Response to Comments from RC1

We appreciate the time and efforts by the editors and referees in reviewing this manuscript. We have addressed each of the concerns indicated in the review reports. Please see the one-to-one response (in blue) following the comments from the reviewer RC1. We believed that the revised version meets the journal publication requirements.

General Comments

'The paper of Feng et al. entitled 'LA Megacity: a high-resolution land-atmosphere modelling system for urban CO_2 emissions' compares different model resolutions and emission maps to identify optimal configurations for simulating CO_2 fields over a megacity. Although this concept of comparing different models or model configurations is not new, urban air quality poses some additional challenges that the authors try to address in this paper. Additionally, they pay attention to monitoring requirements and their new network design methodology can certainly prove useful, also to estimate footprints. However, I believe the authors could stress more the importance and novelty of their study in the context of recent studies, as the summary of current literature lacks an overview of knowledge gaps/remaining challenges and how their study fits into this (except for the paragraph about studies that focused on LA). Other than that I thank the authors for their very nice work.

Specific comments

Why have the authors decided to use one-way nesting? What would be the advantage compared to two-way nesting and what are the consequences?

One-way nesting allows the parent and the nest to exchange information strictly downscale. In this way, the nest solution does not feed back to the parent solution. Two-way nesting allows the information exchange bi-directionally. The nesting feedback impacts the parent domain's solution. This study evaluates the impact of the model physics and grid spacing on the model performance. In this case, one-way nesting is preferable to two-way nesting by which the 4-km model results will just be the smoothed 1.3-km model results.

The authors have chosen to simulate a two-month period per day, rather than doing the whole period in one simulation. This requires reinitialisation of the concentration fields for each day. How do the authors ensure conservation of mass between the simulations? Could you show that this reinitialisation has no impact on the simulated mass fractions?

Reinitialisation is commonly used in weather forecasts and regional modelling methods to prevent simulation drifting too much away from the truth. Running simulations for one-month long without reinitialisation is not proper. However, one should notice that the re-initialisation was only applied to modelled meteorology. The CO_2 fields were carried over from cycle to cycle without any re-initialization. The CO_2 mass therefore was conserved for the entire simulation.

Could the authors clearly specify whether the temporal variations for both emission product are equal? If not, how do they differ and what would be the consequence for the comparison of the products?

Both emission products were developed using "bottom-up" methods. Vulcan quantifies FFCO2 emissions for the entire contiguous United States (CONUS) hourly at approximately 10-km spatial resolution for the year of 2002. The temporal variations are driven by a combination of modeled activity (building energy modeling) and monitoring (power plant emissions). Hestia-LA is a fossil fuel CO2 emissions data product specific in space and time to individual buildings, road segments, and point sources covering the Los Angeles megacity domain for the years of 2011 and 2012. Hestia-LA uses much of the same information for the temporal variations except for the onroad emissions, for which local traffic data is employed as opposed to regional traffic data. Given the similarities, it is unlikely that the small difference in temporal variation could account for the spatial differences, through covariation with atmospheric transport, found here.

Given the limited in-situ GHG measurements that were available for CalNex, we mainly focused on the CO_2 concentration spatial differences over the LA basin caused by the different emission products used. One of the main conclusions of this study is that, driven by the high-resolution emission data product, i.e. Hestia, the model can reproduce the plumes from the point sources. On the other hand, the Vulcan run shows a more smeared-out CO2 distribution over the LA basin (Figure 9b vs. Figure 9c).

The authors state that for the MYNN_UCM configuration the PBL height is better represented for d03 than for d02 and that this is also reflected by other configurations. However, it appears from figure 4 that for some configurations d02 is actually better during the afternoon. This requires some reflection in the text.

Thanks for the suggestions. The text has been modified for reflecting this concern (Page 13 Line 30).

Are the biases shown in Figure 6 for the whole period, including night time?

Figure 6 shows the statics over daytime only. The clarification has been added in the figure caption in the revised paper. Thanks for the comments.

If so, how do the authors reach the conclusion that the dryness in the model causes a lower PBL height (Figure 4) in the afternoon, while the PBL height is actually higher a bit earlier during the day? I would like to see a clear explanation for this, as generally I would think that dryness would cause a higher PBL height.

Yes, dryness usually leads to the higher PBL height. Thanks for pointing out this error. The model overall dry the LA basin but with some exceptions, such as Pasadena area where the ceilometer was deployed, where the model actually moisteners the air. The moistness is consistent with the lower PBL the model simulated in Pasadena.

Page 15, ln. 23-24: 'However, during daytime, with well-mixed conditions, the discrepancy between the WRF-Hestia and WRF-Vulcan runs becomes smaller.'; and similarly: Page 16, ln. 15-17: 'For the same reason, we show that $FFCO_2$ emissions do not play a dominant role around 1400 PST unless there are strong local signals...'. This is an interesting note. Usually, well-mixed daytime concentrations are sampled for inverse modelling, as these conditions are usually better represented by models. That leads to the question how well we could estimate posterior fluxes if a 40% increase in $FFCO_2$ emissions only leads to an increase of less than 1% in the total CO_2 concentration (which is a rough estimate from your Figure 8 at 1400 PST using both 1.3 km simulations). Could the authors digress a bit on the consequences of this note for inverse modelling?

True. Well-mixed daytime concentrations are sampled for inverse modelling, as these conditions are usually better represented by models. However, it should be borne in mind that removing the upwind background value is required in atmospheric inversion (Lauvaux et al., 2012); only ΔCO_2 is used in atmospheric inversion, not total CO_2 concentration (CO_2tot). How to derive ΔCO_2, or, say, determine the background CO_2 (CO_2bkg), from the interested location remains challenges (e.g., Turnbull et al., 2015; Schuh et al., 2010). One of the common ways is subtracting the upwind CO_2 from the downwind location. Figure 8 shows the diurnal variation of CO_2tot. Roughly, if we consider CO_2 concentration at the PV site as CO_2bkg (396 ppm for 1.3-km WRF-Hestia and 397 pm for 1.3-km WRF-Vulcan), with 408 ppm and 405 ppm of CO_2tot at Pasadena, ΔCO_2 for Pasadena is 12 ppm and 8 ppm for 1.3-km WRF-Hestia and 1.3-km WRF-Vulcan, respectively. In this case, the increase of $FFCO_2$ (mixing ratio) for 1.3-km WRF-Hestia vs. 1. .3-km WRF-Vulcan is about 50%, which is close to your estimation.

Section 5 introduces a new network design method. Although mentioned before that this would be discussed, I would like to see a few sentences discussing the need for such new method and the limitations of other methods. Currently, this is only briefly mentioned in the discussion. Could the authors also make a recommendation on which method would be most suitable for future use?

Thank you for your suggestions. We have added more sentences discussing the need and limitation of the correlation method in section 6. See Page 25 Line 8-18.

The new method assesses the correlation of “observed CO_2” with the neighbouring CO_2 concentration based on the forward model simulation. First of all, this method is computationally economical relative to the footprint method. Secondly, the method doesn’t require adjoint models, which can avoid the complexity. Most importantly, it brings extreme flexibility without complexity for various platforms (i.e., in-situ, satellite, etc.) and especially outpaces the analysis for the dense sampling techniques, such as remote sensing dataset. Applying the footprint methods to satellite data at the regional scale modelling is extremely computationally time-consuming and complex.

However, as mentioned in the text, both transport and emissions play a role in the correlation method. The footprint method, in contrast, indicates the influence of the atmospheric transport to the location of the observation only. Hence, the correlation method is subject to overestimation of the influence area versus the footprint method, due to the complicated nature of the atmospheric integrator.

Technical corrections

In Section 3.1 the authors list five criteria for profile selection. The difference between point 4 and 5 should be made more clear.

These two criteria have been merged. See Page 12 Line 18. Thanks!

In Section 3.4, the third paragraph, the authors mention the temperature difference between Granada Hills and downtown LA in F. I would suggest to use Kelvin to make comparison with the other temperature results in Kelvin easier.

Changed. See Page 16 Line 4.

In Section 5, please mention clearly whether you used any data selection or that all data was included for the correlation maps.

There are no data used in Section 5. See Page 21 Line 21 for clarification.

The discussion now starts with new results based on flask samples of radiocarbon. Please move this to the results section. Also I would suggest to introduce the use of radiocarbon earlier, as this not mentioned previously in the paper.

The comparison with the flask samples and the introduction of radiocarbon have been moved to section 3.6 following the comparison to in-situ measured total CO_2. Thanks!

Reference:

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337-354, 10.5194/acp-12-337-2012, 2012.

Schuh, A. E., Denning, A. S., Corbin, K. D., Baker, I. T., Uliasz, M., Parazoo, N., Andrews, A. E., and Worthy, D. E. J.: A regional high-resolution carbon flux inversion of North America for 2004, Biogeosciences, 7, 1625-1644, 10.5194/bg-7-1625-2010, 2010.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., Patarasuk, R., and Razlivanov, I.: Toward quantification and source sector identification of fossil fuel CO_2 emissions from an urban area: Results from the INFLUX experiment, Journal of Geophysical Research: Atmospheres, 120, 2014JD022555, 10.1002/2014JD022555, 2015.

Response to Comments from RC2

We appreciate the time and efforts by the editors and referees in reviewing this manuscript. We have addressed each of the concerns indicated in the review reports. Please see the one-to-one response (in blue) following the comments from the reviewer RC2. We believe that the revised version meets the journal publication requirements.

Overview: The manuscript presents simulated carbon dioxide fields for 2 months centered over Los Angeles. The work demonstrates and tests the ability of a high-resolution meso-scale model to reproduce observed meteorological and carbon dioxide dynamics, with a focus on urban areas, LA in particular. The paper presents a valuable modelling approach in order to understand the temporal and spatial variability of weather variables and CO_2 mixing ratio in urban and background sites. This work is appropriately placed in ACP, and contributes to the burgeoning area of studying carbon emissions from urban areas. I have some general and specific concerns delineated below, that need to be addressed before its publication.

