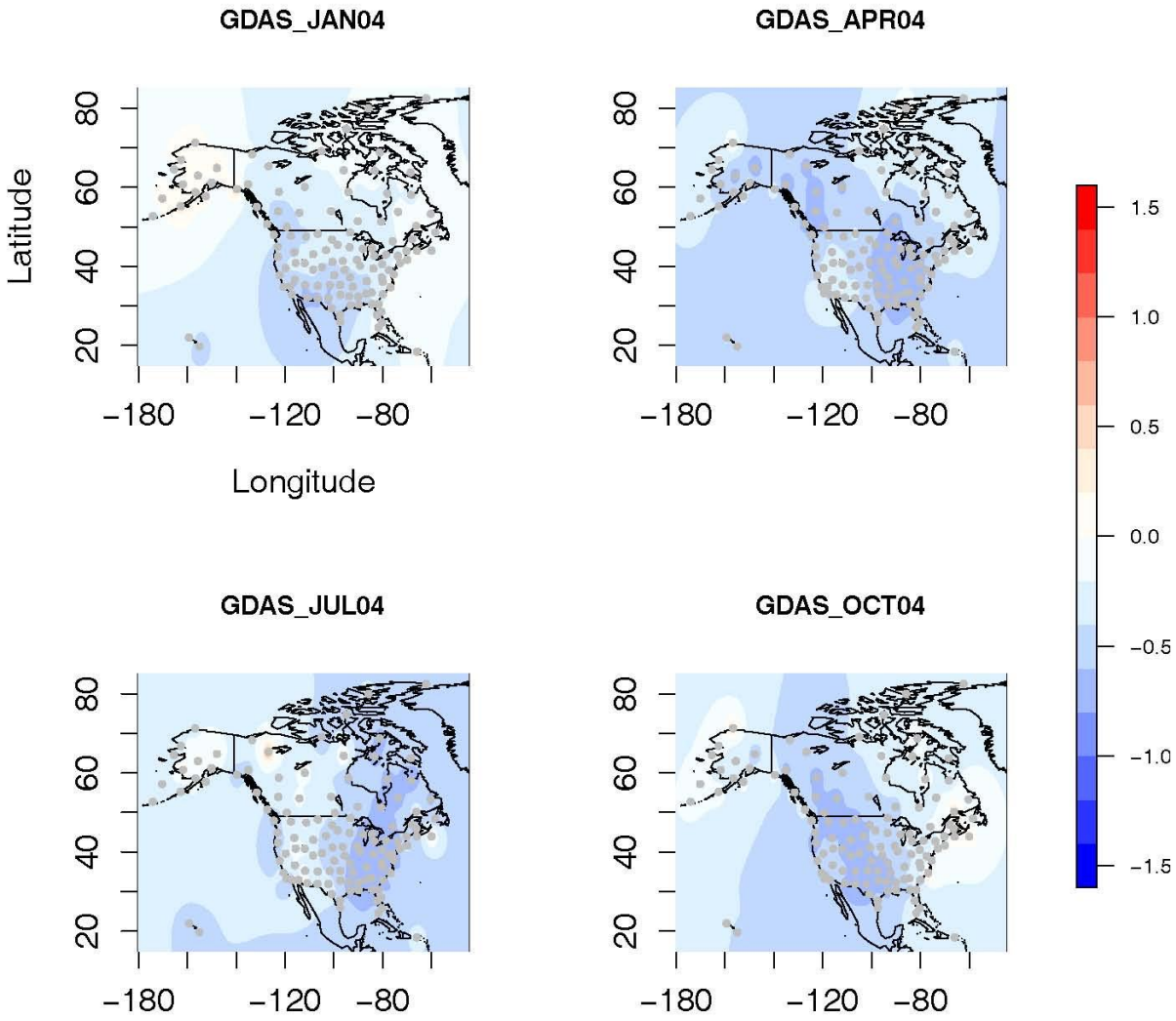

Figure 3 Monthly averaged biases (Normalized Gross Error, NGE) in GDAS-derived mixing height at 00UTC in January, April, July, and October of 2004.

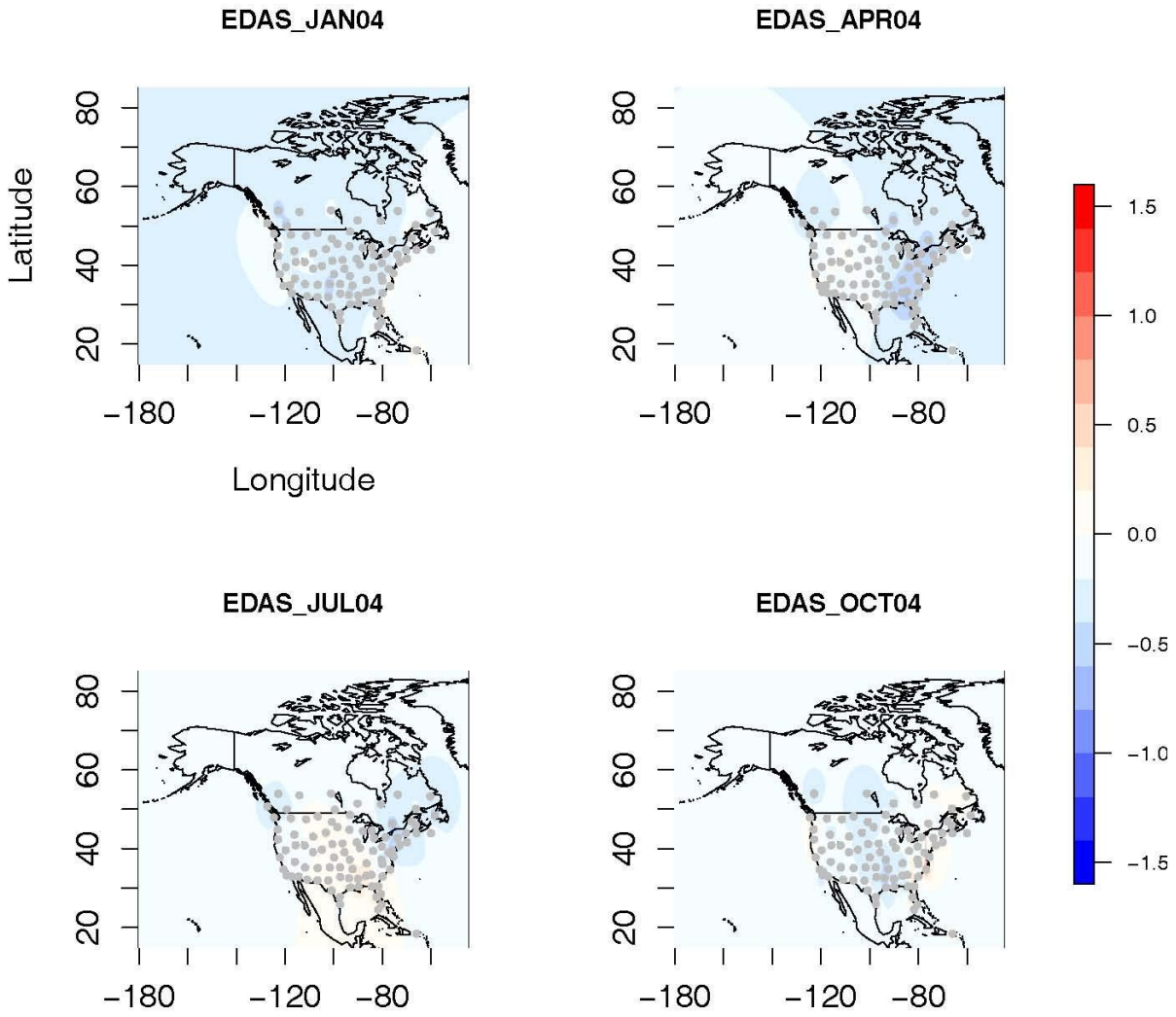

Figure 4 Monthly averaged biases (Normalized Gross Error, NGE) in EDAS-derived mixing height at 00UTC in January, April, July, and October of 2004.

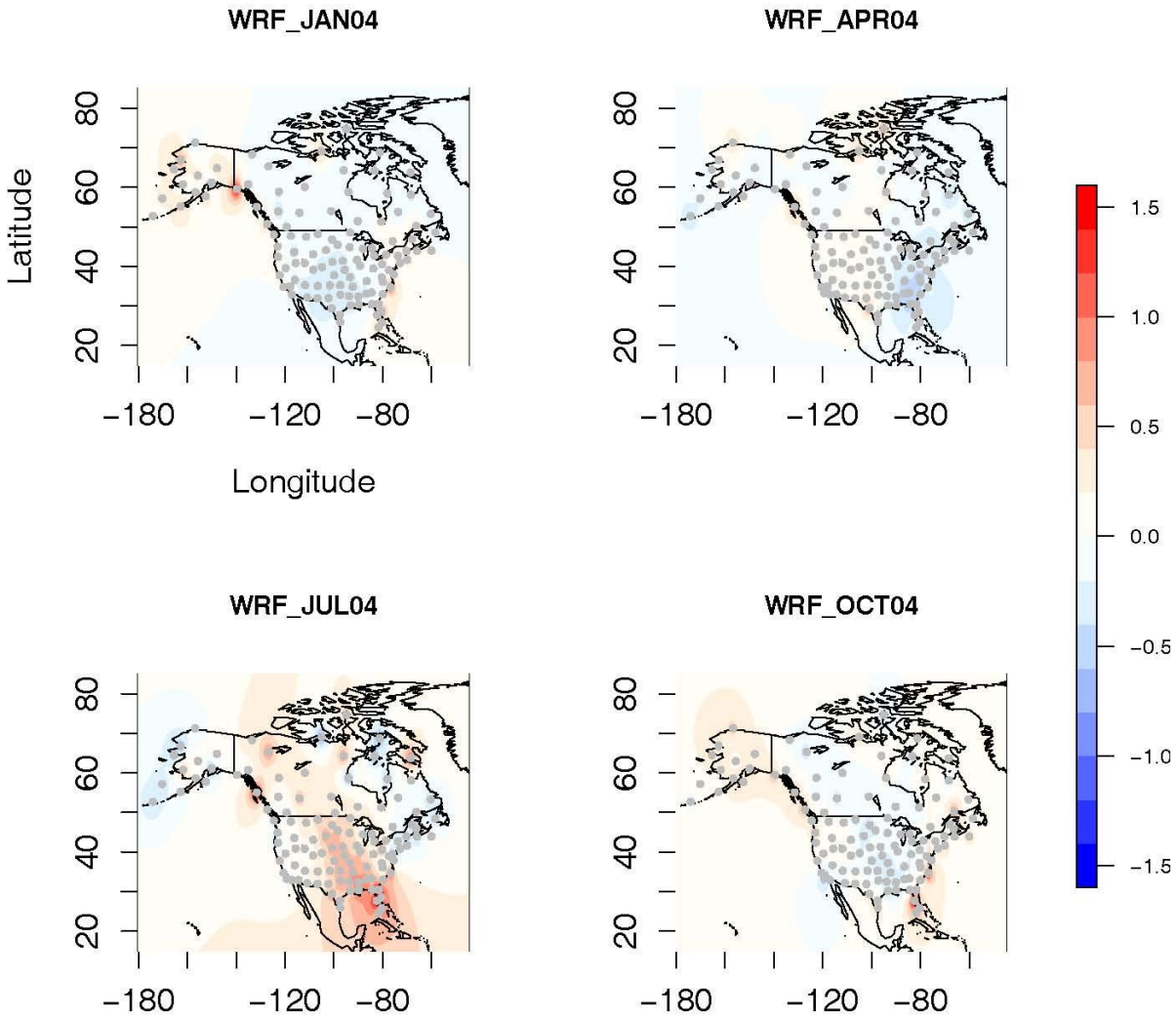

Figure 5 Monthly averaged biases (Normalized Gross Error, NGE) in WRF-derived mixing height at 00UTC in January, April, July, and October of 2004.

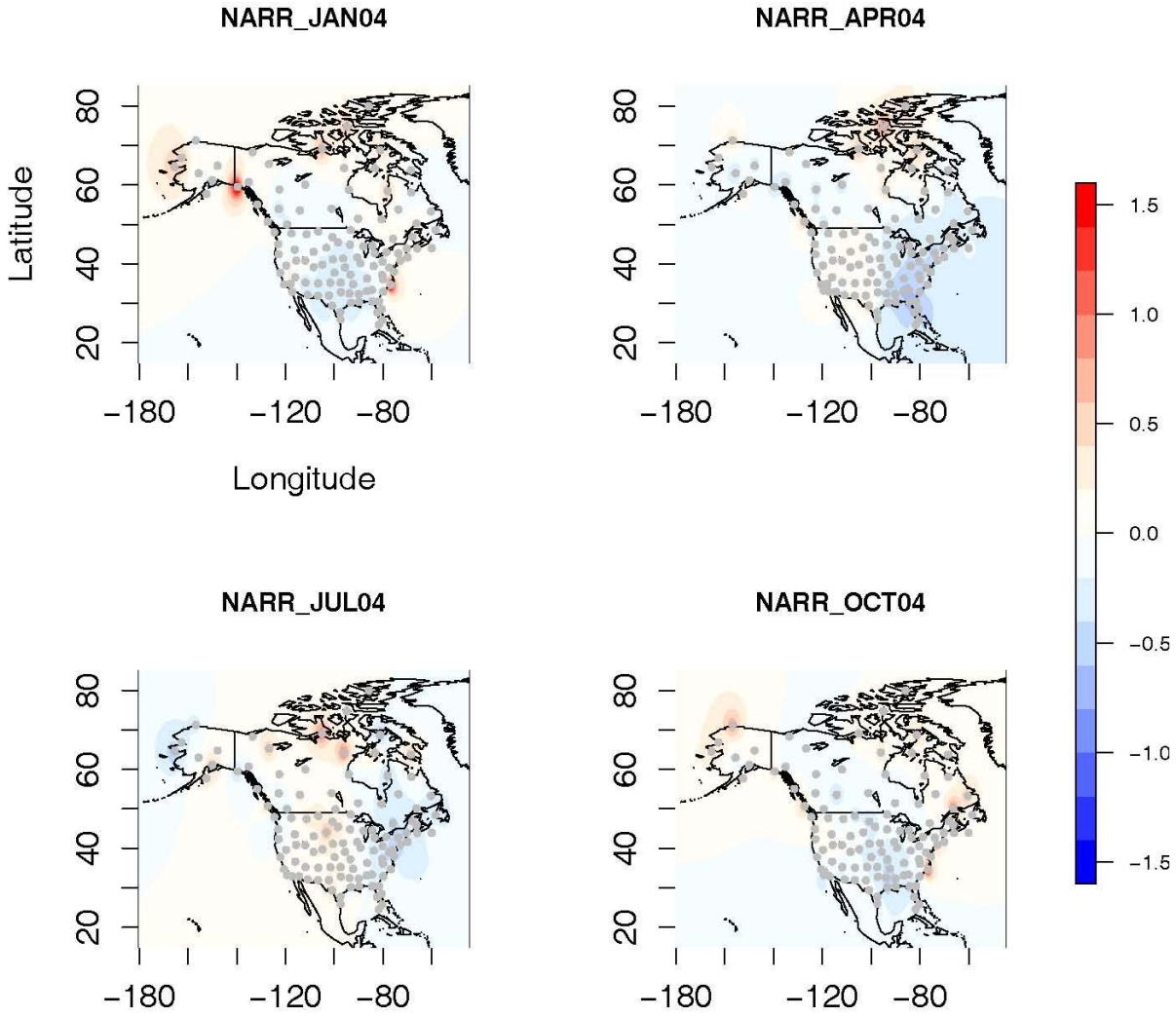

Figure 6 Monthly averaged biases (Normalized Gross Error, NGE) in NARR-derived mixing height at 00UTC in January, April, July, and October of 2004.

