

We thank Reviewer #1 for the constructive criticism of the Discussion paper. The reviewer's comments are shown below in *italics*, while our point-by-point responses are indicated as un-italicized.

Reviewer #1

General Comments

The WRF-STILT model results at 4 and 1.3 km resolutions compared well with observation. But the differences between WRF-STILT model results and observations became unexpected large at 12 km resolution (Fig. 3), with poorer agreement than the much lower resolution models (GDAS-STILT and CarbonTracker, both at 1 degree resolution). The problem seems to be the modeled PBL. The 12 km WRF-STILT model mean PBL shown in Figs. 7, 8 and 9 are quite different from the WRF-1.3, WRF-4 and GDAS-1 deg results. Similarly the GDAS-STILT results exhibit odd behavior (see below).

The results from the modeled PBL scheme actually do not differ appreciably in the three WRF configurations. This was made much clearer by our addition of new figures for the Discussion paper (Figs. 8, 10, 12) showing the time series of PBL heights (and trajectories) with respect to above-ground level (AGL). The key difference is the smoothing of the terrain in the coarser simulations. We have added new text to clarify this point:

" An alternative perspective is to view the trajectory and PBL heights relative to the ground surface ("AGL") instead of above sea level, at each time step backward in time from the receptor (Figs. 8, 10, 12). These figures highlight the fact that while PBL dynamics in the three WRF configurations are similar, the heights of the trajectories relative to the PBL height differ. The trajectory exits above the nocturnal PBL one hour backward in time, on average, while the WRF-12km trajectory spends several hours within the PBL.

The difference in the trajectory behavior can be explained by the differing terrain. In mountainous terrain, PBL heights generally follow the terrain elevations, albeit with attenuated amplitude (Steyn et al., 2013). Thus in WRF-1.3km and 4km, the more highly resolved terrain produced shallow nocturnal PBL height that descend in the valley (Fig. 7) while the corresponding trajectory hovers above it. Viewed relative to the ground surface (Fig. 8), the trajectory originating from HDP appears to have exited above the nocturnal PBL in WRF-1.3km and 4km. In contrast, due to the significantly "flattened" mountains in WRF-12km and in GDAS, the PBL heights exhibit less spatial variation near the mountaintop receptor, since the terrain itself was smoothed. Consequently, WRF-12km trajectories, unlike the WRF-1.3km or -4km cases, travel closer to the ground surface, within the nighttime PBL, even as it is advected away from the three RACCOON sites (Figs. 7, 8). This resulted in stronger nighttime footprints in WRF-12km as seen in Figs. 4 and 5. Another effect of the proximity of the air parcels to the model's ground surface is the slower windspeeds from surface drag, causing the air parcel trajectories to remain close to the 3 sites until the previous day; for HDP and SPL, the mean trajectories spiral toward the site at the surface, following an "Ekman wind spiral" pattern (Holton, 1992). In WRF-1.3km and WRF-4km, the measurement sites are at significantly higher elevations above the resolved valleys in the area surrounding the sites, and the air parcels are found above the shallow nocturnal boundary layer hugging the valley floor, on average (Fig. 7)."

This raises many questions:

(1) Is there a problem with the modeling of PBL in the WRF-12km model?

No--please see response above.

(2) Is there a problem with the modeling of PBL in the GDAS-STILT model (Figs. 10, S10, S12)? Figs. S10 and S12 show HDP and NWR PBL much higher (5000m) compared to PBL from WRF-STILT runs.

We believe the considerable difference in behavior within GDAS is due to both the significantly coarser vertical and horizontal resolutions.

We now explain the impact of this coarse resolution:

"Another noticeable difference in GDAS-ASL trajectory was the significantly higher daytime PBL heights (Figs. 8, 10, 12). We suspect this is because of the greatly reduced vertical resolution within GDAS (23 levels versus 41 levels in WRF): since STILT diagnoses the PBL height to correspond to a model level, a higher PBL height was chosen for GDAS because of the thicker vertical level. Another subtle artifact of the coarse resolution within GDAS can be seen in the anomalously low daytime PBL height just in the vicinity of HDP (Figs. 13, S10). It appears that the GDAS model set an entire $1^{\circ} \times 1^{\circ}$ grid box near HDP to be water body (the Great Salt Lake), thereby suppressing the PBL height."

Fig. 10 appears mislabeled and not consistent with the discussion in the text, therefore very difficult to understand, see specific comment.

We believe that the impression of Fig. 10 being mislabeled likely resulted from the Reviewer looking at an older version of the manuscript instead of the version published as the Discussion paper. See below ("Specific Comments") for details.

(3) How realistic are the PBL results in these models (WRF-1.3km, WRF-4km, WRF-12km, GDAS)? How were the PBL results evaluated (for mountaintop conditions)?

(4) How realistic is PBL model for mountaintop simulations? How does the PBL model account for the difference between flat plain and mountainous conditions (as the model domain contains such surface conditions)?

The authors should provide more technical details on the PBL model (e.g. dependence on atmospheric conditions and topographic variations, daytime and nighttime variations), particularly how it changes for the different (met data) models (to cause the unusual results for WRF-12km and GDAS) in this study. Then present more analysis and discussion on the uncertainties or errors in the PBL and consequently footprints in this study.

The aforementioned questions (3) and (4) are linked and currently difficult to answer. How to carry out PBL simulations in mountainous terrain and how to evaluate them very much remains unsettled. The requisite meteorological and tracer observations (in

addition to CO₂) are limited for the sites examined in this paper. We fully recognize this difficulty and have already included a recommendation for additional observations in future mountaintop sites:

"We recommend additional tracers to be measured in conjunction with the mountaintop CO₂ sites. For instance, combustion tracers such as C¹⁴ and CO (Levin and Karstens, 2007) have been measured alongside CO₂ at mountaintop sites in Europe. Another promising tracer is Rn²²² (Griffiths et al., 2014), which provides a measure of surface exchange and would help provide constraints on the exchange of air measured at the mountaintop with the surface. Co-located meteorological observations—whether in-situ or remotely-sensed (e.g., radar, sodar, lidar)—to probe atmospheric flows and turbulent mixing would also be of significant value in helping to interpret the tracer observations (Rotach et al., 2014; Banta et al., 2013)."

Specific comments

Lines 413-414: 'Due to the significantly "flattened" mountains in WRF-12km and in GDAS, the PBL height exhibits less spatial variation.' This is not true for WRF-12km in Figure 7.

This actually still holds for WRF-12km in Fig. 7, but is true, strictly speaking, only for the key area near the receptor. We have now clarified this point to:

"In contrast, due to the significantly "flattened" mountains in WRF-12km and in GDAS, the PBL heights exhibit less spatial variation near the mountaintop receptor, since the terrain itself was smoothed."

Lines 469-472: 'Focusing on the three-dimensional plots at the hours of 0800 and 1100 MST (Fig. 10), when the simulated peaks are found at SPL and both NWR/HDP, respectively, the peaks coincide with times when average trajectories are found within a relatively shallow morning PBL.' This statement seems to conflict with Fig. 10, which shows the HDP site with a deep morning average PBL. There seems to be error(s) in Fig. 10. The HDP plot domain is the same as shown for NWR in Fig. 9. The plot domain for NWR is as for HDP in Fig. 7. The plotted curves for the two upper plots with nearby domains seem to show drastic difference for PBL, the PBL for the HDP plot is definitely not 'shallow morning PBL' at 6000m.

Reviewer #1 appears to be looking at an earlier version of paper and not the version published as the Discussion paper. Lines 469-472 have different text now, and we believe the text in the published Discussion paper matches what is shown in the figures.

Lines 734-736: 'Red portions of the trajectory refer to the nighttime (1900_0700 MST), while pink portions indicate the daytime (0700_1900 MST).' The red and pink portions are hard to distinguish, use more contrasting colors.

We thank the Reviewer for pointing this out. We have followed the suggestion and changed the colors to show more contrast.

Lines 738-739: 'Fig. 8 Similar to Fig. 7, but for the Storm Peak Laboratory (SPL) site.' The blue PBL line appears to have dark and light portions not explained. Again, the

contrast is hard to distinguish, use more contrasting colors. This may apply to Figs. 7, 9, 10 too.

We thank the Reviewer for pointing this out. We have followed the suggestion and changed the colors to show more contrast.

Figure 7 (also 8, 9, 10, and corresponding figures in supplemental material): add color scale to facilitate figure comparisons.

We do not feel that the color scale is necessary, since the colors represent elevation and the same information is also represented by the z-axis, as well as the three-dimensional terrain representation. Furthermore, these figures already incorporate a lot of information, so adding another colorscale could make the figures look too "busy".

Supplemental material

Figure S3: it is not clear if the flux reversal correction takes into account the longer daytime than nighttime during the summer.

Yes--the longer daytime can be accommodated in our flux reversal correction algorithm. The algorithm preserves the area between the flux time series and the flux = 0 line. So to the extent that the diurnal pattern in the uncorrected CarbonTracker fluxes shows a longer period of release or uptake, this would show up in the corrected fluxes too. In other words, the algorithm preserves the relative length of carbon uptake or release in the original CarbonTracker fluxes, but any nighttime uptake would be moved to the daytime.

Figure S5: 'CO.obs...' to 'CO2.obs...'

We thank the reviewer for pointing out this error. This figure has now been fixed.

Figs. S11, S12: PBL for GDAS higher than 6000m seems unusually high.

We believe the high PBL within GDAS is mainly due to the coarse vertical grid spacing. We have added an explanation of this point:

"Another noticeable difference in GDAS-ASL trajectory was the significantly higher daytime PBL heights (Figs. 8, 10, 12). We suspect this is because of the greatly reduced vertical resolution within GDAS (23 levels versus 41 levels in WRF): since STILT diagnoses the PBL height to correspond to a model level, a higher PBL height was chosen for GDAS because of the thicker vertical level."

We thank Reviewer #2 for the constructive criticism of the Discussion paper. The reviewer's comments are shown below in *italics*, while our point-by-point responses are indicated as un-italicized.

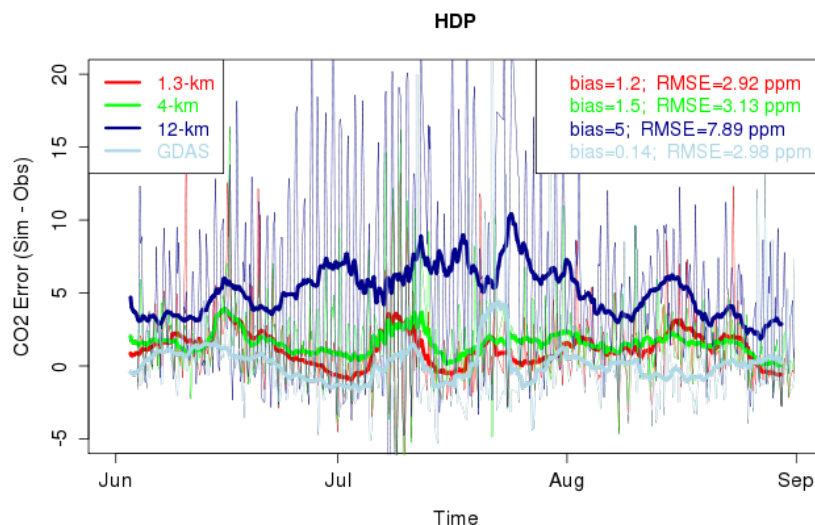
Reviewer #2

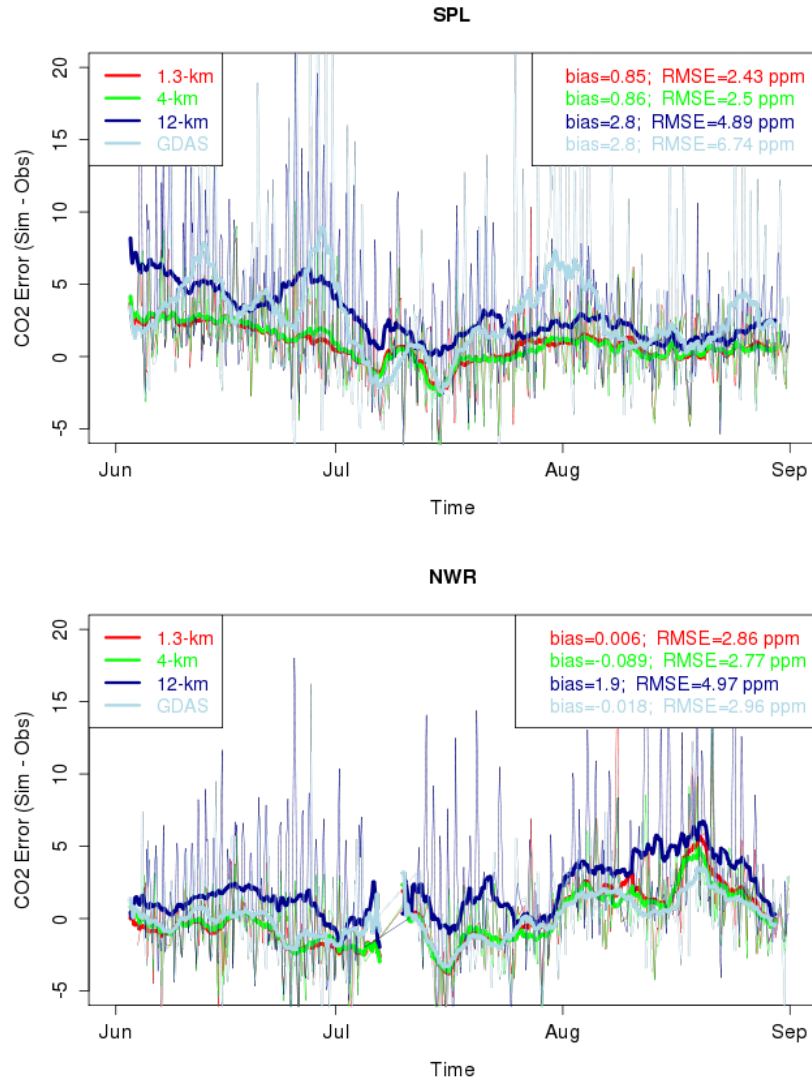
General comments:

The paper is well-written and the observational datasets, models and methods are properly described and referenced. The authors address an urgent question about the future and current use of observations from mountainous areas or within complex terrains more general. Their analysis of mean daily cycle allows an easy interpretation of mean behaviour, but it would be critical to also investigate if there are episodes (meteorological conditions) for which the mismatch between the different transport model configuration is minimized/maximized. I would suggest to add an analysis of the time-dependent offset of each model compared to the observations.

We agree with the Reviewer that beyond the diurnal timescale, there is day-to-day variability in the model behavior that lead to variations in CO₂ model errors. We have now included in the revised paper the time series of CO₂ model errors at SPL, NWR and HDP (see below) with correlations of the CO₂ errors with various meteorological variables (geopotential height and its E↔W gradient, U- and V-winds, and the windspeed). The time series plots and a Table of the correlations of the CO₂ errors with these meteorological variables will be added to the Supplement.

Errors at multi-day timescales are the most strongly correlated with different meteorological variables, depending on which is being examined: V-wind, U-wind, and geopotential height gradient for HDP, SPL, and NWR, respectively. The fact that errors are correlated with different meteorological variables depending on the site location points to a complexity that can only be unraveled with a substantial expansion of the paper. This could potentially be a subject for a future paper. This complexity is in contrast to the average diurnal biases that can in large part be linked to the underlying resolution of modeled terrain, which the current paper focuses on.





In general it would be valuable to have more quantitative results and a discussion how generalizable the findings of this study are for other hilly/mountainous areas.

We do attempt to broaden the scope of this study. In particular, we outline various approaches in the Discussion section with regards to making use of mountaintop CO₂ data that can be considered for other mountainous areas as well. However, we believe detailed quantitative results for other mountainous areas would require dedicated modeling efforts. This need for dedicated efforts for individual mountainous sites is hinted at by differences in results between the three sites examined in this study: HDP, SPL, and NWR (Figs. 3, 4). We have pointed to the contrast in elevation at the mountaintop site to the surrounding terrain as a key factor in explaining differences observed at the three sites (Figs. 7, 9, 11). Hopefully these results would stimulate other researchers around the world to also examine the same approaches and factors mentioned in this paper, but the dedicated modeling efforts necessary to do so would be outside the purview of this paper.

Overall this paper also does not fully address the question of "constraining carbon fluxes", but rather how well different model setups can reproduce the atmospheric concentrations of CO₂.

To be able to really judge if the models (even the best, 1.3km resolution) are able to e.g. distinguish different prior carbon flux estimates the authors would need to perform a sensitivity study using multiple carbon flux data sets and demonstrate a significant impact at the three sites. Potentially one can consider this study rather a step towards a better use of such data rather than already addressing the question of regional carbon fluxes.

We agree that the paper is the first step towards addressing the question of regional carbon fluxes. However, we believe that it is a critical first step that needs to be taken in order to use mountaintop CO₂ data to constrain regional carbon fluxes.

After the above and below comments are addressed I would definitely recommend this study for publication in ACP as it will help to better understand limitations of such observations and were future model developments should/could be focussed to eventually be able to constrain carbon fluxes in such regions.

Specific comments:

Line 44ff: The claim that nearly 70% of the earth land surface is covered by hills or mountains needs to be better validated. This surely depends on the definition for hill or mountain, which is not given here and the cited publication is hard to access (due to the journal it was published in) and the journal has an impact factor below 1.

We thank the Reviewer for pointing this out. The claim of ~70% of the Earth's land surface as being covered by hills or mountains was attributed to Rotach et al. [2008], but we traced this claim to a book written in the German language. Furthermore, there appears to be no clear definition for what is meant by "hills". Therefore, we decided to revise the statement to just referring to "mountains", which cover about one quarter of the Earth's mountains, citing a readily-accessible UNEP report for this purpose (Blyth et al., 2002).

Blyth, S., B. Groombridge, I. Lysenko, L. Miles, and A. Newton, Mountain Watch: Environmental Change and Sustainable Development in Mountains, UNEP World Conservation Monitoring Centre, 2002.

The authors also mention that carbon fluxes in complex terrain need to be better understood to quantify carbon flux. It seems you are implying that all mountainous or hilly areas are (too) hard to model?

We were not necessarily implying all mountainous or hilly areas are difficult to model. We were suggesting that because mountainous areas cover a large fraction of the Earth's land surface and significant amounts of biomass can be found in mountains (e.g., Fig. 1), a better understanding of this under-sampled region is necessary.

Line 193: The authors refer to a previous publication – nonetheless the key parameters e.g. vertical mixing scheme used should be explicitly given in this publication (e.g. by adding a table in this section).

We have revised the paper to include key parameters regarding the WRF configuration. The revised sentence reads:

"Comprehensive testing of different WRF settings have been carried out as part of a previous publication (Mallia et al., 2015), and these settings were adopted here: i.e., the MYJ, Grell-

Devenyi Ensemble, and Purdue Lin schemes for parameterizing the planetary boundary layer (PBL), cumulus convection, and microphysics, respectively.”

Line 216: Please specify if the system allows for two-way nesting or not

Following our testing in Mallia et al. (2015), we have implemented two-way nesting within WRF. This is clarified in the revised text in Sect. 2.2:

“For this study, we ran WRF in a two-way nested mode centered between Utah and Colorado where the RACCOON sites are located (Fig. 2).”

Line 277: A “fix” is mentioned, but not explained at all. Please consider giving a brief description here rather than referring to the supplement. It seems the daily cycle has just been shifted or were there any more complicated adjustments performed?

Yes—only the diurnal pattern has been shifted while preserving the 24-hour integrated carbon flux. We have added more information for the reader in the revised text:

“For this paper, we implemented a fix that removed this artifact by detecting these reversed diurnal patterns, adjusting them while preserving the 24-hour integrated carbon flux. See the Supplement and Fig. S3 for details.”

Line 304: Please consider referring to table 1 here so the reader can easily find the height CT data was extracted from.

We have added a reference to Table 1.

Line 503ff: The first question is repeated here “How can mountaintop CO₂ observations be used to constrain regional scale carbon fluxes, : : :.” But the 5 subsections following rather discuss IF such data can be used or how they can be better used. It remains unclear if there is a definitive answer on how to use them.

Due to the fact that the errors incurred depends on the model resolution, the relationship of the mountaintop site relative to surrounding terrain and emissions, and the quality of the prior fluxes, the definitive answer depends upon each specific situation. Thus we were hesitant to suggest an answer that overgeneralizes. Also see above for the response regarding statements for other mountainous areas.

Line 530ff: Choosing the appropriate model layer to extract CO₂ does indeed introduce a significant additional degree of freedom. The authors suggest other parameters to avoid creating a fudge factor but do not give specific advice here on which tracers could be useful (²²²Rn?). Meteorological data is mentioned but looking at table 2 it seems not at all clear that this would be good parameter or what a cut-off would be. Could you suggest how a suitable proxy could be found?

These tracers and meteorological data were mentioned later on in the Discussion section:

“We recommend additional tracers to be measured in conjunction with the mountaintop CO₂ sites. For instance, combustion tracers such as C¹⁴ and CO (Levin and Karstens, 2007) have been measured alongside CO₂ at mountaintop sites in Europe. Another promising tracer is Rn²²² (Griffiths et al., 2014), which provides a measure of surface exchange and would help provide constraints on the exchange of air measured at the mountaintop with the surface. Co-located meteorological observations—whether in-situ or remotely-sensed (e.g., radar, sodar, lidar)—to probe atmospheric flows and turbulent mixing would also be of significant value in helping to interpret the tracer observations (Rotach et al., 2014; Banta et al., 2013).”

Line 581ff: Here the authors report on the practice of not using mountaintop data but it is unclear how this is linked to this specific study as no suggestion is made how to e.g. better use Schauinsland data. This section should be considered for the introduction to motivate why Approach 1, 3, 4, 5 need to be improved.

We thank the Reviewer for this valuable suggestion. We have followed the Reviewer's suggestion and moved Approach 2 ("Reject mountaintop data") to the Introduction to provide further motivation for the paper.

Line 595ff: When setting up an inversion system it is common (good) practice to assign proper model errors. This seems not specific to this study and the authors fail to give an estimate of the model error for the three sites discussed here. Please consider removing this section or giving quantitative results for the sites and models investigated here. Of course, the model data mismatch calculated here also depends on flux errors, but the authors can surely use this study to give an upper limit of this combined error (and the difference for different model resolutions).

We agree with the Reviewer that quantitative results for the sites and models investigated here would be useful. An estimate of these errors is the RMSE (root-mean-square-error) calculated for each site and model setup. We now direct the reader to the RMSE values shown in the top right-hand corner of the time series plots found at the beginning of this Response.

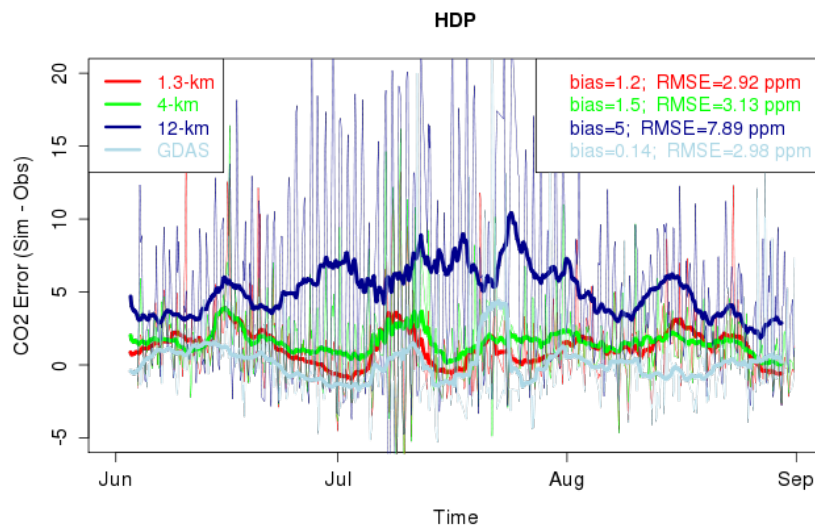
We thank Reviewer #3 for the constructive criticism of the Discussion paper. The reviewer's comments are shown below in *italics*, while our point-by-point responses are indicated as un-italicized.

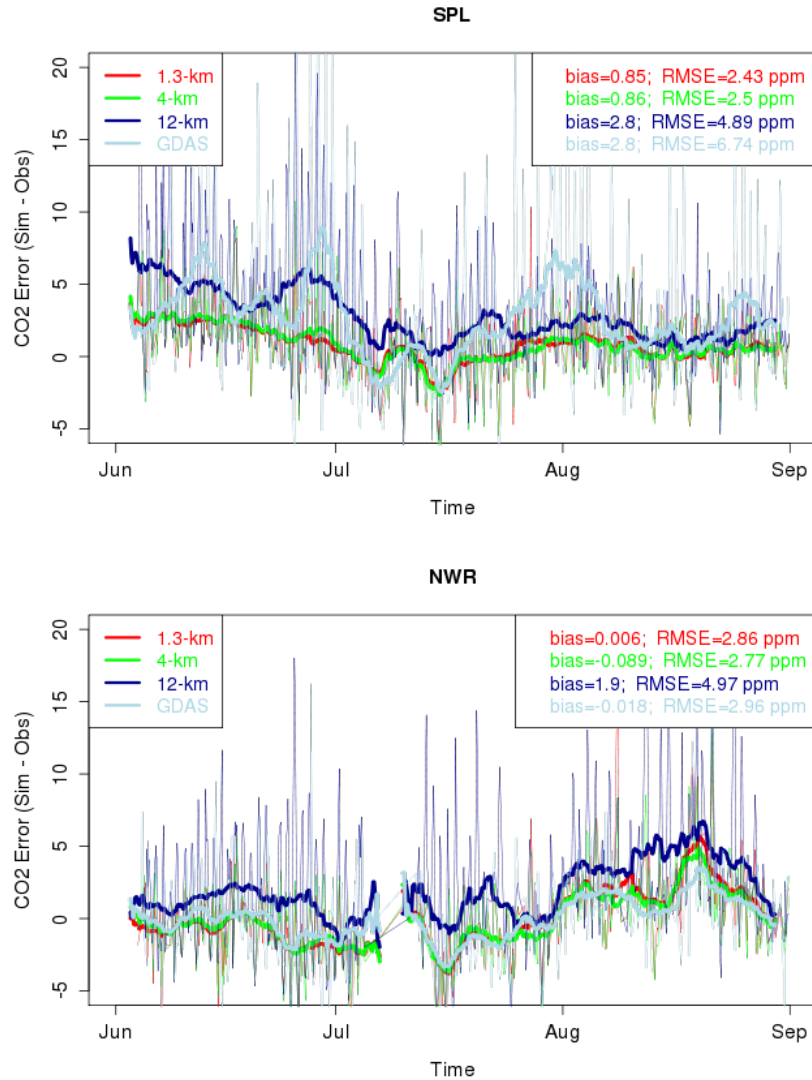
Reviewer #3

The major concern is that how different the CO₂ simulations and observations are compared from day to day, particularly when finer scale models are used as suggested? Averaged results might miss some crucial information hidden in the discrepancies between model simulations and site observations in present study.

We agree with the Reviewer that beyond the diurnal timescale, there is day-to-day variability in the model behavior that lead to variations in CO₂ model errors. We have now included in the revised paper the time series of CO₂ model errors at SPL, NWR and HDP (see below) with correlations of the CO₂ errors with various meteorological variables (geopotential height and its E⇌W gradient, U- and V-winds, and the windspeed). The time series plots and a Table of the correlations of the CO₂ errors with these meteorological variables will be added to the Supplement.

Errors at multi-day timescales are the most strongly correlated with different meteorological variables, depending on which is being examined: V-wind, U-wind, and geopotential height gradient for HDP, SPL, and NWR, respectively. The fact that errors are correlated with different meteorological variables depending on the site location points to a complexity that can only be unraveled with a substantial expansion of the paper. This could potentially be a subject for a future paper. This complexity is in contrast to the average diurnal biases that can in large part be linked to the underlying resolution of modeled terrain, which the current paper focuses on.





The related studies focusing on mountaintop such as this one are important and rarely explored. I would recommend this manuscript for publication after above concern is addressed.

Specific comments:

Line 146: How is the vegetation covered at these mountain stations? Are these monitoring sites far above the tree line? Local influences on CO₂ due to the surrounding plant cover at these sites may need to be considered.

Both HDP and NWR are above the tree line; SPL has a few sparse trees around the site. The RACCOON mountaintop observations have been filtered to remove local influences and to extract values that are more regionally representative, following the work in Brooks et al. (2012). These filtered observations were the ones used in the paper. We have added this reference and modified the text to clarify this point:

“We applied filtering to the mountaintop CO₂ observations to remove local influences and to extract values that are more regionally representative (Brooks et al., 2012). Observations were filtered out in which the within-hour standard deviation is greater than 1.0 ppm or when the differences between the top two inlets are greater than 0.5 ppm, which indicate periods when significant influences that are highly localized to the site are affecting the observations.”

Brooks, B.-G. J., Desai, A. R., Stephens, B. B., Bowling, D. R., Burns, S. P., Watt, A. S., Heck, S. L. and Sweeney, C.: Assessing filtering of mountaintop CO₂ mole fractions for application to inverse models of biosphere-atmosphere carbon exchange, *Atmos. Chem. Phys.*, 12, 2099–2115, 2012.