General Comments: Overall things look quite nice and interesting, but I have a couple of reservations that require more explanation and must be addressed. There needs to be better presentation of modelled vs observed fields in terms of table of scores and 1:1 plots. As currently presented, it is difficult to assess model performance. The second point is that discussions on the physical reasons why a parametrized scheme is better, or on the performance of the modelling, are missing. The last parts that study correlations of the simulated CO_2 fields with GHG measurements is interesting, and well oriented to further inverse modelling studies. I do not have specific remarks on this part.

1) **CO_2 initial and boundary condition**. This is only briefly touched upon in section 2.1, and it is unclear. From what I understand the model is initialized and coupled with CO_2 concentrations coming from observations. The simulations run for 36h. Do you use the predicted CO_2 field from the end of the previous day to start the following day ? Or do you only use CO_2 observations at the beginning of each run ? In the 2nd case, what is the spin-up time ? Is there a significant horizontal and vertical variability in the CO_2 observations ? What impact do varying boundary condition choices make on simulations? We know that in regional studies boundary

conditions play a tremendously important role (Lauvaux et al. TELLUS 2012). The authors must better described what they've done for boundary conditions, and make quantitative assessments of impacts of boundary condition choices on simulations.

We initialized CO_2 fields from the NOAA curtain dataset at the beginning of the first cycle. The simulation runs for 36 hour for each cycle with 12-hour setback for spin-up. For each cycle, only the meteorology is re-initialized; CO_2 fields are carried over from the last cycle. For instance, the first simulation cycle is 00 UTC 15 May to 12 UTC 16 May 2010, and the second cycle is 00 UTC 16 May to 12 UTC 17 May 2010. The initial conditions for 00 UTC 15 May include NARR, NCEP SST and NOAA curtain (CO_2). The initial conditions for 00 UTC 16 May include NARR, NCEP SST and WRF-modelled CO_2 on 00 UTC 16 May from the previous cycle. Briefly, we did not re-initialize CO_2 for each cycle to assure mass conservation over the model domain. The clarification for CO_2 IC and BC has been added in the revised paper (see Page 11 Line 27-29).

We agree that the boundary conditions (BCs) are critical for the CO_2 simulations. In this study, we found there is no significant horizontal and vertical variability in the NOAA curtain dataset; semi-constant BC was used. We have also applied CO_2 modelled by GEOS-Chem BC for our region of interest, which introduced ~+10 ppm model-data mismatch in the WRF model results. This is similar to the findings by (Lauvaux et al., 2012), who found the model-data mismatch was more than 20 ppm in summer over the corn belt area. It also reflects the challenges in determining CO_2 background values for regional scale simulations. We therefore end up with using semi-constant values ("NOAA Curtain") as the model BC in the paper. The NOAA Curtain dataset mainly represents oceanic clean air. In May – June, west to southwest clean marine flow prevails over the Los Angeles Megacity. Using a semi-constant dataset is fairly close to the reality, introducing lower errors to the regional, modelled CO_2 relative to global models, such as GEOS-Chem. However, during October to March, Santa Ana wind events occur frequently, during which easterly to north-easterly winds predominate over the LA basin, and the oceanic air is polluted. In this case, using constant values is no longer feasible.

2) As a large part of the simulated domains is on the sea, and as LA is largely influenced by maritime air masses, is it not a problem to ignore ocean fluxes ? Classically, **oceanic CO_2 fluxes** are parameterised following Takahashi et al. (1997). A sensitivity test with ocean

parametrized fluxes would be appreciated.

The LA megacity is one of the top three fossil fuel emitters in the U.S. Roughly estimated from Hestia at the Pasadena site, the order of fossil fuel emission is about 10-20 umol/m^2/s. The typical oceanic CO_2 flux -0.15 umol/m^2/s (Torres et al., 2011), 0.2 umol/m^2/s (Mu et al., 2014), represents only 1-2% of $FFCO_2$ fluxes and even less compared to CO_2tot. Because of that, we have ignored the oceanic CO_2 signal for simplicity in this study. Yet we do agree that a sensitivity test with oceanic flux would be interesting and should be included in future work. This explanation has been added to the revised paper. See Page 11 Line 18-23.

3) One objective of the paper is to assess the **PBL schemes**, but they are not physically described and the differences between the schemes are not presented . Therefore the conclusions are only limited to WRF technical configuration and physical aspects are not addressed. The 3 PBL schemes have to be described properly (closure, mixing lengths ...) to highlight the differences. Then strengths and weaknesses of each scheme need to be highlighted relating to their characteristics.

In this study, we have selected three most commonly used TKE-driven PBL schemes for comparison, including MYJ, MYNN, and BouLac. MYJ (Janjić, 1994) determines the PBL from the TKE where the PBL top is defined as the height where the TKE profile decreases to the threshold of 0.2 m^2s^{-2}. MYNN2 (Nakanishi and Niino, 2006) is tuned to a database of large eddy simulations (LES) in order to overcome the typical biases associated with other MY-type schemes, such as insufficient growth of convective boundary layer and under-estimated TKE. Additionally, MYNN also considers sub-grid TKE terms, and it determines the PBL top as the height at which the TKE falls below 1.0×10^{-6} m^2 s^{-2}. BouLac (Bougeault and Lacarrere, 1989) has an option designed for use with BEP multi-layer and UCM. It determines PBL top at which TKE reaches 0.005 m^2 s^{-2}. They all are 1.5 order local closure schemes that only consider immediately adjacent vertical levels in the model, which may not fully account for deeper vertical mixing associated with larger eddies and associated countergradient flux correction terms and, thus, tends to prevent the PBL from mixing as deeply to produce cooler and moister conditions. On the contrary, the non-local closure schemes considering a deeper layer account for countergradient fluxes and, thus, generally represent deep PBL circulation better than local schemes. The PBL schemes were reviewed by Cohen et al. (2015).

The main reason that we focus on the TKE-driven PBL schemes only is that the explicitly estimated turbulence fluxes can be used to drive Lagrangian particle dispersion models to computer influence footprints for subsequent atmospheric inversions. Through the model evaluation, we aimed to determine an optimal model configuration for modelling urban CO_2 over the LA megacity, and eventually to use the same system for synthesis analysis in future. In this study, we concluded that MYNN in combination with UCM is optimal for the LA modelling framework, which is consistent with the findings of Coniglio et al. (2013) who showed MYNN supports deep convection springtime.

The strengths and weaknesses of each scheme with their characteristics have been added to the revised paper. See Page 9 Line 3-15. Thanks!

4) In the same way, 2 **urban surface schemes** are tested without having presented their physical differences. The scientific interest is therefore limited. We need to know the scientific reasons why UCM seems better.

UCM is a single-layer urban canopy model, representing urban geometry and 3-D urban surfaces such as walls, roofs and roads. Furthermore, the sensible heat fluxes from the surface are calculated with Monin-Obukhov similarity theory and Jurges formula. The important factor of anthropogenic heat (AH) and its diurnal profiles are included and added to the sensible heat flux from the street canyon (Chen et al., 2011). BEP allows a direct interaction between the buildings and the PBL. BEP considers the 3-D urban surface and the vertical distribute source of buildings and momentum sinks throughout the whole canopy layer. The effects of vertical and horizontal surfaces on momentum, TKE and potential temperature are included. However, BEP requires very high vertical resolution within the PBL and is only compatible with MYJ and BouLac PBL schemes. Given that BEP is computationally expensive, we only test it with BouLac in this study.

The scientific reasons to explain the urban schemes' characteristics have been added to the revised paper. See Page 14 Line 9-12. Thanks!

5) In the **comparison to aircraft PBL height**, the method to determine PBL height is based on the vertical virtual potential temperature gradient. Among the existing methods to determine this parameter (Ri number, parcel method ...), none is perfect. What is the impact of the choice of the

method on the results ?

We have used the vertical virtual potential temperature gradient and Ri number methods to determine PBL top (see Figures R1 and R2 below). Compared to the vertical virtual potential temperature gradient method, the Ri method shows larger bias in the modelled PBL top, deeper for daytime, shallower for nighttime, but the overall conclusion remains the same in terms of model inter-comparison, namely MYNN_UCM shows better agreement with ceilometer measured PBL height. We therefore show only the vertical virtual potential temperature gradient determined PBL in the text.

Figure R1. Absolute difference between the aircraft-determined and modelled PBL height for each profile (flt_yyyymmdd, blue bars) using virtual potential temperature gradient (top) and Richardson number (bottom). The pink bars in the last column represent the averaged bias over all of the profiles for each configuration. Note that the shorter the bar is, the better agreement the model has with the observations.