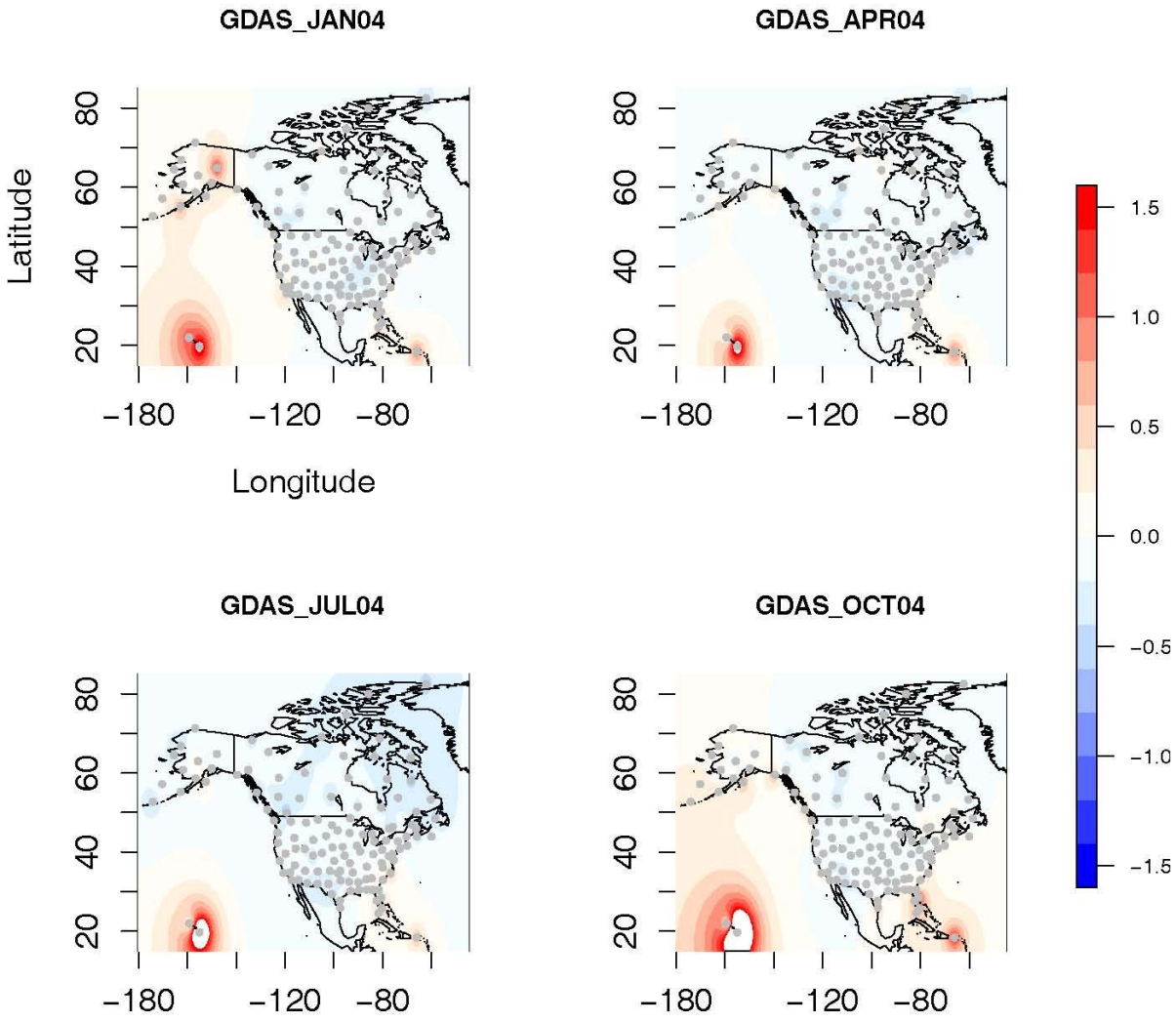

Figure 7 Same as Figure 3, for 12UTC.

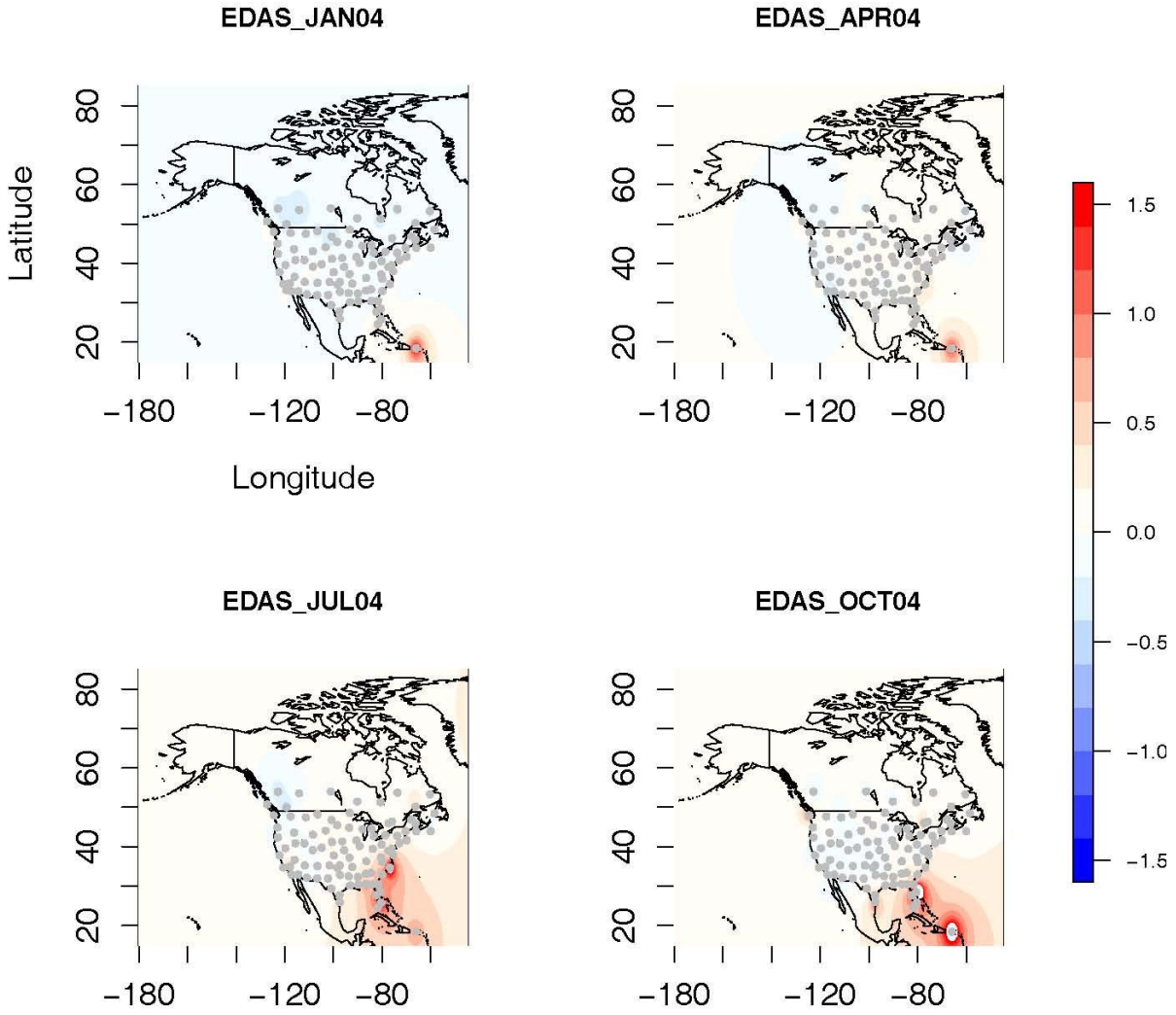

Figure 8 Same as Figure 4, for 12UTC.

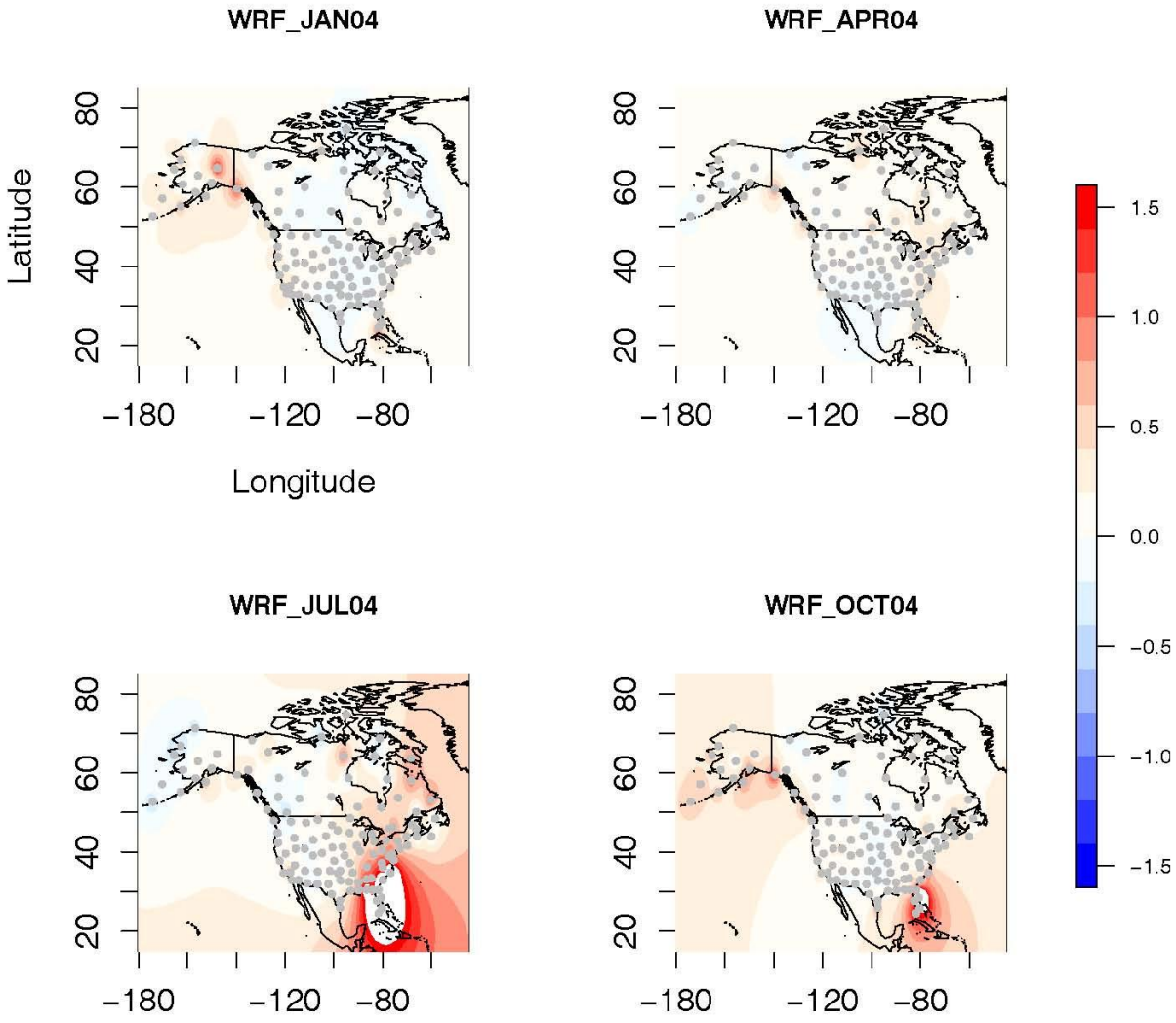

Figure 9 Same as Figure 5, for 12UTC.

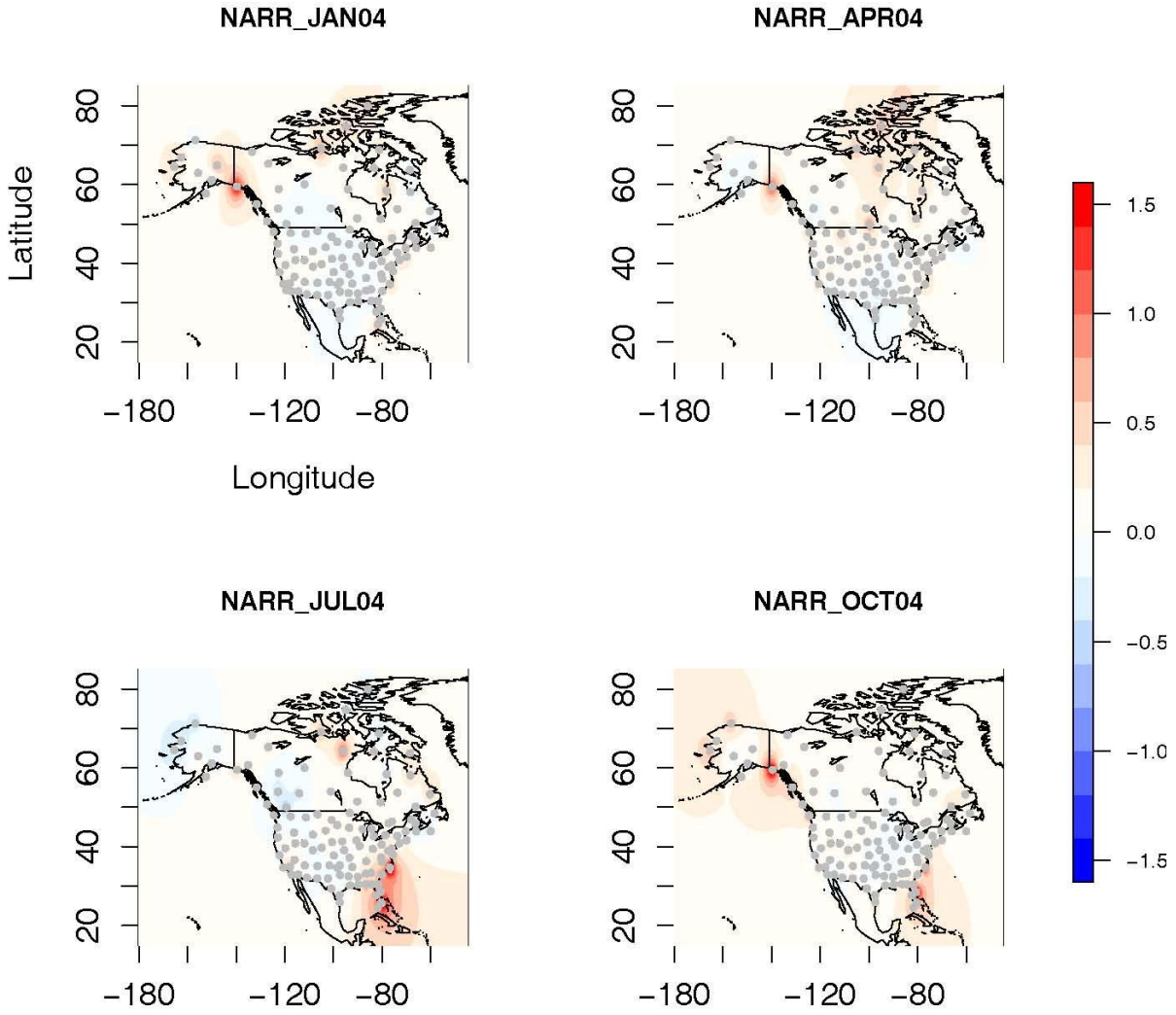

Figure 10 Same as Figure 6, for 12UTC.