Lines 296-297: How much percent is contributed to the wildfires during the study period?

The exact percentage depends on the quantify under consideration: gross versus net fluxes, and the percentage would also depend upon the time of the day. Regardless, the percent is very small and is visually apparent from Fig. S4. We added an additional reference to Fig. S4 to clarify this point:

“Contributions from anthropogenic and wildfire emissions, on average, to the mean CO₂ diurnal cycle observed at all the mountain sites were secondary in comparison to the biosphere (Fig. S4). In particular, the wildfire contributions were episodic and averaged out to negligible contributions over Jun~Aug 2012 (Fig. S4).”

Lines 301-302 and Fig. 3: Fig. 3 illustrates the 3-month averaged diurnal cycles of the results from individual models and observations. Do they show similar diurnal cycles every day during these months? Consider addressing the variabilities of the diurnal cycles (e.g. error bars) in the supplement.

To illustrate the variability in the diurnal cycles we have added error bars to the average diurnal cycles shown in Fig. 3.

Fig. 4: Has the GDAS_1-deg_ASF in Fig. 4 been adjusted with the biospheric fluxes? If not, I would recommend drawing it as gray line and dot as presented in Fig. 3.

Fig. 4 is showing the average diurnal cycle of the footprint totals, which result solely from the simulated atmospheric transport. Thus biospheric fluxes are not incorporated into Fig. 4 and therefore the dashed + dot scheme paralleling Fig. 3 is not necessary here.

Fig. S5: Two CT Level 1 (dark blue) and two CT Level 2 (blue)?

We are not sure about the exact meaning of the Reviewer’s comment here. The CT levels are color-coded as a gradation from dark blue (Level 1) to dark green (Level 8); Level 2 is indeed colored as blue.

Fig. S5 and Table 1: The differences between the site and model altitude (Table 1) seem to be generally associated with the CO₂ biases. For instance, larger differences such as WRF-12km, GDAS, and Carbon Tracker showed larger discrepancies in diurnal cycles of CO₂ at HDP. What

if the model altitude fixed as same as the mountain site? Could it be better correlated to the observational CO₂ data?

We are in full agreement with the Reviewer that CO₂ biases are associated with differences in site versus model altitudes, stemming from increasing discrepancies with terrain as the spatial resolution is degraded. This point is discussed extensively in Sect. 3.2.2. Adjusting the vertical level within the model is an approach that we explored Sect. 4.1 (“Approach 1: Adjust vertical level of simulations from which to compare against observed values”)

Lines 336-380: Are these trajectories in a good agreement for each day during the study period both in nighttime and afternoon hours?

We have generated the time series of CO₂ error at different days during the study and examined the correlation between the errors with other meteorological variables to address this question. See details at the top of this response for more details.

How can mountaintop CO₂ observations be used to constrain regional carbon fluxes?

John C. Lin¹, Derek V. Mallia¹, Dien Wu¹, Britton B. Stephens²

¹Department of Atmospheric Sciences, University of Utah, Salt Lake City, Utah 84112, USA

²Earth Observing Laboratory, National Center for Atmospheric Research, Boulder, Colorado 80301, USA

Correspondence to: John C. Lin (John.Lin@utah.edu)

Manuscript Submitted to Atmospheric Chemistry and Physics

Abstract.

Despite the need for researchers to understand terrestrial biospheric carbon fluxes to account for carbon cycle feedbacks and predict future CO₂ concentrations, knowledge of these fluxes at the regional scale remains poor. This is particularly true in mountainous areas, where complex meteorology and lack of observations lead to large uncertainties in carbon fluxes. Yet mountainous regions are often where significant forest cover and biomass are found—i.e., areas that have the potential to serve as carbon sinks. As CO₂ observations are carried out in mountainous areas, it is imperative that they are properly interpreted to yield information about carbon fluxes. In this paper, we present CO₂ observations at 3 sites in the mountains of the Western U.S., along with atmospheric simulations that attempt to extract information about biospheric carbon fluxes from the CO₂ observations, with emphasis on the observed and simulated diurnal cycles of CO₂. We show that atmospheric models can systematically simulate the wrong diurnal cycle and significantly misinterpret the CO₂ observations, due to erroneous atmospheric flows as a result of terrain that is misrepresented in the model. This problem depends on the selected vertical level in the model and are exacerbated as the spatial resolution is degraded, and our results indicate that a fine grid spacing of ~4 km or less may be needed to simulate a realistic diurnal cycle of CO₂ for sites on top of the steep mountains examined here in the American Rockies. In the absence of higher resolution models, we recommend coarse-scale models to focus on assimilating afternoon CO₂ observations on mountaintop sites over the continent to avoid misrepresentations of nocturnal transport and influence.

1. Introduction

Scientific consensus among climate scientists points to carbon dioxide (CO₂) as the main greenhouse gas leading to climate change (IPCC, 2014). Therefore, a strong need exists to quantify and understand global carbon fluxes, among which the terrestrial biospheric component is the most dynamic, potentially even reversing signs on an annual basis from year to year (Le Quéré et al., 2015; Sarmiento et al., 2010). Yet quantifying and predicting terrestrial biospheric carbon fluxes continue to pose a challenge to researchers, as seen in the large divergence between models in projections of biospheric fluxes into the future (Cox et al., 2000; Friedlingstein et al., 2003; Arora et al., 2013) as well as in hindcast mode, particularly at the regional scale (Sarmiento et al., 2010; Stephens et al., 2007; Fisher et al., 2014).

Because mountains cover approximately a quarter of the Earth's land surface (Blyth et al., 2002), it is imperative to quantify and understand carbon fluxes over "complex terrain". Case in point is the Western U.S., where significant amounts of biomass are found above 1000 m elevation (Fig. 1). Similarly, much of the biomass and potential for terrestrial carbon storage in other parts of the world are found in hills or mountains, partly due to the fact that historical deforestation and biomass removal have been most pronounced in easier-to-access, flat regions (Ramankutty and Foley, 1999).

Despite the importance of regions with complex terrain in regional to global carbon cycling, these areas have hitherto been under-sampled due to logistical difficulties, harsh environmental settings, and violation of flat terrain assumptions in eddy covariance. Recently, Rotach et al. (2014) argued that current difficulties to balance the terrestrial carbon budget are due to inabilities to handle atmospheric circulations in complex terrain. While these authors presented a strong case for the consideration of flows over complex terrain, they did not quantify the implications of neglecting such flows for interpreting CO₂ observations. However, the significance of complex terrain has led to efforts to start closing this gap, in regions such as Europe (Pillai et al., 2011) and the American Rockies ((Schimel et al., 2002); see below).

The American Rockies will be the focus region of this study, which attempts to show how CO₂ concentrations in mountain regions can be properly linked, through atmospheric transport, to biospheric fluxes. While the objective of this paper is to use the American Rockies as a case study to illustrate general aspects of interpreting CO₂ observations in mountainous regions, several other compelling reasons exist for studying this region. Both models and observations have suggested that significant carbon storage can occur in the American Rockies (Fig. 1) (Schimel et al., 2002; Monson et al., 2002; Wharton et al., 2012), albeit this storage is highly sensitive to environmental drivers such as temperature and water availability (Monson et al., 2006; Schwalm et al., 2012; Wharton et al., 2012; Potter et al., 2013) as well as disturbances such as insect infestation (Negron and Popp, 2004) and wildfires (Wiedinmyer and Neff, 2007). These disturbances are also coinciding with rapid population increases in this region (Lang et al., 2008), with concomitant rise in urban CO₂ emissions (Mitchell et al., In Review), urban-wildland interfaces (Mell et al., 2010), and demands for water resources (Reisner, 1993; Gollehon and Quinby, 2000).

Recently, several research efforts have attempted to improve the understanding of carbon fluxes in the American Rockies. Direct eddy covariance-based measurements of carbon fluxes have been carried out in the mountains (Blanken et al., 2009; Yi et al., 2008); however, the eddy covariance technique characterizes fluxes only over a small area of ~1 km² (Baldocchi et al., 2001) and requires careful attention to potential biases from local advection. Ground-based ecological measurements (Anderegg et al., 2012; Tkacz et al., 2008) yield detailed information regarding the ecosystem, but such observations are also limited in spatial coverage and temporal resolution. Atmospheric CO₂ observations can characterize fluxes over hundreds of km (Gerbig et al., 2009), providing important regional scale constraints. Aircraft-based CO₂ measurements in this region have had some success in characterizing regional scale fluxes (Desai et al., 2011), albeit on a sporadic, campaign-based setting. More significantly, a network of accurate CO₂ observations has been maintained on mountaintops in the Rockies for the past decade (Stephens et al., 2011). These observations have been assimilated by

Deleted: hills and

Deleted: almost

Deleted: 70%

Deleted:

Deleted: (Rotach et al., 2008)

Moved down [2]: van der Molen et al. (2007) simulated CO₂ variability near a Siberian observational site and showed that even modest terrain variations of ~500 m over 200 km could lead to considerable CO₂ gradients.

Moved down [3]: Recently, Rotach et al. (2014) argued that current difficulties to balance the terrestrial carbon budget are due to inabilities to handle complex terrain. While these authors presented a strong case for the consideration of flows over complex terrain, they did not quantify the implications of neglecting such flows for interpreting CO₂ observations.

Deleted: -

Moved (insertion) [1]

Deleted: However, the significance of complex terrain has led to efforts to start closing this gap, in regions such as Europe (Pillai et al., 2011) and the American Rockies ((Schimel et al., 2002); see below).

Moved (insertion) [3]

Deleted: -

Deleted: -

sophisticated global carbon data assimilation systems such as “CarbonTracker” (Peters et al., 2007) to retrieve biospheric carbon fluxes over the mountainous regions and the rest of the globe.

Due to the expanding number of CO₂ observations in mountainous areas and the need to understand carbon fluxes in such regions, a strong motivation exists to evaluate existing methods in which CO₂ observations are used in atmospheric models to retrieve carbon fluxes. van der Molen et al. (2007) simulated CO₂ variability near a Siberian observational site and showed that even modest terrain variations of ~500 m over 200 km could lead to considerable CO₂ gradients. Due to the dangers of mis-representing terrain/flows and introducing biases into the carbon inversion system, mountaintop CO₂ observations have often been omitted from carbon inversion systems (Rodenberg, 2005; Geels et al., 2007; Peters et al., 2010). In fact, the most recent release of CarbonTracker (“CT-2015”) stopped assimilating the three RACCOON sites (<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/>).

However, the absence of mountaintop CO₂ observations to constrain carbon inversion systems is, in effect, throwing away valuable information that could inform carbon exchange in potentially important areas of the world (Fig. 1). Case in point is the Schauinsland CO₂ time series on a mountain in the middle of Western Europe, which as of this writing has collected over 40 years of continuous CO₂ data (Schmidt et al., 2003) but remains excluded from numerous carbon inversion systems (Rodenberg, 2005; Geels et al., 2007; Peters et al., 2010).

As a means to evaluate models’ ability to interpret mountaintop CO₂ observations, we specifically adopt the observed diurnal cycle during the summer growing season as a key diagnostic. This is because the diurnal cycle during the growing season, with nighttime respiratory release and daytime photosynthetic drawdown of CO₂, is a prominent feature in the coupling between biospheric fluxes and the atmosphere and one of the dominant modes in the CO₂ time series (Bakwin et al., 1998; Denning et al., 1996). Furthermore, models tend to either use CO₂ data from the nighttime (Keeling et al., 1976) (to sample subsiding air in the mid-troposphere) or from the daytime (during well-mixed conditions), and aspects of the diurnal cycle can provide clues as to whether the model is capturing the link between fluxes and concentrations right at either, both, or neither of these times.

The diurnal pattern of CO₂ observed at the Storm Peak Laboratory, Colorado, was examined by one of the first mesoscale modeling studies that investigated the impact of mountain flows on CO₂ concentrations (De Wekker et al., 2009). Although this study adopted an idealized simulation covering only a single day of observations, it nonetheless underscored the role of daytime upslope winds. A common approach is to assimilate mountain observations at night (Peters et al., 2007), favoring subsidence conditions characterizing free tropospheric concentrations and avoiding the need to resolve daytime upslope flows (Keeling et al., 1976).

Recently, Brooks et al. (2016) used pseudo-observations to examine the detectability of a regional flux anomaly by three mountaintop CO₂ sites in the American Rockies (including Storm Peak Laboratory). For the atmospheric model they adopted a time-reversed Lagrangian particle dispersion model (LPDM), which yields the “footprint”, or source region, of the observation sites (Lin et al., 2012). Although this study investigated whether the three mountaintop sites could detect signals of ecosystem

Moved (insertion) [2]

Deleted: .

Deleted: .

... [1]

Deleted: an obvious way to deal with this problem is to neglect the

Formatted: Subscript

Deleted: data

Deleted: altogether. This is already commonly practiced within

Formatted: Subscript

Deleted: w

Deleted:

Deleted: to evaluate models

disturbance, Brooks et al. (2016) did not specifically examine issues related to erroneous atmospheric transport in complex terrain nor compare modeled CO₂ against observed values.

In this paper, we will focus on the same 3 mountaintop CO₂ sites in the American Rockies and specifically examine the implications of using nocturnal versus daytime data within models, in light of atmospheric models at various grid spacings—from high resolution regional simulations to coarser global scale simulations. More specifically, we will drive a time-reversed LPDM with various meteorological fields and receptor heights. We will probe the implications on the footprint, transport, and the resulting CO₂ concentrations as the driving meteorological fields are degraded with coarser grid spacing and also as different vertical levels within the model are used.

The guiding questions of this paper are, as follows:

1. How do atmospheric flows in mountainous areas affect CO₂ concentrations and their representation in models?
2. What are the errors incurred due to the use of coarse-scale atmospheric simulations?
3. How can mountaintop CO₂ observations be used in an effective manner to constrain regional carbon fluxes in complex terrain?

2. Methodology

2.1 RACCOON Observations

The Regional Atmospheric Continuous CO₂ Network (RACCOON, <http://raccoon.ucar.edu>) was established in 2005 and has collected in situ CO₂ measurements at up to six sites over the past decade (Stephens et al., 2011). Here we present and simulate observations from the three longest running high-alpine sites (Fig. 2; Table 1). The easternmost site (NWR) is at 3,523 m elevation near the treeline on Niwot Ridge, just west of Ward, CO. Niwot Ridge is a LTER site and there is an AmeriFlux tower run by the University of Colorado 3 miles east and 500 m lower on the ridge. The instrumentation reside in the “T-Van” where the U.S. National Oceanic and Atmospheric Administration (NOAA)’s Global Monitoring Division has collected weekly flask samples for measurement of CO₂, isotopes, and other species for over 40 years, and daily flasks since 2006. The middle site (SPL) is at the Desert Research Institute’s Storm Peak Lab (3,210 m on Mt. Werner near Steamboat Springs, CO). This mountaintop observatory has a long history of measurements related to cloud physics, cloud-aerosol chemistry, and air quality. The westernmost site (HDP) is on Hidden Peak (3,351 m, above the Snowbird ski resort, Utah). This mountaintop site generally experiences regionally well-mixed or free-tropospheric air, but with influences from Salt Lake City during boundary-layer growth and venting periods.

The RACCOON measurements are based on a LiCor LI-820 single-cell IRGA with frequent calibrations. The instruments sample air from one of three inlet lines on a tower (two at HDP) and use a suite of four calibration gases plus a fifth surveillance gas. All reference gases are rigorously tied to the WMO CO₂ Calibration Scale with use of the NCAR CO₂ and O₂ Calibration Facility. 100-second average measurement precision is ± 0.1 ppm (1 σ), and intercomparability is estimated from several methods to be 0.2 ppm (Stephens et al., 2011).

We applied filtering to the mountaintop CO₂ observations to remove local influences and to extract values that are more regionally representative (Brooks et al., 2012). Observations were filtered out in which the within-hour standard deviation is greater than 1.0 ppm or when the differences between the top two inlets are greater than 0.5 ppm, which indicate periods when significant influences that are highly localized to the site are affecting the observations. This filtering removed 15%, 16%, and 27% of the hourly observations at HDP, SPL, and NWR, respectively. Regardless, filtering made negligible differences in the observed diurnal cycles in CO₂ (see Supplement). Henceforth, we will refer to the filtered observations when discussing the observed CO₂.

2.2 WRF-STILT Atmospheric Model

The atmospheric modeling framework adopted in this study is a Lagrangian time-reversed particle dispersion model, the Stochastic Time-Inverted Lagrangian Transport (STILT) model (Lin et al., 2003), driven by a mesoscale gridded model, the Weather Research and Forecasting (WRF) model (Skamarock and Klemp, 2008). STILT is a Lagrangian model that simulates the effects of turbulent dispersion using the stochastic motions of air parcels. It has been widely applied to the interpretation of CO₂ and trace gases in general (Lin et al., 2004; Hurst et al., 2006; Gockede et al., 2010; Kim et al., 2013; Mallia et al., 2015; Jeong et al., 2012). WRF is a state-of-the-art non-hydrostatic mesoscale atmospheric model that can simulate a variety of meteorological phenomena (Skamarock and Klemp, 2008), gaining widespread acceptance and usage among the atmospheric science community. Careful coupling between WRF and STILT has been carried out, with an emphasis towards physical consistency and mass conservation (Nehrkorn et al., 2010).

For this study, we ran WRF in a two-way nested mode centered between Utah and Colorado where the RACCOON sites are located (Fig. 2). The grid spacing was refined in factors of 3, from 12 km grid spacing covering the entire Western U.S. to 4 km and then to 1.3 km in the innermost domain that covers all of the RACCOON sites. 41 vertical levels were adopted, with 10 of these levels within 1 km of the ground surface, following Mallia et al. (2015). Comprehensive testing of different WRF settings have been carried out as part of a previous publication (Mallia et al., 2015) and adopted here: i.e., the MYJ, Grell-Devenyi Ensemble, and Purdue Lin schemes for parameterizing the planetary boundary layer (PBL), cumulus convection, and microphysics, respectively. In addition to the testing reported in Mallia et al. (2015), we have also carried out evaluation of the WRF fields specifically using meteorological measurements on mountaintops, near the RACCOON sites. These evaluations reveal that errors in the simulated meteorological fields are reasonable when compared against other atmospheric simulations evaluated in less complex terrain (Mallia et al., 2015), and biases are especially small for the WRF-1.3km fields (Table 2). In this paper, we will examine the resulting differences in meteorological and CO₂ simulations when STILT is driven by WRF fields at three different grid spacings.

In addition to the three WRF domains, we drove STILT with a fourth meteorological field, from NCEP's Global Data Assimilation System (GDAS). GDAS is archived at 1°×1° grid spacing, at 6 hourly intervals and at 23 vertical pressure levels. Driving STILT with GDAS was a means by which we attempted to construct an atmospheric model to resemble the NOAA CarbonTracker product, which was also at

Formatted: Subscript

Deleted: observations

Deleted: —e.g., convection, microphysical, and planetary boundary layer (PBL) schemes—

Deleted: ,

Deleted: we refer the reader to that paper for details regarding the WRF configuration

Deleted: these settings were

1°×1° resolution (and 25 vertical levels) over North America. More details about CarbonTracker can be found in the next section.

Driven by the various meteorological fields, STILT released 2000 air parcels every 3 hours (00, 03, 06, ...21 UTC) for the months of June, July, and August 2012 from the RACCOON sites and transported for 3 days backward in time. An example of STILT-simulated air parcel trajectories can be found in Fig. S1. The choice of 2000 parcels followed from results from sensitivity tests in a previous study, also over the Western U.S. (Mallia et al., 2015). In the case of WRF, STILT has the capability to transport the parcels in a nested fashion. So when we refer to “WRF 1.3km simulations”, it actually means that the atmosphere in the innermost domain (Fig. 2) was simulated at 1.3 km, switching to 4km grid spacing when the parcel left the 1.3km domain; likewise, the 12km winds were used when the parcel left the 4km domain. For the “WRF 4km simulations” we started with the 4 km fields as the innermost domain, and then 12 km in the outer domain.

For each site, we released STILT parcels using two different ways to determine starting levels. When we refer to “AGL”, we mean that the starting height was set at the level of the inlet above the ground surface (Table 1), following the local terrain as resolved in the meteorological model (whether at 1.3-, 4-, 12-km, or 1° grid spacing). The alternative method, referred to as “ASL”, means that the starting level was set to the elevation above sea level. For instance, the HDP site is located at 3351 m above sea level. The ground height as resolved by the 12km WRF model is at 2357 m, so the starting height was placed at 994 (=3351 – 2357) m above the resolved terrain. CarbonTracker, as well as many other global-scale models (Geels et al., 2007; Peters et al., 2010) places the observation site at an internal model level following the ASL method, so the “GDAS-ASL” runs were a means by which we attempted to mimic the global model configuration and to illuminate potential errors that could result from such a configuration. We also tested the AGL height for GDAS, at HDP only. As shown later, these runs were highly erroneous, so we did not carry them out for the other two sites.

The STILT-simulated air parcels were tracked as they were transported backwards in time from the RACCOON receptors (see example in Fig. S1); when they were in the lower part of the PBL, the locations of the parcels and amount of time the parcels spend in the lower PBL were tallied. This information was used in calculation of the “footprint”—i.e., the sensitivity of the receptor to upwind source regions (in units of concentration per unit flux). For more details, see Lin et al. (2003). The footprints, encapsulating the atmospheric transport information, were then combined with gridded fluxes from the biosphere and anthropogenic emissions, which are described in the next sections.

2.3 CarbonTracker CO₂ Concentrations and Biospheric Fluxes

CarbonTracker is a carbon data assimilation system covering the whole globe that retrieves both oceanic and terrestrial biospheric carbon fluxes (Peters et al., 2007). Observed atmospheric CO₂ concentrations are assimilated by CarbonTracker, which adjusts carbon fluxes to minimize differences with the observed CO₂ using an ensemble Kalman filter methodology.

We took three-dimensional CO₂ fields from CarbonTracker to initialize CO₂ concentrations at the end of the 3-day back trajectories from STILT. CarbonTracker-derived biospheric fluxes, along with anthropogenic and fire emissions (Sect. 2.4), were

also multiplied with STILT-derived footprints and combined with the initial CO₂ concentrations to yield simulated CO₂ at the RACCOON receptors.

CarbonTracker is maintained and continues to be developed by the NOAA's Earth System Research Laboratory. For this paper, we adopt the "CT-2013b" version. CT-2013b provides multiple prior estimates of the oceanic, terrestrial, and fossil carbon fluxes, with each combination yielding separate posterior fields of carbon fluxes and CO₂ distributions. CT-2013b results are presented as an average across the suite of prior fluxes and CO₂ fields.

CT-2013b resolves atmospheric transport and fluxes at 1°×1° over North America and 3°-lon × 2°-lat in the rest of the globe, with 25 vertical levels. The driving meteorological fields come from the European Centre for Medium-Range Weather Forecasts's ERA-interim reanalysis. The ensemble Kalman filter system within CarbonTracker solves for scaling factors on weekly timescales to adjust upward or downward biospheric carbon fluxes. The adjustments were made over "ecoregions" on land, rather than attempting to adjust fluxes within individual gridcells, as way to reduce the dimensions of the inversion problem within CarbonTracker. More details regarding the CarbonTracker system can be found in Peters et al. (2007) and on-line at <http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2013B/>.

Since CarbonTracker was designed for global carbon cycle analyses to retrieve large-scale fluxes, the adjustment to biospheric carbon fluxes could result in artifacts at the local to regional scales. More specifically, the attempt to match CO₂ observations with a single scalar can result in flipped diurnal cycles, causing carbon uptake during the night that is partly offset by enhanced respiration in a nearby ecoregion (Fig. S2). For this paper, we implemented a fix that removed this artifact by detecting these reversed diurnal patterns, adjusting them while preserving the 24-hour integrated carbon flux. See the Supplement and Fig. S3 for details.

2.4 Anthropogenic and Fire Emissions

Anthropogenic CO₂ emissions were obtained from the Emission Database for Global Atmospheric Research (EDGAR) (European Commission, 2009), which resolves emissions globally at 0.1°×0.1° annually. In order to temporally downscale the annual emissions, hourly scaling factors were obtained from the Vulcan emission inventory (Gurney et al., 2009) and applied to the EDGAR annual emissions. Lastly, CO₂ emissions from EDGAR for Year 2010 were extrapolated to 2012 using population growth rates across the U.S. since 2010, as this was the last year in which EDGAR emissions were available.

Wildfire emissions for CO₂ were obtained from the Wildland Fire Emissions Inventory (WFEI) (Urbanski et al., 2011). Since these emissions were only reported daily, three-hourly diurnal scaling factors were obtained from Global Fire Emissions Database v3.1 and applied to the daily WFEI emissions to downscale the emissions to sub-daily timescales (Mu et al., 2011; van der Werf et al., 2010).

Contributions from anthropogenic and wildfire emissions, on average, to the mean CO₂ diurnal cycle observed at all the mountain sites were secondary in comparison to the biosphere (Fig. S4). In particular, the wildfire contributions were episodic and averaged out to negligible contributions over Jun~Aug 2012 (Fig. S4). Because of this, we will not touch upon wildfires in the remainder of the paper.

Deleted: (

Deleted:)

3. Results

3.1 Observed versus Simulated Diurnal Cycle

The observed and simulated diurnal cycles of CO₂ for the three selected RACCOON sites are shown in Fig. 3. [These diurnal patterns were calculated from averaging the 3-hourly simulated time series from different model setups, which exhibit significant variability at multi-day synoptic timescales and correlations with different meteorological variables between HDP, SPL, and NWR \(Fig. S5, Table S1\). Due to this complexity we are focussing the analysis on the average diurnal pattern.](#)

The observed diurnal cycle exhibits an amplitude of ~2 ppm, on average, with more elevated concentrations at night and depleted values during the day. In contrast to the observed diurnal cycles, the simulated CO₂ extracted from [the site's altitude within CarbonTracker's output \(Table 1\)](#) exhibits a different cycle. Instead of peaking at night, CO₂ in CarbonTracker reaches its maximum during the afternoon at HDP. At SPL and NWR, the diurnal cycle is significantly attenuated, with nighttime values barely elevated over the background instead of the nighttime enhancement in the observed values.

It appears that the erroneous diurnal pattern at HDP within CarbonTracker can partly be due to the diurnal reversal in the original biospheric fluxes, which showed strong uptake of CO₂ even at night for the gridcell where HDP is located (Fig. S2). This resulted in erroneous diurnal patterns at all of the lowest 8 levels of CarbonTracker (Fig. S6), with the bottom 2 levels exhibiting strong depletions in CO₂ at night and enhancements during the day, pointing to unrealistic nighttime uptake and daytime release.

However, the diurnal reversal in biospheric fluxes alone does not completely explain the erroneous diurnal pattern. Differences in the diurnal pattern between GDAS-ASL simulations after introducing the diurnal fix in biospheric fluxes were not as pronounced at SPL and NWR.

The GDAS-ASL simulations show a pronounced peak of CO₂ in the morning that is missing from observations at all three sites (Fig. 3). We will discuss this feature, also seen in other coarse-scale simulations of mountaintop CO₂ (Geels et al., 2007), in Sect. 3.2 below.

In contrast to GDAS-ASL and CarbonTracker, the WRF-driven simulations better reproduce the shape of the observed diurnal cycle (Fig. 3), with nighttime enhancements and daytime depletions of CO₂. Considerable differences in nocturnal CO₂ concentrations are found, however, in the WRF-STILT runs at various grid spacings. WRF-12km significantly overestimates CO₂ at night, while WRF-1.3km and -4km produced similar CO₂ concentrations that correspond much more closely to observed values. While GDAS simulations started near the ground ("GDAS-AGL") also exhibit nighttime enhancements and daytime depletions of CO₂, the nighttime values are grossly estimated, exceeding even the values in WRF-12km. Therefore, we do not present GDAS using the AGL configurations at the other two sites.

Part of the error in all the simulations against the observations could arise from errors in the CarbonTracker boundary condition imposed at the end of the STILT back trajectories. Evaluations of CT-2013b against aircraft vertical profiles (which were not assimilated into CarbonTracker) at the Trinidad Head and Estevan Point sites on the West Coast of the North American continent carried out by the CarbonTracker team (<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2013B/profiles.php>) indicate that

Deleted: .

Deleted:

Deleted: 5

CT-2013b overestimates CO₂ concentrations by at most 1.0 ppm, on average, during the summer season. Thus, the fact that GDAS and CarbonTracker underestimate CO₂ at night likely cannot be attributed solely to a biased boundary condition.

3.2 Differences in Simulated Transport to Mountaintop Sites

Deleted: .

3.2.1 Footprint Patterns

In order to isolate the impact of differences in atmospheric transport on the simulated CO₂, we examine the average diurnal pattern of the footprint strength over Jun~Aug 2012 (Fig. 4). At each hour of the day we summed the spatially explicit map of the average footprint that marks out the source region of each RACCOON site—shown in Figs. 5~6 for HDP and in the Supplemental Information for the other 2 sites. The result shows the diurnal pattern of the sensitivity of the receptor concentration to upwind fluxes.