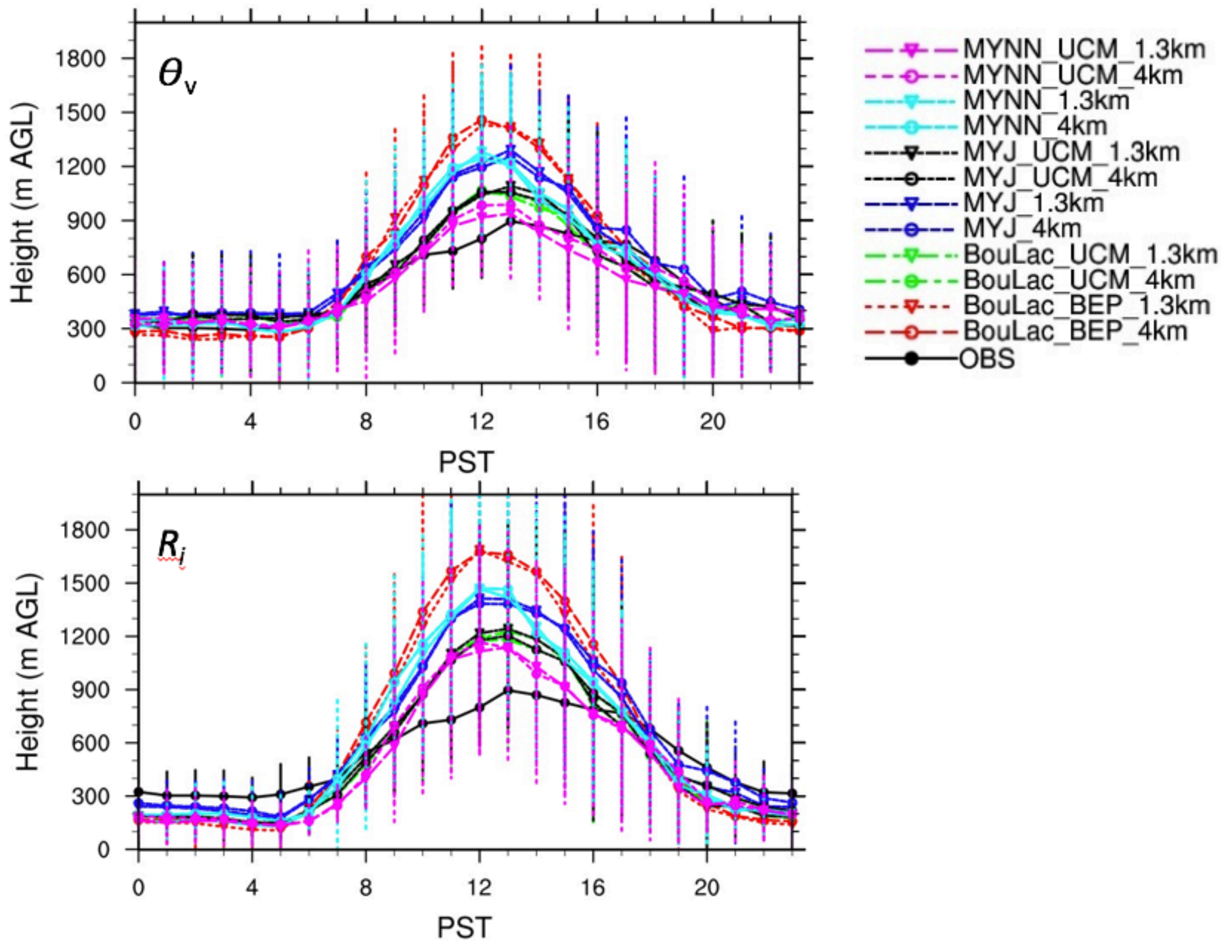

Figure R2. Average diurnal variation of the ceilometer-measured (obs) and modelled PBL heights at California Institute of Technology (Caltech) in Pasadena, CA during 15 May through 15 June 2010. Error bars indicate one standard deviations. Upper: the vertical virtual potential temperature (θ_v) gradient determined PBL. Lower: the Ri number determined PBL. Note that the ceilometer-measured PBL top (black solid line) is the same in these two panels.

For the 3 PBL schemes, biases on PBL heights are significant : errors of 160m in PBL height are not small by any measure. You can see for instance Riette and Lac (2016) for evaluation of PBL height over 1 year with an operation NWP model, with more satisfying values. Qualitative statements should be toned down. What is the error standard deviation? Figure 3 is not appropriate as only biases are represented without standard deviation, and without length scale. How do you also explain that biases are smaller at 4km than at 1.3km, and that the results are different than the comparison to ceilometer?

Please note that Figure 3 (in the manuscript) and Figure R1 (in the response) show the absolute difference between the observation and model for each aircraft profile we selected, so the error of 160 m in PBL height is the mean over seven aircraft profiles only (small sample). We did not intend to make any specific conclusion based on seven profiles. The take-home message of Figure 3 is that the differences between the modelled and aircraft-determined PBL height differ case by case, and none of the model physics options is systematically better than the others. To further define the optimal physics for the PBL height simulation, we presented the all-hours statistics with the ceilometer data in section 3.2 and Figure 4.

Given the relative large number of the ceilometer measurements, similar model evaluation (Table R1) to that of Riette and Lac (2016) has been done and been added to the revised paper (Table 3). Compared to the values evaluated by Riette and Lac (2016), -9.17 m for bias and 115 m for RMSE (PMMC09), the scores of MYNN_UCM fall in a comparable range.

Table R1. Comparison Statistics of model performance relative to the ceilometer data over 1100 – 1700 PST (unit: m AGL)

	Mean	Bias	Standard deviation	RMSE
OBS	835.7	-	223.8	-
MYNN_UCM_d03	828.8	-6.9	82.7	89.7
MYNN_UCM_d02	820.4	-15.3	66.1	94.5
MYNN_d03	1055.6	219.9	205.8	278.2
MYNN_d02	1029.4	193.7	200.0	254.3
MYJ_UCM_d03	961.4	125.8	154.9	168.8
MYJ_UCM_d02	971.4	135.7	109.3	157.7
MYJ_d03	1115.3	279.7	174.4	308.7
MYJ_d02	1105.1	269.5	150.9	291.6
BouLac_UCM_d03	936.1	100.5	147.3	149.9
BouLac_UCM_d02	958.7	123.1	104.8	148.7
BouLac_BEP_d03	1233.9	398.3	239.0	442.2
BouLac_BEP_d02	1244.3	408.6	219.5	446.0

6) **Dynamics** : why do you use one-way nested domains and not 2-way ?

One-way nesting allows the parent and the nest to exchange information strictly downscale. In this way, the nest solution does not feed back to the parent solution. Two-way nesting allows the information exchange bi-directionally. The nesting feedback impacts the parent domain's solution. This study evaluates the impact of the model physics and grid spacing on the model performance. In our experiment, one-way nesting is preferable to two-way nesting for which the 4-km model results will just be smoothed 1.3-km model results.

Advection and temporal schemes should be specified in Table 1, with the time steps for the different resolutions. Page 7 line 16 : what is the height of the 1st level ?

5th and 3rd order differencing for horizontal and vertical advection respectively are used. 3rd order Runge-Kutta is used for time integration with 45, 24, and 5 s for outermost, middle, innermost domains, respectively. These specifications have been added to Table 1 in the revised paper. The first level of the model setup is about 8 m above ground level (see Page 8 Line 21).

7) **Comparison to radar wind profiler** : what is the period of evaluation ? Is it 2 months ? Tables of scores for wind speed and duration would be useful and easier to read than scores included in the text.

The evaluation was done over daytime for the entire one-month simulation. Our intent in this paper is to present the model errors varying with height. For this pupose a figure is preferable.

Also, in Fig.5, if it is related to a 2 months period, it would be better to normalize the vertical coordinate by the PBL height.

We appreciate your suggestion and will take it into account in future work.

8) **Comparison to NWS surface stations** : all the stations are not represented on Fig.S1 and the domain is not the same. As a complement to Fig.6, a table with scores for MYNN_UCM is necessary, not only with biases but also with rmse. As a complement to Fig.6, it would be useful to provide two figures with the orography and the urban fraction for 1.3km resolution, and to discuss if the scores are related to orography, urban area... At 1.3km, what is the resolution of the orography database ?

Figure S1 is a map showing the location of all of the GHG measurement over the LA basin, which matches the triangles in Figure 1b, 6, 9a, 9b, and 12. Figure 1b shows the orography for the 1.3 km domain. Figure S1 shows the orography as well, although the domain is not exactly the same as other figures. Usually to the locations of NWS station relative to the location of the GHG measurements (triangles), we estimated the relevant orography. We choose to keep the figures as in the original manuscript to avoid redundancy.

We tried to explain the model bias with orography at the beginning, but could find no clear correlation. The RMSE maps below have been added to the revised paper (Figure 7).