Average biases of z_i in different months are shown as colored bars in Figures 11 and 12. The monthly mean biases for each model are averages over all locations, and error bars indicate one standard deviation from the mean. Diurnal variation of biases in model estimated z_i seems to reveal biphasic patterns: negative biases at 00UTC and positive values at 12UTC are noticeable. Also, standard deviation of the biases is larger at 12UTC when the global model (GDAS) looks to perform better than the mesoscale models, in contrast to 00UTC. Considering the spatial maps (Figs. 3 ~ 6 and 7 ~ 10) with Figures 11 ~ 12, we can see that scattered biases of different signs over the regions cancel each other when taking the mean. This cancellation of biases in z_i has implications when dealing with biases in regional scale flux estimations, because averaging biases blurs signs and magnitude of spatial distribution of the error in mixed layer heights. Given that the estimation is averaged, estimated flux cannot assume that biases would be appropriately treated over certain spatial scales.

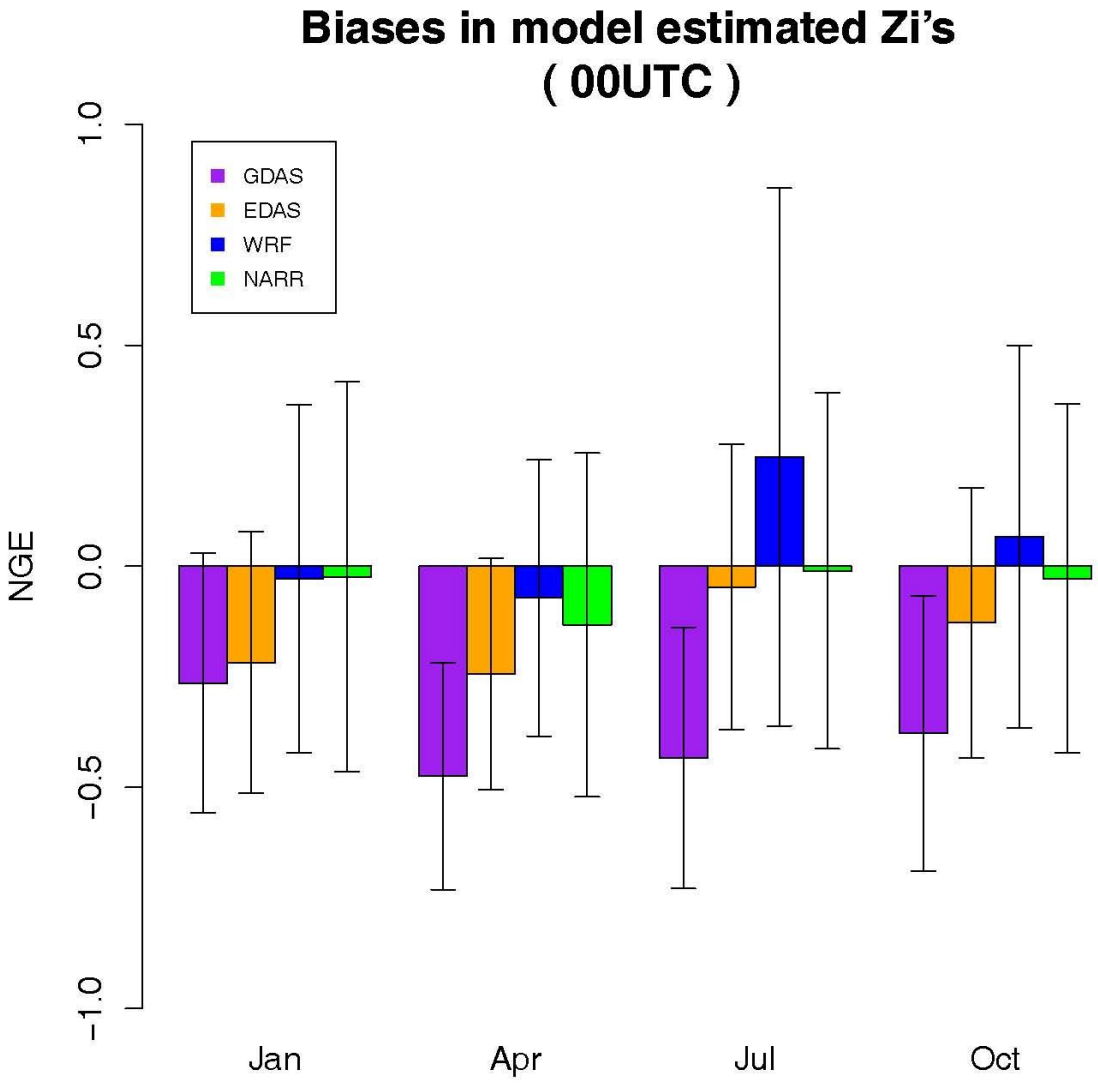

Figure 11 Biases of model estimated mixing heights at 00UT. Error bars indicate one standard deviation from the monthly mean value.

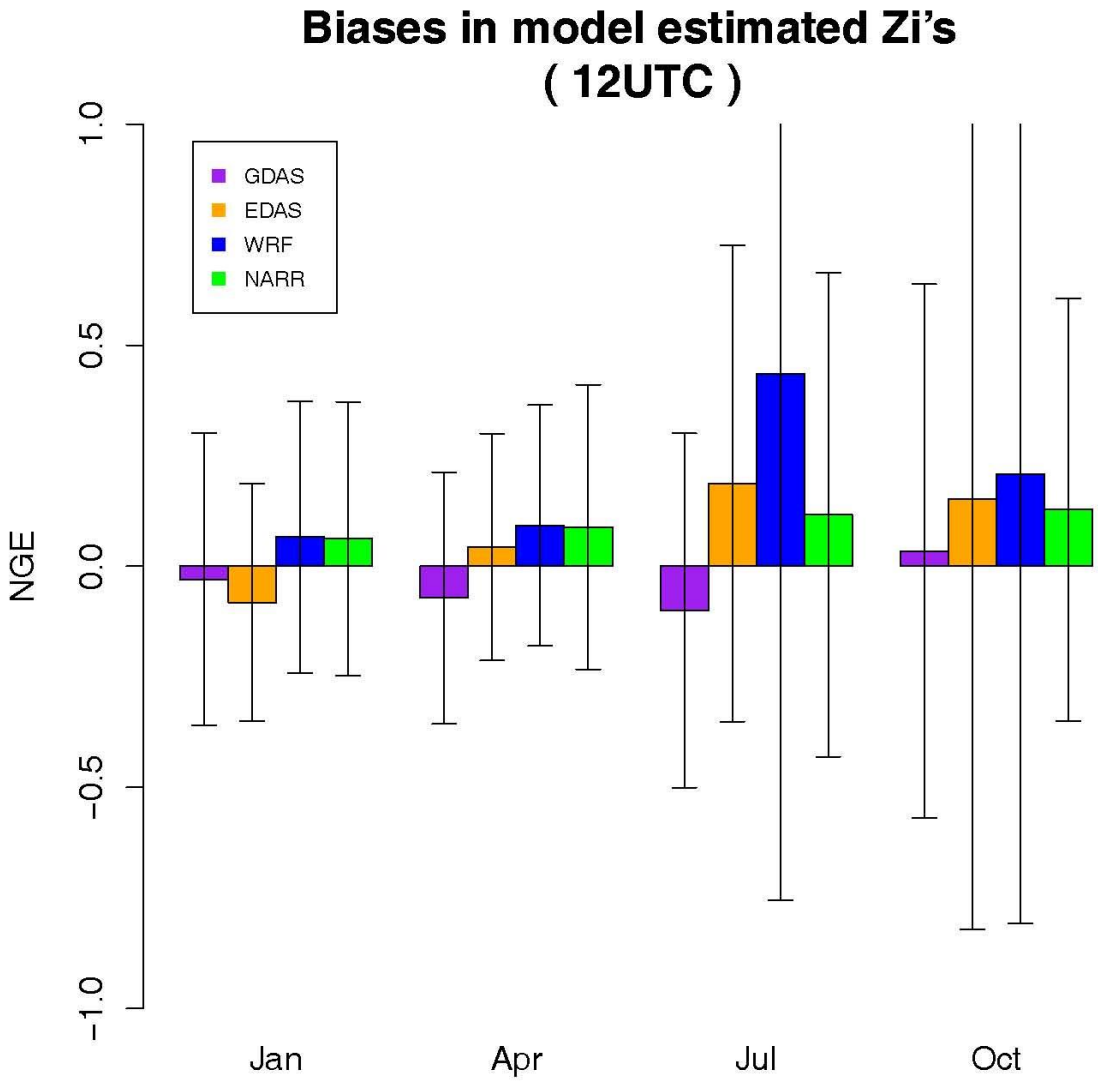

Figure 12 Same as Figure 11, for 12UTC.

Spatial variability of RRE reveals that the random errors are large either at locations where the measurements are sparse or at the edge of the continents where there are no nearby measurements (Figures 13 ~ 20). For example, locations in subtropical Atlantic coast of the U.S. seem prone to larger random errors for all meteorological models. High Arctic sites also seem to suffer from poor model representation of mixed layer heights. The random error increases between 00 to 12UTC, and the largest value appears at 12UTC when it is spatially more variant (Figures 21 and 22).

It is intriguing to note that areas with large random errors such as the Gulf Coast often seem to correspond to locations with significant biases. The spatial correspondence between the random errors and the biases raises the possibility that the two errors over North America might be linked. However, linear correlation between the two errors in fact is negligible, and predictability with a linear regression of the random error using the biases, or the vice versa, cannot be established (Figures 23 and 24). This suggests that simple scaling of biases in a transport model would not suffice as a remedy for deriving more reliable fluxes in the lower atmosphere over North America, because random errors would still persist. Nevertheless, it is clear from the error maps that at several "problem sites" modeled z_i exhibit both the random and systematic errors. As shown in Figures 25 and 26, larger errors appear to be associated with sites that are located at lower elevations, typically within a couple hundred meters above sea level.

The mixed layer height at a site responds in general to solar radiation, large scale topography, surface albedo, soil moisture, and land cover. Hence, it is reasonable that geographic location inherently introduces systematic uncertainties to any measured mixed layer height. Those uncertainties can explain in part the spatio-temporal distribution of the errors. For example, mixing layers near large water bodies will experience different diurnal evolution from a mixing layer over the prairies, for instance due to sea-breeze effect and varying soil moisture. Unfortunately, however, those meso-scale variabilities are challenging to fully capture with models adopting grid spacing of 30 to 40 km, as is the case for EDAS, WRF, and NARR here [*Pielke*, 2002].

When the ratio of NGE to RRE is calculated at each location where models-to-RAOB comparisons are made, it is seen that the random errors surpass the biases in most cases (NGE:RRE <1), except for GDAS in April and July (Figures 27 and 28). In general, magnitude of biases for most cases is about a half of the random error. A previous study carried out in Europe also suggested that biases are about a half or less of random errors [*Gerbig et al*, 2008]. The NGE to RRE ratio for 12UTC in Figure 11b is somewhat smaller than for 00UTC in most cases, which probably stems from the fact that NGE is smaller while RRE is larger at 12UTC (Figures 11 ~ 12 and 21 ~22).

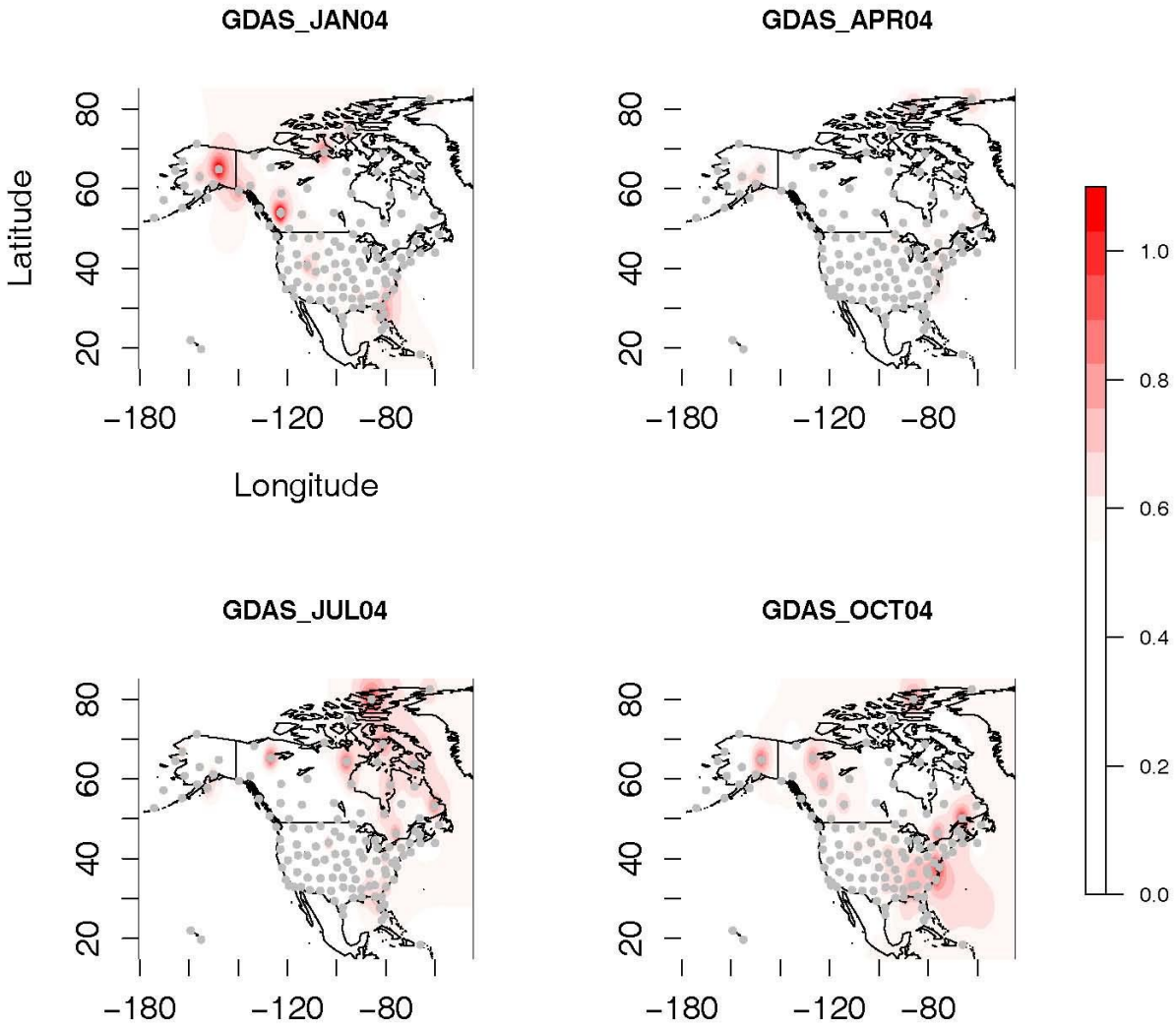

Figure 13 Monthly averaged random error (Relative Random Error, RRE) in GDAS-derived mixing height at 00UTC in January, April, July, and October of 2004.