To a large extent, the diurnal variation in footprint strength mirrors the simulated CO₂ concentrations. Nocturnal enhancements in the footprints are seen in the WRF-driven simulations, with the WRF-12km exhibiting the strongest nocturnal footprints. Footprints from WRF-1.3km and WRF-4km are weaker at night than from WRF-12km and closely resemble each other. GDAS-AGL footprints (only shown at HDP) are the highest among all models at night, leading to the drastic overestimation in CO₂ in Fig. 3. In contrast, GDAS-ASL footprints exhibit a peak in the morning and are generally smaller in value than their WRF counterparts at other times of the day at HDP and NWR. At SPL, the GDAS-ASL footprint strengths are stronger and more in line with values from the other models.

Footprints are weaker during the daytime, and in contrast to the nighttime, differences between footprint strengths simulated by different models are significantly smaller. In particular, the differences are minimized in the afternoon.

These patterns are also seen in the footprint maps. We further examine differences in the spatial patterns of average footprints produced from the various WRF and GDAS configurations. The spatial patterns are contrasted at two different times of the day, associated with the nighttime and afternoon hours: 0200MST (0900UTC) and 1400 MST (2100UTC), respectively. Only HDP is shown for these two hours of the day in Figs. 5 and 6; similar figures for SPL and NWR can be found in the Supplementary Information (Figs. S7~S10). The footprint maps show marked differences at night (Fig. 5): the WRF-12km footprints are clearly stronger than their counterparts from the other 3 model configurations, with higher values covering the Wasatch Range near the HDP site. Meanwhile, the GDAS-ASL footprint at 0200 MST shows a striking contrast, with very low values around HDP and the Wasatch area in general.

Deleted: 6

Deleted: 9

The afternoon footprints at 1400MST (Fig. 6) display much more similarity with each other. Not only do the spatial patterns between the WRF and GDAS runs resemble one another; the significant differences in footprint strengths, with overestimation by WRF-12km and underestimation by GDAS-ASL, are no longer found. The aforementioned nighttime divergence and afternoon correspondence between footprint patterns are repeated at the SPL and NWR sites (Figs. S7~S10).

Deleted: 6

Deleted: 9

To further understand the nighttime divergence between model configurations, we now examine the average air parcel trajectories within Figs 5 and 6. It is worth noting that these trajectories differ from conventional mean wind trajectories that do not incorporate effects from turbulent dispersion (Lin, 2012). Instead, these mean trajectories

are determined by averaging the 2000 stochastic air parcel trajectories from STILT used for simulating transport arriving at a specific hour at a particular site, and then averaging over the ~90 days spanning June~August 2012. Thus there are ~180,000 stochastic trajectories averaged into generating the mean trajectory, thereby incorporating the net effect of turbulence on atmospheric transport. An example showing a subset of stochastic air parcels giving rise to the average trajectory is given in Fig. S1 for NWR, for 1400 MST.

Similar to the footprints, average trajectories differ much more at night than in the afternoon. Differences in average air parcel trajectories and the underlying resolved mountainous terrain are further examined in the next section.

3.2.2 Three-dimensional Terrain and Trajectories

The 3D terrain plots in Figs. 7, 9, and 11 illustrate the degradation in terrain resolved by coarser grid spacings and the resulting differences in average STILT-derived stochastic air parcel trajectories started at night (0900UTC) from the three sites. The afternoon (2100UTC) plots are shown in the Supplementary Information (Figs. S1~S13). The PBL heights, which determine whether air parcels are affected by surface fluxes (and lead to nonzero footprint values) are also plotted as blue lines in the same plots. Note that the apparent intersection of the PBL height with the ground in Figs. 7 and 9 is an artifact from averaging of multiple PBL heights along stochastic trajectories (Fig. S1).

Despite terrain smoothing compared against WRF-1.3km, WRF-4km produced STILT trajectories that are very similar to those from WRF-1.3km at all three sites, suggesting that salient features of the mountain flows resolved with 1.3km spacing are also found in the 4km spacing. In contrast, WRF-12km and GDAS-ASL both differed significantly from the more finely-gridded WRF simulations. Not only did the trajectories deviate from the higher resolution counterparts; the relationships between the trajectory vis-à-vis the PBL height, critical for determining footprints and simulating CO₂ changes (Sect. 2.2), also differ. The WRF-12km trajectories spend more time within the PBL, while GDAS-ASL trajectories are found much less within the PBL, because they start at a greater height above ground level.

An alternative perspective is to view the trajectory and PBL heights relative to the ground surface ("AGL") instead of above sea level, at each time step backward in time from the receptor (Figs. 8, 10, 12). These figures highlight the fact that while PBL dynamics in the three WRF configurations are similar, the heights of the trajectories relative to the PBL height differ. The trajectory exits above the nocturnal PBL one hour backward in time, on average, while the WRF-12km trajectory spends several hours within the PBL.

The difference in the trajectory behavior can be explained by the differing terrain. In mountainous terrain, PBL heights generally follow the terrain elevations, albeit with attenuated amplitude (Steyn et al., 2013). Thus in WRF-1.3km and 4km, the more highly resolved terrain produced shallow nocturnal PBL height that descend in the valley (Fig. 7) while the corresponding trajectory hovers above it. Viewed relative to the ground surface (Fig. 8), the trajectory originating from HDP appears to have exited above the nocturnal PBL in WRF-1.3km and 4km. In contrast, due to the significantly "flattened" mountains in WRF-12km and in GDAS, the PBL heights exhibit less spatial variation near the mountaintop receptor, since the terrain itself was smoothed. Consequently, WRF-12km trajectories, unlike the WRF-1.3km or -4km cases, travel closer to the ground surface,

Deleted: 0

Deleted: 2

within the nighttime PBL, even as it is advected away from the three RACCOON sites (Figs. 7, 8). This resulted in stronger nighttime footprints in WRF-12km as seen in Figs. 4 and 5. Another effect of the proximity of the air parcels to the model's ground surface is the slower windspeeds from surface drag, causing the air parcel trajectories to remain close to the 3 sites until the previous day; for HDP and SPL, the mean trajectories spiral toward the site at the surface, following an "Ekman wind spiral" pattern (Holton, 1992). In WRF-1.3km and WRF-4km, the measurement sites are at significantly higher elevations above the resolved valleys in the area surrounding the sites, and the air parcels are found above the shallow nocturnal boundary layer hugging the valley floor, on average (Fig. 7).

Deleted: or

Although both WRF-12km and GDAS poorly resolve the mountains, a key difference in the case of GDAS-ASL is that the air parcels were released at a site's elevation above sea level (following what is generally done in CarbonTracker, and other global models), much higher above ground than the release used in WRF-12km, which was selected to be the height in AGL above the flattened mountain. Therefore, the GDAS-ASL trajectories were significantly higher than the PBL height in the model (particularly at HDP and NWR), which followed the flattened ground surface in the $1^\circ \times 1^\circ$ grid spacing. Another noticeable difference in GDAS-ASL trajectory was the significantly higher daytime PBL heights (Figs. 8, 10, 12). We suspect this is because of the greatly reduced vertical resolution within GDAS (23 levels versus 41 levels in WRF): since STILT diagnoses the PBL height to correspond to a model level, a higher PBL height was chosen for GDAS because of the thicker vertical level. Another subtle artifact of the coarse resolution within GDAS can be seen in the anomalously low daytime PBL height just in the vicinity of HDP (Figs. 13, S11). It appears that the GDAS model set an entire $1^\circ \times 1^\circ$ grid box near HDP to be water body (the Great Salt Lake), thereby suppressing the PBL height.

Deleted: 0

The three-dimensional plots can explain the higher nighttime footprint strengths at SPL (Figs. 4, S7). This result appears to be a consequence of the relative elevation of the site and surrounding terrain. The elevation of the surrounding valley floor at SPL is closer to that of the mountaintop location of SPL (Fig. 9); therefore, air parcels released from SPL would have a stronger tendency to reside within the PBL even over the surrounding valleys, unlike the steeper dropoff--i.e., deeper valley--upwind of HDP (Fig. 7) and NWR (Fig. 11).

Deleted: 6

As already found in the footprints (Fig. 5), the afternoon (2100 UTC) differences in air parcel trajectories are much smaller (Figs. S11~S13). We suspect that this is due to the fact that the deeper daytime PBL height causes the trajectories to reside within the PBL, and stronger mixing within the daytime PBL minimize the relative terrain differences. A previous modeling study focusing on the SPL area has also suggested the daytime afternoon PBL depth to extend above the mountaintop (De Wekker et al., 2009), indicating that differences between terrain resolution and the resulting flows could be reduced due to the strong mixing taking place within the deep afternoon PBL. Consequently, simulations in the afternoon show much smaller divergence between various model configurations, resulting in similar footprint strengths and CO_2 values (Figs. 3 and 4). More evidence of the convergence in afternoon simulated CO_2 can be found in the small differences in CO_2 modeled at CarbonTracker's different levels during this time (Fig. S6).

Deleted: 0

Deleted: 2

Deleted: 5

A few studies have specifically focused on the flows and atmospheric transport around the NWR site. These authors have pointed to thermally driven flows, particularly downslope drainage flow events at night (Sun et al., 2007; Sun and De Wekker, 2011; Blanken et al., 2009). Daytime upslope events, while weaker, were also noted (Sun and De Wekker, 2011; Blanken et al., 2009; Parrish et al., 1990). It may seem that the 3D trajectories in Fig. 11 and Fig. S13 run counter to the presence of such thermally driven flows. We suspect that this is because the thermally driven flows induced by the terrain cannot be discerned in the mean trajectories, which also reflect the larger scale flows that can be stronger than the local scale thermally driven flows (Zardi and Whiteman, 2013). When one examines the stochastic trajectories from which the mean trajectories are based (Fig. S1), it is clear that some upslope trajectories can be detected.

We now examine the reason for the erroneous daytime peak in simulated CO₂ from GDAS-ASL that does not show up in the observations (Fig. 3). We specifically focus on this feature because the daytime peak was also found in other coarse-scale simulations of CO₂ for mountaintop sites--e.g., in Europe (Geels et al., 2007). Focusing on the three-dimensional plots at the hours of 0800 and 1100 MST (Fig. 13), when the simulated peaks are found at SPL and both NWR/HDP, respectively, the peaks coincide with times when average trajectories are found within a relatively shallow morning PBL. As the air parcels move backward in time, when the morning transitions backward in time to the nighttime, many of them would still be found within the shallow nighttime PBL. Due to the shallowness of the nocturnal PBL, the footprint values for the air parcels found there would be high. These parcels would also be sampling the nighttime CO₂ release and therefore lead to enhancements in CO₂. In other words, the erroneous daytime peak reflects enhanced CO₂ that is vented up to the observing height within the model during the day. We suspect that something similar is taking place in other global models, leading to similar erroneous daytime CO₂ peaks (Geels et al., 2007).

4. Discussion

This study has sought to answer the question: how can mountaintop CO₂ observations be used to constrain regional scale carbon fluxes, given the complex terrain and flows in the vicinity of mountaintop sites? To address this question, we have driven a Lagrangian particle dispersion model simulating the transport of turbulent air parcels arriving at 3 mountaintop CO₂ sites in the Western U.S. We then examined potential differences in simulated results as the atmospheric simulations are driven by meteorological fields resolved with differing grid spacings and at different vertical levels.

We found that the observed average diurnal CO₂ pattern is better reproduced by simulations driven by WRF-1.3km and WRF-4km ("AGL" configuration), with minimal differences between the two configurations (Fig. 3). The coarser-scale models (WRF-12km_AGL, GDAS-1°, and CarbonTracker) fail to reproduce the observed diurnal pattern at all 3 sites. The problem is especially severe at night, when both GDAS-ASL and CarbonTracker lack the nocturnal enhancements. In contrast, WRF-12km (AGL) shows nocturnal CO₂ buildup that is clearly too strong. The overestimation problem is exacerbated when both coarser grid spacing and "AGL" configuration are adopted, as seen in GDAS-AGL at HDP (Fig. 3).

The overestimate in nighttime CO₂ from WRF-12km (AGL) is due to the preponderance of simulated air parcels found within the nocturnal PBL (Figs. 7~9),

Deleted: 2

which can be traced to the fact that air parcels are closer to the ground surface when mountains are flattened. Conversely, when released at “ASL” levels air parcels are found much higher above the nocturnal PBL due to the flattening of mountains in a coarse-scale global model like GDAS, resulting in minimal sensitivity to nighttime biospheric fluxes and lack of CO₂ buildup. Such large errors in estimated carbon fluxes due to lack of ability to resolve patterns have also been found in earlier studies in Europe (Pillai et al., 2011; Peters et al., 2010).

The natural question, then, is what can researchers do with mountaintop CO₂ observations, given the difficulty in resolving the terrain and flows in complex terrain?

4.1 Approach 1: Adjust vertical level of simulations from which to compare against observed values

The diurnal cycle simulated within CarbonTracker varies significantly as a function of the vertical level (Fig. S6) from which CO₂ is extracted, particularly at night. The strongly attenuated diurnal cycle in the interpolated level corresponding to the ASL elevation of the mountaintop sites (orange dashed) is found at higher levels within CarbonTracker too, away from the first few levels near the ground. At HDP, the nighttime depletion of CO₂ at lower levels appears to be due to the erroneous nighttime photosynthetic uptake in the gridcell where HDP is located (Fig. S2).

Interestingly, at SPL and NWR the diurnal pattern at a level between Levels 2 and 3 appears to correspond more closely to the overall observed CO₂ diurnal cycle, perhaps due to the presence of nighttime enhancements closer to the model surface that is absent from the higher levels closer to the ASL elevation. The closer correspondence to observed patterns may call for researchers to adjust the vertical level to maximize resemblance to observations. This was carried out at Jungfraujoch (Folini et al., 2008), where the authors simulated carbon monoxide (CO) at multiple heights and arrived at a height of 80 m above the model’s ground surface as the best correspondence with the observed CO, which was measured closer to the ground (Rinsland et al., 2000). Instead, a different study simulating observations at the same site adopted a height of 830 m above the model ground surface (Tuzson et al., 2011). This example illustrates the divergence in researchers’ choices for the vertical level in the midst of mountainous terrain.

It is worth noting that the introduction of additional degrees of freedom in the vertical level in “fitting” the measured CO₂ diurnal cycle within a carbon assimilation system is potentially problematic. The reason is that the assimilation system seeks to solve for carbon fluxes by examining the mismatch between observed versus simulated CO₂ concentrations. If the mismatch is due to erroneous fluxes, the introduction of additional degrees of freedom in the vertical level would compensate for erroneous fluxes. For instance, if the nighttime carbon fluxes are overestimated in the model, this should show up as an enhanced CO₂ concentration that is larger than observed values. However, this overestimation in CO₂ would be reduced by picking a higher vertical level rather than fixing the overly large efflux in the model. The optimal level could differ between night and day as well; for instance, a level higher than Level 2 would fit better against observations during the daytime at SPL and NWR (Fig. S6). If different levels are adopted at different times of the day, the degrees of freedom that can be adjusted would be even larger, and model-data mismatches would be used in vertical level adjustments instead of correcting erroneous biospheric fluxes.

Deleted: 5

Deleted: 5

652 Regardless, there is some role for vertical level adjustments to remove the gross
 653 mismatch in the observed vs simulated diurnal cycles. If the vertical level is indeed
 654 adjusted in a carbon inversion system, we suggest that additional information (e.g.,
 655 comparisons to meteorological observations or other tracers) is used rather than
 656 maximizing the match to the target species (i.e., CO₂, in the case of a carbon inversion
 657 system).

658 The CO₂ values at multiple levels within CarbonTracker show that unlike the
 659 nighttime, differences between vertical levels are much smaller during the afternoon at
 660 SPL and NWR (Fig. S6), suggesting that the simulated CO₂ values are not as sensitive to
 661 the choice of vertical level. We suspect that the large differences between vertical levels
 662 at HDP are due to the flipped diurnal cycle in biospheric fluxes within CarbonTracker
 663 (Fig. S3). Otherwise, the lack of sensitivity to the choice of vertical level suggests that
 664 coarse-scale models should assimilate afternoon observations, rather than nighttime
 665 observations (see “Approach 3” below).

666 4.3 Approach 2: Assign errors to account for model errors

667 Instead of neglecting the mountaintop CO₂ observations altogether, an alternative
 668 approach is to make use of the observations, but assigning them errors within the model-
 669 measurement discrepancy error covariance matrix to account for model deficiencies (Lin
 670 and Gerbig, 2005; Gerbig et al., 2008). One estimate of the model-measurement
 671 discrepancy error is the root-mean-square error (RMSE), which ranges from less than 3
 672 ppm for WRF-1.3km to over 7 ppm for WRF-12km (Fig. S5). In this way, the inversion
 673 system would assign less weight to observations that the model has difficulties simulating.
 674 Given the systematic misrepresentation of the diurnal cycle in coarse-scale models,
 675 particularly at night (Fig. 3), this approach will effectively throw away much of the data
 676 as noise, due to inadequacies in the model. This naturally leads to the next possible
 677 approach of just having coarse-scale models assimilate afternoon observations.

678 4.4 Approach 3: Have coarse-scale models assimilate afternoon observations 679 instead of nighttime

680 Our results show that the simulated CO₂ values are more in accordance with observed
 681 values in the afternoon (Fig. 3). This follows from the fact that afternoon trajectories and
 682 footprints match their higher resolution counterparts (Figs. 6, S8, S10, S11~S13), likely
 683 due to the deeper afternoon PBL depth and the reduction of terrain effects (Steyn et al.,
 684 2013). In other words, relative differences in PBL depth associated with flattening of
 685 mountains are lessened when the PBL is deeper; thus the impact on whether an air parcel
 686 sampled by the mountaintop site falls within the PBL is also attenuated under afternoon
 687 vigorous mixing conditions.

688 Based on these results, and in lieu of better transport, we suggest coarse-scale
 689 models may be better served to assimilate afternoon observations over the continent at
 690 their above sea level elevation. This is contrary to what has been commonly practiced by
 691 researchers, when nighttime mountaintop observations were assimilated (Peters et al.,
 692 2007; Keeling et al., 1976) to avoid daytime upslope flows and when nocturnal
 693 observations that represent free tropospheric conditions would better match coarse
 694 resolution models. We have found that sampling coarse-scale (1 deg) models at the
 695 corresponding ASL height have significant difficulties simulating nighttime CO₂, since it
 696 appears that the model failed to represent the strength of the nocturnal footprint at the 3

Deleted: 5

Deleted: is

Moved up [1]: Due to the dangers of mis-representing terrain/flows and introducing biases into the carbon inversion system, an obvious way to deal with this problem is to neglect the mountaintop data altogether. This is already commonly practiced within carbon inversion systems (Rodenbeck, 2005; Geels et al., 2007; Peters et al., 2010). In fact, the most recent release of CarbonTracker (“CT-2015”) stopped assimilating the three RACCOON sites (<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/>). ... [2]

Deleted: 4

Deleted: .

Deleted: 4.2 Approach 2: Reject mountaintop data .

Deleted: .

Deleted: .

Deleted: 3

Deleted: .

Deleted: 4

Deleted: 7

Deleted: 9

Deleted: 10

Deleted: 12

RACCOON mountaintop sites (Figs. 4, 5). Thus the inability of coarse-scale models to simulate the transport and PBL depths result in the lack of nocturnal enhancements and thereby the wrong diurnal cycle (Fig. 3). Conversely, sampling the 12-km simulation at the AGL height also has significant difficulties simulating nighttime CO₂, because it overestimates the nocturnal footprint.

However, careful attention needs to be paid to upslope flows in the afternoon and the potential mis-interpretation of more localized biospheric signals or anthropogenic signals from below the mountain. A study from Jungfraujoch in Europe suggested that as much as ~40% of the days in a year are influenced by thermally driven flows (Griffiths et al., 2014). During the afternoon, the mountaintop site would then be influenced by thermally driven upslope winds, as also pointed out by a number of studies around NWR, along the Colorado Front Range (Sun et al., 2010; Sun and De Wekker, 2011; Parrish et al., 1990) as well as SPL (De Wekker et al., 2009). For sites like HDP and NWR, which have large nearby urban areas at lower elevation, upslope conditions can be of particular concern if not properly accounted for. If these sites experience elevated CO₂ in the afternoon from pollution sources, and this transport is not captured by the models, then natural CO₂ sources can be significantly overestimated.

We found it encouraging that despite the proximity of significant population and anthropogenic emissions from the Salt Lake and Denver area to the HDP and NWR sites, respectively, the WRF-1.3km model suggests that the additional contribution of anthropogenic CO₂ in the afternoon, over and beyond the nighttime signal is less than 1ppm, on average (Fig. S4). Presumably this is because of the high elevation of HDP and NWR in relation to the urban area and the dilution of signals as they move up slope; the afternoon urban signal would be enhanced if the sites were placed at lower peaks.

Regardless, it is prudent to consider mountaintop sites as not necessarily “pristine” sites and to consider potential contributions from surrounding anthropogenic emissions on these observations. It has been estimated that as of the year 2000, over 10% of the world population live in mountainous areas (Huddleston et al., 2003), meaning that any mountaintop site could very well see anthropogenic signatures. We recommend additional tracers to be measured in conjunction with the mountaintop CO₂ sites. For instance, combustion tracers such as C¹⁴ and CO (Levin and Karstens, 2007) have been measured alongside CO₂ at mountaintop sites in Europe. Another promising tracer is Rn²²² (Griffiths et al., 2014), which provides a measure of surface exchange and would help provide constraints on the exchange of air measured at the mountaintop with the surface. Co-located meteorological observations—whether in-situ or remotely-sensed (e.g., radar, sodar, lidar)—to probe atmospheric flows and turbulent mixing would also be of significant value in helping to interpret the tracer observations (Rotach et al., 2014; Banta et al., 2013).

4.5 Approach 4: Adopt high-resolution modeling frameworks

The least problematic, though potentially costly in terms of computational time, approach to reduce modeling errors when interpreting mountaintop CO₂ observations is to adopt a high resolution modeling framework. This conclusion was also arrived at by previous studies (Pillai et al., 2011; van der Molen and Dolman, 2007; De Wekker et al., 2009). From our results, it appears that meteorological fields from WRF at 4-km grid spacing, driving a Lagrangian particle dispersion model, can reproduce most features from a 1.3-km simulation, and generate a CO₂ diurnal cycle that qualitatively matches the observed

Deleted: 5

pattern. Once the WRF fields are degraded to 12-km grid spacing, the model fails to capture such features.

While at least 4-km resolution in the meteorological fields is needed for the sites examined here in the American Rockies, we anticipate that the minimum resolution would depend on the level of complexity in the terrain, the height of the observational site, and relationship with surrounding sources/sinks.

5. Conclusions

Given the large extent of the Earth's surface covered by hills and mountains and the large amount of biomass and potential for carbon storage in complex terrain (Fig. 1), we call for expanded efforts in observing and modeling CO₂ and other tracers on mountaintop sites. This study has illustrated the potential for even coarse-scale models to extract information from these observations when focusing on the daytime, afternoon values, and the ability of high resolution models to simulate the general features of the summertime diurnal CO₂ cycle even in the midst of significant terrain complexity. However, we acknowledge that even the highest resolution model adopted in this paper undoubtedly is subject to limitations of its own, and that deviations between simulated versus observed CO₂ diurnal cycles arise from errors in both atmospheric transport as well as the biospheric fluxes. Due to the focus on atmospheric transport in this paper, errors in the simulations caused by shortcomings in the biospheric fluxes remain outside the scope of this study (except for corrections to the flipped diurnal cycle; Fig. S3)

Even though current models remain imperfect, we call for sustained and expanded observations of CO₂ and other tracers (e.g., CO, ²²²Rn, and the isotopes of CO₂) co-located with meteorological observations on mountaintop sites to create enhanced datasets that can be further utilized by modeling frameworks of the future. Finally, we call for testing and gathering of three-dimensional CO₂ observations over complex terrain, as revealed by intensive airborne campaigns like the Airborne Carbon in the Mountains Experiment (Sun et al., 2010).

Acknowledgements

This study was supported by NOAA Climate Program Office's AC4 program, award # NA13OAR4310087 (UU). The RACCOON network has been supported by NSF (EAR-0321918), NOAA (NA09OAR4310064) and DOE (DE-SC0010624 and DE-SC0010625). The National Center for Atmospheric Research is sponsored by the National Science Foundation. We thank site collaborators at the RACCOON sites: Ian and Gannet Hallar at SPL, Dean Cardinale at HDP, and the CU Mountain Research Station and staff at NWR. We thank J. Knowles for the meteorological data at NWR and S. Urbanski for sharing the WFEI fire emission inventory. A. Andrews and A. Jacobson are gratefully acknowledged for helpful discussions.

Deleted:

Figure Captions

Fig. 1

Aboveground biomass [mega-tons of carbon] from the North American Carbon Program baseline dataset for year 2000 (Kelldorfer et al., 2013) overlaid on topographic surface in the Western U.S., resolved at $0.5^{\circ} \times 0.5^{\circ}$ grid spacing.

Fig. 2

The WRF simulation domain, covering the Western U.S. with a series of nests with 12-, 4-, and 1.3-km grid spacing. These WRF meteorological fields are used to drive air parcel trajectories within the STILT model.

Fig. 3

The average diurnal CO_2 pattern during June–August 2012 as observed at the 3 mountaintop sites in the RACCOON network: Hidden Peak (HDP), Storm Peak Laboratory (SPL), and Niwot Ridge (NWR). Compared against the observations are simulated diurnal CO_2 patterns from different models: CarbonTracker, STILT driven with WRF at different grid spacings, and STILT driven with GDAS. Multiple GDAS-driven STILT model configurations are shown, including runs without fixes to the biospheric fluxes (“biofluxorig”; see Supplemental Information), as well as releasing air parcels at the elevations of the sites above mean seal level (“ASL”) or, for HDP only, at the inlet height (Table 1) above the model’s ground level (“AGL”). All of the WRF-driven STILT runs place the release point of air parcels following the AGL configuration. Error bars denote standard errors of the diurnal averages.

Fig. 4

The average diurnal footprint strengths at HDP, SPL, and NWR over June–August 2012 from STILT, driven with different meteorological fields and release heights (ASL vs AGL). The footprint strength was derived by summing over the spatial distribution of footprint values (Fig. 5).

Fig. 5

The average footprint (shown in \log_{10}) for the Hidden Peak (HDP) site in Utah, at night: 0200 MST (0900 UTC), gridded at $0.1^{\circ} \times 0.1^{\circ}$. The site is denoted as a triangle. The average back trajectory (averaged over the stochastic STILT trajectories) is drawn as a line, with points indicating trajectory locations every hour, as the trajectory moves back from the site indicated as points. Magenta parts of the trajectory refer to the nighttime (1900–0700 MST), while pink portions indicate the daytime (0700–1900 MST). Parts of the trajectory are shaded with blue when it is found below the average height of the PBL along the trajectory.

Fig. 6

Similar to Fig. 5, but for the afternoon: 1400 MST (2100 UTC) at HDP.

Fig. 7

Three dimensional plots of the terrain over a domain of $\sim 1^{\circ} \times 1^{\circ}$ surrounding HDP, as resolved by the WRF and GDAS models at various grid spacings. The HDP site is denoted as a triangle. Also shown is the average back trajectory, derived by averaging

Deleted: Red

857 locations of the numerous stochastic trajectories simulated by STILT, driven by the
858 various WRF meteorological fields and the global GDAS field. Back trajectories were
859 started from HDP at 0200 MST (0900 UTC). Points indicate trajectory locations every
860 hour, as the trajectory moves back from the site indicated as points. Magenta portions of
861 the trajectory refer to the nighttime (1900–0700 MST), while pink portions indicate the
862 daytime (0700–1900 MST). In addition, the PBL heights averaged along the
863 backtrajectory are shown as the blue line.

Deleted: Red

864 **Fig. 8**

865 Time series of the average back trajectory and PBL heights relative to the ground surface
866 ("AGL") instead of above sea level, at each time step backward in time from the receptor
867 (triangle). Magenta portions of the trajectory refer to the nighttime (1900–0700 MST),
868 while pink portions indicate the daytime (0700–1900 MST). The PBL heights averaged
869 along the backtrajectory are shown as the blue line. The nighttime PBL height is
870 indicated in dark blue, while the daytime portion is in light blue. The height of the site is
871 indicated by the black triangle at the starting time of the back trajectory.

Deleted: Red

872 **Fig. 9**

873 Similar to Fig. 7, but for the Storm Peak Laboratory (SPL) site.

874 **Fig. 10**

875 Similar to Fig. 8, but for the Storm Peak Laboratory (SPL) site.

876 **Fig. 11**

877 Similar to Fig. 9, but for the Niwot Ridge (NWR) site.

878 **Fig. 12**

879 Similar to Fig. 10, but for the Niwot Ridge (NWR) site.

880 **Fig. 13**

881 Similar to three-dimensional terrain and trajectory plots as shown in Figs. 7, 9, and 11,
882 but for just the GDAS 1 deg. ASL simulations and for the morning hours of 0800 MST
883 and 1100 MST.