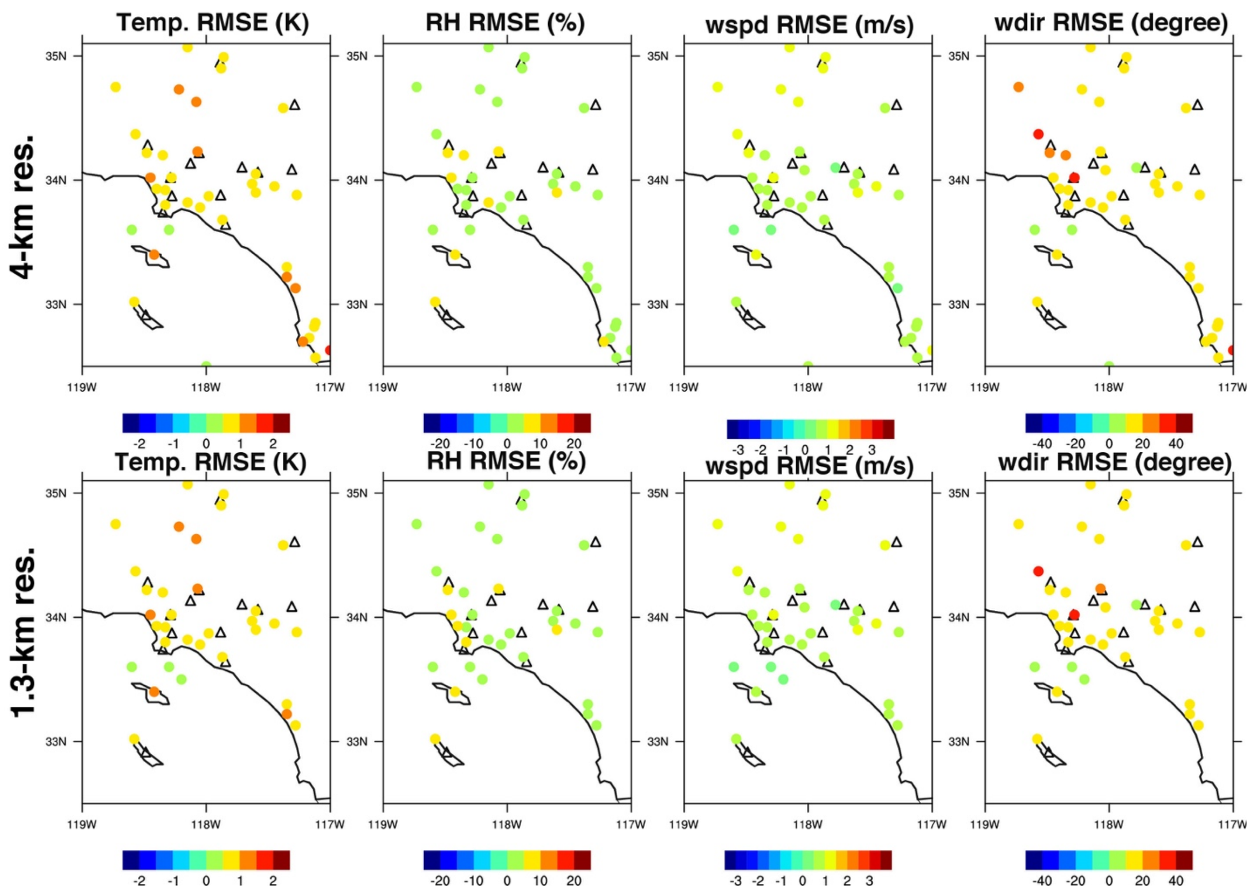

Figure R3. RMSE maps of the MYNN_UCM runs versus National Weather Stations (NWS) over the LA megacity (Model – NWS): (a1-a4) 4-km run; (b1 – b4) 1.3-km run. Black triangles indicate the locations of the GHG measurement sites.

9) **Comparison to in-situ CO_2** : once again, a table of scores (bias and rmse) with the 4 simulations, as a complement to Fig.7, is missing.

Thanks for the suggestion. We have added the two tables (Table 4 and 5) below as complements to Figure 7 in the revised paper.

Table R2. Statistics of modelled CO_2 (unit: ppm) with different configurations relative to in-situ CO_2 between 1300 – 1700 PST

	Pasadena		Palos Verdes	
	bias	RMSE	bias	RMSE
1.3 km WRF-Hestia	8.91	18.43	2.57	17.00
4 km WRF-Hestia	7.03	14.50	8.09	19.64
1.3 km WRF-Vulcan	1.20	11.10	5.03	10.62
4 km WRF-Vulcan	-1.38	9.13	4.20	9.40

Table R3. Statistics of daily afternoon averaged modelled CO_2 (unit: ppm) with different configurations relative to in-situ CO_2*

	Pasadena		Palos Verdes	
	bias	RMSE	bias	RMSE
1.3 km WRF-Hestia	-1.39	6.21	-0.75	4.71
4 km WRF-Hestia	0.58	4.38	-1.77	4.59
1.3 km WRF-Vulcan	-3.43	5.51	1.37	5.21
4 km WRF-Vulcan	-4.41	6.12	0.58	4.38

*Averaged over 1300 – 1700 PST

10) This study focuses only on **two months** of modelling and observations (May-June 2010). Conclusions thus must be quite limited, as one cannot extrapolate to generalized model performance from such a limited duration comparison, which could be particularly favourable or

unfavourable. The limited duration of model/observations must be presented, and its impact on conclusions should be discussed.

The Los Angeles basin is surrounded to the north and east by mountain ranges with summits of 2-3 km, with the ocean to the west and the desert to the north. From April to September, LA is in a warm, dry, and stable air mass. Alongshore steady wind flow predominates this area. In contrast, from October to March, moist onshore flows bring precipitation to LA.

Details about LA climate can be found in the study of Conil and Hall (2006).

The focus of this study is from the middle of May to the middle of June, which is representative of the dry season. We agree that the study based on a one-month simulation has its limitations. The model has to be evaluated and verified as the time period and spatial region of interest change. The limitation of this study has been added to the revised paper (see Page 26 Line 16-25).

One element of this is discussing time/computation to simulate one-month, and whether the current model construct could be expected to run for years to compare w/ the observational record being recorded in LA & USA.

This one-month high-resolution simulation with 288x288x50 grids and 5-s time steps has taken 11520 CPU hours (45 hours x 256 processors) on NAS High performance supercomputer Pleiades. See Page 26 Line 22-25. Using the same number of processors on Pleiades, a one-year simulation will take about 23 days to complete, which is still reasonable. It is, however, not practical for the large scale, i.e., the contiguous United States.

Specific comments :

P8 line 5 : It can be added that the coupling between mesoscale meteorological model and lagrangian particle model can be used in an operational framework to deal with accidental release (Lac et al., 2008).

Added. See Page 9 Line 18-20.

Table 1 : There could be probably a mistake for shortwave radiation scheme : does RRTMG deal with SW radiation ?

The RRTMG shortwave scheme has been included in version 3.1 and above.

Abstract : The acronym $FFCO_2$ is used before being presented.

Thanks for catching this. The full name has been added in the revised paper.

References :

Riette and Lac : A New Framework to Compare Mass-Flux Schemes Within the AROME Numerical Weather Prediction Model, Boundary-Layer Meteorology, 2016, 1—29.

Lac, C., F. Bonnardot, O. Connan, C. Camail, D. Maro, D. Hebert, M. Rozet, and J. Pergaud, Evaluation of a mesoscale dispersion modelling tool during the CAPITOUL experiment, Meteorol. Atmos. Phys., 102, 263-287, 2008.

Bougeault, P., and Lacarrere, P.: Parameterization of Orography-Induced Turbulence in a Mesobeta--Scale Model, Monthly Weather Review, 117, 1872-1890, 10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2, 1989.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., and Zhang, C.: The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems, International Journal of Climatology, 31, 273-288, 10.1002/joc.2158, 2011.

Cohen, A. E., Cavallo, S. M., Coniglio, M. C., and Brooks, H. E.: A Review of Planetary Boundary Layer Parameterization Schemes and Their Sensitivity in Simulating Southeastern U.S. Cold Season Severe Weather Environments, Weather and Forecasting, 30, 591-612, doi:10.1175/WAF-D-14-00105.1, 2015.

Coniglio, M. C., Jr., J. C., Marsh, P. T., and Kong, F.: Verification of Convection-Allowing WRF Model Forecasts of the Planetary Boundary Layer Using Sounding Observations, Weather and Forecasting, 28, 842-862, doi:10.1175/WAF-D-12-00103.1, 2013.

Conil, S., and Hall, A.: Local Regimes of Atmospheric Variability: A Case Study of Southern California, Journal of Climate, 19, 4308-4325, 10.1175/JCLI3837.1, 2006.

Janjić, Z. I.: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes, Monthly Weather Review, 122, 927-945, 10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2, 1994.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337-354, 10.5194/acp-12-337-2012, 2012.

Mu, L., Mu, L., Stammerjohn, S. E., Lowry, K. E., and Yager, P. L.: Spatial variability of surface pCO_2 and air-sea CO_2 flux in the Amundsen Sea Polynya, Antarctica, Elementa (Washington, D.C.), 2, 000036, 10.12952/journal.elementa.000036, 2014.

Nakanishi, M., and Niino, H.: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog, Boundary-Layer Meteorol, 119, 397-407, 10.1007/s10546-005-9030-8, 2006.

Torres, R., Pantoja, S., Harada, N., González, H. E., Daneri, G., Frangopulos, M., Rutllant, J. A., Duarte, C. M., Rúiz-Halpern, S., Mayol, E., and Fukasawa, M.: Air-sea CO_2 fluxes along the coast of Chile: From CO_2 outgassing in central northern upwelling waters to CO_2 uptake in southern Patagonian fjords, Journal of Geophysical Research: Oceans, 116, 10.1029/2010JC006344, 2011.

LA Megacity: a High-Resolution Land-Atmosphere Modelling System for Urban CO_2 Emissions

Sha Feng[1,2*], Thomas Lauvaux[3,2], Sally Newman[4], Preeti Rao[2], Ravan Ahmadov[5,6], Aijun Deng[3], Liza I. Díaz-Isaac[3], Riley M. Duren[2], Marc L. Fischer[7], Christoph Gerbig[8], Kevin R. Gurney[9], Jianhua Huang[9], Seongeun Jeong[7], Zhijin Li[2], Charles E. Miller[2], Darragh O'Keeffe[9], Risa Patarasuk[9], Stanley P. Sander[2], Yang Song[9], Kam W. Wong[4,2], Yuk L. Yung[4]

[1] JIFRESSE, University of California, Los Angeles, Los Angeles, CA

[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

[3] Department of Meteorology and Atmospheric Science, Pennsylvania State University, State College, PA

[4] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA

[5] Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, CO

[6] Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA

[7] Lawrence Berkeley National Laboratory, Berkeley, CA

[8] Max Planck Institute for Biogeochemistry, Hans-Knöll-Str.10, 07745 Jena, Germany

[9] School of Life Science, Arizona State University, Tempe, AZ

[*] now at Department of Meteorology, Pennsylvania State University, University Park, PA 16802, USA

Correspondence to: Sha Feng (sfeng@psu.edu)