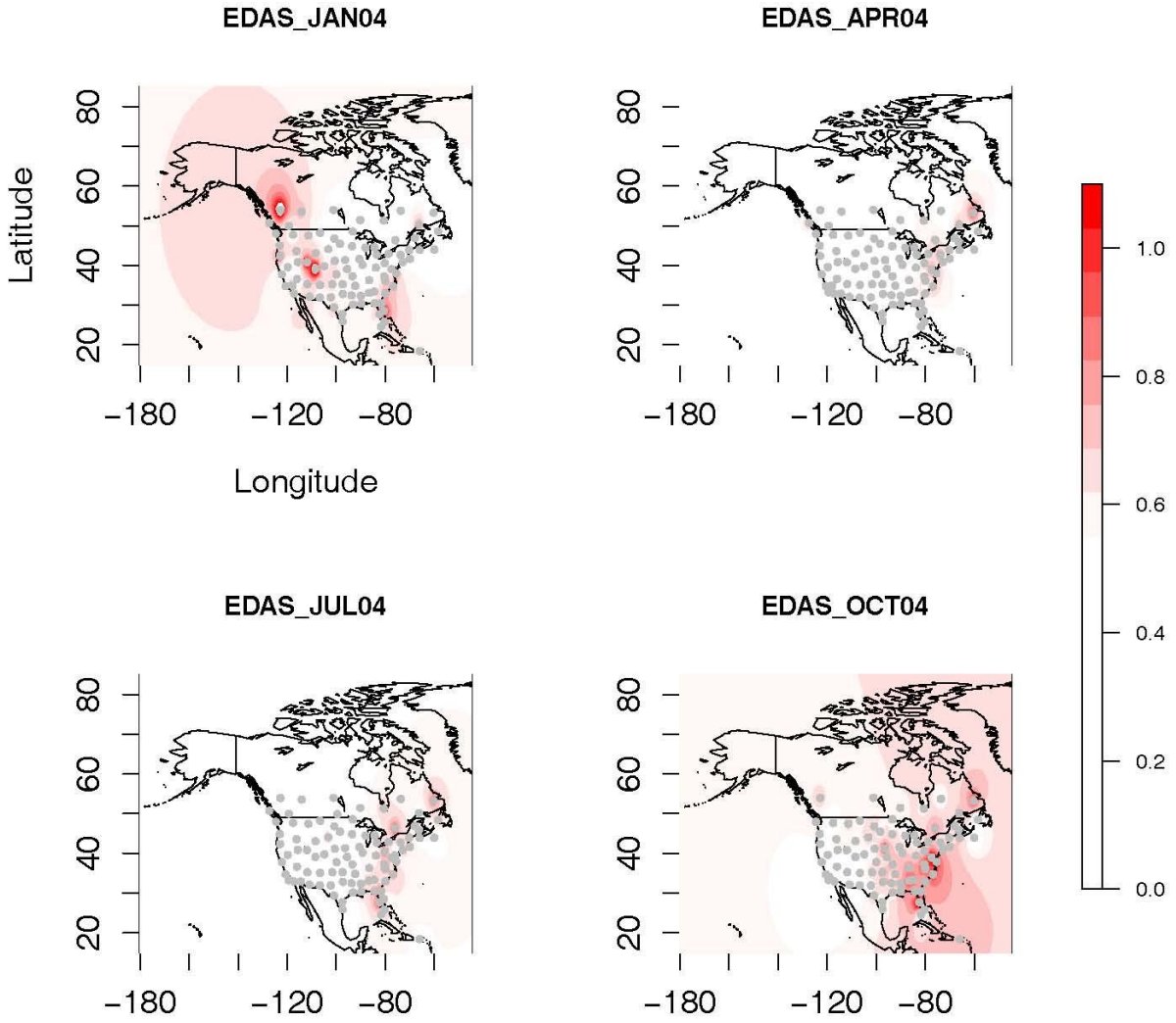

Figure 14 Monthly averaged random error (Relative Random Error, RRE) in EDAS-derived mixing height at 00UTC in January, April, July, and October of 2004.

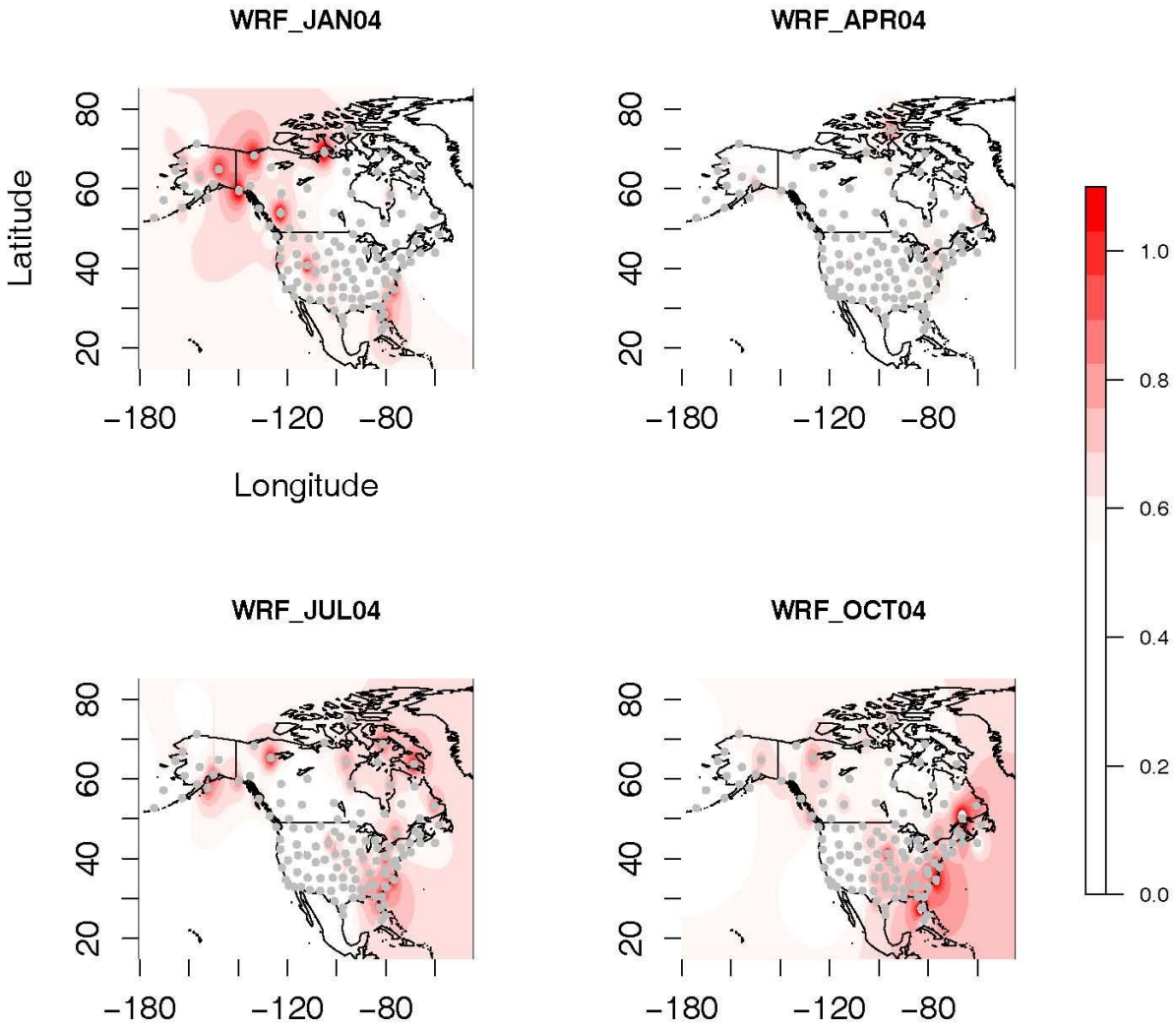

Figure 15 Monthly averaged random error (Relative Random Error, RRE) in WRF-derived mixing height at 00UTC in January, April, July, and October of 2004.

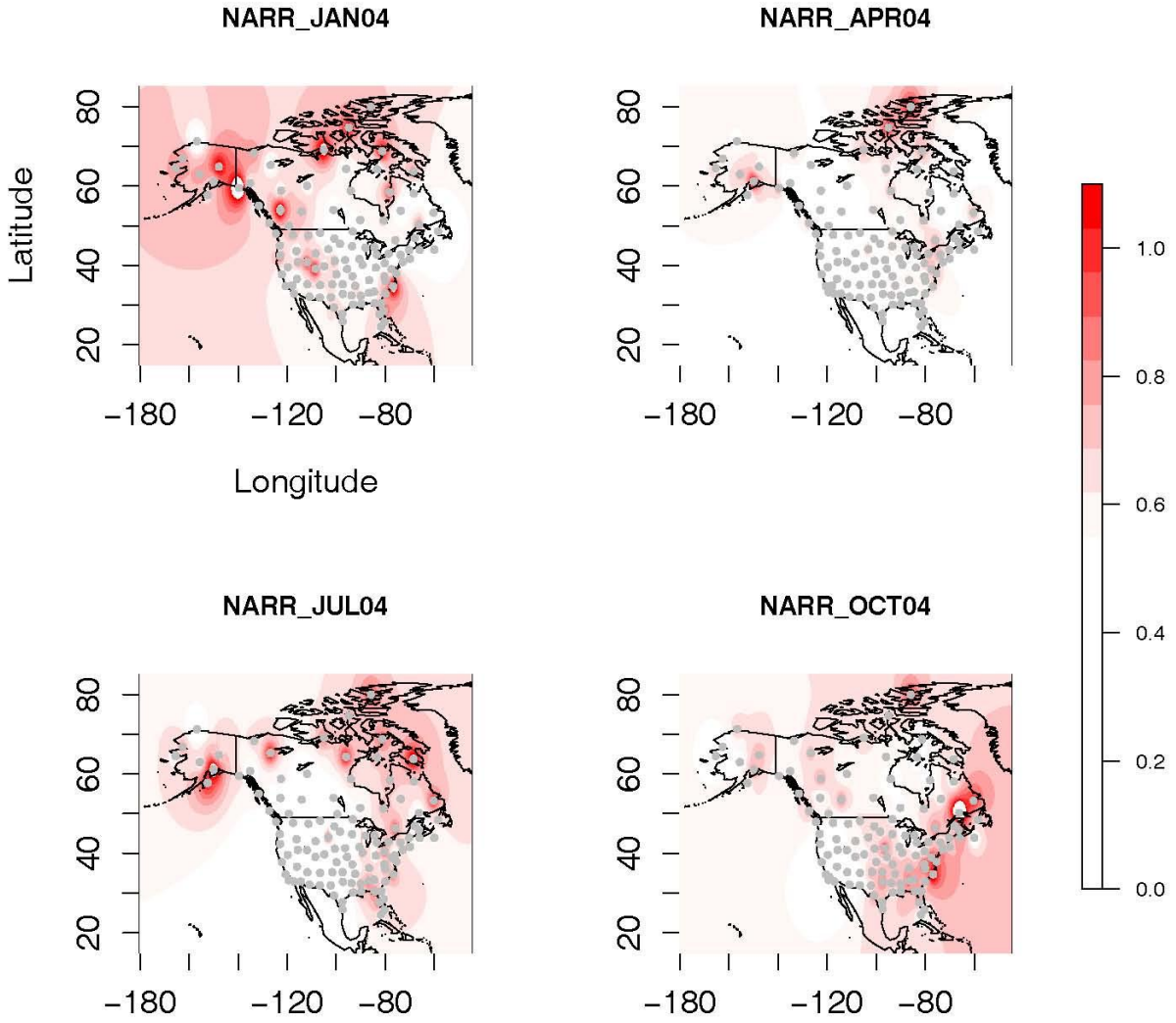

Figure 16 Monthly averaged random error (Relative Random Error, RRE) in NARR-derived mixing height at 00UTC in January, April, July, and October of 2004.

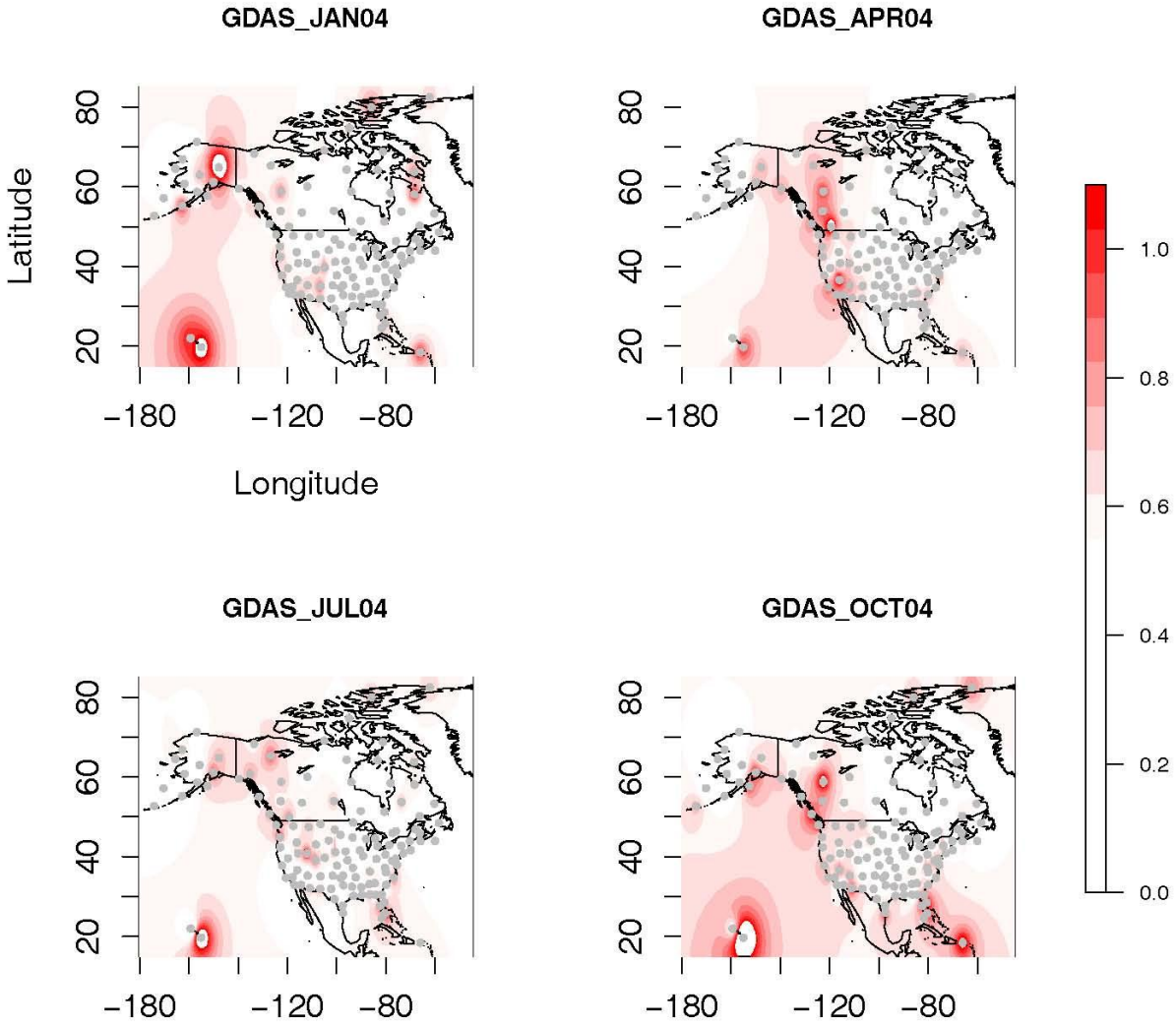

Figure 17 Same as Figure 13, for 12UTC.