884

887 **Tables**

888

	Hidden Peak (HDP)	Storm Peak Lab (SPL)	Niwot Ridge (NWR)
Latitude/Longitude	40° 33' 38.80" N 111° 38' 43.48" W	40° 27' 00" N 106° 43' 48" W	40° 03' 11" N 105° 35' 11" W
Top Inlet Height	17.7 m	9.1 m	5.1 m
Site Altitude [m above sea level]	3351 m	3210 m	3523 m
Site Altitude as Resolved by Models [m above sea level]:			
<i>WRF-1.3km</i>	2996 m	3038 m	3411 m
<i>WRF-4km</i>	2918 m	2818 m	3382 m
<i>WRF-12km</i>	2357 m	2724 m	3076 m
<i>GDAS</i>	1856 m	2757 m	2333 m
<i>CarbonTracker</i>	2004 m	2582 m	2276 m

Deleted: Model

889 **Table 1.** Characteristics of RACCOON mountaintop sites examined in this paper, as
890 well as the representation of terrain in different meteorological files at these sites.
891
892

894

Site	SPL				NWR			
Run type	1.3-km WRF	4-km WRF	12-km WRF	GDAS	1.3-km WRF	4-km WRF	12-km WRF	GDAS
u-wind BIAS [m/s]	-0.5	-1.5	-0.9	2.3	0.1	-0.3	-1.4	-0.2
v-wind BIAS [m/s]	-0.6	-0.3	-0.2	1	0.2	0.4	0.9	1.1
u-wind RMSE [m/s]	3.1	3.8	3.2	3.7	3.5	3.4	3.4	3.2
v-wind RMSE [m/s]	2.7	2.7	2.3	2.5	2.2	2.1	2.2	3

895

896

897

898

899

900

901

902

Table 2. Comparisons of different meteorological files driving STILT against hourly-averaged wind observations at Storm Peak Laboratory (-106.74 W; 40.45 N) and at Niwot Ridge (-105.586 W; 40.053 N; 3502 m ASL) (Knowles, 2015), near the RACCOON CO₂ site. Meteorological observations were not available at the Hidden Peak site. Error statistics are presented separately for the west-to-east component ("u-wind") and south-to-north component ("v-wind") of the wind velocity vector.

References

- Anderegg, W. R. L., Berry, J. A., Smith, D. D., Sperry, J. S., Anderegg, L. D. L., and Field, C. B.: The roles of hydraulic and carbon stress in a widespread climate-induced forest die-off, *Proceedings of the National Academy of Sciences*, 109, 233-237, 10.1073/pnas.1107891109, 2012.
- Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon-Concentration and Carbon-Climate Feedbacks in CMIP5 Earth System Models, *Journal of Climate*, 26, 5289-5314, 10.1175/jcli-d-12-00494.1, 2013.
- Bakwin, P. S., Tans, P. P., Hurst, D. F., and Zhao, C.: Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program, *Tellus*, 50B, 401-415, 1998.
- Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor and energy flux densities, *Bulletin of the American Meteorological Society*, 82, 2415-2435, 2001.
- Banta, R. M., Shun, C. M., Law, D. C., Brown, W., Reinking, R. F., Hardesty, R. M., Senff, C. J., Brewer, W. A., Post, M. J., and Darby, L. S.: Observational Techniques: Sampling the Mountain Atmosphere, in: *Mountain Weather Research and Forecasting: Recent Progress and Current Challenges*, edited by: Chow, K. F., De Wekker, F. J. S., and Snyder, J. B., Springer Netherlands, Dordrecht, 409-530, 2013.
- Blanken, P. D., Williams, M. W., Burns, S. P., Monson, R. K., Knowles, J., Chowanski, K., and Ackerman, T.: A comparison of water and carbon dioxide exchange at a windy alpine tundra and subalpine forest site near Niwot Ridge, Colorado, *Biogeochemistry*, DOI:10.1007/s10533-10009-19325-10539, 2009.
- [Blyth, S., B. Groombridge, I. Lysenko, L. Miles, and A. Newton, Mountain Watch: Environmental Change and Sustainable Development in Mountains, UNEP World Conservation Monitoring Centre, 2002.](#)
- [Brooks, B.-G. J., Desai, A. R., Stephens, B. B., Bowling, D. R., Burns, S. P., Watt, A. S., Heck, S. L. and Sweeney, C.: Assessing filtering of mountaintop CO₂ mole fractions for application to inverse models of biosphere-atmosphere carbon exchange, *Atmos. Chem. Phys.*, 12, 2099–2115, 2012.](#)
- Brooks, B. G. J., Desai, A. R., Stephens, B. B., Michalak, A. M., and Zscheischler, J.: Feasibility for detection of ecosystem response to disturbance by atmospheric carbon dioxide, *Biogeosciences Discuss.*, 2016, 1-29, 10.5194/bg-2016-223, 2016.
- Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., and Totterdell, I. J.: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, 408, 184-187, 2000.
- De Wekker, S. F. J., Ameen, A., Song, G., Stephens, B. B., Hallar, A. G., and McCubbin, I. B.: A Preliminary Investigation of Boundary Layer Effects on Daytime

Deleted: .

949 Atmospheric CO₂ Concentrations at a Mountaintop Location in the Rocky
 950 Mountains, *Acta Geophysica*, 57, 904-922, 2009.
 951 Denning, A. S., Randall, D. A., Collatz, G. J., and Sellers, P. J.: Simulations of terrestrial
 952 carbon metabolism and atmospheric CO₂ in a general circulation model. Part 2:
 953 Simulated CO₂ concentrations, *Tellus*, 48B, 543-567, 1996.
 954 Desai, A. R., Moore, D. J. P., Ahue, W. K. M., Wilkes, P. T. V., De Wekker, S. F. J.,
 955 Brooks, B.-G., Campos, T., Stephens, B. B., Monson, R. K., Burns, S. P., Quaife,
 956 T., Aulenbach, S. M., and Schimel, D. S.: Seasonal pattern of regional carbon
 957 balance in the central Rocky Mountains from surface and airborne measurements,
 958 *Journal of Geophysical Research*, 116, doi:10.1029/2011JG001655, 2011.
 959 European Commission: Joint Research Centre/Netherlands Environmental Assessment
 960 Agency, Emission Database for Global Atmospheric Research (EDGAR), in,
 961 release version 4.0 ed., 2009.
 962 Fisher, J. B., Sikka, M., Oechel, W. C., Huntzinger, D. N., Melton, J. R., Koven, C. D.,
 963 Ahlström, A., Arain, M. A., Baker, I., Chen, J. M., Ciais, P., Davidson, C., Dietze,
 964 M., El-Masri, B., Hayes, D., Huntingford, C., Jain, A. K., Levy, P. E., Lomas, M.
 965 R., Poulter, B., Price, D., Sahoo, A. K., Schaefer, K., Tian, H., Tomelleri, E.,
 966 Verbeeck, H., Viovy, N., Wania, R., Zeng, N., and Miller, C. E.: Carbon cycle
 967 uncertainty in the Alaskan Arctic, *Biogeosciences*, 11, 4271-4288, 10.5194/bg-11-
 968 4271-2014, 2014.
 969 Folini, D., Uhl, S., and Kaufmann, P.: Lagrangian particle dispersion modeling for the
 970 high Alpine site Jungfraujoch, *Journal of Geophysical Research: Atmospheres*, 113,
 971 n/a-n/a, 10.1029/2007JD009558, 2008.
 972 Friedlingstein, P., Dufresne, J. L., Cox, P. M., and Rayner, P.: How positive is the
 973 feedback between climate change and the carbon cycle?, *Tellus*, 55B, 692-700,
 974 2003.
 975 Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R.,
 976 Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U.,
 977 Rödenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing
 978 atmospheric transport models for future regional inversions over Europe –
 979 Part 1: mapping the atmospheric CO₂ signals, *Atmos. Chem. Phys.*, 7, 3461-3479,
 980 10.5194/acp-7-3461-2007, 2007.
 981 Gerbig, C., Kerner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport
 982 models: error characterization and propagation, *Atmospheric Chemistry and*
 983 *Physics*, 8, 591-602, 2008.
 984 Gerbig, C., Dolman, A. J., and Heimann, M.: On observational and modelling strategies
 985 targeted at regional carbon exchange over continents, *Biogeosciences*, 6, 1949-1959,
 986 2009.
 987 Göckede, M., Michalak, A. M., Vickers, D., Turner, D. P., and Law, B. E.: Atmospheric
 988 inverse modeling to constrain regional - scale CO₂ budgets at high spatial and
 989 temporal resolution, *Journal of Geophysical Research*, 115,
 990 doi:10.1029/2009JD012257, 2010.
 991 Gollehon, N., and Quinby, W.: Irrigation in the American West: Area, Water and
 992 Economic Activity, *International Journal of Water Resources Development*, 16,
 993 187-195, 10.1080/07900620050003107, 2000.

- Griffiths, A. D., Conen, F., Weingartner, E., Zimmermann, L., Chambers, S. D., Williams, A. G., and Steinbacher, M.: Surface-to-mountaintop transport characterised by radon observations at the Jungfraujoch, *Atmos. Chem. Phys.*, 14, 12763-12779, 10.5194/acp-14-12763-2014, 2014.
- Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakuma, S., and de la Rue du Can, S.: High resolution fossil fuel combustion CO₂ emission fluxes for the United States, *Environmental Science and Technology*, 43, 5535-5541, 2009.
- Holton, J. R.: An introduction to dynamic meteorology, Academic Press, San Diego, 1992.
- Huddleston, B., Ataman, E., de Salvo, P., Zanetti, M., Bloise, M., Bel, J., Franceschini, G., and d'Ostiani, L. F.: Towards a GIS-based analysis of mountain environments and populations, Food and Agriculture Organization, Rome, 26, 2003.
- Hurst, D., Lin, J. C., Romashkin, P., Gerbig, C., Daube, B. C., Matross, D. M., Wofsy, S. C., and Elkins, J. W.: Continuing emissions of restricted halocarbons in the USA and Canada: Are they still globally significant?, *Journal of Geophysical Research*, 111, doi:10.1029/2005JD006785-doi:006710.001029/002005JD006785, 2006.
- IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change IPCC, Geneva, Switzerland, 151 pp., 2014.
- Jeong, S., Zhao, C., Andrews, A. E., Bianco, L., Wilczak, J. M., and Fischer, M. L.: Seasonal variation of CH₄ emissions from central California, *Journal of Geophysical Research: Atmospheres*, 117, n/a-n/a, 10.1029/2011JD016896, 2012.
- Keeling, C. D., Bacastow, R. B., Bain-Bridge, A. E., Ekdahl, C. A., Guenther, P. R., Waterman, L. S., and Chin, J. F. S.: Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, *Tellus*, 28, 538-551, 1976.
- Kellndorfer, J., Walker, W., Kirsch, K., Fiske, G., Bishop, J., Lapoint, L., Hoppus, M., and Westfall, J.: NACP Aboveground Biomass and Carbon Baseline Data, V.2 (NBCD 2000), U.S.A., 2000, in, ORNL Distributed Active Archive Center, 2013.
- Kim, S. Y., Millet, D. B., Hu, L., Mohr, M. J., Griffis, T. J., Wen, D., Lin, J. C., Miller, S. M., and Longo, M.: Constraints on carbon monoxide emissions based on tall tower measurements in the U.S. Upper Midwest, *Environmental Science and Technology*, 47, 8316-8324, 2013.
- Knowles, J. F.: Spatio-temporal patterns of soil respiration and the age of respired carbon from high-elevation alpine tundra, Ph.D., Department of Geography, University of Colorado, Boulder, Boulder, Colorado, USA, 119 pp., 2015.
- Lang, R. E., Sarzynski, A., and Muro, M.: Mountain megas: America's newest metropolitan places and a federal partnership to help them prosper, Brookings Institute, Washington, D.C., 64, 2008.
- Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneeth, A., Boden, T. A., Bopp, L., Bozec, Y., Canadell, J. G., Chini, L. P., Chevallier, F., Cosca, C. E., Harris, I., Hoppema, M., Houghton, R. A., House, J. I., Jain, A. K., Johannessen, T., Kato, E., Keeling, R. F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C. S., Landschützer, P., Lenton, A., Lima, I. D., Marland, G., Mathis, J. T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck,

- C., Saito, S., Salisbury, J. E., Schuster, U., Schwinger, J., Séférian, R., Segsneider, J., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Werf, G. R., Viovy, N., Wang, Y. P., Wanninkhof, R., Wiltshire, A., and Zeng, N.: Global carbon budget 2014, *Earth Syst. Sci. Data*, 7, 47-85, 10.5194/essd-7-47-2015, 2015.
- Levin, I., and Karstens, U. T. E.: Inferring high-resolution fossil fuel CO₂ records at continental sites from combined 14CO₂ and CO observations, *Tellus B*, 59, 245-250, 10.1111/j.1600-0889.2006.00244.x, 2007.
- Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: the Stochastic Time-Inverted Lagrangian Transport (STILT) model, *Journal of Geophysical Research*, 108, doi:10.1029/2002JD003161, 2003.
- Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Grainger, C. A., B.B., S., Bakwin, P. S., and Hollinger, D. Y.: Measuring fluxes of trace gases at regional scales by Lagrangian observations: application to the CO₂ Budget and Rectification Airborne (COBRA) study, *Journal of Geophysical Research*, 109, doi:10.1029/2004JD004754, doi:10.1029/2004JD004754, 2004.
- Lin, J. C., and Gerbig, C.: Accounting for the effect of transport errors on tracer inversions, *Geophysical Research Letters*, 32, doi:10.1029/2004GL021127- doi:021110.021029/022004GL021127, 2005.
- Lin, J. C.: Lagrangian modeling of the atmosphere: an introduction, in: *Lagrangian Modeling of the Atmosphere*, edited by: Lin, J. C., Brunner, D., Gerbig, C., Stohl, A., Luhar, A., and Webley, P., *Geophysical Monograph*, American Geophysical Union, 1-11, 2012.
- Lin, J. C., Brunner, D., Gerbig, C., Stohl, A., Luhar, A. K., and Webley, P. W.: *Lagrangian Modeling of the Atmosphere*, in: *Geophysical Monograph*, American Geophysical Union, 349, 2012.
- Mallia, D. V., Lin, J. C., Urbanski, S., Ehleringer, J., and Nehrkorn, T.: Impacts of upstream wildfire emissions on CO, CO₂, and PM_{2.5} concentrations in Salt Lake City, Utah, *Journal of Geophysical Research*, 120, doi:10.1002/2014JD022472, 2015.
- Mell, W. E., Manzello, S. L., Maranghides, A., Butry, D., and Rehm, R. G.: The wildland-urban interface fire problem--current approaches and research needs, *International Journal of Wildland Fire*, 19, 238-251, 2010.
- Mitchell, L. E., Lin, J. C., Bowling, D. R., Pataki, D. E., Strong, C., Schauer, A. J., Bares, R., Bush, S., Stephens, B. B., Mendoza, D., Mallia, D. V., Holland, L., Gurney, K. R., and Ehleringer, J. R.: Long-term urban carbon dioxide observations reveal spatial and temporal dynamics related to urban form and growth, *Proceedings of the National Academy of Sciences*, In Review.
- Monson, R. K., Turnipseed, A. A., Sparks, J. P., Harley, P. C., Scott-Denton, L. E., Sparks, K., and Huxman, T. E.: Carbon sequestration in a high-elevation, subalpine forest, *Global Change Biology*, 8, 459-478, 2002.
- Monson, R. K., Lipson, D. L., Burns, S. P., Turnipseed, A. A., Delany, A. C., Williams, M. W., and Schmidt, S. K.: Winter forest soil respiration controlled by climate and microbial community composition, *Nature*, 711-714, 2006.

- Mu, M., Randerson, J. T., van der Werf, G. R., Giglio, L., Kasibhatla, P., Morton, D., Collatz, G. J., DeFries, R. S., Hyer, E. J., Prins, E. M., Griffith, D. W. T., Wunch, D., Toon, G. C., Sherlock, V., and Wennberg, P. O.: Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide, *Journal of Geophysical Research-Atmospheres*, 116, 2011.
- Negron, J. F., and Popp, J. B.: Probability of ponderosa pine infestation by mountain pine beetle in the Colorado Front Range, *Forest Ecology and Management*, 191, 2004.
- 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, *Meteorology and Atmospheric Physics*, 107, 51-64, 2010.
- Parrish, D. D., Hahn, C. H., Fahey, D. W., Williams, E. J., Bollinger, M. J., Hübner, G., Buhr, M. P., Murphy, P. C., Trainer, M., Hsie, E. Y., Liu, S. C., and Fehsenfeld, F. C.: Systematic variations in the concentration of NO_x (NO Plus NO₂) at Niwot Ridge, Colorado, *Journal of Geophysical Research: Atmospheres*, 95, 1817-1836, 10.1029/JD095iD02p01817, 1990.
- Peters, W., Jacobson, A., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K. A., Miller, J. B., Bruhwiler, L., Petron, G., Hirsch, A. I., Worthy, D., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proceedings of the National Academy of Sciences*, 104, 18925-18930, 2007.
- Peters, W., Krol, M. C., Van Der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., Di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. A. J., Van Der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morgu, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Global Change Biology*, 16, 1317-1337, 10.1111/j.1365-2486.2009.02078.x, 2010.
- Pillai, D., Gerbig, C., Ahmadov, R., Rodenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B., and Lavrie, J. V.: High-resolution simulations of atmospheric CO₂ over complex terrain – representing the Ochsenkopf mountain tall tower, *Atmospheric Chemistry and Physics*, 11, 7445-7464, 2011.
- Potter, C., Fladland, M., Klooster, S., Genovese, V., Hiatt, S., and Gross, P.: Satellite Data Analysis and Ecosystem Modeling for Carbon Sequestration Assessments in the Western United States, in: *Carbon Sequestration and Its Role in the Global Carbon Cycle*, American Geophysical Union, 89-99, 2013.
- Ramankutty, N., and Foley, J. A.: Estimating historical changes in global land cover: Croplands from 1700 to 1992, *Global Biogeochemical Cycles*, 13, 997-1027, 10.1029/1999GB900046, 1999.
- Reisner, M.: *Cadillac desert: The American West and its disappearing water*, Penguin, 1993.

- Rinsland, C. P., Mahieu, E., Zander, R., Demoulin, P., Forrer, J., and Buchmann, B.: Free tropospheric CO, C₂H₆, and HCN above central Europe: Recent measurements from the Jungfraujoch station including the detection of elevated columns during 1998, *Journal of Geophysical Research: Atmospheres*, 105, 24235-24249, 10.1029/2000JD900371, 2000.
- Rodenbeck, C.: Estimating CO₂ sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Max-Planck Institut für Biogeochemie, Jena, Germany, 53, 2005.
- Rotach, M. W., Andretta, M., Calanca, P., Weigel, A. P., and Weiss, A.: Boundary layer characteristics and turbulent exchange mechanisms in highly complex terrain, *Acta Geophysica*, 56, 194-219, 2008.
- Rotach, M. W., Wohlfahrt, G., Hansel, A., Reif, M., Wagner, J., and Gohm, A.: The world is not flat: implications for the global carbon balance, *Bulletin of the American Meteorological Society*, 95, 1021-1028, 2014.
- Sarmiento, J. L., Gloor, M., Gruber, N., Beaulieu, C., Jacobson, A. R., Mikaloff-Fletcher, S. E., Pacala, S., and Rodgers, K.: Trends and regional distributions of land and ocean carbon sinks, *Biogeosciences*, 7, 2351-2367, 2010.
- Schimmel, D., Running, S., Monson, R., Turnipseed, A., and Anderson, D.: Carbon sequestration in the mountains of the Western US, *Eos*, 83, 445-456, 2002.
- Schmidt, M., Graul, R., Sartorius, H., and Levin, I.: The Schauinsland CO₂ record: 30 years of continental observations and their implications for the variability of the European CO₂ budget, *Journal of Geophysical Research: Atmospheres*, 108, n/a-n/a, 10.1029/2002jd003085, 2003.
- Schwalm, C. R., Williams, C. A., Schaefer, K., Baldocchi, D., Black, T. A., Goldstein, A. H., Law, B. E., Oechel, W. C., Paw U, K. T., and Scott, R. L.: Reduction in carbon uptake during turn of the century drought in western North America, *Nature Geoscience*, doi:10.1038/NGEO1529, 2012.
- Skamarock, W. C., and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, *Journal of Computational Physics*, 227, 3465-3485, 2008.
- Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., and Nakazawa, T.: Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO₂, *Science*, 316, 1732-1735, 2007.
- Stephens, B. B., Miles, N. L., Richardson, S. J., Watt, A. S., and Davis, K. J.: Atmospheric CO₂ monitoring with single-cell NDIR-based analyzers, *Atmospheric Measurement Techniques*, 4, 2737-2748, 2011.
- Steyn, D., De Wekker, S. J., Kossmann, M., and Martilli, A.: Boundary Layers and Air Quality in Mountainous Terrain, in: *Mountain Weather Research and Forecasting*, edited by: Chow, F. K., De Wekker, S. F. J., and Snyder, B. J., Springer Atmospheric Sciences, Springer Netherlands, 261-289, 2013.
- Sun, J., Burns, S. P., Delany, A. C., Oncley, S. P., Turnipseed, A. A., Stephens, B. B., Lenschow, D. H., LeMone, M. A., Monson, R. K., and Anderson, D. E.: CO₂ transport over complex terrain, *Agricultural and Forest Meteorology*, 145, 1-21, <http://dx.doi.org/10.1016/j.agrformet.2007.02.007>, 2007.

- Sun, J., Oncley, S. P., Burns, S. P., Stephens, B. B., Lenschow, D. H., Campos, T., Monson, R. K., Schimel, D. S., Sacks, W. J., De Wekker, S. F. J., Lai, C. T., Lamb, B., Ojima, D., Ellsworth, P. Z., Sternberg, L. S. L., Zhong, S., Clements, C., Moore, D. J. P., Anderson, D. E., Watt, A. S., Hu, J., Tschudi, M., Aulenbach, S., Allwine, E., and Coons, T.: A multiscale and multidisciplinary investigation of ecosystem-atmosphere CO₂ exchange over the Rocky Mountains of Colorado, *Bulletin of the American Meteorological Society*, 209-230, 2010.
- Sun, J., and De Wekker, S. F. J.: Atmospheric carbon dioxide transport over mountainous terrain, in: *Mountain Ecosystems*, edited by: Richards, K. E., Nova Science Publishers, 101-121, 2011.
- Tkacz, B., Moody, B., Castillo, J. V., and Fenn, M. E.: Forest health conditions in North America, *Environmental Pollution*, 155, 409-425, <http://dx.doi.org/10.1016/j.envpol.2008.03.003>, 2008.
- Tuzson, B., Henne, S., Brunner, D., Steinbacher, M., Mohn, J., Buchmann, B., and Emmenegger, L.: Continuous isotopic composition measurements of tropospheric CO₂ at Jungfraujoch (3580 m a.s.l.), Switzerland: real-time observation of regional pollution events, *Atmos. Chem. Phys.*, 11, 1685-1696, 10.5194/acp-11-1685-2011, 2011.
- Urbanski, S. P., Hao, W. M., and Nordgren, B.: The wildland fire emission inventory: western United States emission estimates and an evaluation of uncertainty, *Atmospheric Chemistry and Physics*, 11, 12973-13000, 2011.
- van der Molen, M. K., and Dolman, A. J.: Regional carbon fluxes and the effect of topography on the variability of atmospheric CO₂, *Journal of Geophysical Research: Atmospheres*, 112, n/a-n/a, 10.1029/2006JD007649, 2007.
- van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions, and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009), *Atmospheric Chemistry and Physics*, 10, 11707-11735, 2010.
- Wharton, S., Falk, M., Bible, K., Schroeder, M., and Paw U, K. T.: Old-growth CO₂ flux measurements reveal high sensitivity to climate anomalies across seasonal, annual and decadal time scales, *Agricultural and Forest Meteorology*, 161, 1-14, 2012.
- Wiedinmyer, C., and Neff, J. C.: Estimates of CO₂ from fires in the United States: implications for carbon management, *Carbon Balance and Management*, 2, 10.1186/1750-0680-2-10, 2007.
- Yi, C., Anderson, D. E., Turnipseed, A. A., Burns, S. P., Sparks, J. P., Stannard, D. I., and Monson, R. K.: The contribution of advective fluxes to net ecosystem CO₂ exchange in a high-elevation, subalpine forest, *Ecological Applications*, 18, 1379-1390, 2008.
- Zardi, D., and Whiteman, C. D.: Diurnal Mountain Wind Systems, in: *Mountain Weather Research and Forecasting*, edited by: Chow, F. K., De Wekker, S. F. J., and Snyder, B. J., Springer Atmospheric Sciences, Springer Netherlands, 35-119, 2013.

Due to the dangers of mis-representing terrain/flows and introducing biases into the carbon inversion system, an obvious way to deal with this problem is to neglect the mountaintop data altogether. This is already commonly practiced within carbon inversion systems (Rodenbeck, 2005; Geels et al., 2007; Peters et al., 2010). In fact, the most recent release of CarbonTracker (“CT-2015”) stopped assimilating the three RACCOON sites (<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/>).

However, the absence of mountaintop CO₂ observations to constrain carbon inversion systems is, in effect, throwing away valuable information that could inform carbon exchange in potentially important areas of the world (Fig. 1). Case in point is the Schauinsland CO₂ time series on a mountain in the middle of Western Europe, which as of this writing has collected over 40 years of continuous CO₂ data (Schmidt et al., 2003) but remains excluded from numerous carbon inversion systems (Rodenbeck, 2005; Geels et al., 2007; Peters et al., 2010).

Above-ground Biomass in the Western U.S.