Abstract

Megacities are major sources of anthropogenic fossil fuel CO_2 ($FFCO_2$) emissions. The spatial extents of these large urban systems cover areas of 10,000 km^2 or more with complex topography and changing landscapes. We present a high-resolution land-atmosphere modelling system for urban CO_2 emissions over the Los Angeles (LA) megacity area. The Weather Research and Forecasting (WRF)-Chem model was coupled to a very high-resolution $FFCO_2$ emission product, Hestia-LA, to simulate atmospheric CO_2 concentrations across the LA megacity at spatial resolutions as fine as ~1 km. We evaluated multiple WRF configurations, selecting one that minimized errors in wind speed, wind direction, and boundary layer height as evaluated by its performance against meteorological data collected during the CalNex-LA campaign (May-June 2010). Our results show no significant difference between moderate- (4-km) and high- (1.3-km) resolution simulations when evaluated against surface meteorological data, but the high-resolution configurations better resolved PBL heights and vertical gradients in the horizontal mean winds. We coupled our WRF configuration with the Vulcan 2.2 (10 km resolution) and Hestia-LA (1.3-km resolution) fossil fuel CO_2 emission products to evaluate the impact of the spatial resolution of the CO_2 emission products and the meteorological transport model on the representation of spatiotemporal variability in simulated atmospheric CO_2 concentrations. We find that high spatial resolution in the fossil fuel CO_2 emissions is more important than in the atmospheric model to capture CO_2 concentration variability across the LA megacity. Finally, we present a novel approach that employs simultaneous correlations of the simulated atmospheric CO_2 fields to qualitatively evaluate the greenhouse gas measurement network over the LA megacity. Spatial correlations in the atmospheric CO_2 fields reflect the coverage of individual measurement sites when a statistically significant number of sites observe emissions from a specific source or location. We conclude that elevated atmospheric CO_2 concentrations over the LA megacity are composed of multiple fine-scale plumes rather than a single homogenous urban dome. Furthermore, we conclude that $FFCO_2$ emissions monitoring in the LA megacity requires $FFCO_2$ emissions modelling with ~1 km resolution because coarser resolution emissions modelling tends to overestimate the observational constraints on the emissions estimates.

1 Introduction

Carbon dioxide (CO_2) is a major anthropogenic contributor to climate change. It has increased from its preindustrial (1750) level of 278 ± 2 ppm (Etheridge et al., 1996) to over 400 ppm in recent years, as reported by the National Oceanic and Atmospheric Administration (NOAA) and Scripps Institution of Oceanography [http://co2now.org/]. Clear evidence has shown that the continued increase of the atmospheric CO_2 concentration is dominated by global fossil fuel consumption during the same period (IPCC, 2013) and land use change (Houghton, 1999).

Urban areas are significant sources of fossil fuel CO_2 ($FFCO_2$), representing more than 50% of the world's population and more than 70% of $FFCO_2$ (UN, 2006). In particular, megacities (cities with urban populations greater than 10 million people) are major sources of anthropogenic emissions, with the world's 35 megacities emitting more than 20% of the global anthropogenic $FFCO_2$, even though they only represent about 3% of the Earth's land surface (IPCC, 2013). The proportion of emissions from megacities increases monotonically with the world population and urbanization (UN, 2006, 2010). Developed and developing megacities around the world are working together to pursue strategies to limit CO_2 and other greenhouse gas (GHG) emissions (C40, 2012).

Carbon fluxes can be estimated using "bottom-up" and "top-down" methods. Typically, $FFCO_2$ emissions are determined using "bottom-up" methods, by which fossil fuel usage from each source sector is convolved with the estimated carbon content of each fuel type to obtain $FFCO_2$ emission estimates. Space-time resolved $FFCO_2$ data sets using "bottom-up" methods clearly reveal the fingerprint of human activity with the most intense emissions being clustered around urban centres and associated power plants (e.g., Gurney et al., 2009; Gurney et al., 2012). At the global and annual scale, $FFCO_2$ emission estimates remain uncertain at ±5%, varying widely by country and reporting method (Le Quéré et al., 2014). At the urban scale, the uncertainties of $FFCO_2$ emission estimates are often 50-200 % (Turnbull et al., 2011; Asefi-Najafabady et al., 2014). On the other hand, "top-down" methods could potentially estimate biases in bottom-up emissions, and could also detect trends that cities can use for decision-making, due to changing economic activity or implementation of new emission regulations.

“Top-down” methods involve atmospheric measurements and usually include an atmospheric inversion of CO_2 concentrations, using atmospheric transport models to estimate carbon fluxes (i.e., posterior fluxes) by adjusting the fluxes (i.e., prior fluxes) to be consistent with observed CO_2 concentrations (e.g., Lauvaux et al., 2012; Lauvaux et al., 2015; Tarantola, 2005; Enting et al., 1994; Gurney et al., 2002; Baker et al., 2006; Law et al., 2003). In general, a prior flux is required for estimating the fluxes using an atmospheric inversion. The uncertainties in “top-down” methods can be attributed to errors in the observations (e.g., Tarantola, 2005), emission aggregation errors from the prior fluxes (e.g., Gurney et al., 2012; Engelen et al., 2002), and physical representation errors in the atmospheric transport model (e.g., Díaz Isaac et al., 2014; Gerbig et al., 2008; Kretschmer et al., 2012; Lauvaux et al., 2009; Sarrat et al., 2007). Previous studies showed that regional high-resolution models can capture the measured CO_2 signal much better than the lower resolution global models and simulate the diurnal variability of the atmospheric CO_2 field caused by recirculation of nighttime respired CO_2 well (Ahmadov et al., 2009). Previous studies (Ahmadov et al., 2009; Pillai et al., 2011; Pillai et al., 2010; Rödenbeck et al., 2009) have discussed the advantages of high resolution CO_2 modelling on different domains and applications. Recent efforts to study $FFCO_2$ emissions on urban scales have benefited from strategies that apply in-situ observations concentrated within cities and mesoscale transport models (e.g., Wu et al., 2011; Lauvaux et al., 2015; Strong et al., 2011; Lac et al., 2013; Bréon et al., 2015).

The Los Angeles (LA) megacity is one of the top three $FFCO_2$ emitters in the U.S. The atmospheric CO_2 concentrations show complex spatial and temporal variability resulting from a combination of large $FFCO_2$ emissions, complex topography, and challenging meteorological variability (e.g., Brioude et al., 2013; Wong et al., 2015; Angevine et al., 2012; Conil and Hall, 2006; Ulrickson and Mass, 1990; Lu and Turco, 1995; Baker et al., 2013; Chen et al., 2013; Newman et al., 2013). Past studies exploring CO_2 concentrations over the LA megacity used measurement methods ranging from ground-based to airborne, from in-situ to column. Those studies consistently reported robust enhancements (e.g., 30-100 ppm in-situ and 2-8 ppm column) and significant variability of the CO_2 concentrations for the LA megacity (Newman et al., 2013; Wunch et al., 2009; Wong et al., 2015; Kort et al., 2012; Wennberg et al., 2012; Newman et al, 2016). There have been limited radiocarbon (^{14}C) isotopic tracer studies (Newman et al., 2013;; Djuricin et al., 2010; Riley et al., 2008; Newman et al, 2016). Newman et al. (2013) showed that $FFCO_2$ constituted 10 - 25 ppm of the CO_2 excess observed in the LA basin by averaging the flask samples at 1400 PST during 15 May – 15 June, 2010. Djuricin et al. (2010) demonstrated that fossil fuel combustion contributed approximately 50~70 % of CO_2 sources in LA. Recently, using CO_2 mole fractions and $\Delta^{14}C$ and $\delta^{13}C$ values of CO_2 in the LA megacity observed in inland Pasadena (2006–2013) and coastal Palos Verdes peninsula (autumn 2009–2013), Newman et al. (2016) demonstrated that fossil fuel combustion is the dominant source of CO_2 for inland Pasadena. Airborne campaigns over LA (typically days to weeks in duration) included ARCTAS-CA (Jacob et al., 2010) and CalNex-LA (Brioude et al., 2013). All of these earlier studies were limited in their ability to investigate the spatial and temporal characteristics of LA carbon fluxes given relatively sparse observations. To better understand and quantify the total emissions, trends, and detailed spatial, temporal, and source sector patterns of emissions over the LA megacity requires both a denser measurement network and a land-atmosphere modelling system appropriate for such a complex urban environment. In this paper, we couple the Weather Research and Forecasting (WRF) – Chem model to a high-resolution $FFCO_2$ emission product, Hestia-LA, to study the spatiotemporal variability of urban CO_2 concentrations over the LA megacity.

The mesoscale circulation over the LA megacity is challenging for atmospheric transport models due to a variety of phenomena, such as “Catalina” eddies off the coast of southern California and the coupling between the land-sea breeze and winds induced by the topography (Angevine et al., 2012; Conil and Hall, 2006; Ulrickson and Mass, 1990; Kusaka and Kimura, 2004b; Kusaka et al., 2001). In this paper we present a set of simulations exploring WRF model physics configurations for the LA megacity, evaluating the model performance against meteorological data from the CalNex-LA campaign period, 15 May – 15 June 2010. Angevine et al. (2012) investigated how WRF model performance varied with spatial resolution and PBL scheme, etc., for the CalNex-LA campaign period; however, they focused the model meteorological evaluation on the spatial resolutions of 12- and 4-km. In the present study we focus on three critical aspects of the WRF model configuration – the planetary boundary layer (PBL) scheme, the urban surface scheme, and the model spatial resolution – as well as the effects of the $FFCO_2$ emissions product spatial resolution. Through these four aspects, the impacts of physical representation errors and emission aggregation errors on the modelled CO_2 concentrations across the LA megacity are investigated.