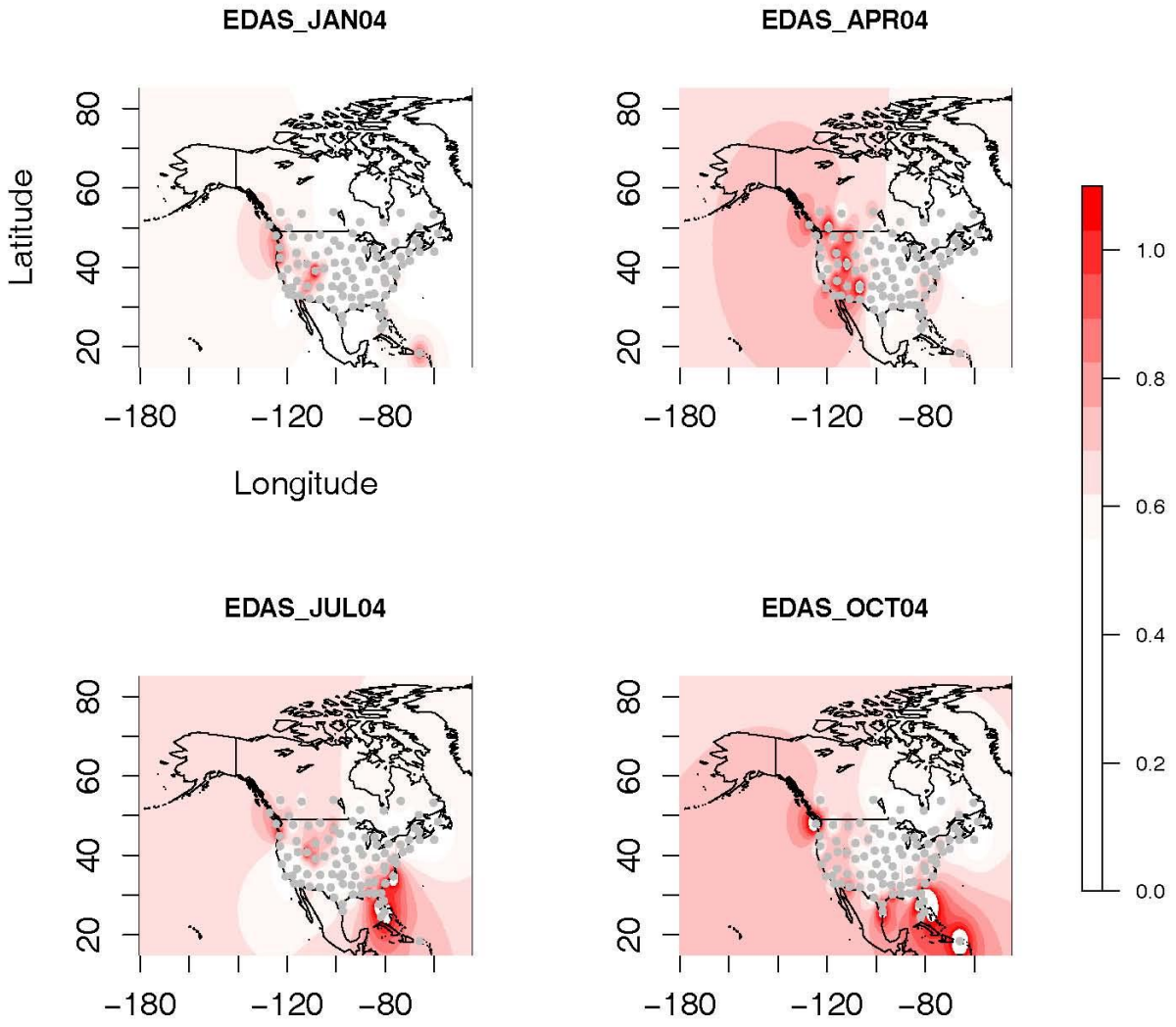

Figure 18 Same as Figure 14, for 12UTC.

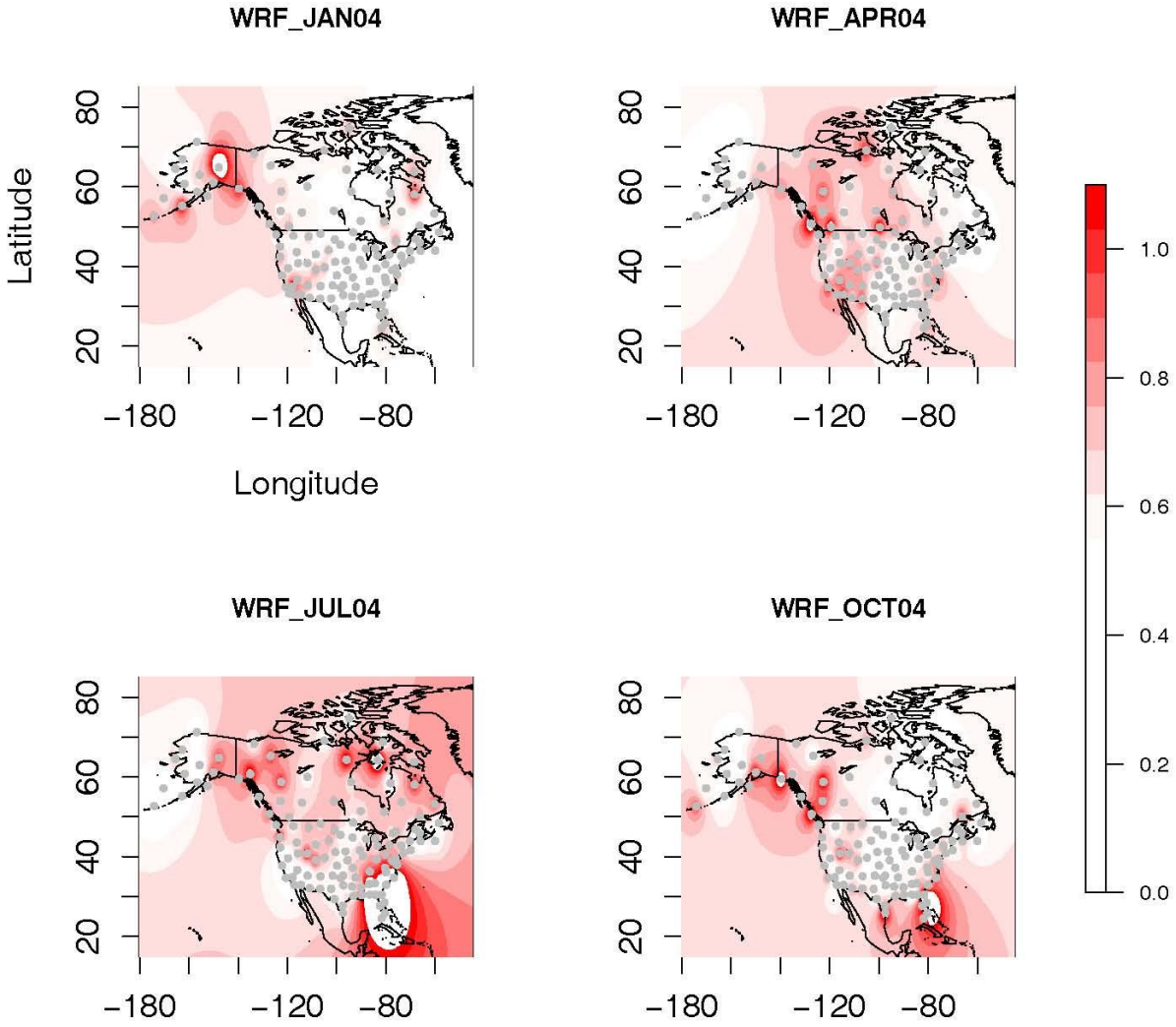

Figure 19 Same as Figure 15, for 12UTC.

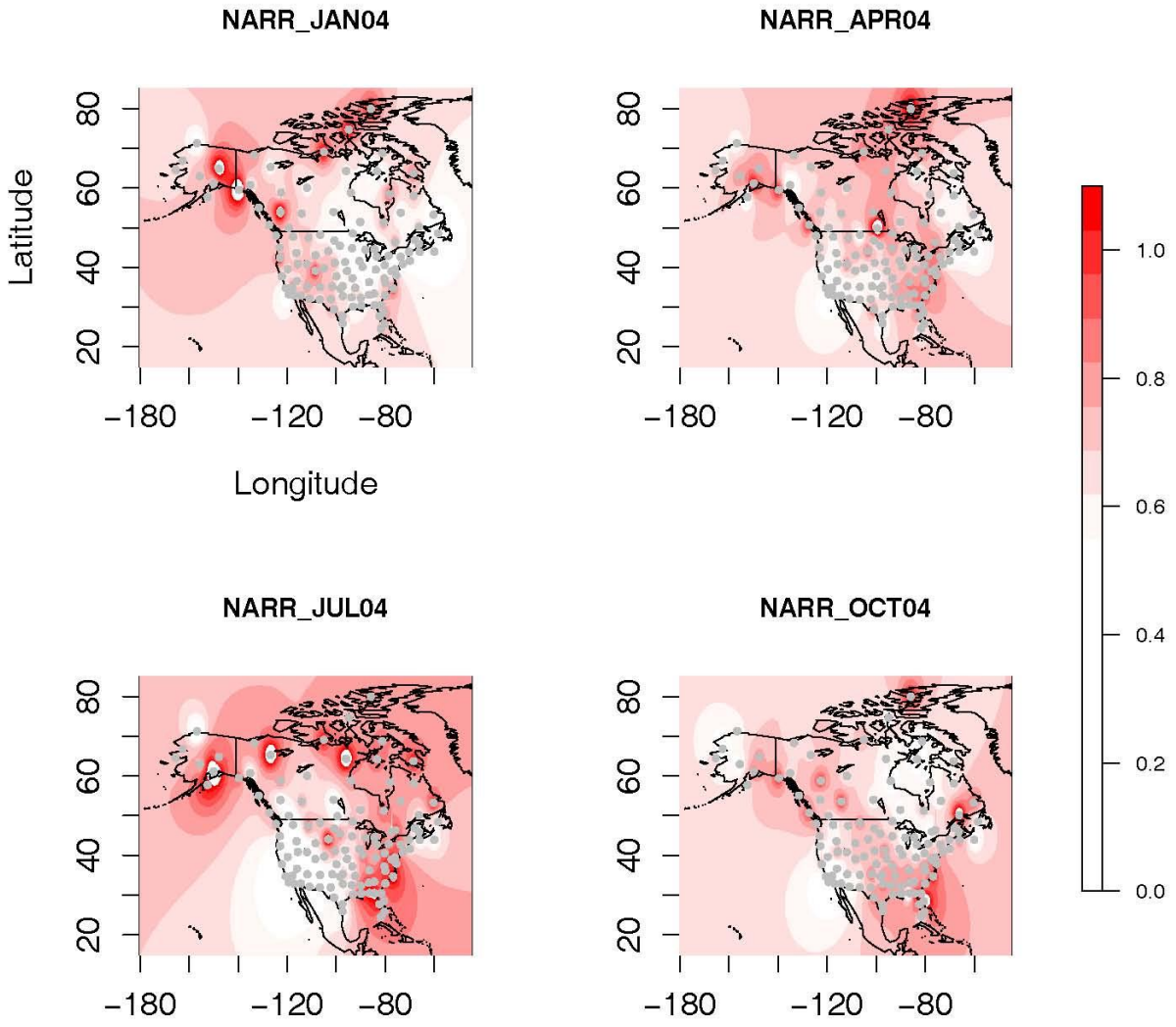

Figure 20 Same as Figure 16, for 12UTC.

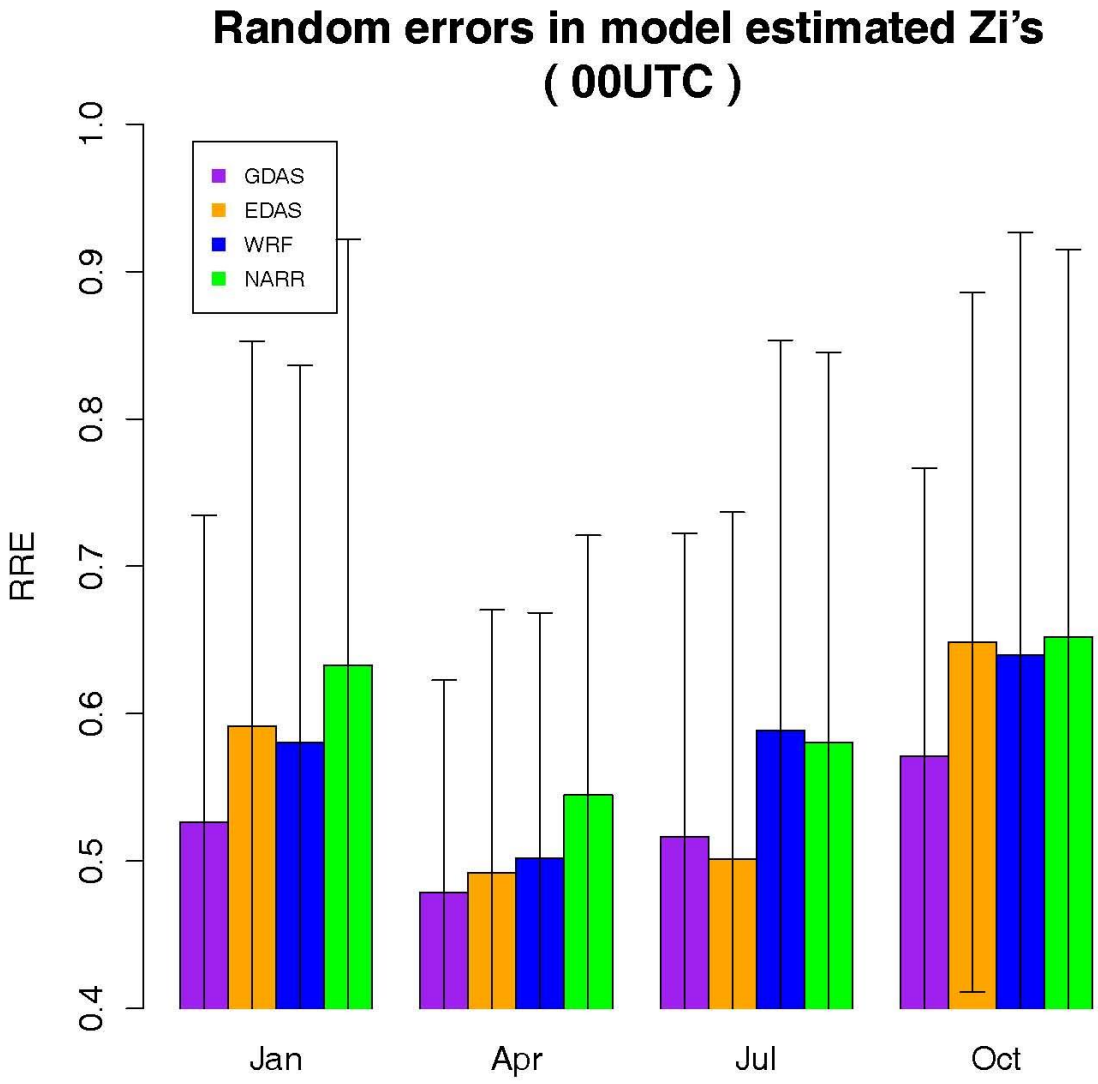

Figure 21 RRE of model estimated mixing heights at 00UTC. Error bars indicate one standard deviation from the monthly mean value.