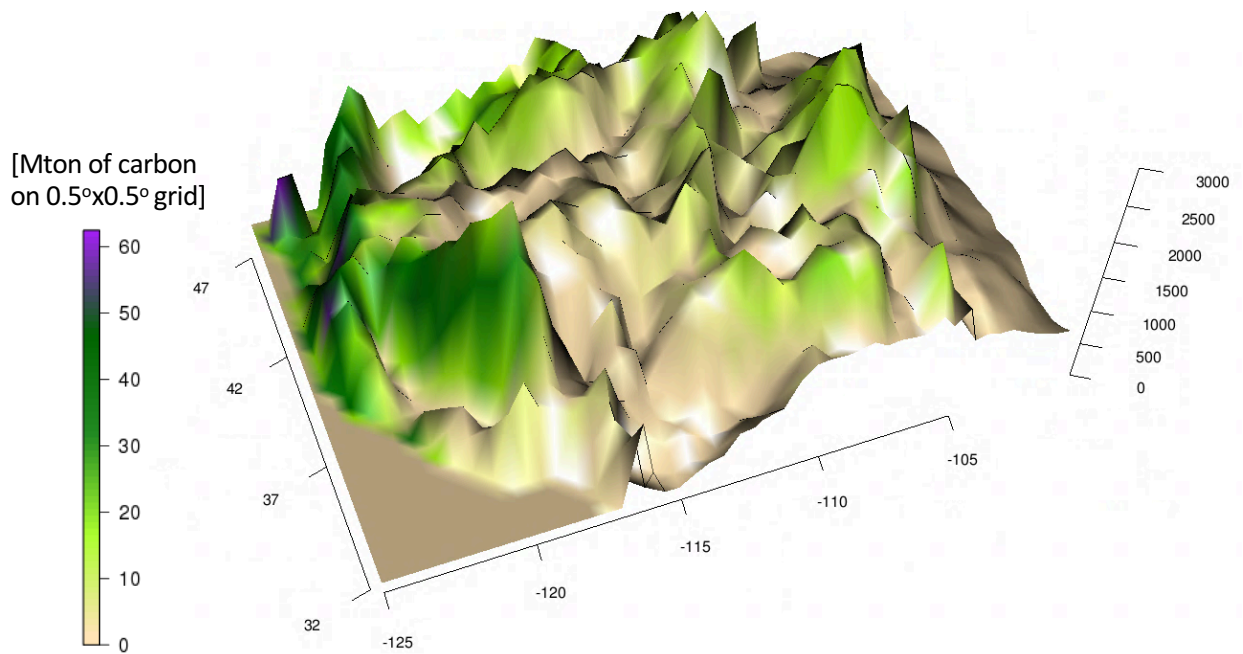


Fig. 1

WRF Domains

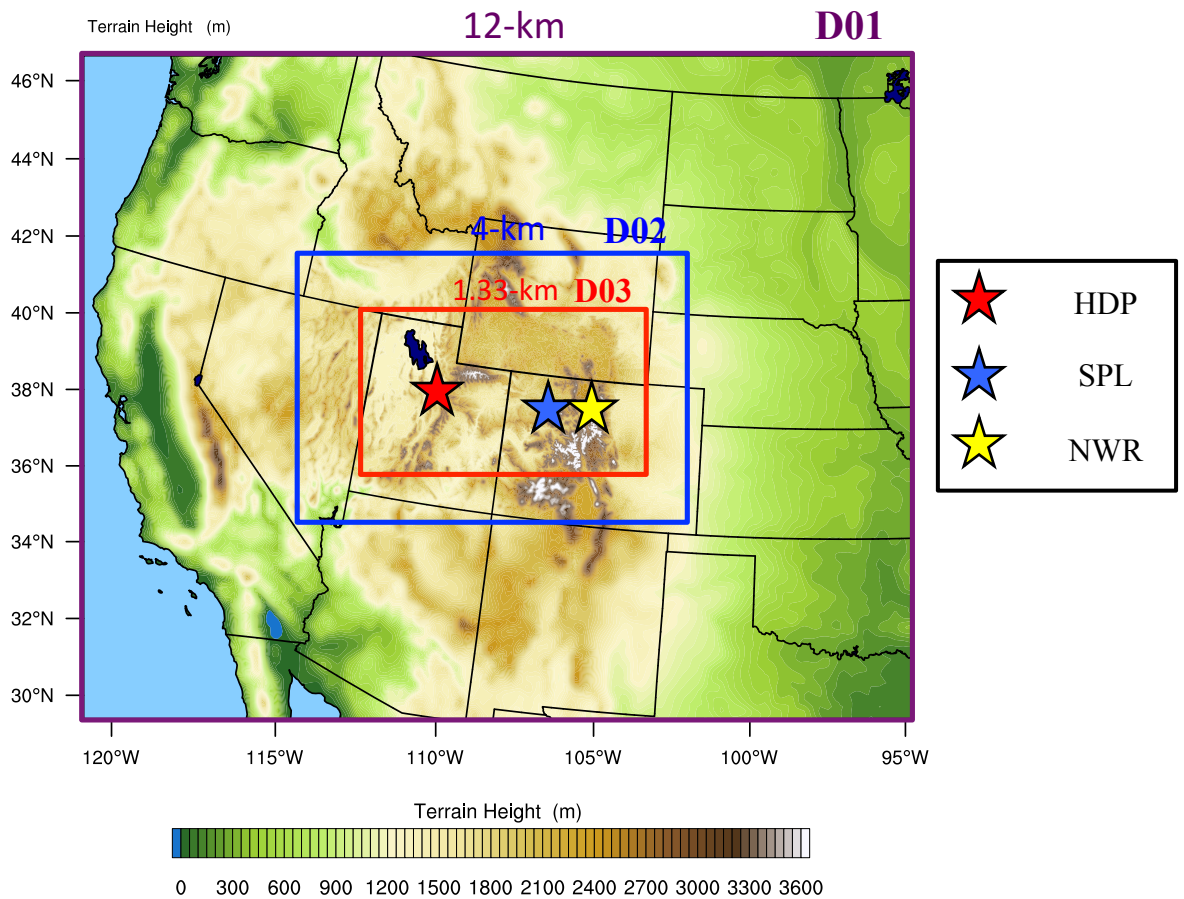


Fig. 2

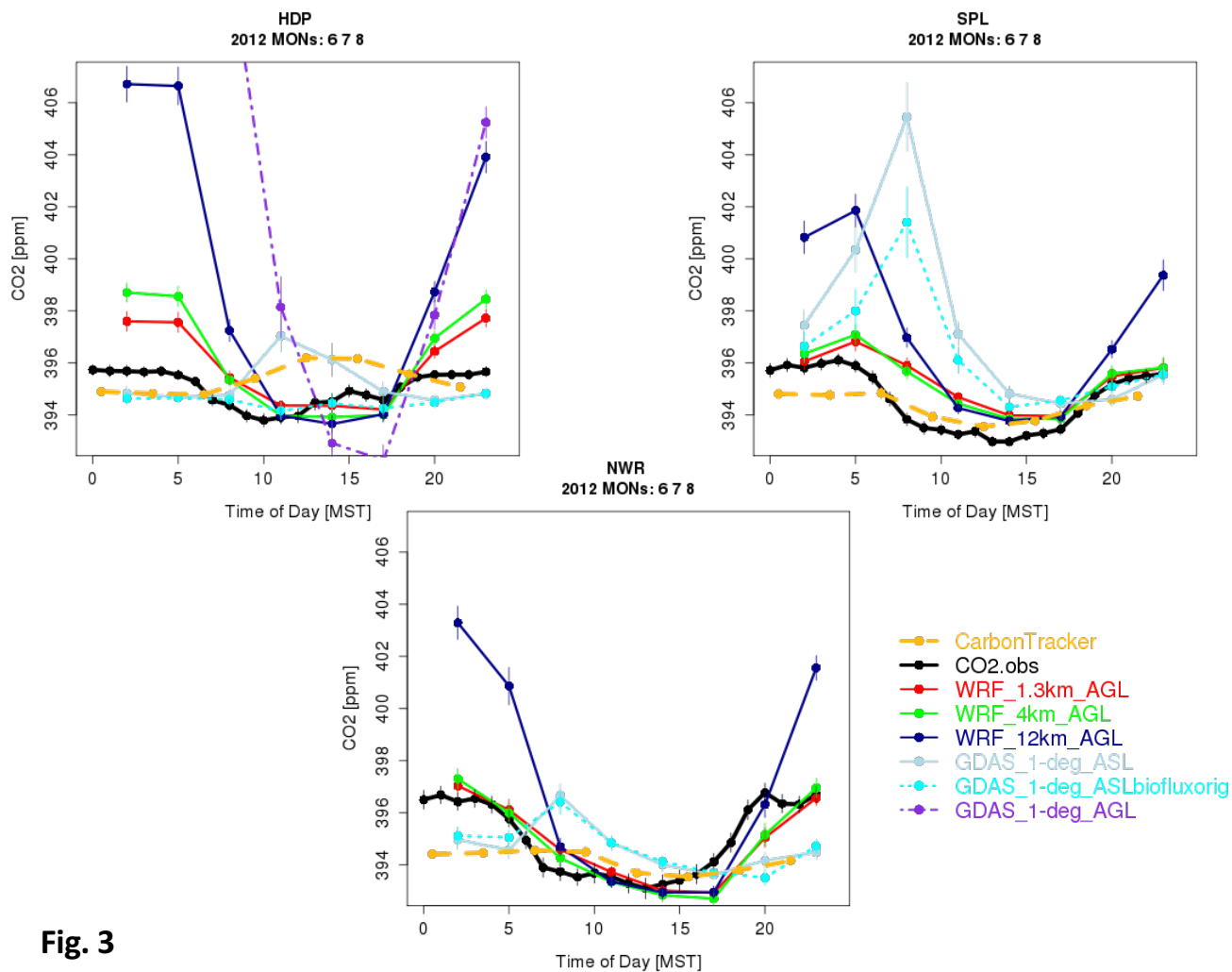


Fig. 3

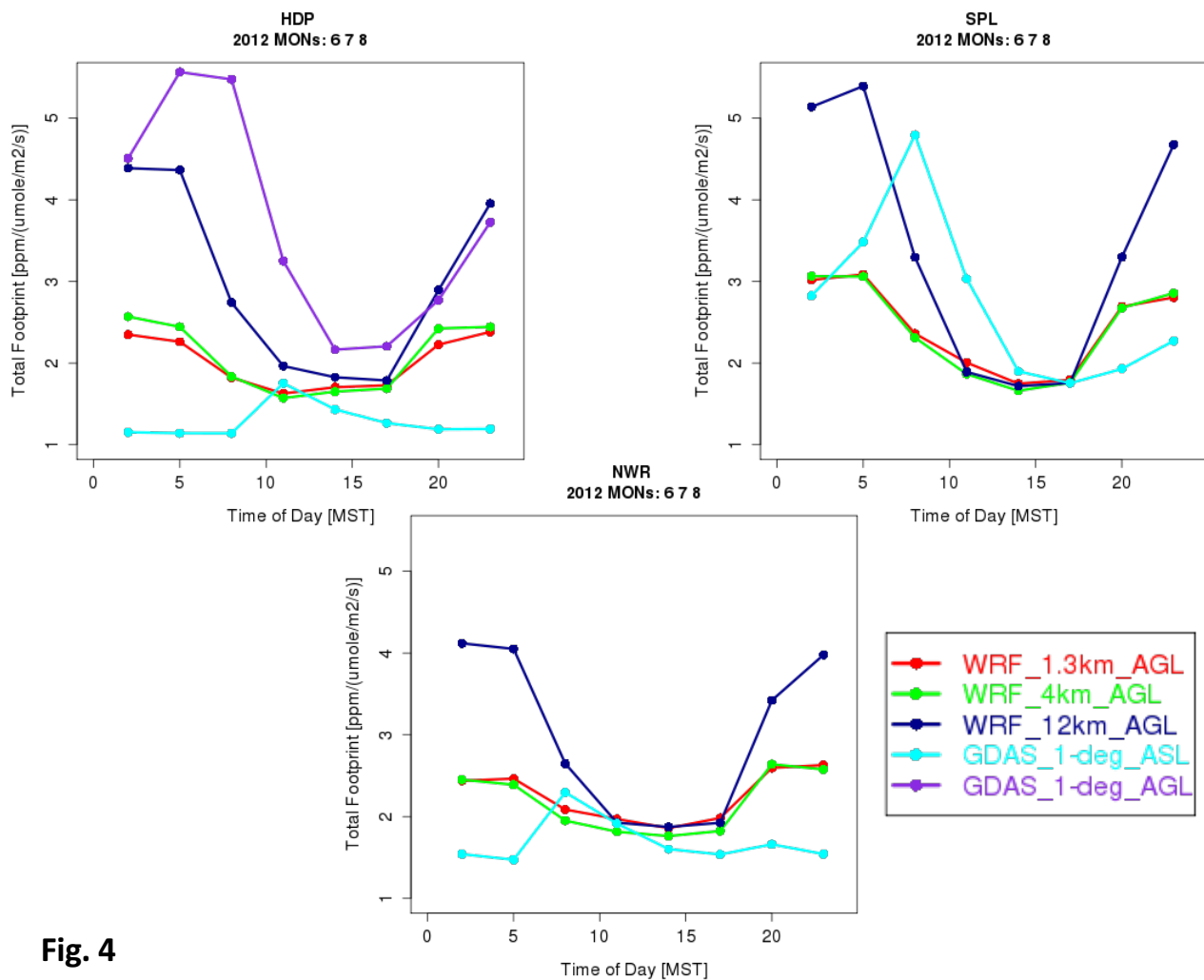
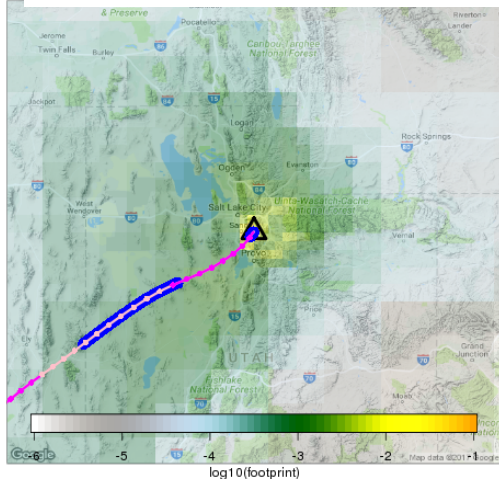
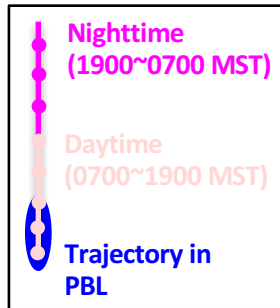


Fig. 4

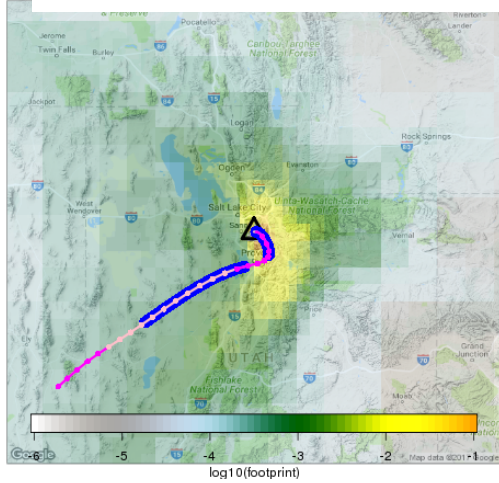
HDP ave footprint: WRF-1.3km (AGL)



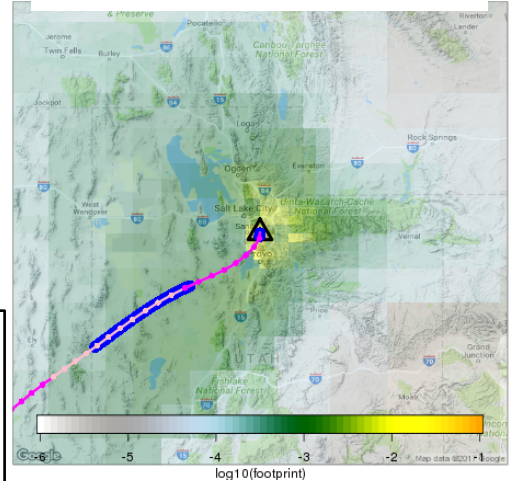
0900 UTC
(0200 MST)



HDP ave footprint: WRF-12km (AGL)



HDP ave footprint: WRF-4km (AGL)



HDP ave footprint: GDAS-1° (ASL)

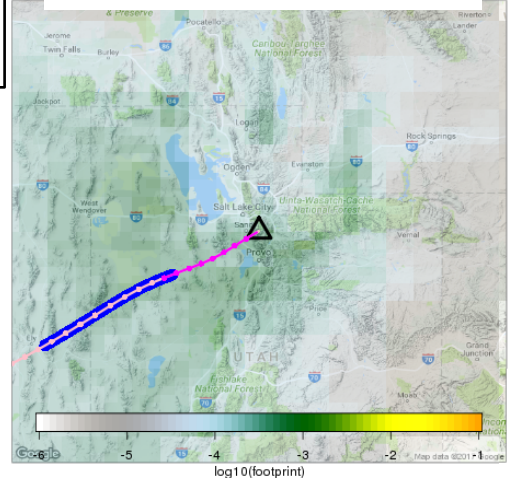
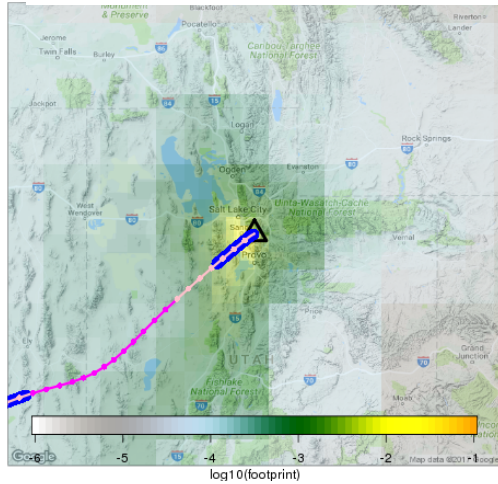


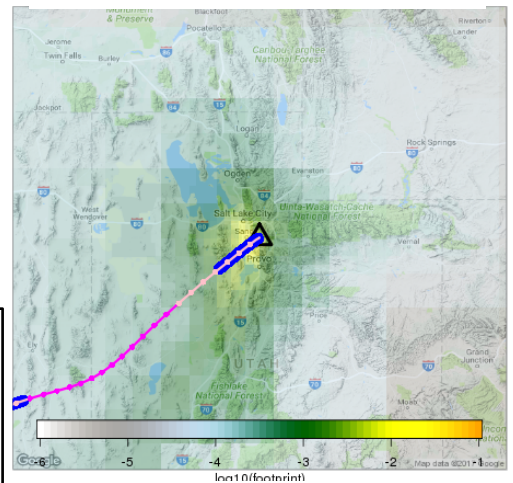
Fig. 5

HDP ave footprint: WRF-1.3km (AGL)

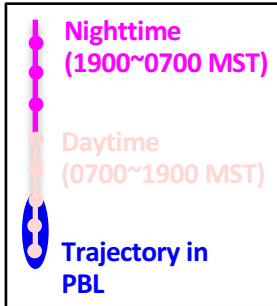
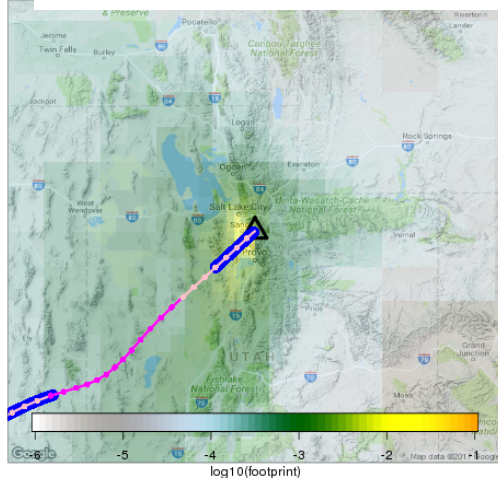


2100 UTC
(1400 MST)

HDP ave footprint: WRF-4km (AGL)



HDP ave footprint: WRF-12km (AGL)



HDP ave footprint: GDAS-1° (ASL)

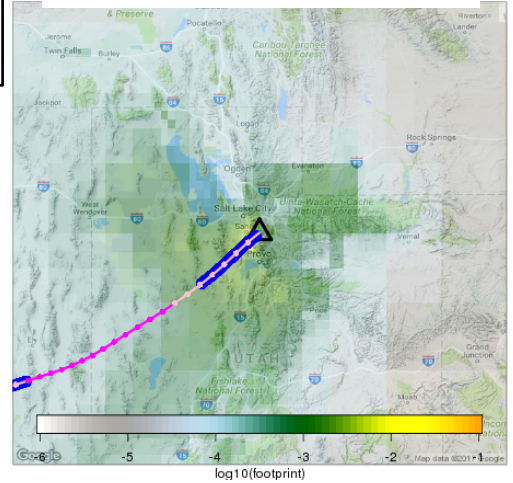


Fig. 6

HDP: Mean 3D Trajectory of Stochastic Particles & PBL ht for Different Runs

0900 UTC (0200 MST)

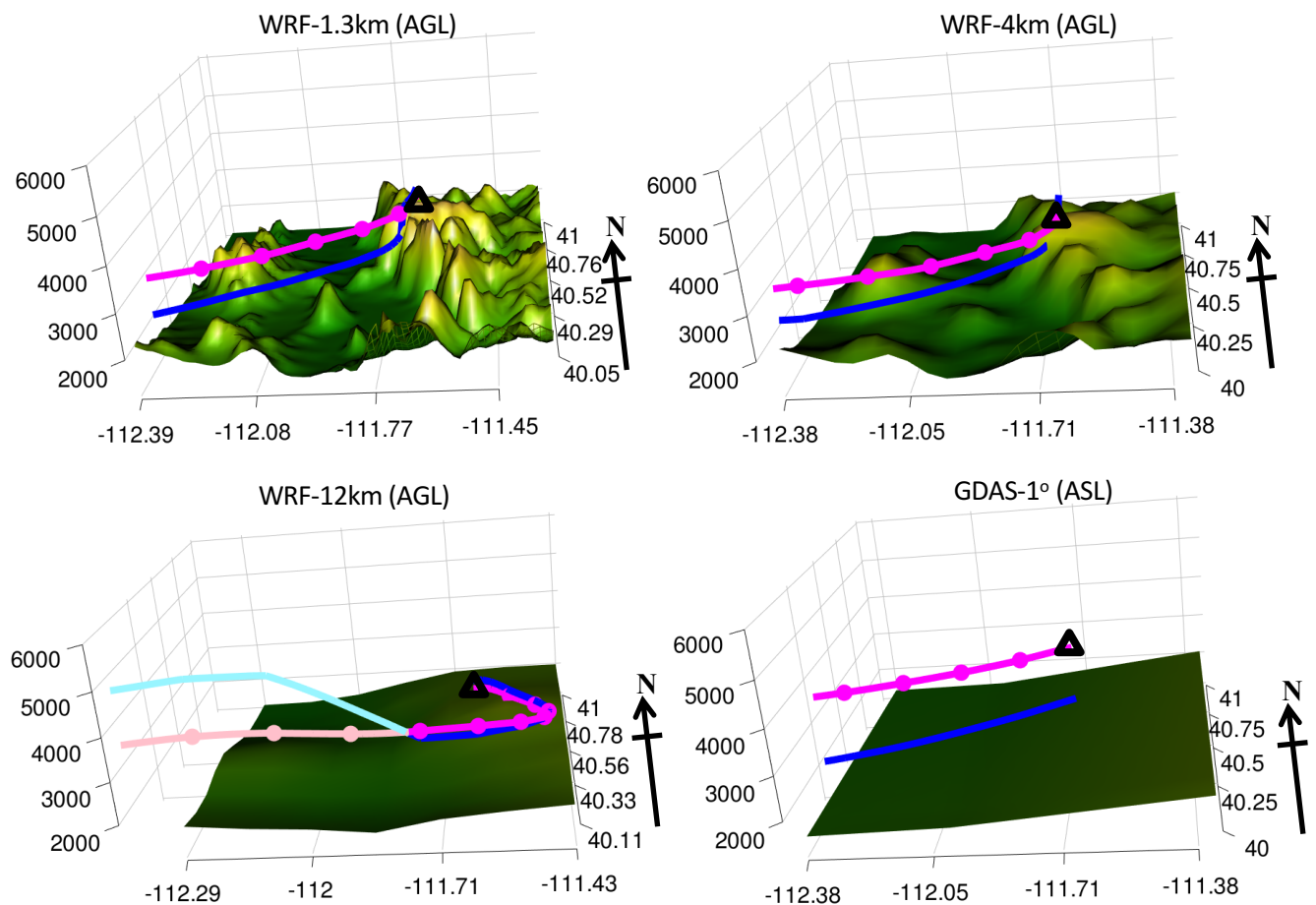


Fig. 7

HDP: Mean 3D Trajectory of Stochastic Particles & PBL ht (above ground level)

0900 UTC (0200 MST)

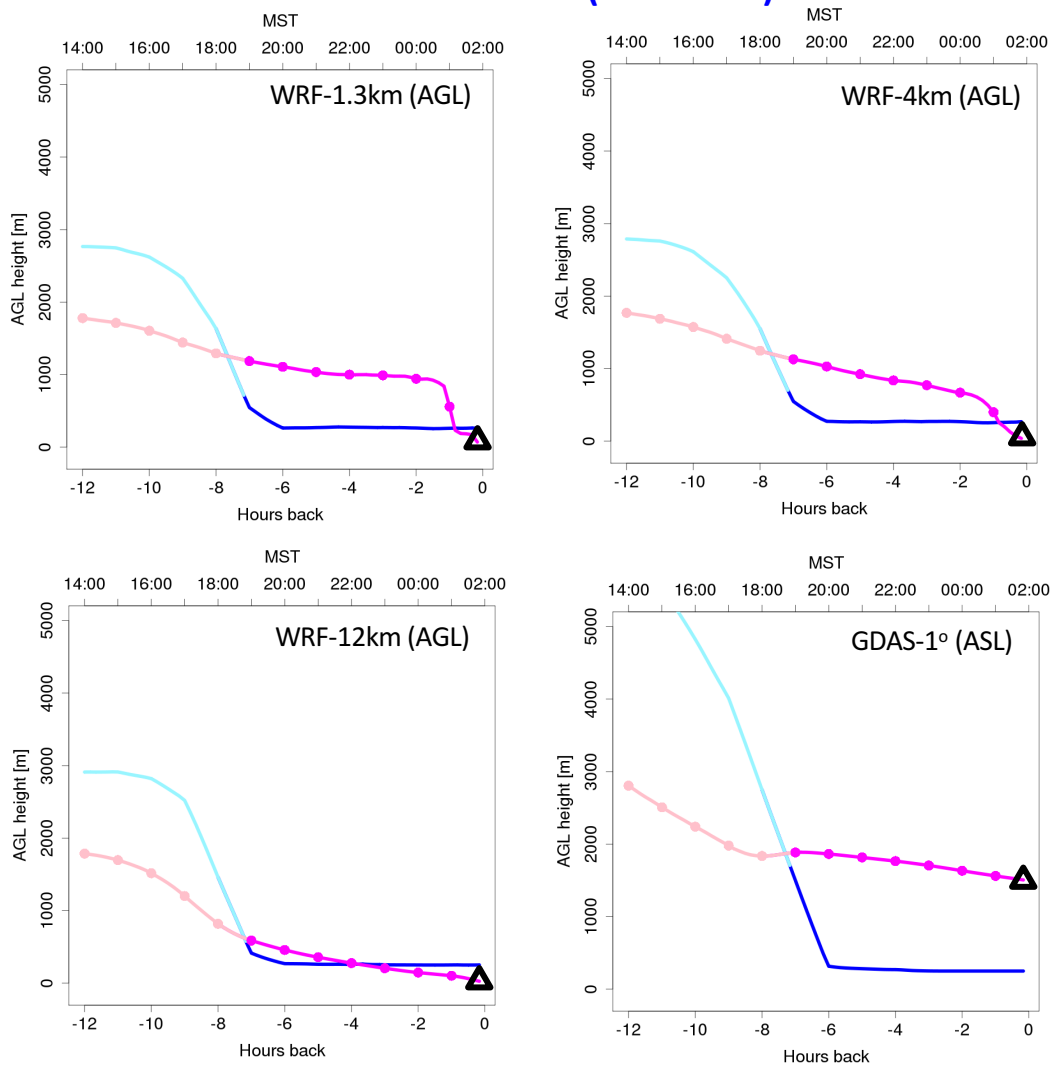


Fig. 8

SPL: Mean 3D Trajectory of Stochastic Particles & PBL ht for Different Runs
0900 UTC (0200 MST)

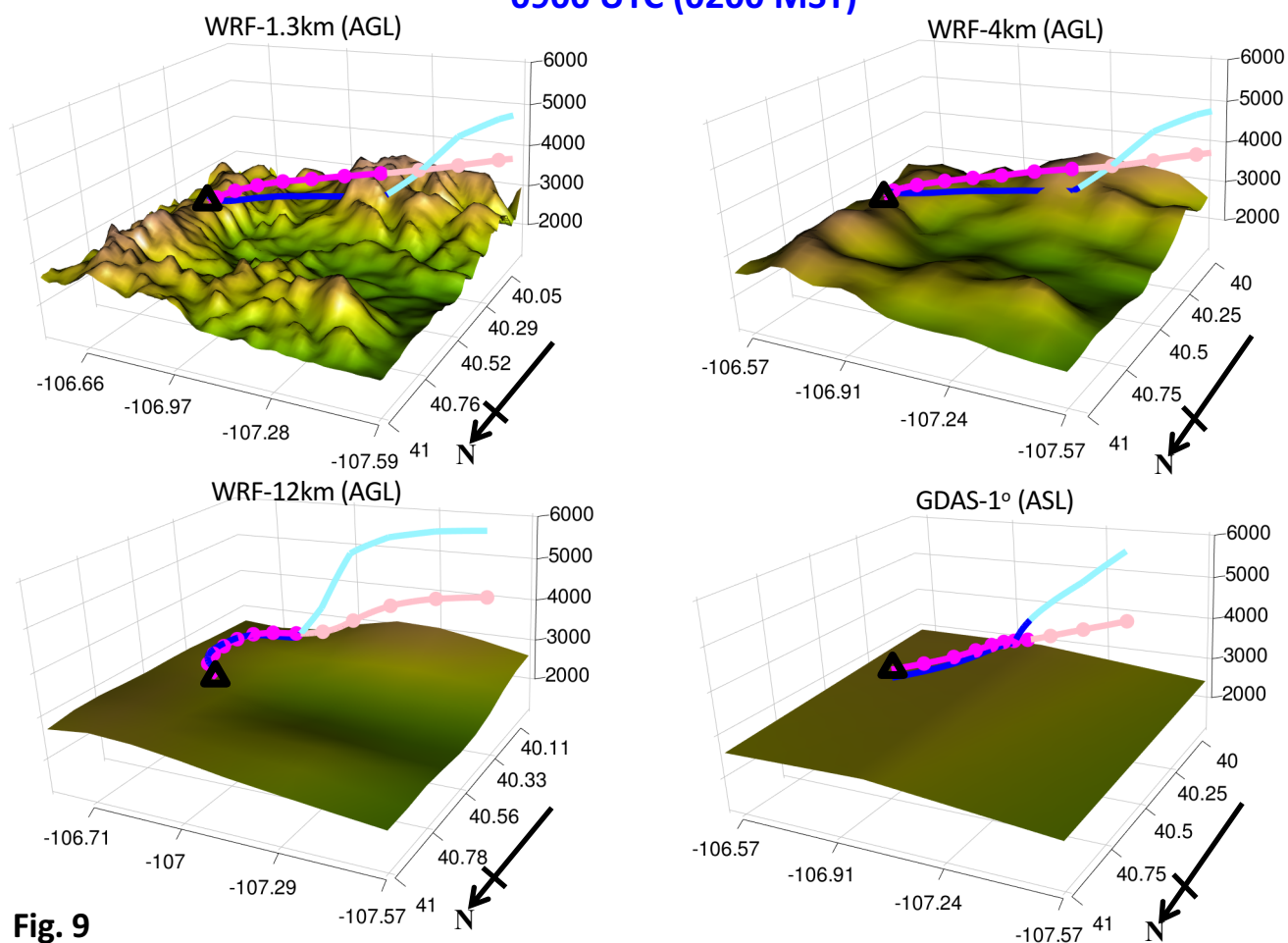


Fig. 9

SPL: Mean 3D Trajectory of Stochastic Particles & PBL ht (above ground level)