Moreover, a novel approach is proposed to evaluate the design of the greenhouse gas (GHG) measurement network for the LA megacity. The LA measurement network consists of 14 observation sites designed to provide continuous atmospheric CO_2 concentrations to assess the anthropogenic carbon emissions distribution and trends. The goal of the network design exploration is to optimize the atmospheric observational constraints on the surface fluxes. Kort et al. (2013) found that a minimum of eight optimally located, in-city surface CO_2 observation sites were required for accurate assessment of CO_2 emissions in LA using the "footprint" method (backward mode) and based on a national $FFCO_2$ emission product Vulcan (Gurney et al, 2009; Gurney et al, 2012). Here we assess the influence of each observation site using spatial correlations in terms of the simulated CO_2 (forward mode) at high-resolution. This method brings flexibility to allow us to evaluate the existing measurement network or to design a measurement network for various observation platforms, i.e., in-situ, aircraft, satellite, etc. In this paper, we will investigate the application to in-situ measurement network design.

The remainder of the paper is organized as follows. Section 2 describes the modelling framework, including initial conditions and boundary conditions for WRF-Chem. In section 3, we assess the quality of the model results, focusing on accurate representation of the PBL height, wind speed and wind direction, and CO_2 concentration. Section 4 presents the spatial and temporal patterns of simulated CO_2 concentration fields over the LA megacity using various $FFCO_2$ emissions products. Section 5 describes the forward mode approach for evaluating the spatial sensitivity of the 2015-era surface GHG measurement sites within the LA megacity. Discussion of model errors, model sampling strategy, and the density of the LA GHG measurement network from the forward model perspective is given in section 6. A summary is given in section 7. Section 8 lists the author contributions.

2 Modelling Framework

Sensitivity experiments were conducted using WRF-Chem version 3.6.1 with various PBL schemes, urban surface schemes, and model resolutions to define an optimized configuration for simulating atmospheric CO_2 concentration fields over the LA megacity. The impact of the resolution of $FFCO_2$ emission products is investigated in section 4.

2.1 WRF model setup

All of the model runs used one-way triple-nested domains with resolutions of 12-, 4-, and 1.3-km. The coarse domain (d01) covers most of the western US; the intermediate domain (d02) covers California and part of Mexico (Figure 1a); the innermost domain (d03) covers the majority of the South Coast Air Basin, a portion of the southern San Joaquin Valley and extends into the Pacific Ocean to include Santa Catalina and San Clemente Islands (Figure 1b). The Los Angeles basin, a subset of South Coast Air Basin, is surrounded to the north and east by mountain ranges with summits of 2-3 km, with the ocean to the west and the desert to the north. The basin consists of the West Coast Basin, Central Basin, and Orange County Coastal Plain. The boundaries of these three regions are the Newport Inglewood Fault and the boundary between Los Angeles County and Orange County. In this study, our analysis is limited to the innermost domain (d03), referred to hereafter as the LA megacity. All three of the model domains use 51 terrain following vertical levels from surface to 100 hPa, of which 29 layers are below 2 km above ground level (AGL) and the first level is about 8 m AGL.

The meteorological fields and surface parameters, such as soil moisture, were initialized by the three-hourly North American Regional Reanalysis (NARR) data set with a horizontal resolution of 32 km (Mesinger et al., 2006) and the six-hourly NCEP sea surface temperature data set with a horizontal resolution of 12 km (ftp://polar.ncep.noaa.gov/pub/history/sst/ophi). A summary of WRF configurations common to all sensitivity runs is shown in Table 1. The impact of varying the PBL parameterization, urban surface, and model resolution was investigated by conducting sensitivity runs summarized in Table 2.

PBL schemes are used to parameterize the unresolved turbulent vertical fluxes of heat, momentum, and constituents within the PBL. There are tens of mesoscale PBL schemes available in the WRF package. The details of PBL schemes can be found in the review paper by Cohen et al. (2015). Briefly, the PBL schemes represent turbulent mixing on the local or non-local basis. The local schemes only consider immediately adjacent vertical levels in the model. This tends to prevent vertical mixing and to produce relatively shallow PBL. Non-local schemes allow for a deeper mixing layer. We selected the three commonly used turbulent kinetic energy (TKE)-driven local PBL schemes (1.5 order) for the sensitivity runs: the Mellor-Yamada-Janjie technique (MYJ), Mellor-Yamada Nakanishi and Niino Level 2.5 (MYNN), and Bougeault-Lacarrère (BouLac). MYJ (Janjić, 1994) defines the PBL top where the TKE profiles decrease to a threshold of 0.2 m^2s^{-2}; MYNN (Nakanishi and Niino, 2006) is tuned to a database of large eddy simulations (LES) and sets the PBL top where the TKE falls below 1.0×10^{-6} m^2 s^{-2}; BouLac (Bougeault and Lacarrere, 1989) defines the PBL top where TKE reaches 0.005 m^2 s^{-2}.

The TKE-driven PBL schemes explicitly estimate the turbulent fluxes from mean atmospheric states and/or their gradients and can be used to drive a Lagrangian particle dispersion model in subsequent atmospheric inversions (e.g., Lauvaux et al., 2008). The coupling between the mesoscale meteorological and Lagrangian particle models can be used in an operational framework to deal with accidental release (Lac et al., 2008) .

For an accurate representation of the LA CO_2 distribution, the necessity of incorporating a urban surface scheme was tested by alternatively including a single-layer urban canopy model (UCM, Kusaka and Kimura, 2004a), a multiple-layer building environment parameterization (BEP, Martilli et al., 2009), and no urban surface scheme. Note that BEP requires very high vertical resolution within the PBL and is only compatible with MYJ and BouLac PBL schemes. Given that BEP is computationally expensive, we only test it with BouLac in this study. A detailed description of urban parameterization schemes available in WRF is provided by Chen et al. (2011).

We chose to test and evaluate our WRF-Chem configuration during the middle of May – middle of June 2010 time period of the CalNex-LA campaign (Ryerson et al., 2013) to take advantage of the extra meteorological measurements recorded during the campaign. Hourly simulations were conducted for 36-h periods starting with a 12-h meteorological spin-up at 12:00 UTC of the previous day. Hence, when concatenating the model output, each new run is introduced at 0000 UTC. All of the analyses in the following sections are limited to the region of the LA megacity.

2.2 Configuration for the CO_2 simulation

This paper analyses the impact of both physical representation errors and emission aggregation errors on the modelled CO_2 concentrations across the LA megacity. WRF-Chem version 3.6.1 allows for online CO_2 tracer transport coupled with the Vegetation Photosynthesis and Respiration Model (VPRM) (Ahmadov et al., 2007; Xiao et al., 2004). VPRM calculates hourly net ecosystem exchange based on MOIDS satellite estimates of the land surface water index and enhanced vegetation index (EVI), short wave radiance and surface temperature. A detailed description of VPRM can be found in Mahadevan et al. (2008). In this study, the defaults of the VPRM parameters were used given limited number of observation available for optimization.

Anthropogenic $FFCO_2$ fluxes were alternatively prescribed from the Vulcan 2.2 and Hestia-LA 1.0 $FFCO_2$ emission products developed at Arizona State University (Gurney et al., 2009; Gurney et al., 2012; Gurney et al., 2015; Rao et al., 2015). Both emission products were developed using "bottom-up" methods. Vulcan quantifies $FFCO_2$ emissions for the entire contiguous United States (CONUS) hourly at approximately 10-km spatial resolution for the year of 2002, The temporal variations are driven by a combination of modelled activity (building energy modelling) and monitoring (power plant emissions) (Gurney et al., 2009). Hestia-LA, by contrast, is a fossil fuel CO_2 emissions data product specific in space and time to the individual building, road segments, and point sources covering the the Los Angeles megacity domain for the years of 2011 and 2012 (Rao et al., 2015; Gurney et al., 2015; Gurney et al., 2012; Zhou and Gurney, 2010). It quantifies hourly $FFCO_2$ emissions for the counties of Los Angeles, Orange, San Bernardino, Ventura, and Riverside, at approximately 1.3 km x 1.3 km. Hestia-LA uses much of the same information for the temporal variations of Vulcan except for the onroad emissions, for which local traffic data is employed as opposed to regional traffic data. Given the similarities, it is unlikely that the small difference in temporal variation between Hestia-LA and Vulcan could account for the spatial differences, through covariation with atmospheric transport, found in this study. For more details about Hestia-LA, see Rao et al. (2015).

Atmospheric CO_2 concentrations in WRF-Chem were alternatively driven by the Vulcan and Hestia-LA emissions at the resolutions of 4 km and 1.3 km. Hence, four different emission datasets were generated – Vulcan 10 km emissions transported at 4-km or 1.3-km resolution, and Hestia-LA 1.3 km emissions transported at 4-km or 1.3-km resolution. The Hestia-LA emissions were aggregated from the native building-level resolution to the 1.3 and 4 km resolutions via direct summation in the specified model grids. Hestia-LA 2011 is temporally shifted for creating the weekday-weekend cycle for the year of 2010. The Vulcan $FFCO_2$ emissions were interpolated by using a bilinear operator and by preserving the value of the integral of data between the source (10-km) and destination (4- and 1.3-km) grid. Additionally, the ratio of the total carbon emissions over the state between the years of 2002 and 2015 from California Air Resource Board (http://www.arb.ca.gov/) was uniformly applied to the Vulcan emissions to temporally scale Vulcan from the 2002 base year to 2010.

No CO_2 ocean fluxes were prescribed in this study. The order of magnitude of oceanic CO_2 fluxes is minus one in the unit of μmol/m^2/s: -0.15 μmol/m^2/s along the coast of Chile calculated by Torres et al. (2011), +0.2 μmol/m^2/s for Southern Ocean by Mu et al. (2014), while fossil fuel emissions are about 20 μmol/m^2/s (roughly estimated from Hestia-LA at the Pasadena site). At regional scales, anthropogenic and biogenic fluxes are much larger than ocean fluxes, so we assume the ocean fluxes are negligible.