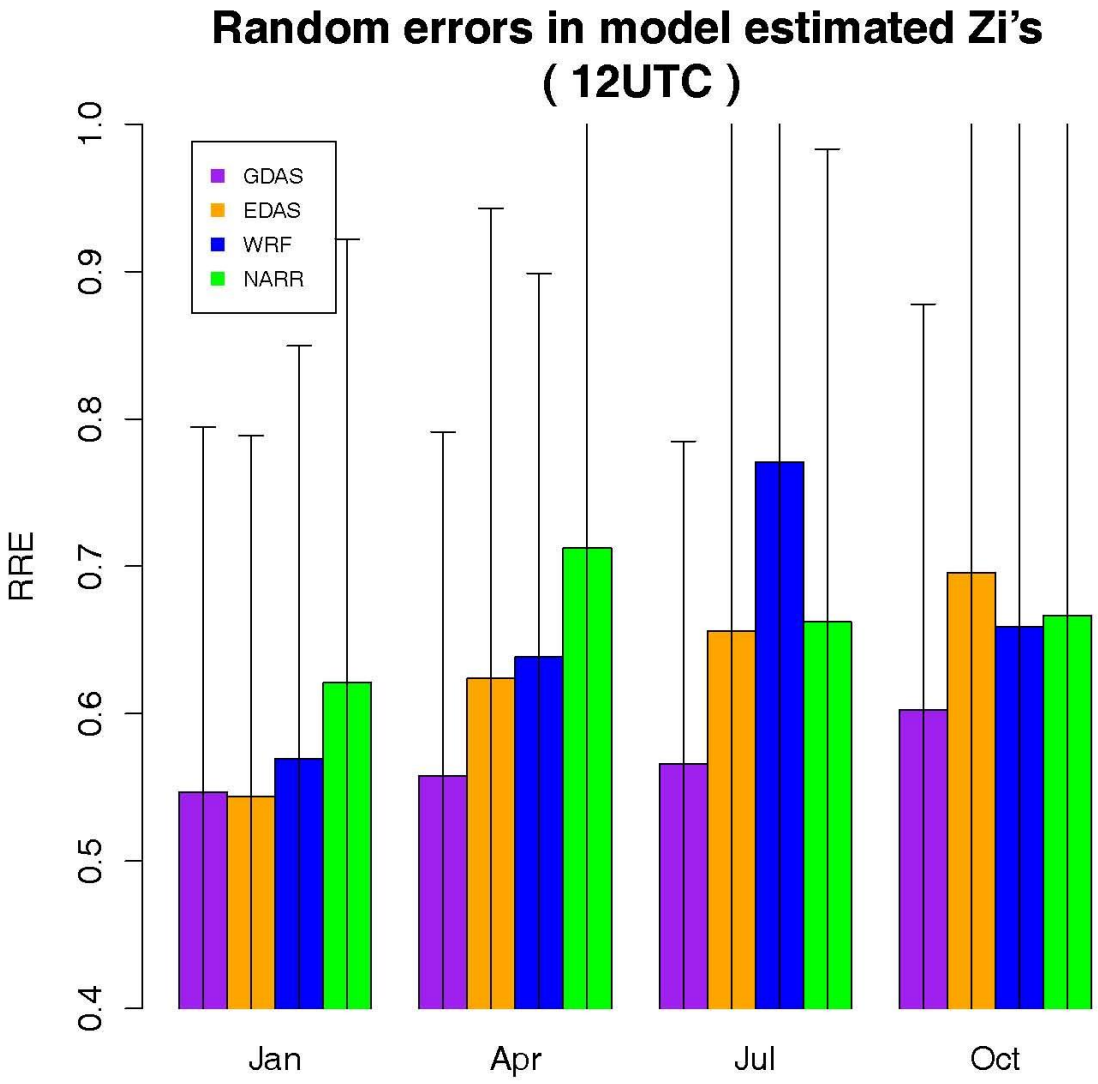

Figure 22 Same as Figure 21, for 12UTC.

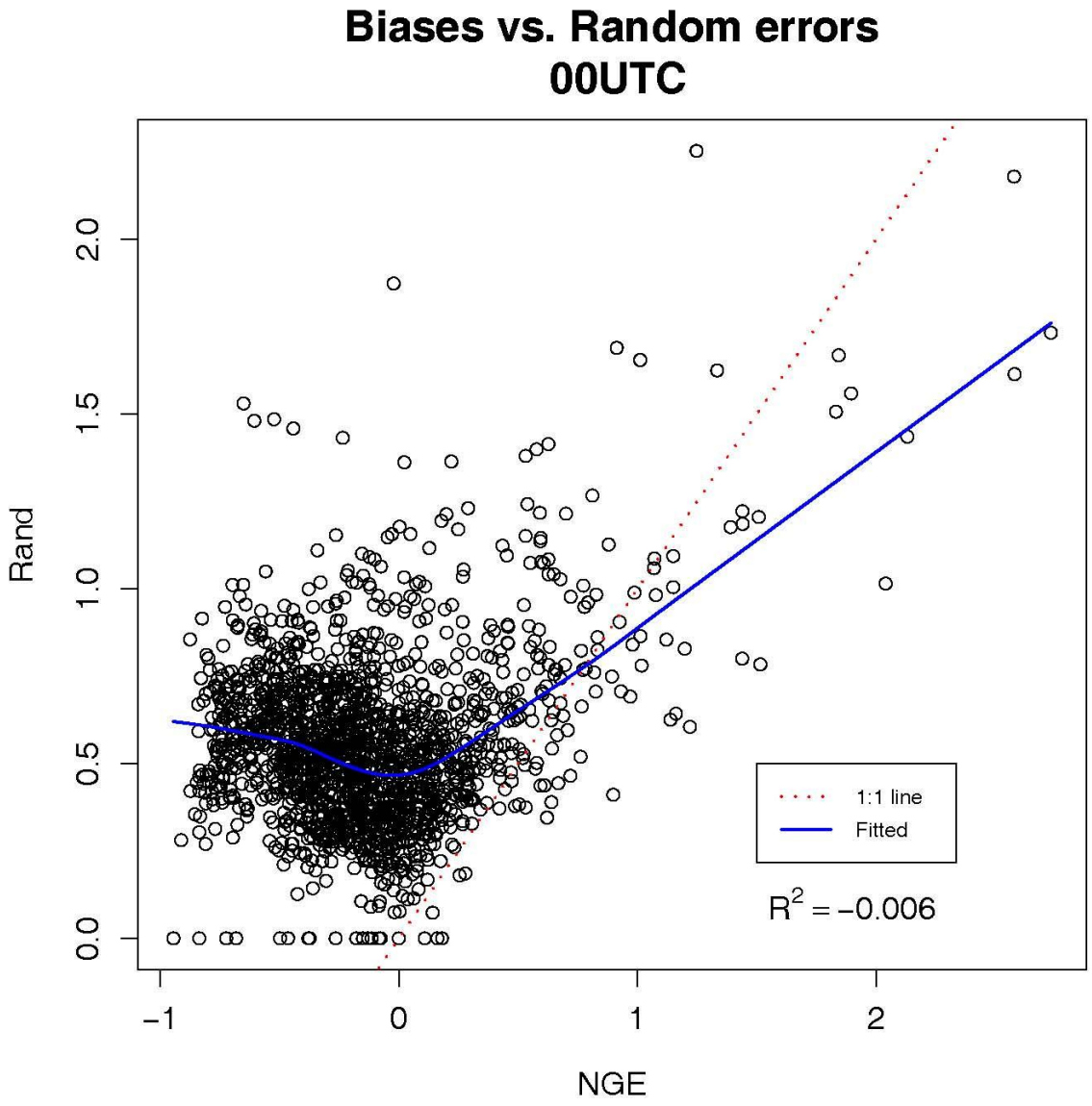

Figure 23 Scatter plots of biases versus Random errors for all models at 00UTC. Solid blue lines indicate statistical fitting with a linear regression, and dotted red lines one to one.

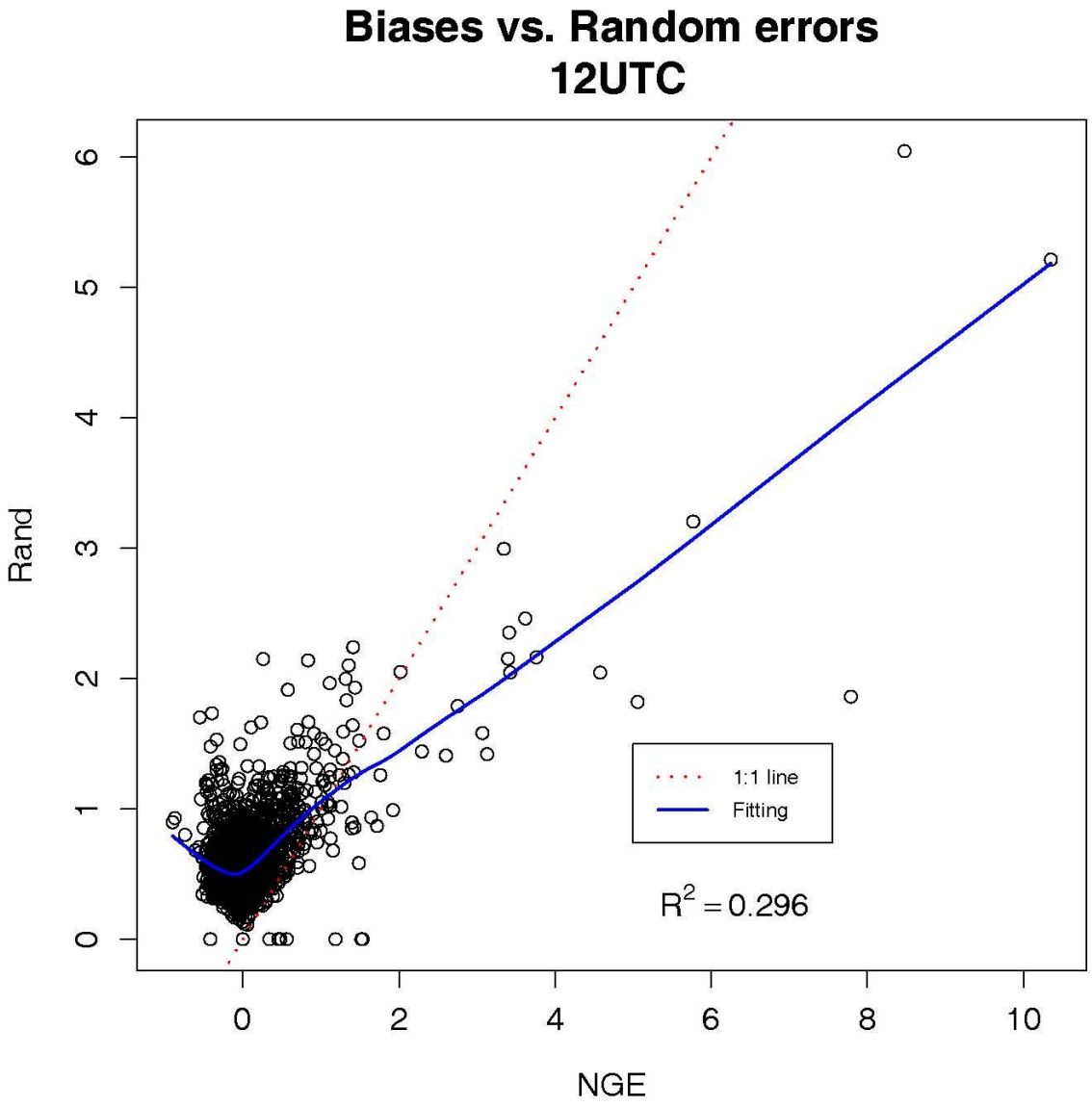

Figure 24 Same as Figure 23, for 12UTC.

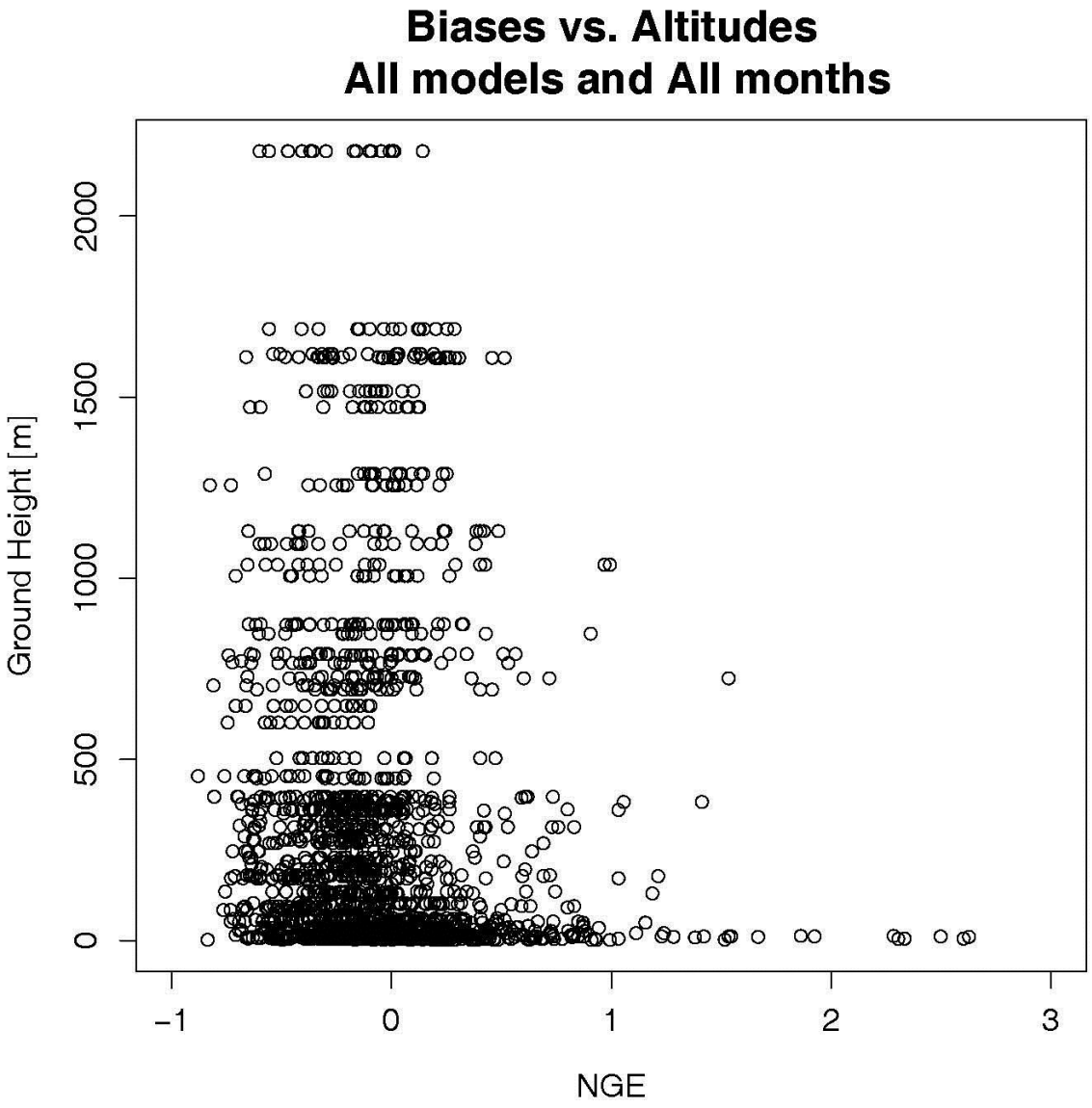

Figure 25 Relationship between biases and ground heights at observation sites.