0900 UTC (0200 MST)

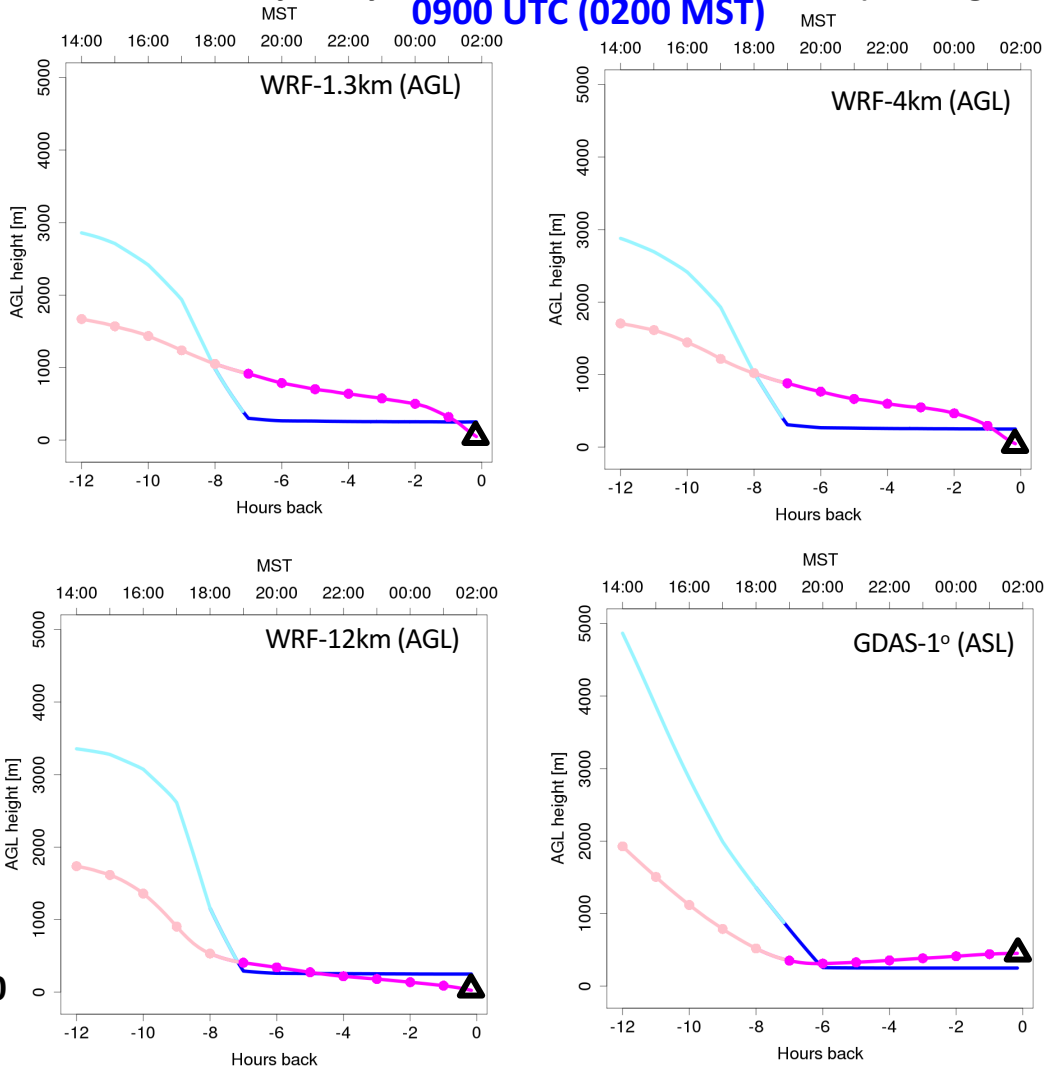


Fig. 10

NWR: Mean 3D Trajectory of Stochastic Particles & PBL ht for Different Runs **0900 UTC (0200 MST)**

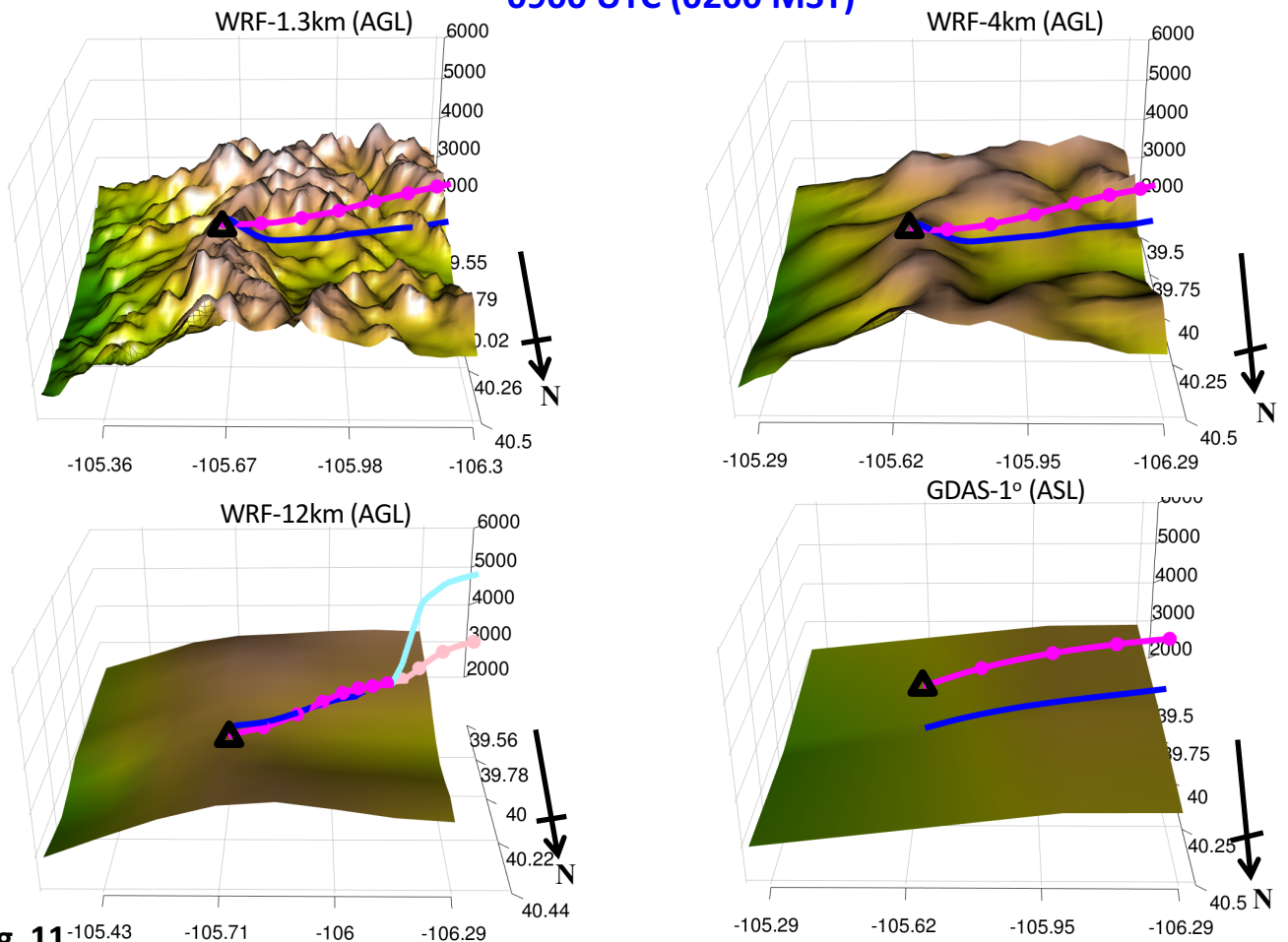


Fig. 11

NWR: Mean 3D Trajectory of Stochastic Particles & PBL ht (above ground level)

0900 UTC (0200 MST)

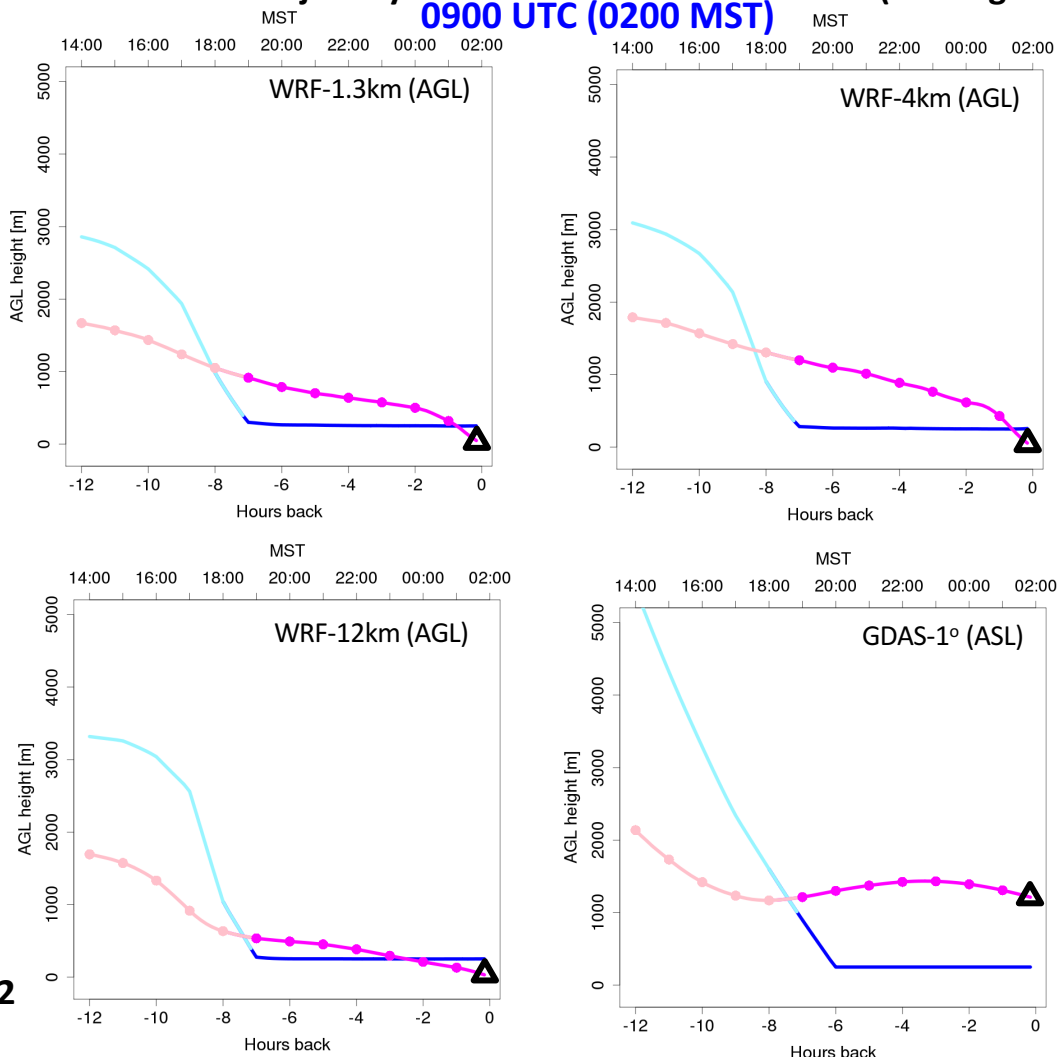


Fig. 12

Mean 3D Trajectory of Stochastic Particles & PBL ht GDAS-1° (ASL)

HDP
(1100 MST)

SPL
(0800 MST)

NWR
(1100 MST)

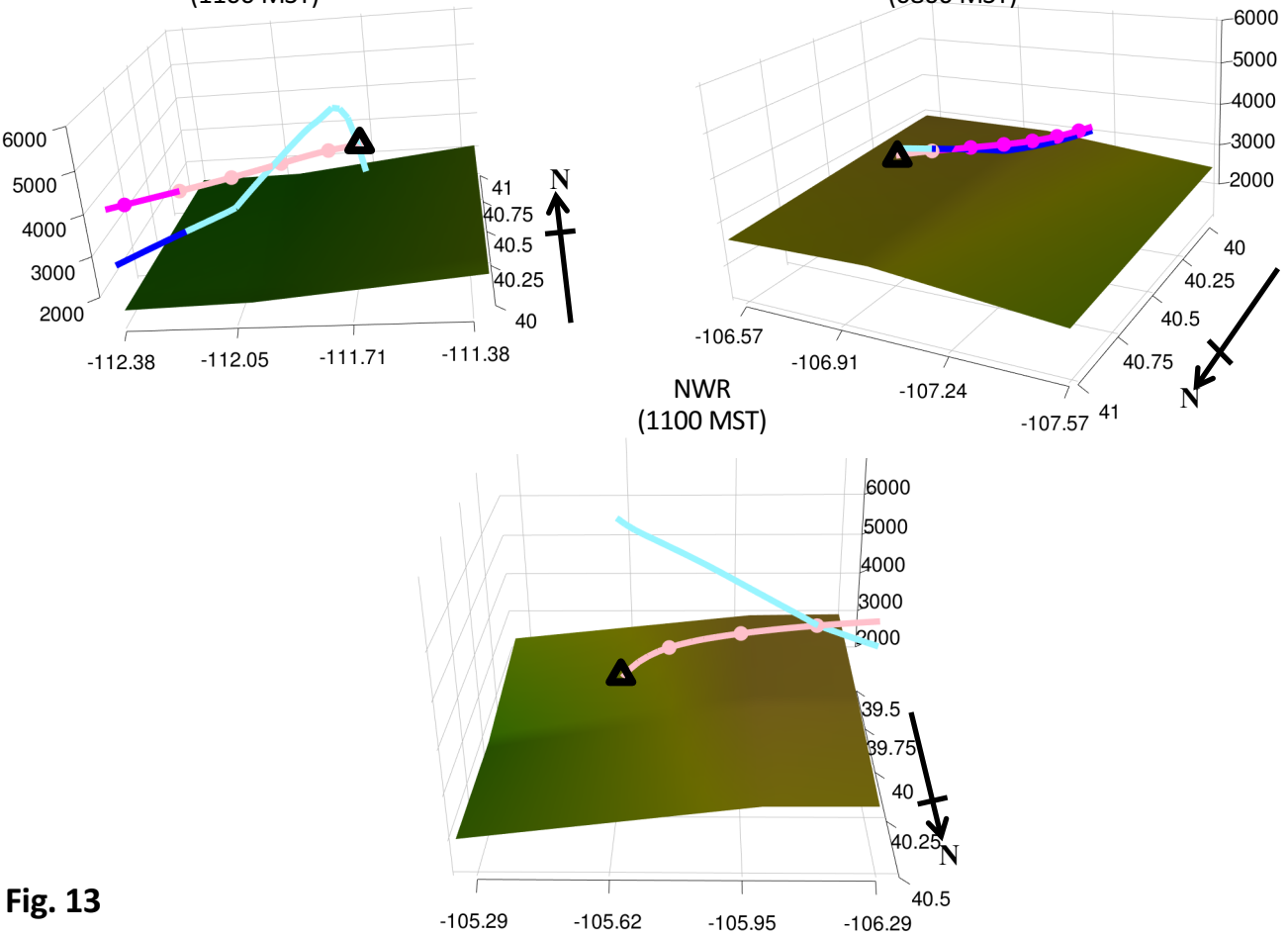


Fig. 13

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13

2
3
4
5
6
7
8
9
10
11
12
13

4
5
6
7
8
9
10
11
12
13

6
7
8
9
10
11
12
13

7
8
9
10
11
12
13

9
10
11
12
13

11
12
13

14
15
16
17
18
19
20
21
22
23

15
16
17
18
19
20
21
22
23

Adjusting the CT-2013b diurnal cycle

In the CarbonTracker assimilation process, attempts to match CO₂ observations could result in “dipoles” in scaling factors between nearby ecoregions, leading to negative fluxes even at night (Fig. S2a). While respiration can occur during the day when vegetation is under stress (e.g., droughts), photosynthetic uptake (negative fluxes) at night, in the absence of sunlight, is biologically unphysical. In order to correct the reversed diurnal cycle seen in CarbonTracker, a reversal had to be first detected within CarbonTracker for the selected grid cell for a given day. Once the reversal was detected, the sign of the biospheric flux was flipped. The positive flux was then adjusted so that the net flux for the selected gridcell for the given day was equal to 0. Finally, the negative flux was adjusted so that the final net flux was equal to the original net flux, which preserved the total net flux for the day (Fig. S3). The resulting biospheric flux pattern can be seen in Fig. S2b.

Fig. S2

Mean biospheric fluxes from Jun~Aug 2012 averaged between 0600~0900 UTC (2300~0200 MST). (a) Biospheric fluxes for the unmodified CarbonTracker flux fields and (b) biospheric fluxes for the adjusted CarbonTracker flux fields. The black circle represents HDP, the black diamond represents SPL, and the black star represents NWR.

Fig. S3

Schematic showing the adjustment of erroneous diurnal pattern in biospheric flux within CarbonTracker (red line), with nighttime uptake, to a corrected biospheric flux (green line). The dashed line represents a flux of 0.

Fig. S4

Average contributions to CO₂ variations at HDP, SPL, and NWR from biospheric, anthropogenic, and wildfire fluxes at different times of the day between Jun~Aug 2012 as simulated by STILT, driven with WRF-1.3km winds. Also shown are the observed variations, calculated by subtracting out the STILT-derived background (see Sect. 2.3).

Deleted: .

Fig. S5

Time series of CO₂ errors (Simulated – Observed) at the HDP, SPL, and NWR sites during the study period (Jun~Aug 2012) for the different model configurations—i.e., WRF-1.3km AGL, 4-km AGL, 12-km AGL, and GDAS-ASL. The thin lines denote CO₂ errors calculated at high frequency, at 3-hourly time spacing. The thick lines represent the CO₂ errors smoothed with a 4-day centered running average that will be correlated with other meteorological variables (Table S1). The bias and root-mean-square error (RMSE) reported on the top-right hand box are calculated based on the 3-hourly time series. The gap in the earlier part of July at NWR is due to missing observations.

Fig. S6

Mean CO₂ concentrations extracted from the bottom 8 levels of CarbonTracker, in the respective gridcells where the HDP, SPL, and NWR sites are located. The mean model heights of the bottom 8 levels are (in meters AGL): 25, 103, 247, 480, 814, 1259, 1822, 2508. The concentrations interpolated to the heights of the 3 sites are indicated by the orange dashed line. The observed values are drawn in black, with unfiltered data (dashed) and after applying the filter for removing local influences (solid; Sect. 2.1).

Deleted: .

Deleted: 5

Fig. S7

The average footprint (shown in log₁₀) for the SPL at 0200 MST (0900 UTC), gridded at 0.1°×0.1°. The site is denoted as a triangle. The average back trajectory (averaged over the stochastic STILT trajectories) is drawn as a line, with points indicating trajectory locations every hour, as the trajectory moves back from the site indicated as points. Magenta parts of the trajectory refer to the nighttime (1900~0700 MST), while pink portions indicate the daytime (0700~1900 MST). Parts of the trajectory are shaded with blue when it is found below the average height of the PBL along the trajectory.

Deleted: 6

Deleted: Red

Fig. S8

Similar to Fig. S7, but for 1400 MST (2100 UTC).

Deleted: 7

Deleted: 6

91

92 **Fig. S2**

93 Similar to Fig. S7, but for the NWR site.

94

Deleted: 8

Deleted: 6

95 **Fig. S10**

96 Similar to Fig. S9, but for 1400 MST (2100 UTC).

97

Deleted: 9

Deleted: 8

98 **Fig. S11**

Deleted: 0

99 Three dimensional plots of the terrain over a domain of $\sim 1^\circ \times 1^\circ$ surrounding HDP, as
100 resolved by the WRF and GDAS models at various grid spacings. Also shown is the
101 average back trajectory, derived by averaging locations of the numerous stochastic
102 trajectories simulated by STILT, driven by the various WRF meteorological fields and
103 the global GDAS field. Back trajectories were started from HDP at 1400 MST (2100
104 UTC). Points indicate trajectory locations every hour, as the trajectory moves back from
105 the site indicated as points. Magenta portions of the trajectory refer to the nighttime
106 (1900~0700 MST), while pink portions indicate the daytime (0700~1900 MST). In
107 addition, the PBL heights averaged along the backtrajectory are shown as the blue line.

Deleted: Red

108

109 **Fig. S12**

110 Similar to Fig. S11, but for SPL.

111

Deleted: 1

Deleted: 0

112 **Fig. S13**

113 Similar to Fig. S11, but for NWR.

114

115

Deleted: 2

Deleted: 0

Table S1

Correlation coefficients between CO₂ errors smoothed with a 4-day centered running average (Fig. S5) and potential explanatory meteorological variables observed near the HDP, SPL, and NWR sites. The smoothing window of 4-days was selected to focus on synoptic scale meteorological changes. The meteorological observations come from radiosondes launched at 00-UTC and 12-UTC from the following airports: Salt Lake City, Grand Junction, and Denver for the HDP, SPL, and NWR sites, respectively. The meteorological variables are extracted at the 500-hPa level and include the observed geopotential height (GPH), geopotential height gradient between NWR and HDP (NWR – HDP), observed windspeed, as well as the U- and V- components of the observed wind vector. The GPH time series is processed by subtracting its 20-day running average (centered) to remove trends and seasonal variations and then smoothed with a 4-day running average. Pearson correlation coefficients are reported here; coefficients with lower statistical significance ($p>0.05$) are not shown and indicated with “-“ in the Table.

	HDP				SPL				NWR			
	WRF 1.3-km	WRF 4-km	WRF 12-km	GDAS	WRF 1.3-km	WRF 4-km	WRF 12-km	GDAS	WRF 1.3-km	WRF 4-km	WRF 12-km	GDAS
GPH	0.35	0.57	-	0.39	-	-	0.17	-	-	-	-	-0.19
GPH gradient	-0.37	-0.39	-	-	0.20	0.25	0.57	-	-0.71	-0.69	-0.64	-0.66
U wind	-	-	-0.30	-	0.60	0.64	0.54	0.49	0.30	0.33	0.20	0.37
V wind	-0.53	-0.49	-	-	-	0.19	0.49	0.24	-0.52	-0.52	-0.44	-0.49
Windspeed	-0.16	-0.24	-0.18	-	0.43	0.49	0.59	0.34	-	-	-	-