Lateral boundary conditions and initial conditions for CO_2 concentration fields were taken from the three-dimensional CO_2 background (often called the "NOAA curtain" for background) estimated from measurements in the Pacific (Jeong et al., 2013). Unlike meteorology, CO_2 fields were initialized only at the start time of the entire simulation and were carried over simulation cycle to cycle (without any re-initialization) until the end of the entire simulation to conserve CO_2 air mass over the model domains.

3 Model – data comparison

Meteorological observations obtained during the CalNex-LA campaign (http://www.esrl.noaa.gov/csd/projects/calnex/) include PBL height sampled by NOAA P-3 flights and aerosol backscatter ceilometer (Haman et al., 2012; Scarino et al., 2013), a radar wind profiler operated by the South Coast Air Quality Management District near Los Angeles International Airport (LAX), and CO_2 in situ measurements (Newman et al., 2013). Additionally, the NWS (National Weather Service, www.weather.gov) surface observations are used.

3.1 Comparison to aircraft PBL height

During CalNex-LA, 17 P-3 research flights sampled the daytime and nighttime PBL, marine surface layer, and the overlying free troposphere throughout California (Ryerson et al., 2013). We imposed four criteria for selecting aircraft profiles of potential temperature for PBL height comparisons:

1) Aircraft profiles sample within the innermost model domain (d03, Figure 1b);
2) Profiles sample during daytime (1100 PST – 1700 PST) when the CO_2 concentrations in PBL is well mixed;
3) Profiles acquired within ±30 min of the model output;
4) Ability to determine the PBL height from the vertical gradient of potential temperature.

Based on these four criteria, we selected seven aircraft profiles collected between 16 May and 19 May 2010. Figure 2 shows a profile acquired on 19 May 2010 when the aircraft was sampling over Pasadena, California.

The model diagnostic PBL height calculated by each PBL scheme can differ due to the Richardson bulk number (R_i) used (e.g., Kretschmer et al., 2014; Hong et al., 2006; Yver et al., 2013). To avoid this difference, we determined modelled PBL height based on the vertical virtual potential temperature gradient. The case in Figure 2 shows that the modelled PBL height agrees within 50 meters of the aircraft-determined and ceilometer-measured PBL height

Figure 3 shows the absolute difference between the modelled and aircraft-determined PBL height for each selected aircraft profile. The differences between the modelled and aircraft-determined PBL height differ case by case, and none of the model physics is systematically better than others. However, BouLac_BEP and MYNN have larger biases than others. The averaged bias of BouLac_BEP is 289 m for d02, 295 m for d03; MYNN bias is 179 m for d02 and 216 m for d03. For other configurations, the averaged biases are smaller than 160 m. The modelled PBL bias appears somewhat smaller in the 4-km runs than the 1.3-km runs. This, however, is based on seven selected aircraft profiles only. To further define the optimal physics for the PBL height simulation, we will present the all-hours statistics with the ceilometer data in section 3.2.

3.2 Comparison to ceilometer PBL height

Accurate simulation of the time evolution of the PBL depth is crucial to properly simulate the vertical mixing and ventilation of CO_2 emitted at the surface. The ceilometer measurements during CalNex-LA (Haman et al., 2012) allow us to evaluate the time evolution of the modelled PBL depth. Compared with the ceilometer-measured PBL height, the maximum discrepancies between model and observations occur from around 1100 PST – 1200 PST when the nocturnal PBL is fully collapsed and 1700 PST when it starts to form again (Figure 4). Among all of the model physics, MYNN_UCM shows the best agreement with the observations, while BouLac_BEP differs from ceilometer the most. The absolute bias of the MYNN_UCM modelled PBL height ranges from 5 to 198 m and 0 to 184 m with mean biases of -15.3 m (d02) and -6.9 m (d03) and root-mean-square error (*RMSE*) of 89.7 m and 94.5 m for 4- and 1.3-km resolution, respectively, which is similar to the range in the study of Riette and Lac (2016). They evaluated the model performance with different model sizes for an operational weather forecast system (AROME, application of Research to Operations at Mesosclae) against the observed PBL height at five observation sites, showing mean bias of -9.17 m and *RMSE* of 115 m for 200 × 200 grids, 6.17 m and 95.5 m for 108 × 108 grids. In our experiences, the statistics of MYNN_UCM_1.3km and MYNN_UCM_4km suggest the 1.3-km model resolution improves the model performance of the PBL simulation as compared with the ceilometer. The improvement in the high-resolution model runs can be seen in the statistics for

MYJ_UCM, BouLac_UCM, and BouLac_BEP, but not MYNN or MYJ (Table 3). Note that the ceilometer measurements were all at Caltech and thus reflect basin interior conditions. These are expected to be very different from coastal conditions in terms of the temporal evolution and eventual height of the mid-day PBL as well as the timing of the nocturnal PBL collapse. The domain is much larger and more varied than captured by a single location.

We also notice that UCM-coupled simulations agree with the ceilometer better than other combinations (Table 3, MYNN_UCM vs. MYNN, MYJ_UCM vs. MYJ, BouLac_UCM vs. BouLac_BEP). The inclusion of UCM yields model simulations with comparably higher relative humidity over the LA megacity (not shown). This corresponds to lower PBL height, which largely reduces the discrepancy of the modelled PBL from the observations (see UCM runs with their counterparts in Figure 4).

3.3 Comparison to radar wind profiler

Atmospheric dynamics has a direct influence on the CO_2 transport. Realistically reproducing the vertical gradient of wind fields is crucial. In Figure 5, we show the average difference in the wind profiles between the models and the radar wind profiler at LAX (Angevine et al., 2012). Most of the simulations show relatively larger wind speed bias near the surface: BouLac_BEP, MYJ, and MYNN with bias of 2.4 ± 2.2 m/s, BouLac_UCM and MYJ_UCM with bias of 2.0 ± 2.3 m/s. In contrast, it is encouraging to see that MYNN_UCM agrees with the radar measurement with mean bias of 1.4 ± 2.0 m/s, a lower mean bias than for the other configurations. As we found in the PBL evaluation, UCM-coupled simulations tend to reduce the wind speed bias at this location.

For wind direction, likewise, MYNN_UCM agrees with the observations slightly better below 800 m (~1.1 m/s for the averaged error), although the model bias is much less pronounced across the configurations. However, we notice that MYNN_UCM shows larger wind direction bias between 800 – 1400 m than others due to relatively lower PBL height simulated (not shown).

Improvement provided by the 1.3-km model resolution is visible near the PBL height (800 – 1400 m). A finer model resolution tends to resolve the vertical gradients of the atmospheric state better.

Angevine et al. (2012) evaluated a set of model configurations with the highest model resolution at 4 km for CalNex-LA using the same radar wind profiler data. The optimal configuration (the total energy–mass flux boundary layer scheme and ECMWF reanalysis) they found showed 1.1 ± 2.7 m/s bias in wind speed and -2.6 ± 67° in wind direction near the surface. Here MYNN_UCM displays similar performance to their optimal configuration. At the 4-km model resolution, the biases of MYNN_UCM are 1.4 ± 2.0 m/s in wind speed and -1.3 ± 20.0° in wind direction.

In summary, the MYNN_UCM configuration showed the best agreement with meteorological observations among the configurations we evaluated at given locations. In section 3.4, we examine the performance of MYNN_UCM across the LA megacity.

3.4 Comparison to NWS surface stations

We introduce the observations from the NWS surface network to demonstrate the model performance across the LA megacity. The objective analysis program OBSGRID is used to remove erroneous data and observations that are not useful (Deng et al., 2009; Rogers et al., 2013).

Figure 6 shows the model bias of temperature, relative humidity, wind speed, and wind direction compared to the NWS surface data across the LA megacity. The locations of the GHG measurement sites are marked (see details in Table 6 and Figure S1). Overall, there is little difference in the simulated surface atmospheric state variables between the 4-km and 1.3-km runs; i.e., the 1.3-km run does not show any significant improvement compared to the 4-km run at the surface (even though it resolves the vertical gradient of atmospheric states and PBL better, Figure 4 and 5).

For temperature (Figure 6a1 and 6b1), the model is colder than the observations by 0.5 - 1.0 K. Larger temperature biases occur in the desert. For relative humidity (Figure 6a2 and 6b2), the model is dryer (teal blue) than the observations but with two exceptions: Santa Monica coastal area and Pasadena to Mt. Wilson area (light green). See Figure S1 for the location. The model dryness is consistent with the findings of Nehrkorn et al. (2012). The model is 5% dryer over the basin with a somewhat larger bias of 5% - 10% near Granada Hills and Ontario. These two locations have the highest temperature in the summer – typically 7 K or more warmer than downtown LA in May-June (77 °F for downtown LA and 84 °F for Ontario. See http://www.intellicast.com/Local/History.aspx). For the Pasadena area, the model is moister than the observations. The moistness tends to cause lower PBL heights, which can be seen in the comparison to the ceilometer-determined PBL height at Caltech in Pasadena, California (Figure 4): MYNN_UCM has a shallower PBL in comparison to the ceilometer during the 1400 PST – 1800 PST time period.

The model overestimates wind speed by ~1.0 m/s (Figure 6a3 and 6b3). The tendency of the model to overestimate wind speed is fully documented in previous studies (e.g., Angevine et al., 2012; Brioude et al., 2013; Nehrkorn et al., 2012; Yver et al., 2013). For surface wind direction, model bias is within ±10° for most of the LA megacity. The larger biases appear near the foothills of Santa Monica Mountains, San Gabriel Mountains, and University of Southern California (USC) due to the topography.

Compared with other model physics (not shown), we notice that USC, located just south of downtown LA, is a challenging location for mesoscale modelling, in particular for wind simulations. All of the model physics consistently show a relatively large wind bias at USC except BouLac_BEP that is not seen in the remainder of the domain. We also noticed that adding UCM to MYNN decreases the modelled temperature, while all of the other models' physics have a warm bias compared to observations.