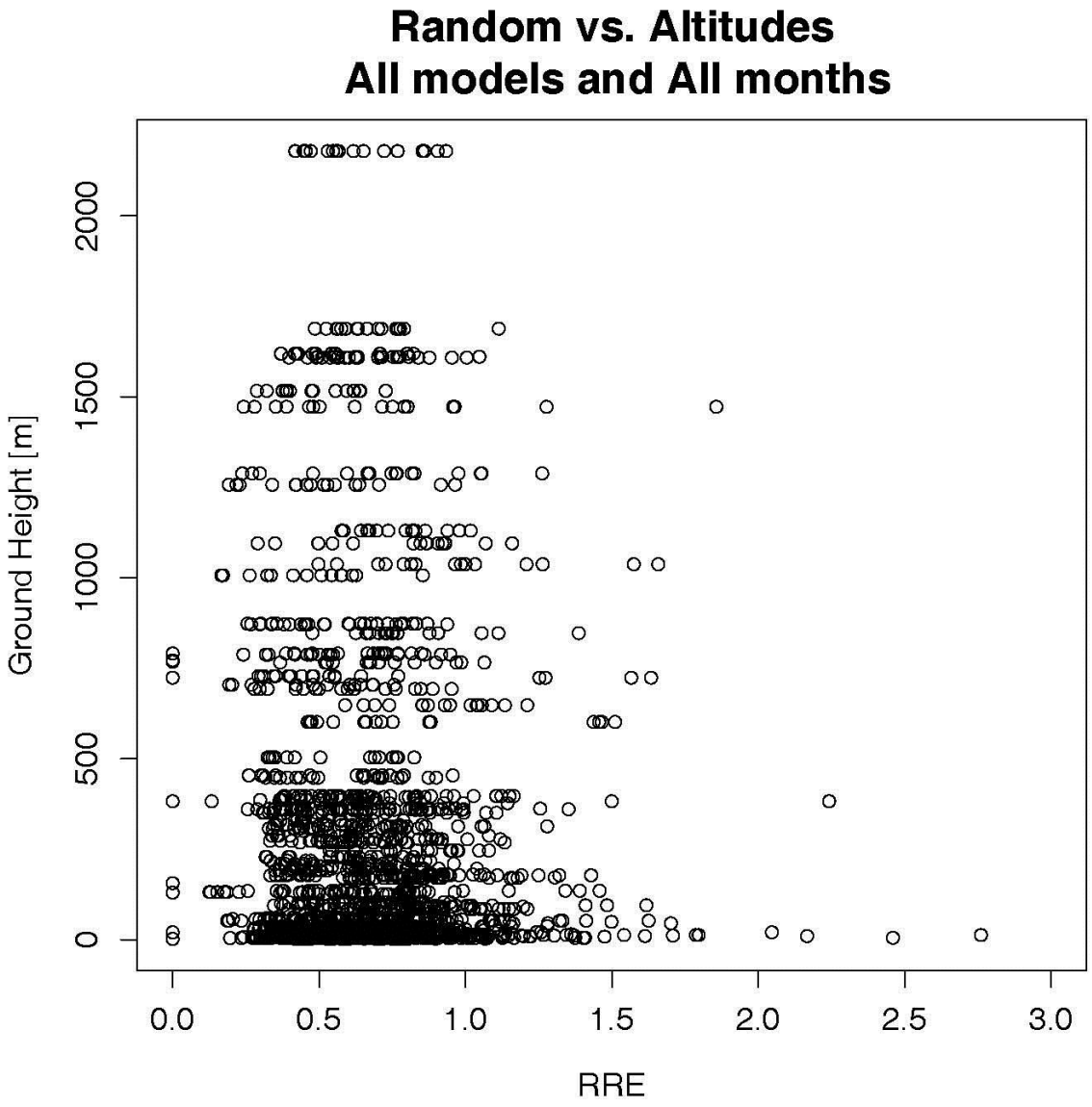

Figure 26 Same as Figure 25, for random errors

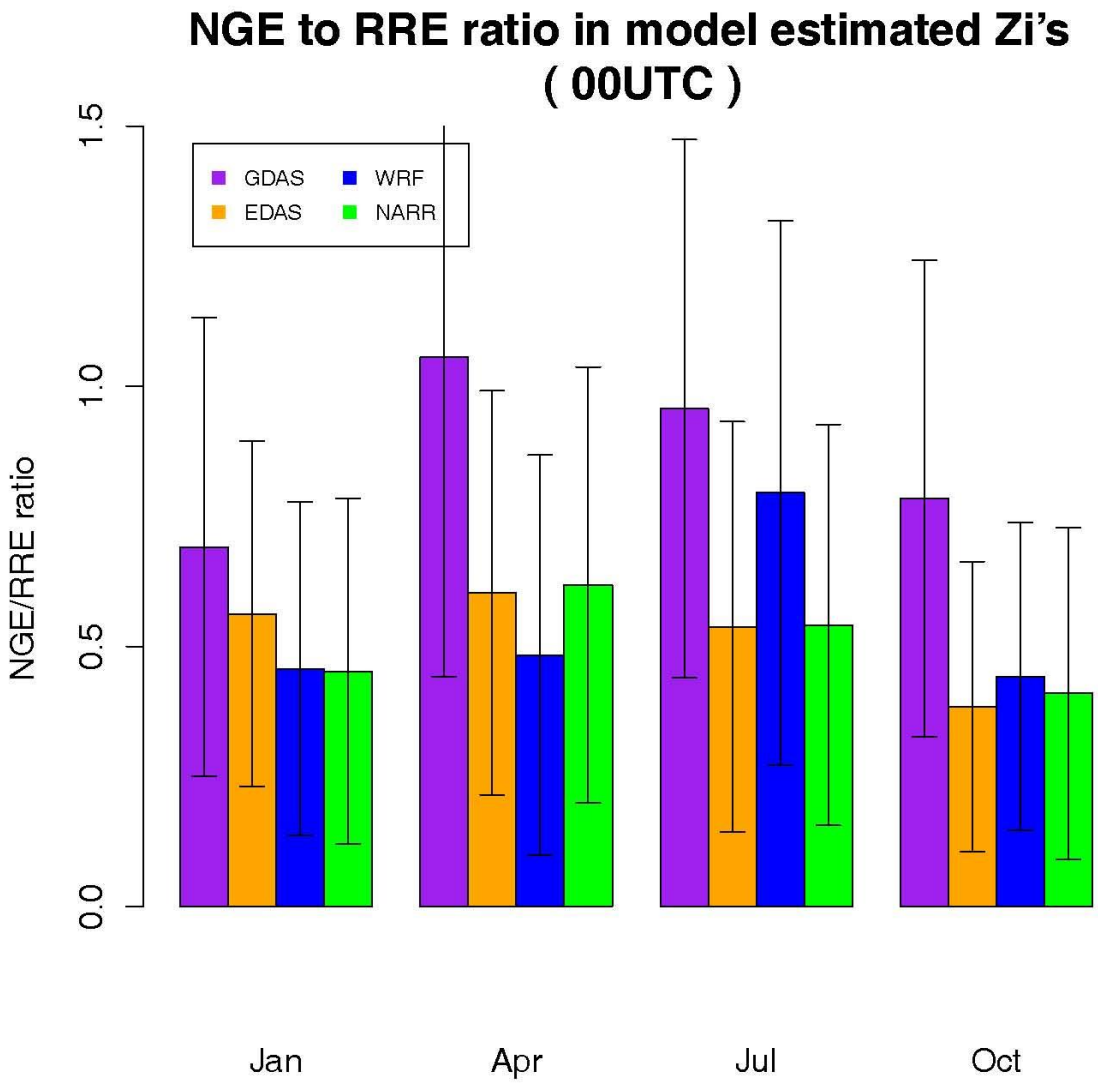

Figure 27 Biases divided by random errors at observation sites for 00UTC. Error bars indicate one standard deviation from the monthly mean value.

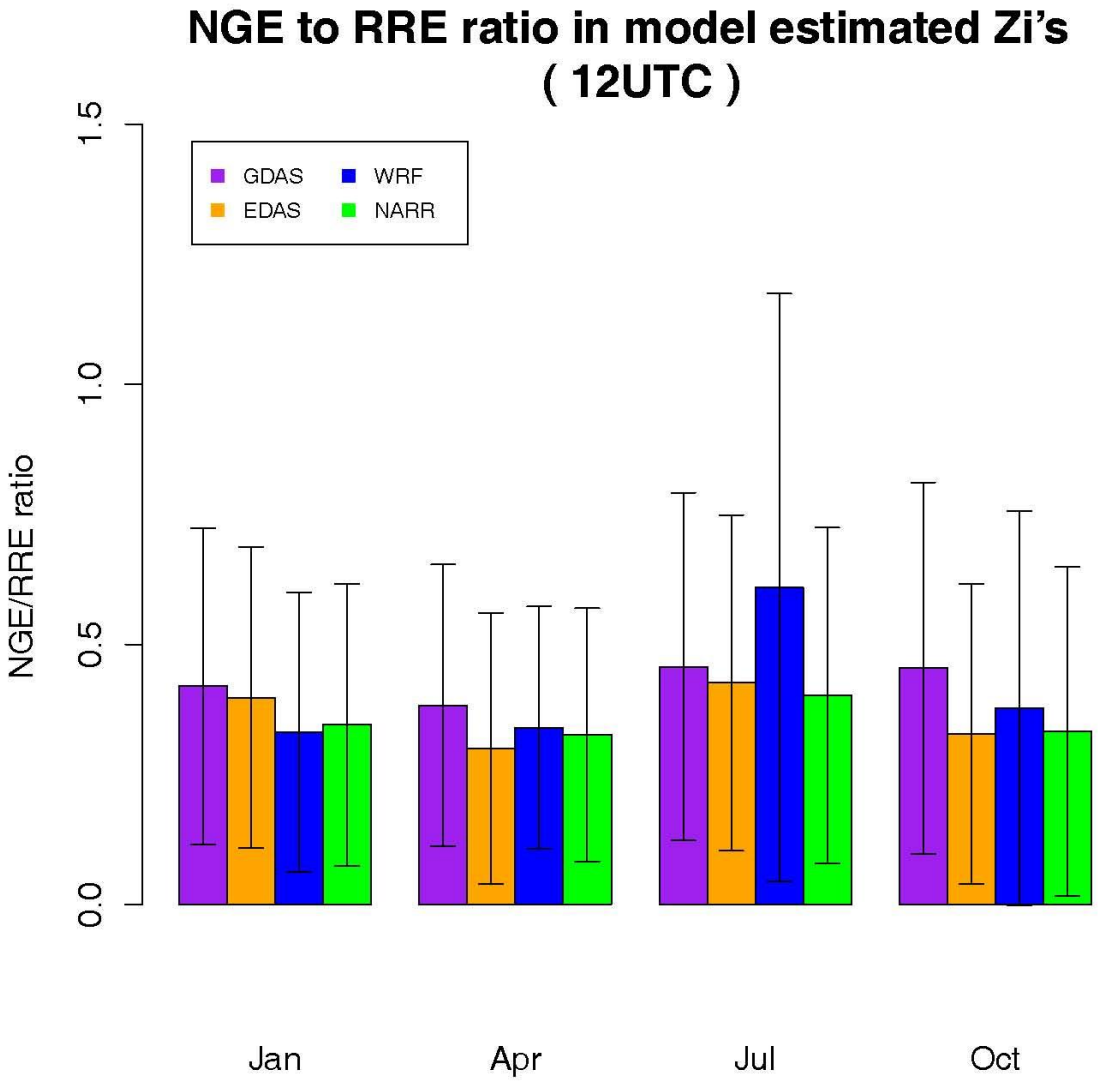

Figure 28 Same as Figure 27, for 12UTC.

3.3 Attribution of errors in z_i

To investigate possible causes of the uncertainties in Ri_b, and consequently z_i, uncertainties in potential temperatures and horizontal winds are examined as the two main variables are used to compute the numbers that have determined the mixed layer heights. Errors of A and B in Ri_b (Eq. 4) are statistically uncorrelated at the 95% confidence interval. The independence of the errors between A and B ensures that the relative uncertainties in the bulk Richardson number can propagate through the relative uncertainties of A and B using equation 5. The model errors in A and B from different models are summarized in Table 2. The table constitutes a summary of relative errors in the components of the bulk Richardson number computed at different altitudes. The largest uncertainties in both A and B of about 20% are found at the bottom 1 km of the atmosphere, and then the relative errors decrease as one ascends to higher levels. Largest errors appearing at the lowest part of the atmosphere suggest that models estimating concentration changes with given surface fluxes may have substantial uncertainties in the prediction results due to the errors in meteorology. Also, the magnitude of relative errors in the temperature term (A) is mostly comparable to those in the windspeed term (B). The comparable magnitudes of errors in A and B mean that windspeed and potential temperature exert more or less the same leverage in causing errors in Ri_b. This implies that the uncertainties in z_i are subject to errors arising from both windspeed and potential temperature, because errors in Ri_b directly translate to errors in model estimated z_i.

3.3.1 Comparisons against CAP profiler network

Although run without data assimilation, WRF performs almost as well in simulating z_i as NARR which ran at similar horizontal resolutions with data assimilation. However, because RAOB was assimilated by the NCEP models (i.e., GDAS, EDAS, and NARR), comparisons against RAOB may underestimate the errors in these models. Since observations from the profiler network (CAP) are not assimilated by any of the models, CAP provides an alternative and independent dataset to be compared against model predictions. The same procedures described in Chapter 2.3.2 were applied to CAP observations to calculate relative uncertainties of A and B. The locations of CAP sites used in this study are shown in the Figure 2, and Table 2 highlights the relative uncertainties in A and B from results. Comparisons to CAP mostly yield larger errors than the ones from RAOB. Increased errors shown in CAP comparisons reflect that values of RAOB data assimilation that might have helped improve the model results. When comparing errors in RAOB to those in CAP for both A and B (Table

2), the magnitudes of uncertainties differ notably from model to model. Both GDAS and EDAS resulted in significantly increased errors for both *A* and *B* while the errors in the other two limited area mesoscale models−WRF and NARR−remain similar. It seems also noticeable that WRF is competitive in its capability to capture Ri_b and z_i even without RAOB assimilation. Since NCEP models that assimilated RAOB (i.e., EDAS and NARR) did not outperform WRF, whether assimilating RAOB would add qualitative advantages to the two mesoscale models looks debatable.