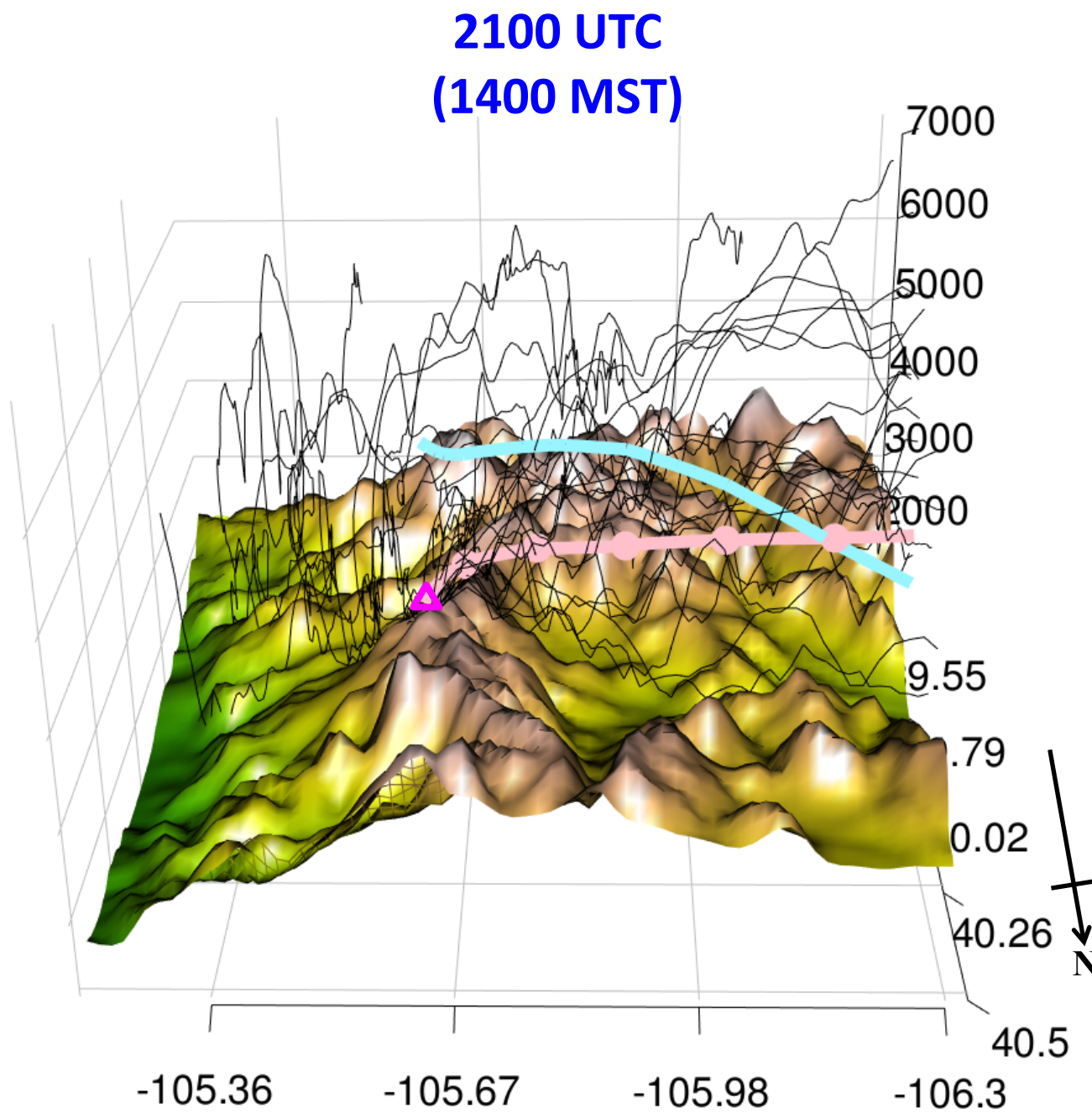


Fig. S1

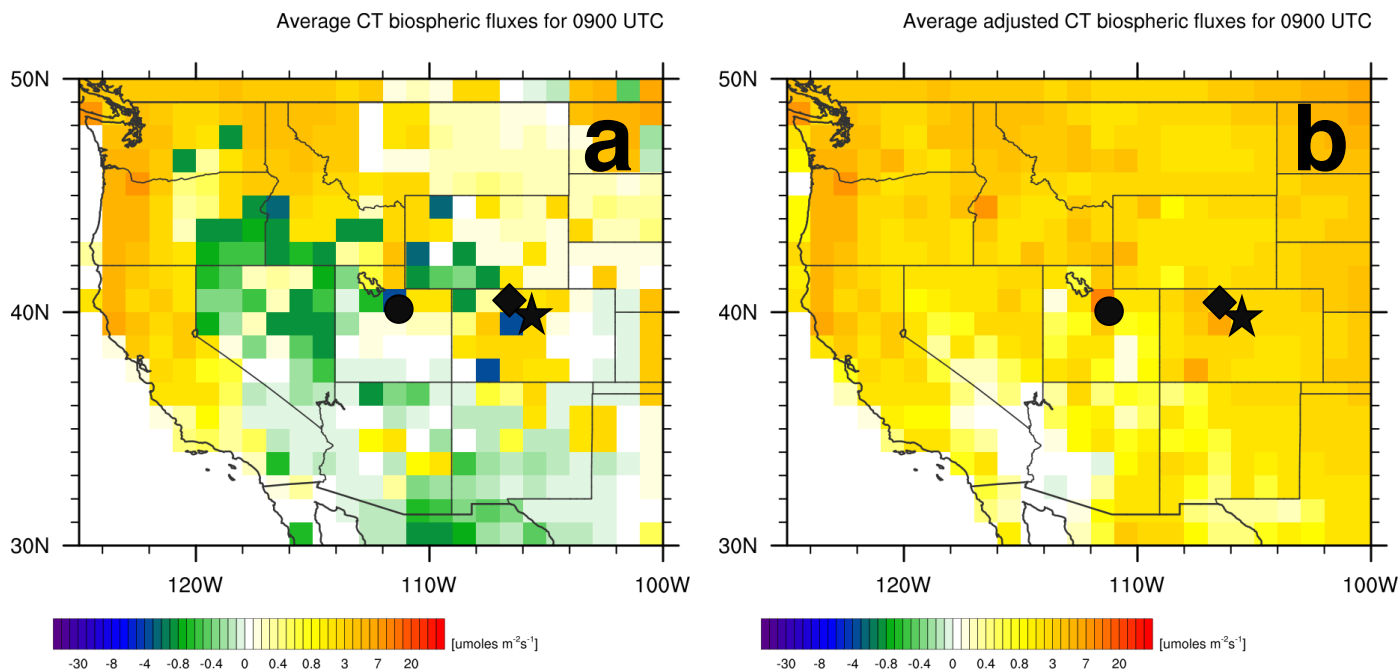


Fig. S2

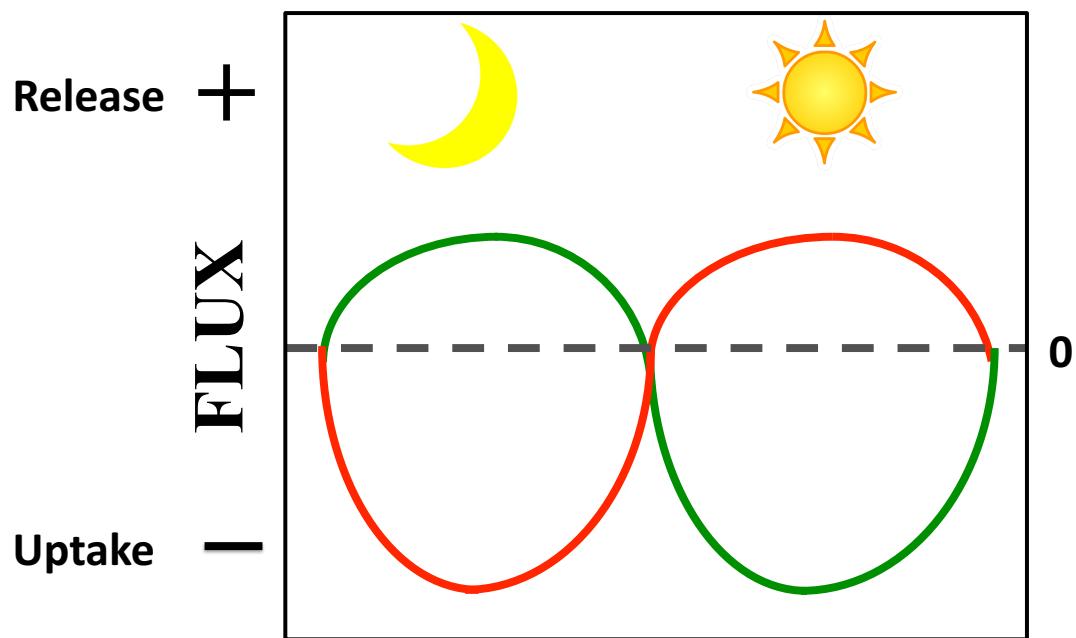


Fig. S3

Average Diurnal Contributions from Different CO₂ Sources

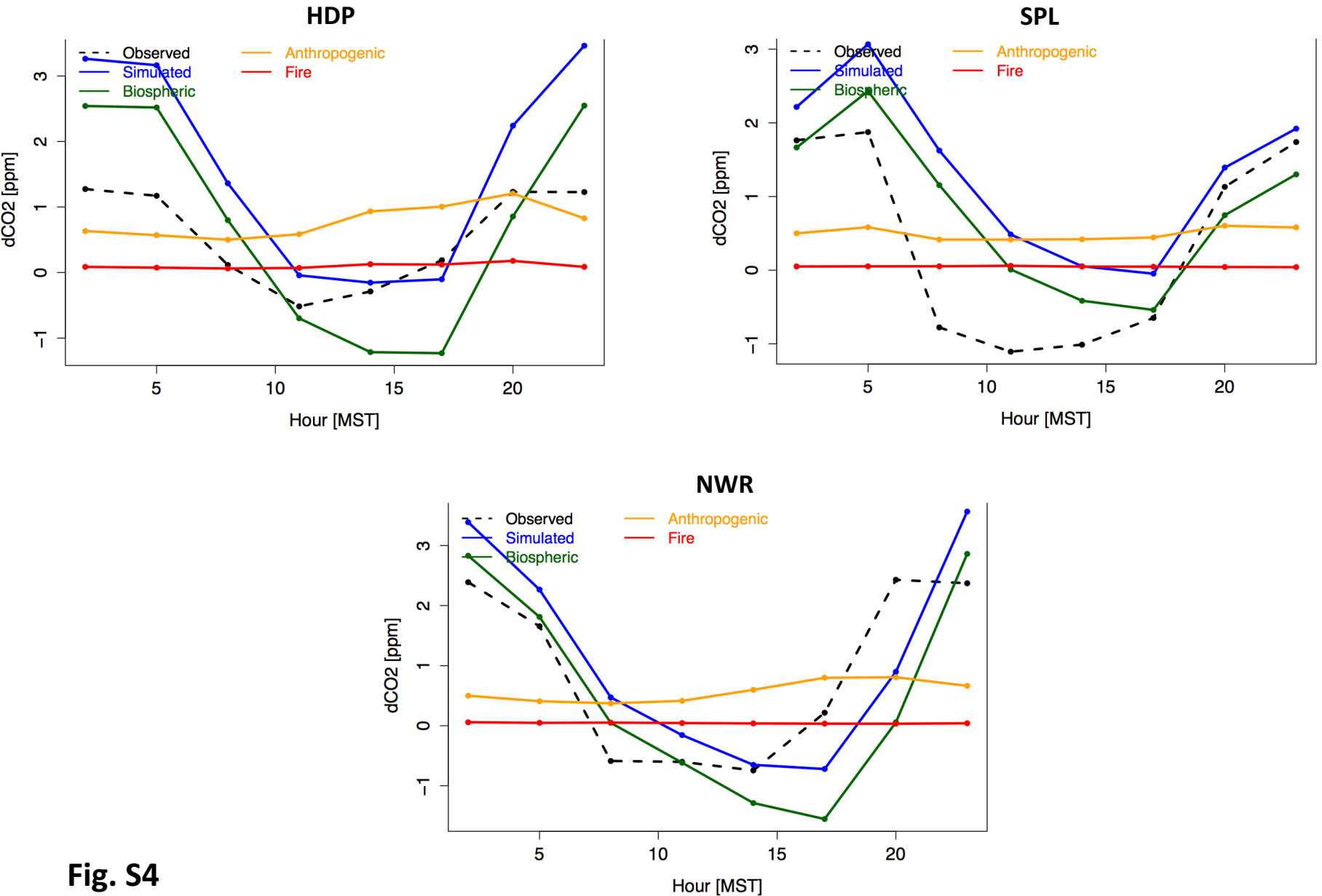


Fig. S4

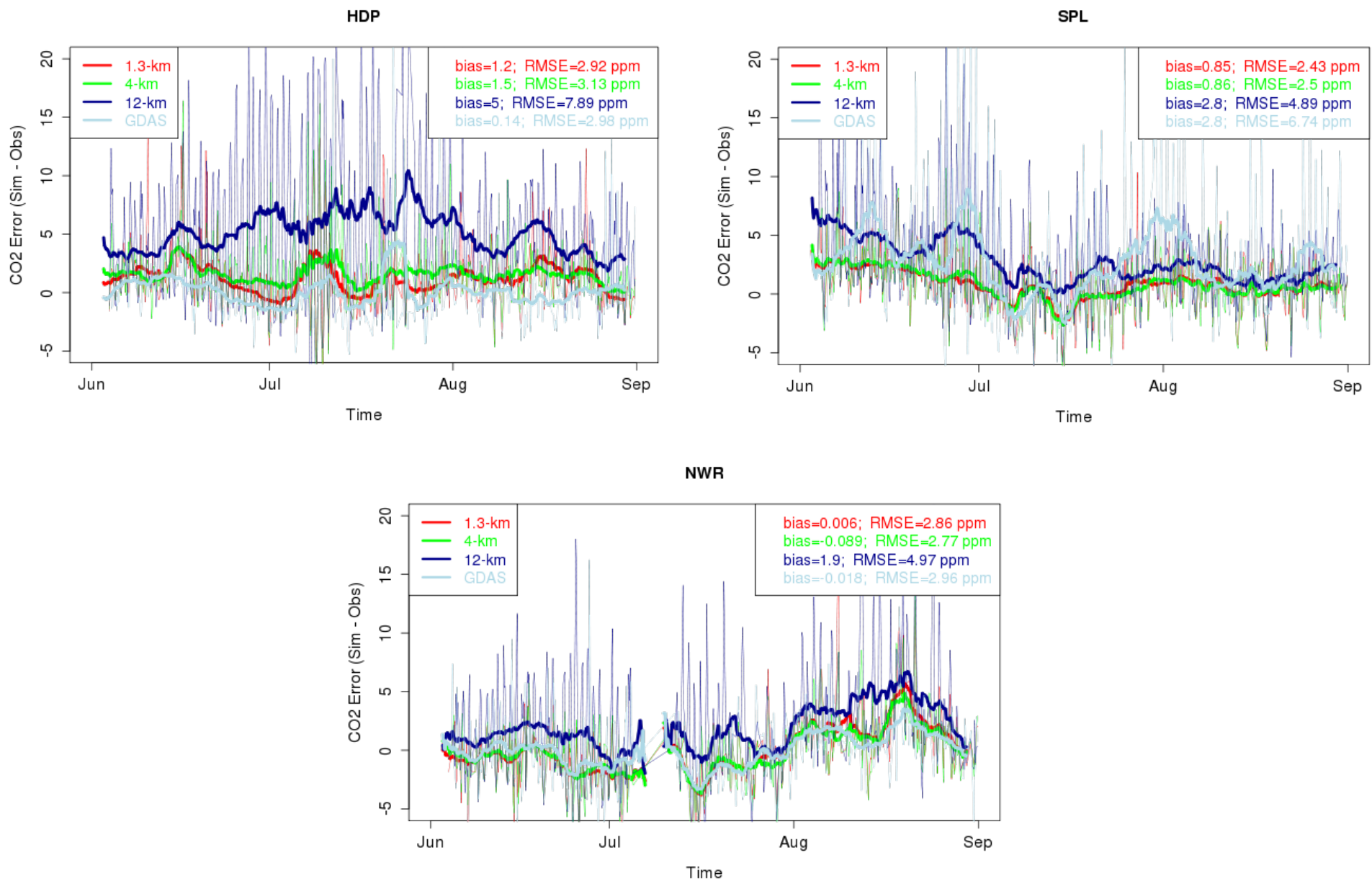
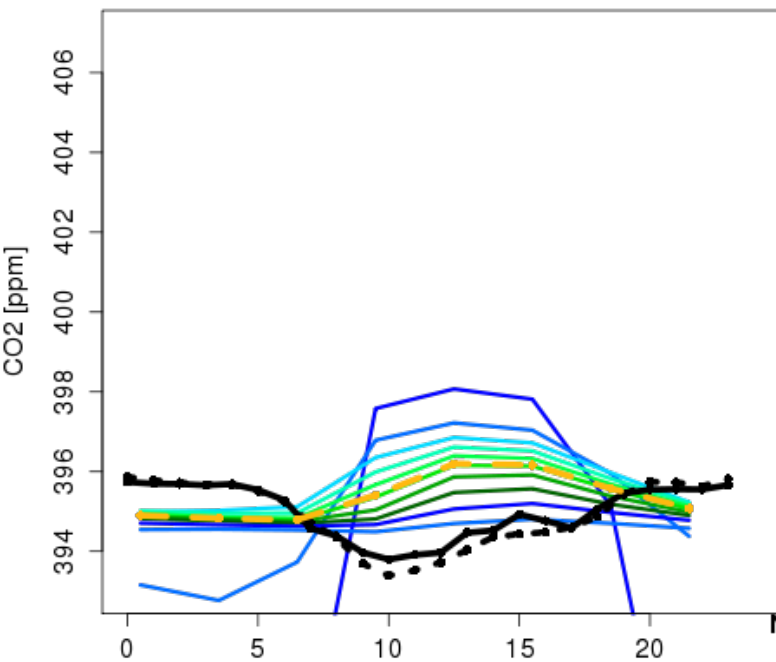
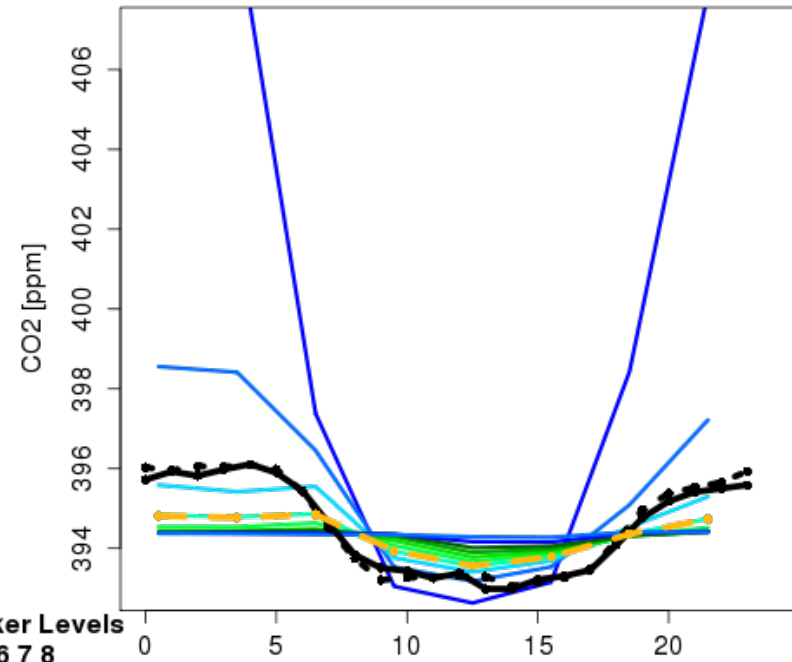


Fig. S5

HDP CarbonTracker Levels
2012 MONs: 6 7 8

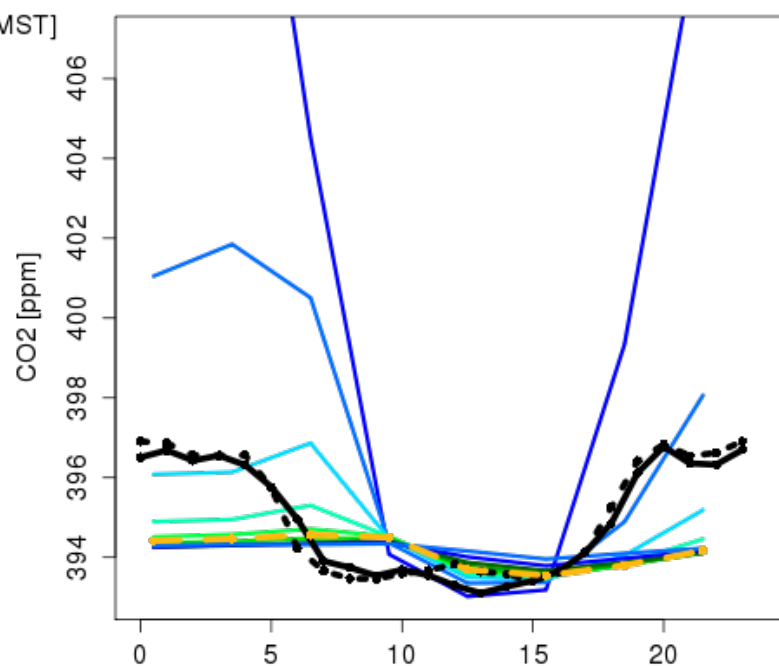


SPL CarbonTracker Levels
2012 MONs: 6 7 8



NWR CarbonTracker Levels
2012 MONs: 6 7 8

Time of Day [MST]



Time of Day [MST]

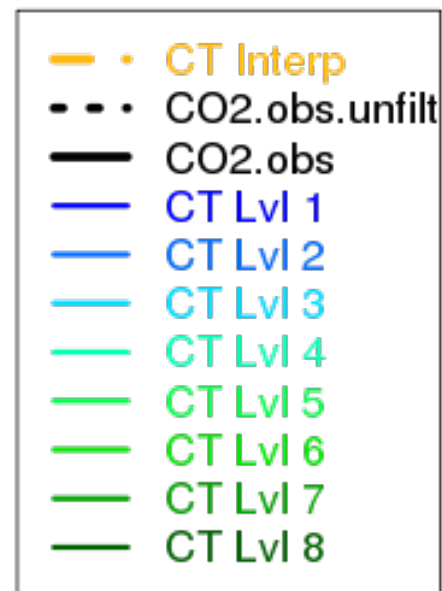
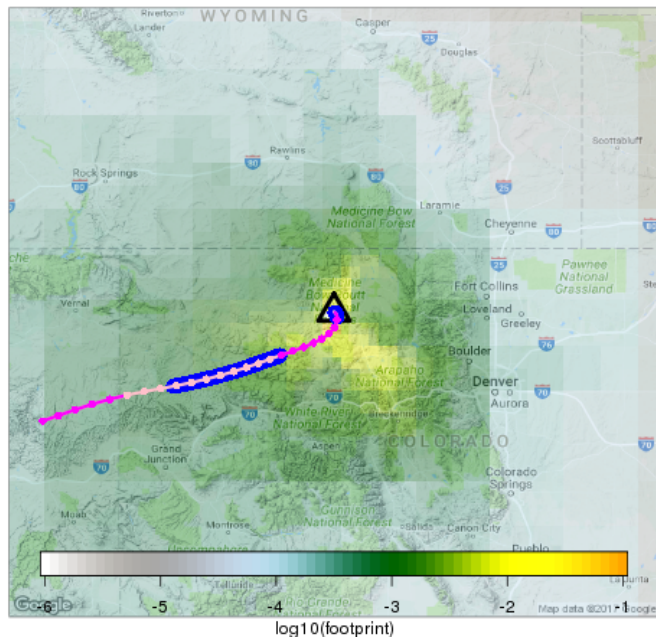
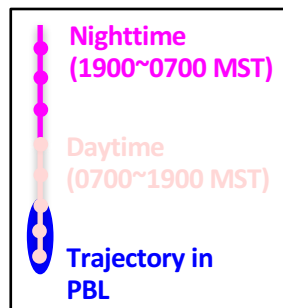


Fig. S6

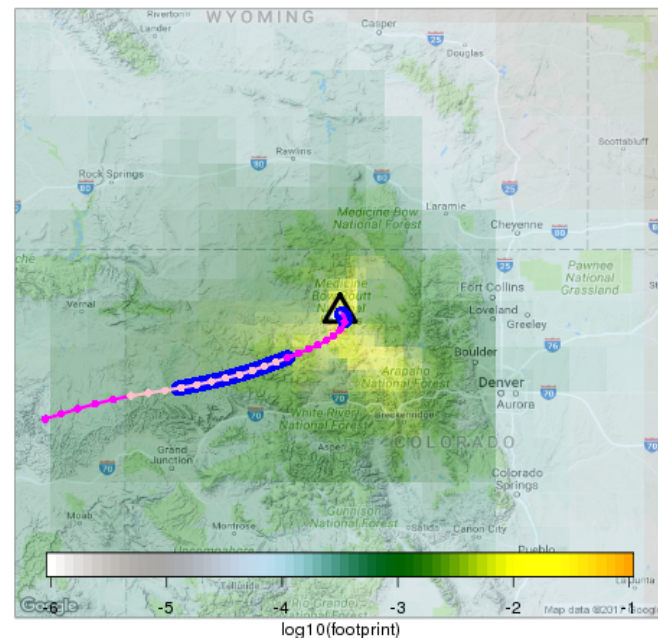
SPL ave footprint: WRF-1.3km (AGL)



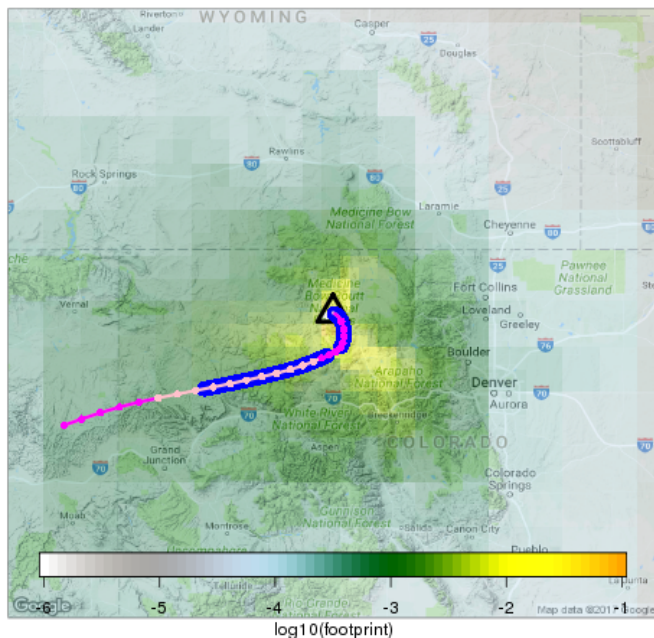
0900 UTC
(0200 MST)



SPL ave footprint: WRF-4km (AGL)



SPL ave footprint: WRF-12km (AGL)



SPL ave footprint: GDAS-1° (ASL)

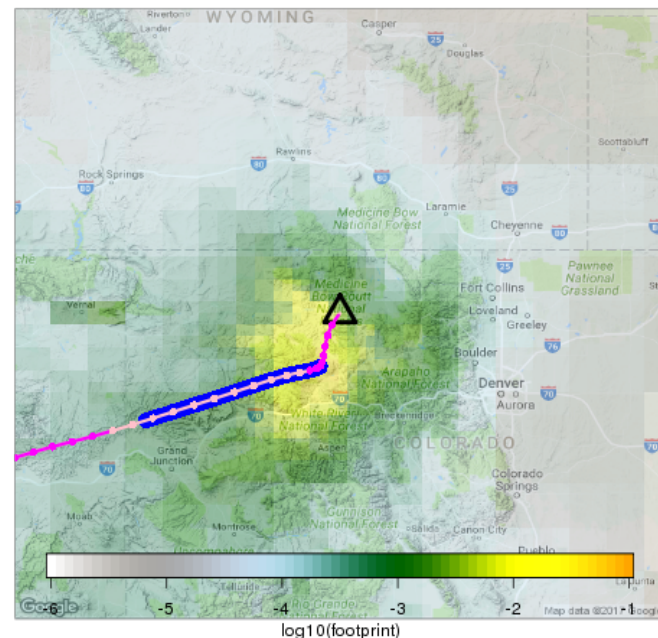
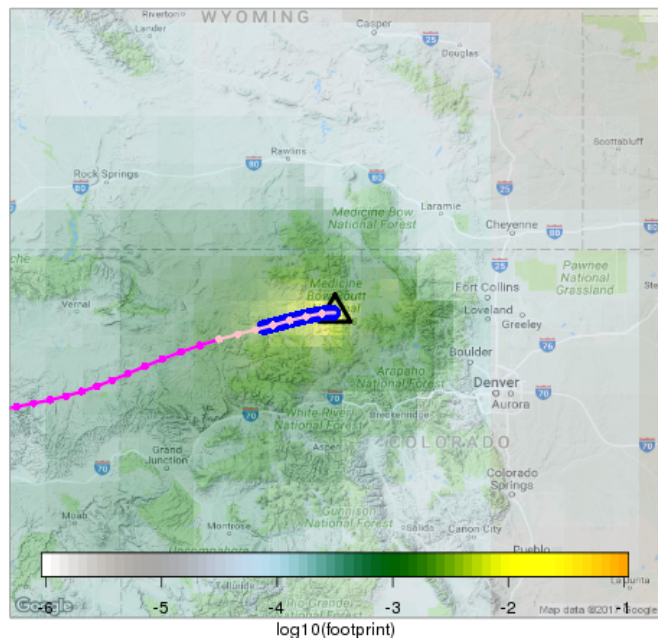
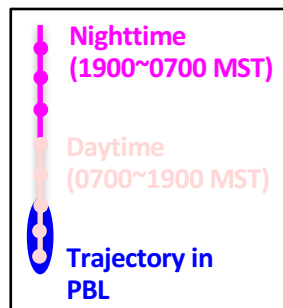


Fig. S7

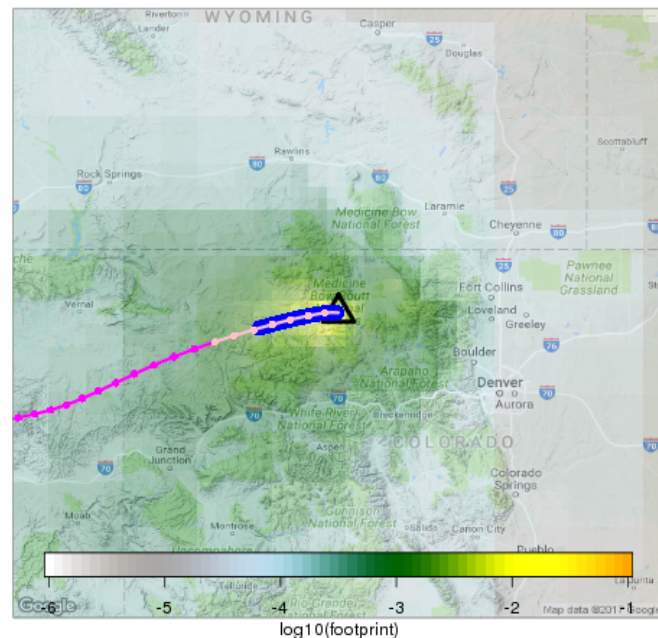
SPL ave footprint: WRF-1.3km (AGL)



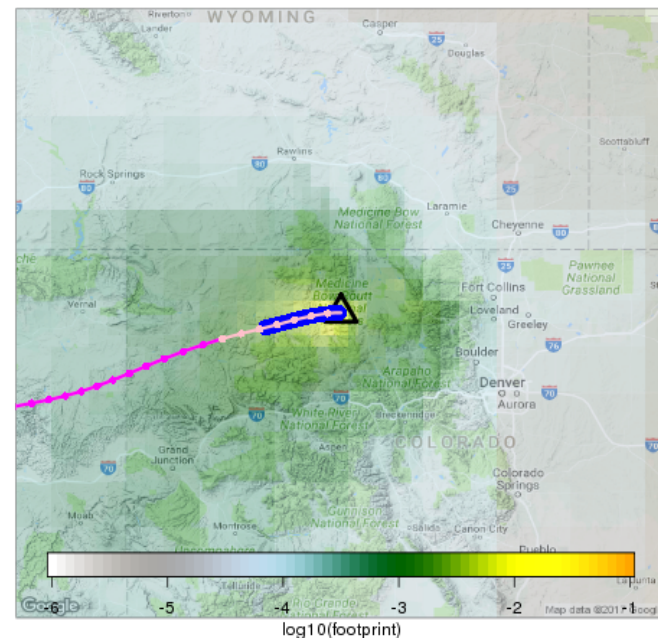
2100 UTC
(1400 MST)



SPL ave footprint: WRF-12km (AGL)



SPL ave footprint: WRF-4km (AGL)



SPL ave footprint: GDAS-1° (ASL)

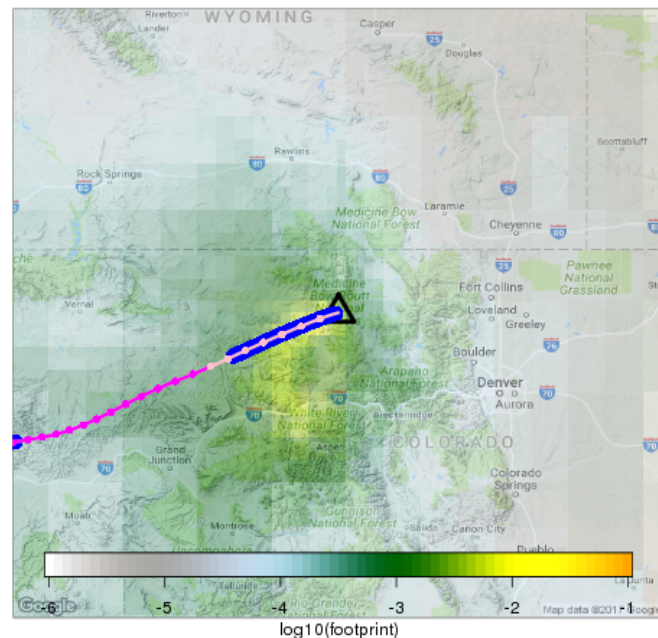
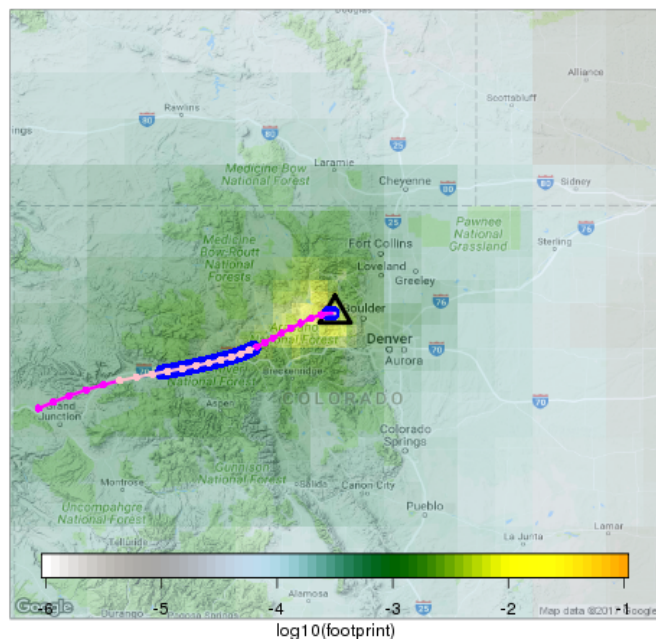
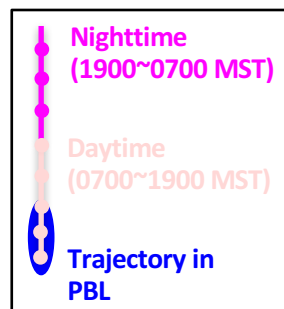


Fig. S8

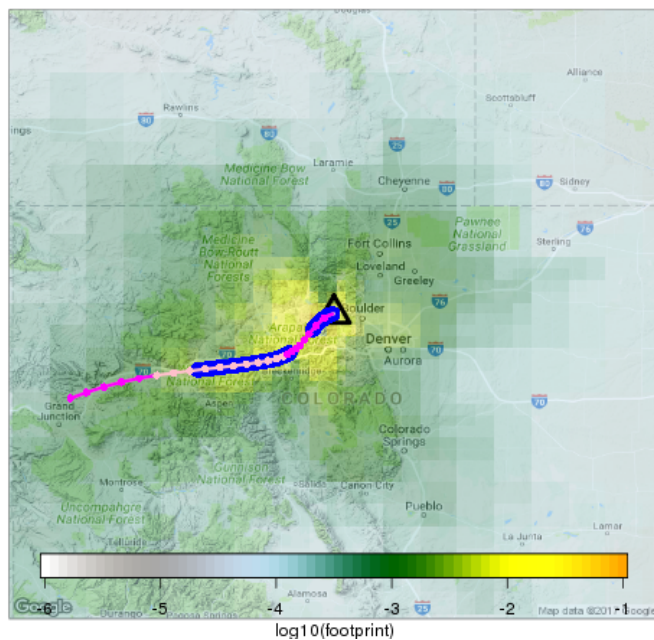
NWR ave footprint: WRF-1.3km (AGL)



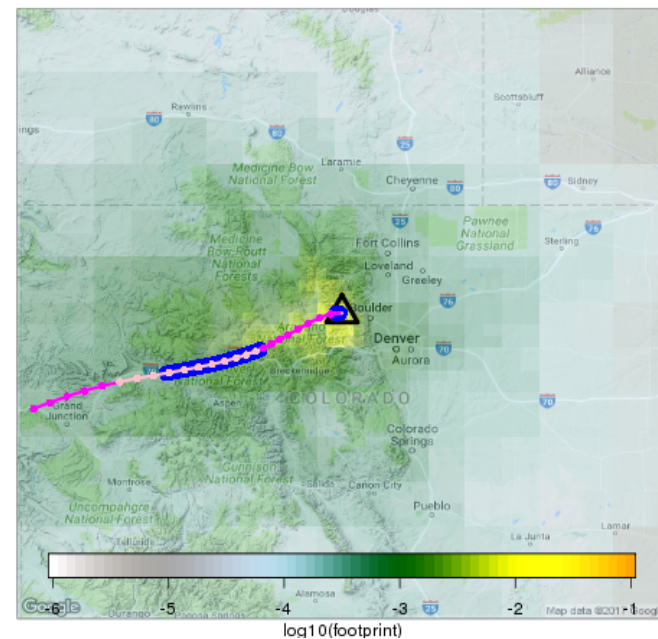
0900 UTC
(0200 MST)



NWR ave footprint: WRF-12km (AGL)



NWR ave footprint: WRF-4km (AGL)



NWR ave footprint: GDAS-1° (ASL)

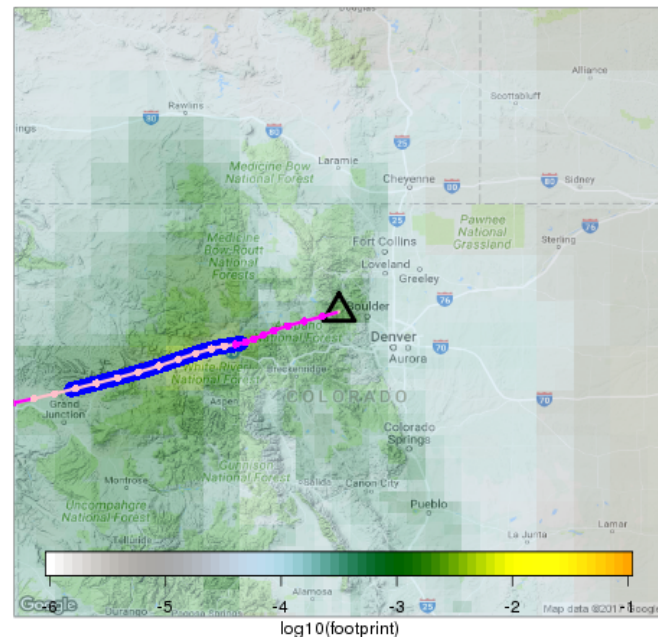
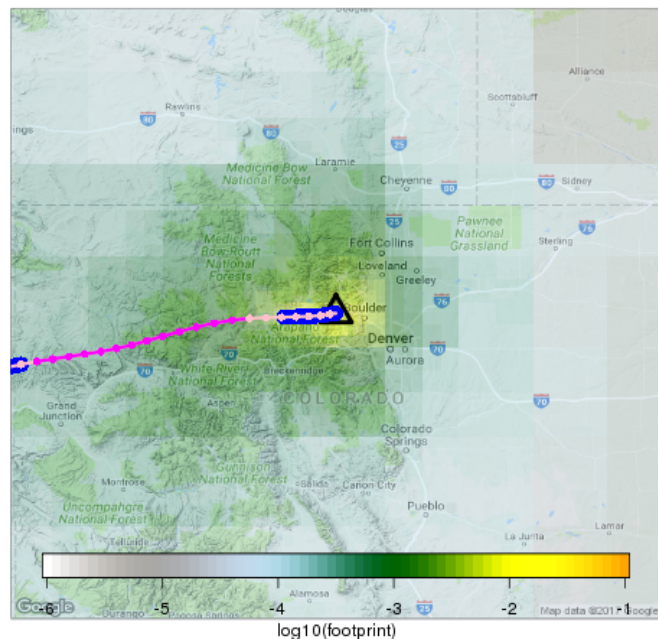
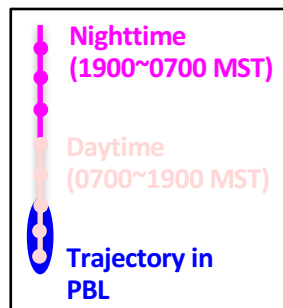


Fig. S9

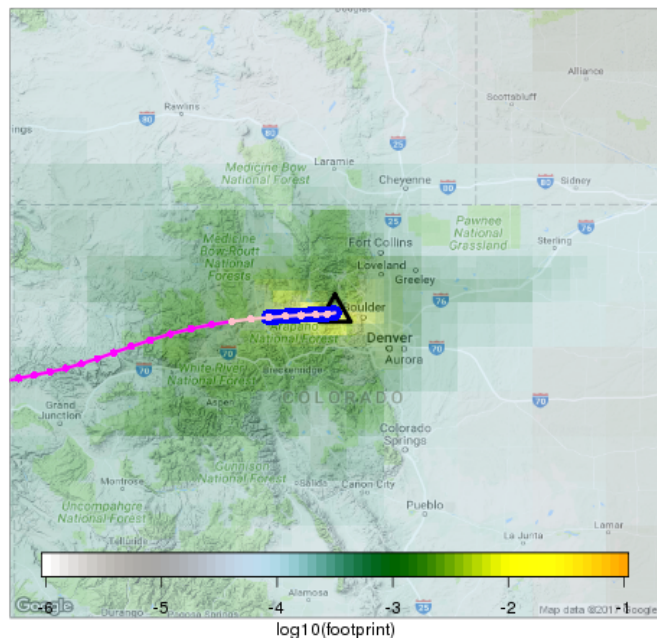
NWR ave footprint: WRF-1.3km (AGL)



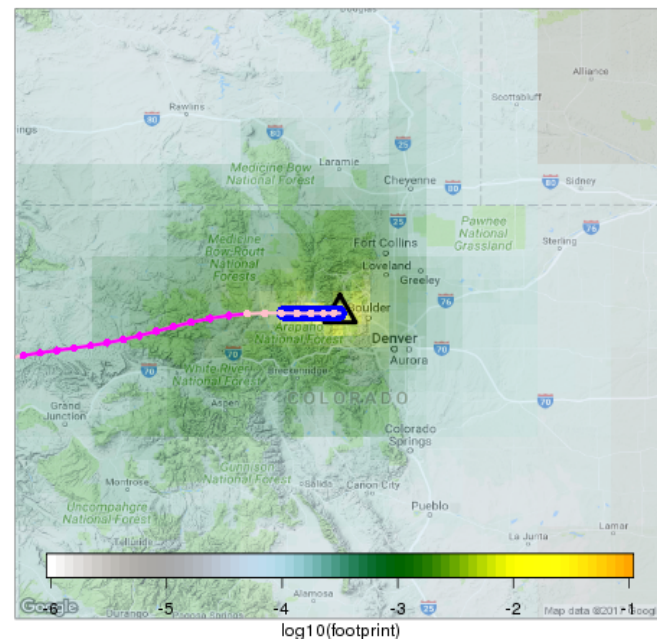
2100 UTC
(1400 MST)



NWR ave footprint: WRF-12km (AGL)



NWR ave footprint: WRF-4km (AGL)



NWR ave footprint: GDAS-1° (ASL)

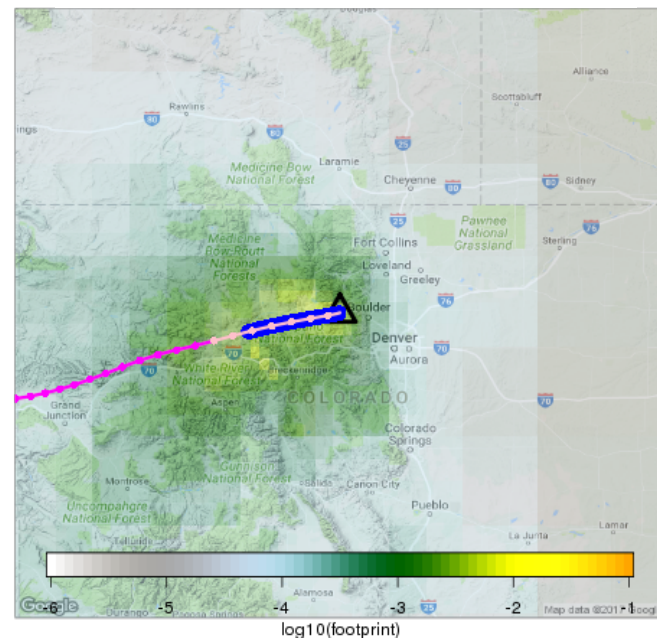
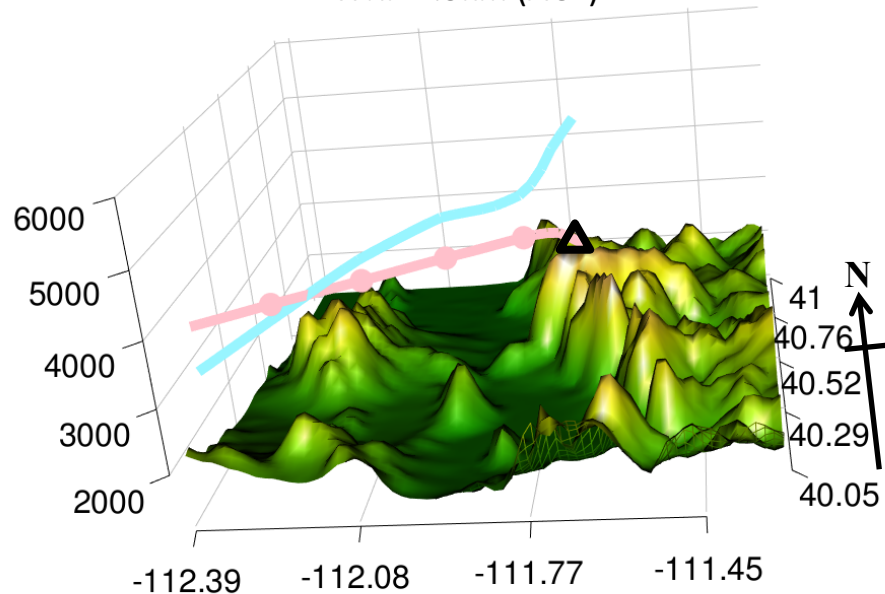


Fig. S10

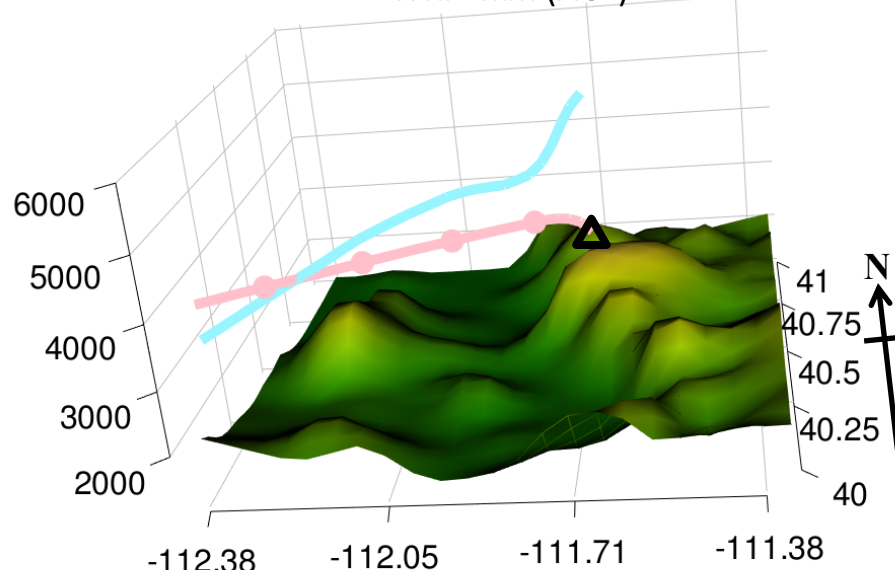
HDP: Mean 3D Trajectory of Stochastic Particles & PBL ht for Different Runs

2100 UTC (1400 MST)

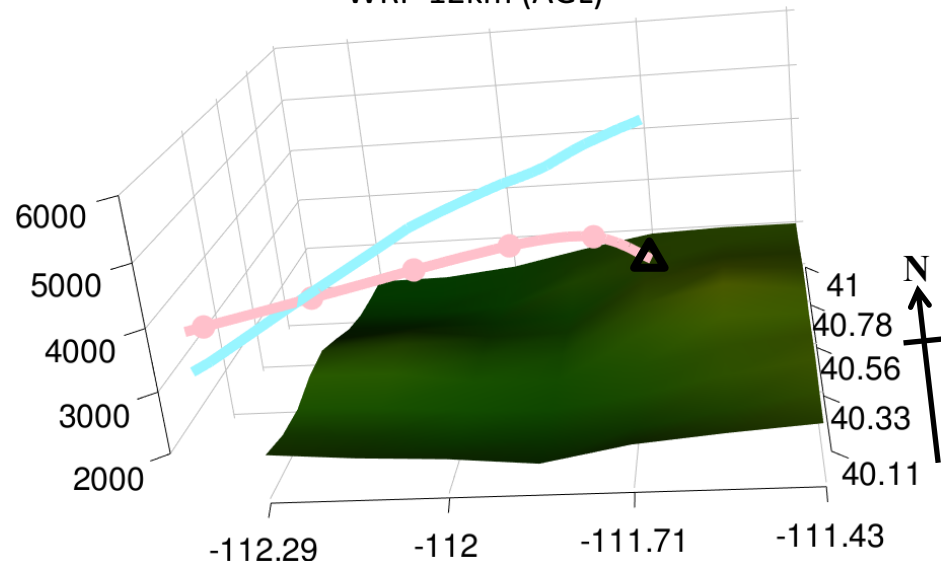
WRF-1.3km (AGL)



WRF-4km (AGL)



WRF-12km (AGL)



GDAS-1° (ASL)

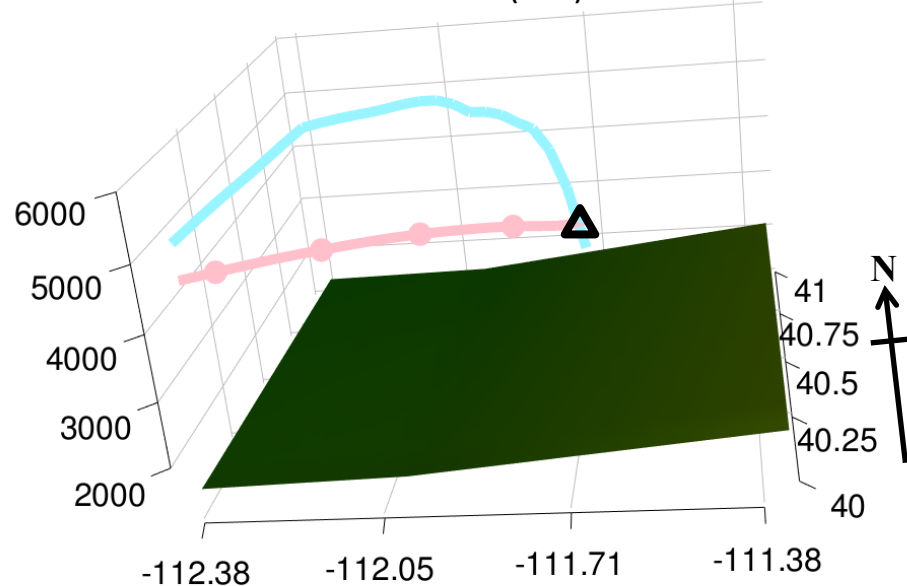
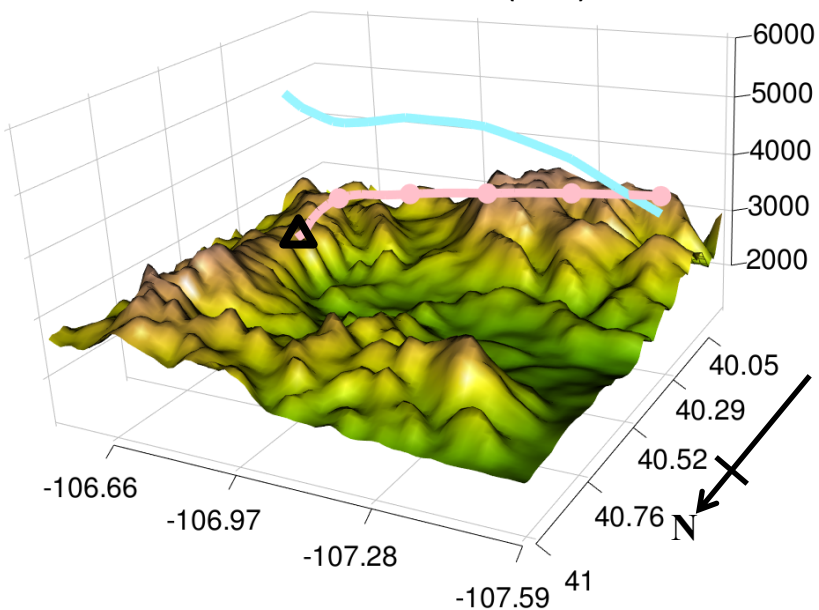


Fig. S11

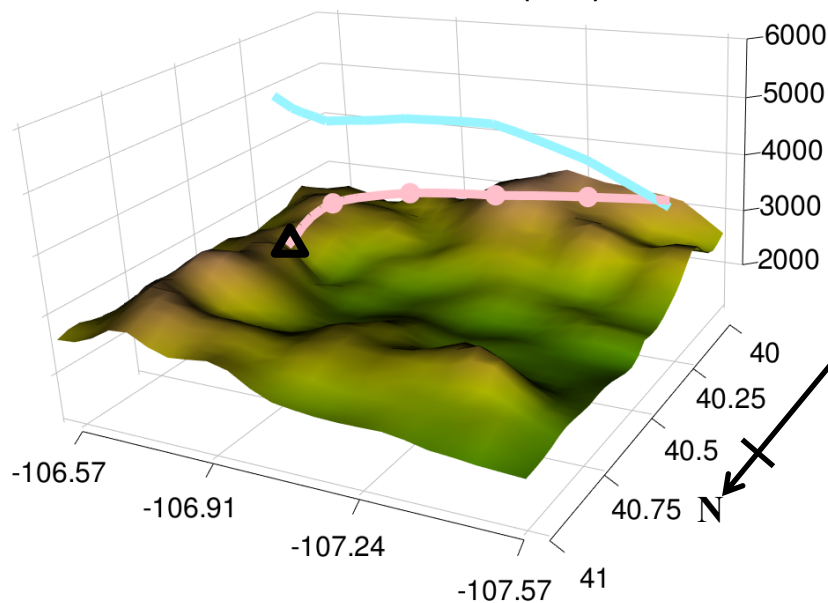
SPL: Mean 3D Trajectory of Stochastic Particles & PBL ht for Different Runs

2100 UTC (1400 MST)

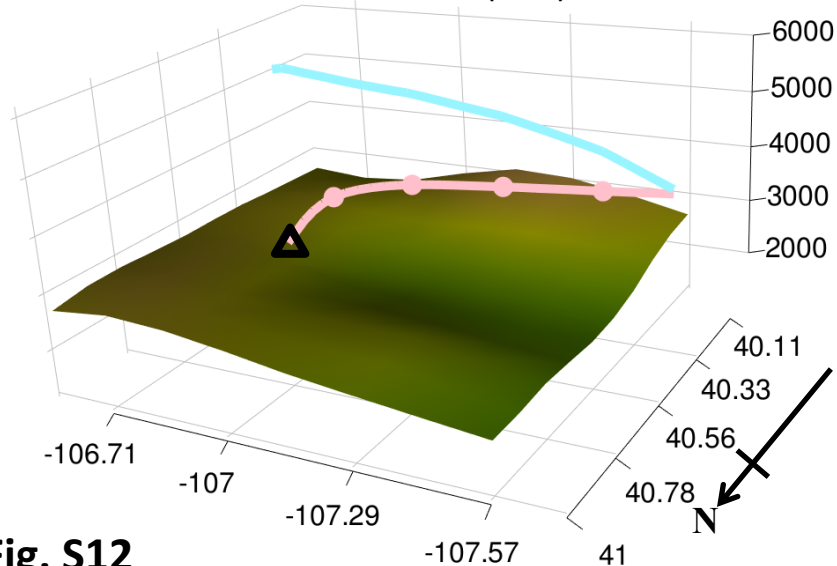
WRF-1.3km (AGL)



WRF-4km (AGL)



WRF-12km (AGL)



GDAS-1° (ASL)

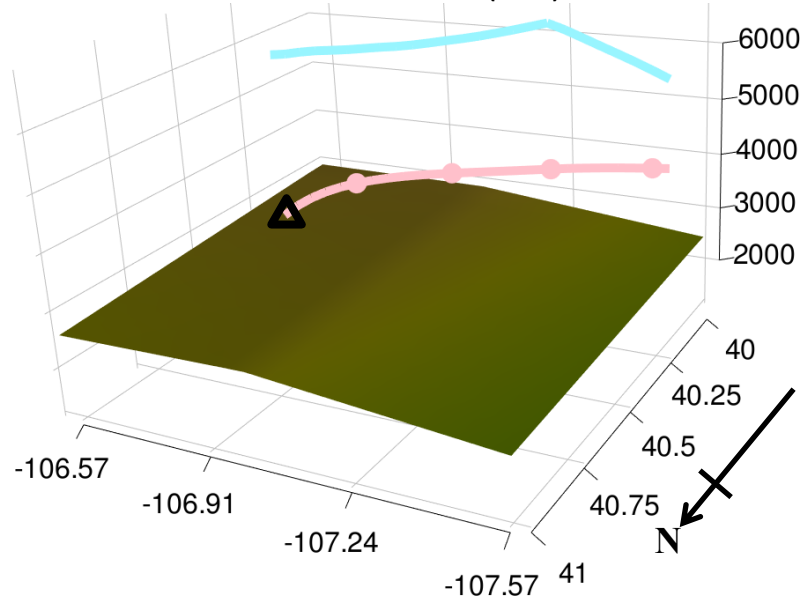
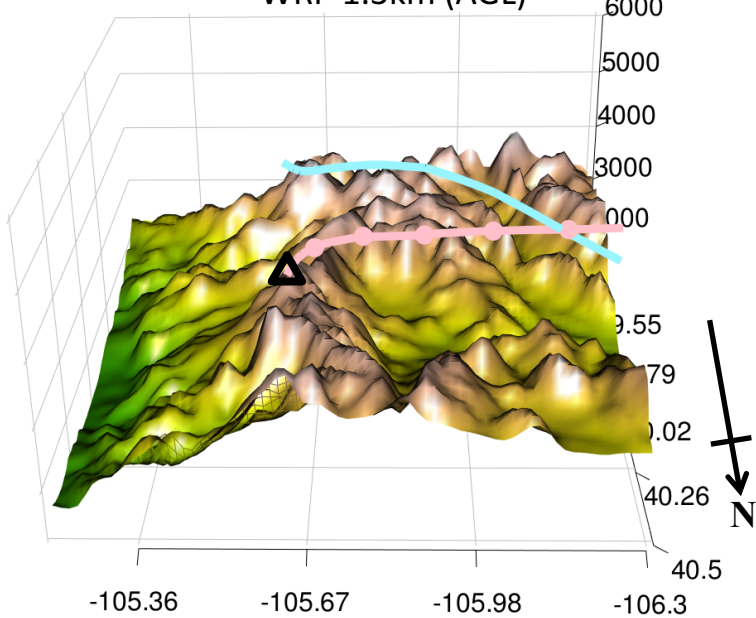


Fig. S12

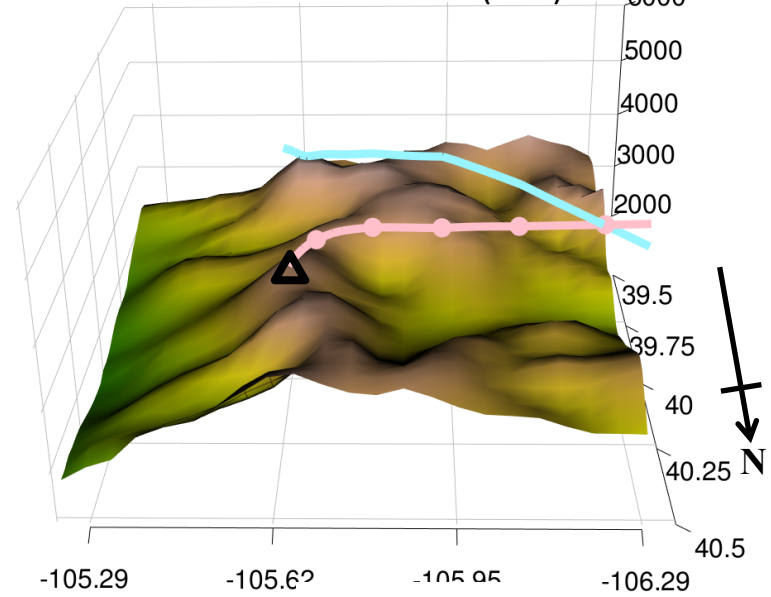
NWR: Mean 3D Trajectory of Stochastic Particles & PBL ht for Different Runs

2100 UTC (1400 MST)

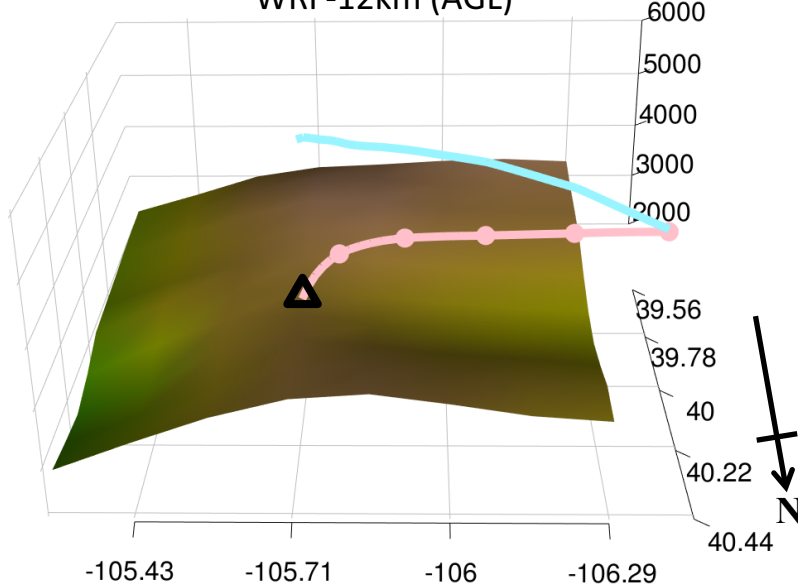
WRF-1.3km (AGL)



WRF-4km (AGL)



WRF-12km (AGL)



GDAS-1° (ASL)

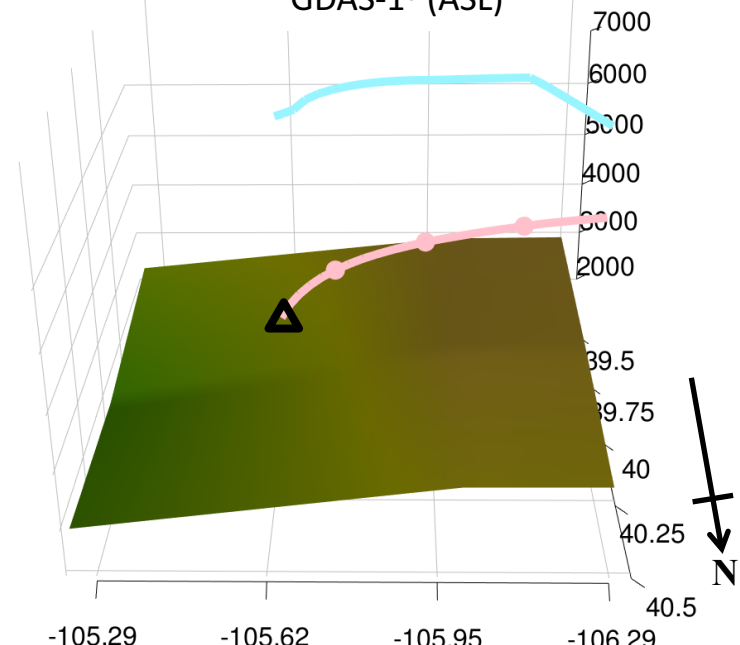


Fig. S13