Table 2 Relative uncertainties in the components of bulk Richardson numbers from comparisons. Values are obtained from averages taken over the four months:

$\overline{\left(\frac{\delta Ri_b}{Ri_b}\right)} = \overline{\left(\frac{\delta A}{A}\right)} - \overline{\left(\frac{\delta B}{B}\right)}$, $\boldsymbol{where\ A \equiv \frac{g(z-z_0)[\theta(z)-\theta(z_o)]}{\theta(z)}\ and\ B \equiv u(z)^2 + v(z)^2.}$

Models	Altitude A.G.L. [km]	$\overline{\left(\delta A/A\right)}$ [%]		$-\overline{\left(\delta B/B\right)}$ [%]	
		RAOB	CAP	RAOB	CAP
GDAS	≤ 1	- 21.5	- 29.6	24.0	29.9
	1 ~ 3	- 3.0	*	12.4	21.5
	3 ~ 5	0.1	*	9.0	10.2
	5 ~ 8	- 1.1	*	5.3	- 2.7
EDAS	≤ 1	- 19.3	- 22.6	18.9	25.3
	1 ~ 3	- 3.5	*	10.6	20.0
	3 ~ 5	0.3	*	7.0	7.4
	5 ~ 8	0.6	*	4.4	- 0.4
NARR	≤ 1	- 22.8	- 22.4	15.6	14.6
	1 ~ 3	- 9.4	*	9.7	17.1
	3 ~ 5	- 2.9	*	5.0	4.5
	5 ~ 8	- 1.7	*	2.4	- 2.0
WRF	≤ 1	- 21.9	- 23.8	15.3	14.9
	1 ~ 3	- 9.7	*	9.4	17.4
	3 ~ 5	- 2.9	*	4.8	4.0
	5 ~ 8	- 1.7	*	2.6	- 1.3

* Temperature profiles from optional RASS are available up to 1 km.

The increment of uncertainties in *A* and *B* shown in models-to-CAP comparisons might have arisen from a couple of factors besides the data assimilation. Firstly, a basic assumption for the profiler's ability to measure winds is that the mean wind carries along turbulent eddies that induce scattering [*U.S. EPA*, 2000]. Therefore, signals can be contaminated when turbulent eddies are large relative to the wavelength of the transmitted signals. Besides, various local phenomena such as presence of migrating birds induce non-atmospheric scattering by rendering signals, which typically result in bogus winds [*Lind*, 2011]. Eliminating bogus winds was accomplished by applying the median absolute deviations estimator (MAD, [*Iglewicz and Hoaglin*, 1993]). MAD is a robust measure of the statistical dispersion of a univariate dataset, and is defined as the median of the absolute deviation of the data's geometric mean. Due to the fact that the median is less affected by extreme values in the tail, MAD is ideal for screening the data for outliers.

Chapter 4 Summary and Conclusions

As discussed in Chapter 1, uncertainties in vertical mixing pose considerable difficulties when interpreting atmospheric tracer concentrations. For instance, errors in model-retrieved mixed layer height (z_i) have implications for flux estimations from inversion studies and for air quality modeling. Biases in z_i would require proper corrections in flux estimations while random errors need to be incorporated in error covariance matrices. Such errors, for example, could be enhanced or compensated by errors in surface fluxes. This study was designed to investigate uncertainties in representation of z_i by different atmospheric models with different grid spacings, model parameterizations, and data assimilation characteristics. By comparing multiple models, this study attempted to warrant a choice of an atmospheric model that provides meteorology with marginal uncertainties in terms of mixed layer heights for transport modeling purposes.

The main findings of this study may be divided into two sub-groups: 1) quantification and 2) attribution of errors in model derived z_i.

1) Direct comparisons of the z_i reveal seasonal variation of biases with significant spatial complexity. In spite of its complexity, some patterns in z_i biases can be identified.
   1.1) First of all, the overall magnitude of biases is smaller in regional scale models than that of the global GDAS model, which predominantly shows underestimation during 00UTC.
   1.2) Second, a distinctive difference in biases exists between days and nights (Figure 5). The models mostly show underestimation of mixed layer height during “daytime” (00UTC) when the mixing is vigorous, and overestimation during “nighttime” (12UTC). This may imply that the amplitude of diurnal evolution of the mixing layer represented by models is smaller than that of RAOB measurements. Negative biases in midday z_i can result in over-prediction of tracer concentrations in the atmospheric boundary layer, requiring weaker surface fluxes to match observed values.
   1.3) Third, areas with the largest RRE seem to be coincident with regions of large positive NGE (Figure 9). The linear correlation between the two error properties, however, is insignificant despite of their visual accordance. Spearman correlation and linear regression did not detect any statistically significant evidence of

correlations. Nonetheless, there is reasonable agreement between outliers of both NGE and RRE; extremely large errors are more frequently found at low land locations including near coastal areas.

1.4) Another important finding is that random errors are about double in size than the magnitude of biases (Figures 27 ~ 28). With no linear correlation between NGE and RRE, the fact that RRE is substantially significantly larger than NGE suggests a way to account for the uncertainties through Bayesian inversions. Previous studies [*Lin and Gerbig*, 2005; *Gerbig et al*, 2008; *Kim et al*, 2010] demonstrated improved uncertainty estimations in regional scale tracer fluxes using a stochastic, Lagrangian approach. This method relied on estimated error covariance properties and the fact that random errors are larger than biases, which is indeed confirmed in this study.

2) Substantial variation of uncertainties in z_i over the regional scale demands examination to attribute causes of the errors. Errors in z_i in this study must be related to atmospheric variables such as potential temperature and windspeed, since z_i is derived from the bulk Richardson number using those exact variables.

2.1) Both windspeed and temperature related terms exert similar influence to cause z_i errors.

2.2) The attribution of uncertainties in Ri_b to the explanatory variables differs from one altitude to another. Relative errors in windspeed and potential temperature are the greatest near the surface and diminish with altitude. Hence, concentration changes in the lower atmosphere can be induced by errors in mixing, not only by influences from surface fluxes.

2.3) Comparisons of model results against an independent measurement were made to probe values of data assimilation. Models-to-CAP comparisons almost always result in increased errors than the ones from comparisons against RAOB, which support that data assimilation can improve model results.

2.4) Error statistics from WRF-to-RAOB have magnitudes similar to other regional models. In addition, comparable error ranges in winds and potential temperatures in WRF-to-CAP suggest that assimilating radiosonde observations did not seem to create disadvantages for WRF. Rather, capability of WRF to reproduce z_i is

competitive with NARR even without assimilating radiosonde observations albeit NARR still outperforms WRF.

4.1 Recommendations for future research

This research has quantified the uncertainties of mixed layer heights in four different atmospheric models. Results provided herein can be used to support applications of NARR meteorology for simulating tracer transport. However, evidence from this study indicates that sizeable z_i errors exist, and the exact magnitudes of the errors differ among models and vary between locations and time. The most obvious research question that emerges from the major finding of the current study would be about the implication of and remedies to those complexities in atmospheric model errors. In this regard, addressing the role of uncertainties in mixed layer heights should therefore accompany error characteristics to expand understanding of regional scale tracer flux estimation with atmospheric modeling.

Bibliography

Alkhaled, A.A., A.M. Michalak, S. R. Kawa, and co-authors (2008), A global evaluation of the regional spatial variability of column integrated CO2 distributions, *J. Geophys. Res.*, 113, D20303, doi:10.1029/2007JD009693.

Aulagnier, C., P. Rayner, P. Ciais, and co-authors (2008), Is the recent build-up of atmospheric CO2 over Europe reproduced by models. Part 2: an overview with the atmospheric mesoscale transport model CHIMERE, *Tellus, Ser. B*, 62, 14-25.doi:10.1111/j.1600-0889.2009.00443.x.

Benjamin, S.G., B.E. Schwartz, E.J. Szoke, and co-author (2004), The value of wind profiler data in U.S. weather forecasting, *Bull. Amer. Meteor. Soc.*, 85(12), 1871-1886, doi: 10.1175/BAMS-85-12-1871

Bivand, R.S., E.J. Pebesma, and V. Gomez-Rubio (2008), *Applied spatial data analysis with R*, Springer.

Denning, A.S., D.A. Randall, G.J. Collatz, and co-author (1996), Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model. Part2: Simulated CO2 concentrations, *Tellus, Ser. B*, 48, 543-567.

Derber J.C., D.F. Parrish, and S.J. Lord (1991), The new global operational analysis system at the National Meteorological Center, *Weather and Forecasting*, 6, 538-547.

Gerbig, C., J.C. Lin, S.C. Wofsy, and co-authors (2003), Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, *J. Geophys. Res.*, 108(D24), 4757, doi:10.1029/2003JD003770.

Gerbig, C., S. Körner, and J.C. Lin (2008), Vertical mixing in atmospheric tracer transport models: error characterization and propagation, *Atmos. Chem. Phys.*, 8, 591-602.

Hanna, S.R. and R. Yang (2001), Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths, *J. of Appl. Meteor.*, 40, 1095-1104.

Holzworth, G. (1964), Estimates of mean maximum mixing depth in the contiguous Unites States, *J. Appl. Meteor.*, 6, 1039-1044.

Holzworth, G. (1967), Mixing Depths, Wind speeds and air pollution potential for selected locations in the United States, *Mon. Weather Rev.*, 92, 235-242.

Iglewicz, B. and D. Hoaglin (1993), *How to detect and handle outliers, The ASQC basic references in Quality Control*, vol. 16, American Society for Quality Control, Milwaukee, Wisconsin.

Ishizawa, M., D. Chan, K. Higuchi, and co-authors (2006), Rectifier effect in an atmospheric model with daily biospheric fluxes: impact on inversion calculation, *Tellus*, 58(B), 447-462.

Jacobson, M.Z. (2005), *Fundamentals of atmospheric modeling*, Cambridge.

Kanamitsu, M. (1989), Description of the NMC Global Data Assimilation and Forecast System, *Weather and Forecasting*, 4, 335-342.

Kim, M., J.C. Lin, D. Worthy (2010), Towards a regional-scale carbon atmospheric inversion system over Canada, paper presented at 2010 Annual General Meeting, Canadian Carbon Program, Montreal, Canada.

Kim, D., and W. Stockwell (2008), An online coupled meteorological and air quality modeling study of the effect of complex terrain on the regional transport and transformation of air pollutants over the western United States, *Atmospheric Environment*, 42, 4006-4021, doi:10.1016/j.atmosenv.2006.11.031.

Lin, J.C., and C. Gerbig (2005), Accounting for the effects of transport errors on tracer inversions, *Geogphys. Res. Lett.*, 32(1), L01802, doi:10.1029/2004GL021127.

Lin, J.C, C. Gerbig, S.C. Wofsy, and co-authors (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.*, 108(D16), 4493, doi:10.1029/2009JD003161.

Lind, R. (2011), personal communication, Naval Postgraduate School, California, U.S.A.

Mahadevan, P., S. C. Wofsy, D.M. Matross, and co-authors (2008), A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochem. Cycles*, 22, GB2005, doi:10.1029/2006GB002735.

Matross, D.M., A. Andrews, M. Pathmathevan, and co-authors (2006), Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, *Tellus(B)*, 58(5), 344-358, doi:10.1111/j.1600-0889.2006.00206.x

Mesinger, F., G. DiMego, E. Kalnay, and co-authors (2006), North American Regional Reanalysis, *Bull. Amer. Meteor. Soc.*, 87(3), 343-360, doi: 10.1175/BAMS-87-3-343.

Miller, P.A., M.F. Barth, and L.A. Benjamin (2005), An update on MADIS observation ingest, integration, quality control, and distribution capabilities. *21st Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, San Diego, CA, American Meteorological Society.

Nehrkorn, T., J. Eluszkiewicz, S.C. Wofsy, and co-authors (2010), Coupled weather research and forecasting–stochastic time-inverted lagrangian transport (WRF–STILT) model, *Meteorol. Atmos. Phys.*, 107, 51-64, doi:10.1007/s00703-010-0068-x.

NOAA, ESRL Global Systems Division (2009), ESRL Radiosonde Database, *http://esrl.noaa.gov/raobs/General_Information.html.*

Peters, W, A.R. Jacobson, C. Sweeney, and co-authors (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proceedings of the National Academy*

of Sciences, 104, 18925-18930.

Pielke, R. A. (2002), *Mesoscale meteorological modeling. 2nd Ed.*, Academic press, San Diego, CA.

Pielke, R.A. and M Uliasz (1998), Use of meteorological models as inputs to regional and mesoscale air quality models-limitations and strengths, *Atmospheric Environment*, 32, 1455-1466.

Saha, S., S. Moorthi, H. Pan, and co-authors (2010), The NCEP climate forecast system reanalysis, *Bull. Amer. Meteor. Soc*, 91(8), 1015-1057, doi: 10.1175/2010BAMS3001.1

Schwartz, B., and M. Govett (1992), A hydrostatically consistent North American radiosonde data base at the Forecast Systems Laboratory, 1946–present, *NOAA Tech. Memo*, ERL FSL-4, NOAA FSL, Natl Oceanic and Atmos. Admin., Boulder, Colo, 81 pp.

Seaman, N.L., D. R. Stauffer, and A. M. Lario (1995), A multiscale four dimensional data assimilation system applied in the San Joaquin Valley during SARMAP. Part I: Modeling design and basic performance characteristics. *J. Appl. Meteor.*, 34, 1739–1761.

Stauffer, D. R., N.L. Seaman, and F. S. Binkowski (1991), Use of four-dimensional data assimilation in a limited-area mesoscale model. Part II**:** Effects of data assimilation within the planetary boundary layer, *Mon. Wea. Rev.***,** 119, 734-754.

Siebert, P., F. Beyrich, S. Gryning, and co-authors (2000), Review and intercomparison of operational methods for the determination of the mixing height, *Atmospheric Environment*, 34, 1001-1027.

Skamarock, W.C. (2006), Positive-definite and montonic limiters for unrestricted-timestep transport schemes, *Mon. Wea. Rev.*, 134, 2241-2250

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, W. Wang, J.D. Powers (2005), A description of the advanced research WRF version 2, *Technical Note 468+STR*, MMM Division, NCAR, Boulder, CO, 88 pp.

Stephens, B.B., K. R. Gurney, P. P. Tans, and co-author (2007), Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO2, *Science*, 316, 1732-1735.

Stull, R. (1988), *An introduction to boundary layer meteorology*, Kluwer Academic Publishers.

Taylor, J.R. (1997), *An introduction to error analysis*, University science books.

Turquety, S., J.A. Logan, D.J. Jacob, and co-authors (2007), Inventory of boreal fire emissions for North America in 2004: Importance of peat burning and pyroconvective injection, *J. Geophys. Res.*, 112, D12S03, doi:10.1029/2006JD007281.

U.S. EPA (2000), *Meteorological monitoring guidance for regulatory modeling applications*, EPA-454/R-99-005.

Vogelezang, D.H.P., and A.A.M. Holtslag (1996), Evolution and model impacts of alternative

boundary layer formulations. *Boundary-Layer Meteorology*, 81, 245-269.

Wallace, J.M., P. Hobbs (2006), *Atmospheric science: an introductory survey*, Elsevier Inc.

Wang, I.T. (1981), The determination of surface-layer stability and eddy fluxes using wind speed and vertical temperature gradient measurements, *J. of Appl. Meteor.*, 20(10), 1241-1248.

Zhang, Jian and S. Trivikrama Rao (1999), The Role of Vertical Mixing in the Temporal Evolution of Ground-Level Ozone Concentrations, *J. of Appl. Meteor.*, 38, 1674-1691.