All of the analyses above focused on the meteorology over the LA megacity. The results indicate little difference horizontally between 4- and 1.3-km runs across the basin. Similarly, there are only small differences in the *RMSE* maps as well (Figure 7). This consistent with the assumption in Angevine et al. (2012) that a finer grid may not give better results. However, the 1.3-km run tends to resolve the vertical gradients of atmospheric state variables and PBL better, which likely improves the vertical mixing and ventilation of modelled atmospheric CO_2 concentrations. In the following sections, we will use the MYNN_UCM configuration with the resolution of 4 km and 1.3 km for the simulations of atmospheric CO_2 concentration fields over the LA megacity.

3.5 Comparisons to in-situ CO_2

We coupled Hestia and Vulcan $FFCO_2$ emission products individually with the MYNN_UCM to generate four sets of simulated CO_2 concentrations: WRF-Hestia 1.3-km, WRF-Hestia 4-km, WRF-Vulcan 1.3-km, and WRF-Vulcan 4-km. The runs with the same model resolution have the same meteorology but differ in emissions, and vice versa.

During CalNex-LA, in-situ observation sites at Pasadena and Palos Verdes continuously measured surface CO_2 concentrations. Measurements were recorded using a Picarro (Santa Clara, CA) Isotopic CO_2 Analyser (cavity ring-down spectrometer), model G1101-i, for Pasadena and an infrared gas analyser from PP Systems (Haverford, MA), model CIRAS-SC for Palos Verdes. In addition, periodic flask samples were collected for analysis of $^{14}CO_2$ for extracting fossil fuel and biogenic signals. See Newman et al. (2016) for details about the sites and sampling information. Figure 8 shows the comparison of the time series of hourly (Figure 8a,b) and daily afternoon (Figure 8c,d) averaged CO_2 concentrations (1300 PST – 1700 PST) between model and observations. Tables 4 and 5 is the comparison statistics of the four CO_2 runs against the in-situ measurements as a complement to Figure 8a,b and Figure 8c,d, respectively. Overall, the model captures the temporal variability of CO_2 but overestimates CO_2 during nighttime. During afternoons, the model agrees with the observations fairly well (Figure 8c and 8d) except for a few events: all simulations underestimate CO_2 concentrations by about 10 ppm around 28 May and 4-6 June for Pasadena and 21 May for Palos Verdes. These events lasting two – three days are likely related to synoptic scale processes. Using the averaged Pacific Ocean CO_2 signal as background may explain the failure to capture these events. Further investigation of the background air would provide insights related to synoptic variability but is beyond the scope of this work.

Inter-comparison of the diurnal patterns among these four runs (Figure 9a) shows WRF-Hestia runs tend to overestimate the CO_2 concentration around noon and underestimate CO_2 in the late afternoon at the Pasadena site, while WRF-Vulcan runs tend to underestimate the CO_2 concentration for the entire period. Hence, WRF-Hestia runs show larger model bias based on the statistics for the daytime afternoon hour but smaller errors based on the daytime afternoon average (Table 4 and 5). Next we focus on this diurnal variability.

Clear diurnal variations of the surface CO_2 concentrations were observed for both sites (Figure 9). The observed CO_2 concentrations increase at night and remain high until sunrise, and they quickly drops as the boundary layer grows after sunrise (Figure 9a and 9b). The amplitude of this diurnal cycle is greater in Pasadena than in Palos Verdes.

For the Pasadena site, during nighttime, when the PBL is shallow, CO_2 is trapped locally: the more fossil fuel is emitted, the higher CO_2 concentration is simulated. Consequently, the WRF-Vulcan runs show considerably lower CO_2 concentration than the WRF-Hestia runs due to the lower emissions in Vulcan at the Pasadena site (Figure 9c). However, during daytime, with well-mixed conditions, the discrepancy between the WRF-Hestia and WRF-Vulcan runs becomes smaller at this site. Among these runs, the 1.3-km WRF-Hestia run successfully captures the diurnal variation of the surface CO_2 concentration, although a noontime peak is in the model not present in the observations. By contrast, the 4-km WRF-Hestia run underestimates the CO_2 concentration during 0200 PST – 0700 PST even though emissions were comparable between Hestia 4-km and Hestia 1.3-km (Figure 9c). The underestimation of the simulated CO_2 concentration likely results from the representation errors in the atmospheric transport due to the coarser model resolution.

For Palos Verdes, however, none of the model results match the observations. All of the runs show a peak in the simulated CO_2 concentration around 0800 PST, which very likely corresponds to the failure to simulate the eastward marine flow as a part of the Catalina eddy (e.g., Bosart, 1983; Davis et al., 2000). This CO_2 concentration peak is incorrectly reproduced by the model advecting the $FFCO_2$ emitted from the strong point sources in Long Beach, California (Figure 1d) and in turn contaminating the air of Palos Verdes.

3.6 Comparisons to flask-sampled CO_2

The isotopic tracer radiocarbon (^{14}C) can be used for distinguishing between fossil fuel and biogenic sources of CO_2 (Djuricin et al., 2010; Newman et al., 2013; Newman et al., 2016; Pataki et al., 2006; Pataki et al., 2007; Levin et al., 2003; Miller et al., 2012; Turnbull et al., 2006; Turnbull et al., 2009). During CalNex-LA, flask samples collected on alternate afternoons at 1400 PST were combined to produce two CO_2 samples samples per month in Pasadena (weekly samples were combined to produce one radiocarbon sample per month in Palos Verdes) for extracting anthropogenic and biogenic signals from the total CO_2 concentration. Note that the two samples for Palos Verdes were sampled from 1 May to 31 May and from 1 June to 30 June, not exactly overlapping the CalNex-LA period; the two for Pasadena were sampled from 15 May to 31 May and from 1 June to 15 June, overlapping the CalNex-LA period. See Newman et al. (2016) for details about the sites and sampling information. Figure 10 presents the comparisons of the modelled and flask-sampled anthropogenic fossil fuel and biogenic CO_2. From both the flask samples and model simulations, the CO_2 signal from the biosphere is much weaker than $FFCO_2$ in the LA megacity. The two-week flask sampled biogenic CO_2 is about 2 ppm on average. We note that the 1.3-km WRF-Vulcan run overestimates the $FFCO_2$ concentrations by 20 ppm over the second half of the month (Figure 10d), implying that low-resolution CO_2 emissions can be very critical for a coastal site (complex terrain) with strong point sources nearby.

Strong temporal variability of the simulated biogenic and $FFCO_2$ can be seen for both sites (Figure 10a,10c,10e,10g). For the Pasadena site, the 1.3-km run shows nearly flat biogenic CO_2 concentrations during 15 May to 30 May when the 4-km run has more variability (Figure 10e). A large botanical garden covering 207 acres (The Huntington Library, Art Collections, and Botanical Gardens) is about 1.6 km away from the Pasadena site, which may suggest that higher model resolution (1.3 km vs. 4 km) could resolve the land cover better. However, there is still up to about 3-ppm discrepancy in the modelled biogenic CO_2 from the flask samples (Figure 10f). Similar discrepancy can be seen for Palos Verdes as well (Figure 10h). Reasonably determining CO_2 from biogenic sources remains challenging. Additional measurements are needed to constrain biogenic fluxes.

4 Spatial pattern of the surface CO_2

The spatial pattern of surface CO_2 concentration exhibits diurnal variability over the LA megacity due to the complexity of the topography and the variability of circulation patterns, PBL heights, and $FFCO_2$ emissions. Each plays an important role in sequence or at the same time. Here, we only focus on the pattern at 1400 PST when the atmospheric CO_2 concentration is well mixed in the PBL. At 1400 PST, there is a close relationship between CO_2 concentration and atmospheric transport; the error due to the PBL height determination is at a minimum. For the same reason, we assume that $FFCO_2$ emissions do not play a dominant role around 1400 PST unless there are strong local signals from point sources, such as power plants, refineries, airports etc.

In this section, we define the 1.3-km WRF-Hestia run as the reference simulation. For simplicity, all of the relevant CO_2 spatial patterns we present are selected from the second model layer (about 24 m AGL). Figure 11a and 11b display the topography and the average CO_2 concentration at 1400 PST overlaid with the first empirical orthogonal function (EOF1) of the surface wind pattern, respectively. The locations of the 13 GHG measurement sites in the LA megacity domain are marked in the figures (see Table 6 and Figure S1 for details about the observation sites). Note that the 2015-era surface GHG measurement network includes 14 sites in total, while 13 sites are embedded in the innermost model domain. According to the geography mentioned in section 2.1, the Granada Hills (GH), Compton, and USC sites are located in the West Coast Basin, the Pasadena and Mt. Wilson (MWO) sites are in the Central Basin, and California State University Fullerton (CSUF), Ontario, and San Bernardino (SB) sites are in the Orange County Coastal Plan. Additionally, the Dryden and Victorville (VV) sites are located in deserts; the Palos Verdes (PV), University of California Irvine (UCI), and San Clemente Island (SCI) are on the coast. Although the Dryden site is actually a TCCON (Total Carbon Column Observing Network, Wunch et al., 2011) site, in the analysis, we assume it provides near-surface point measurements like the other sites, for simplicity.

Blocked by the mountains, the emitted CO_2 is trapped in the basin; the desert is usually as clean as the upwind ocean. Specifically, Dryden (not shown on the figure), VV, SCI (not shown on the figure), Palos Verdes and UCI are much cleaner than other sites (Figure

11b). At 1400 PST, sea breeze prevails over the LA megacity. Affected by the geometry of Palos Verdes Peninsula, the sea breeze is divided into west and southwest onshore flows that then converge in the Central Basin. Strong CO_2 signals emitted from electricity production and industry (with annual emission of 86.9 million kgC, Figure 1d) are trapped in a limited area. We notice that the south-western flow, which appears stronger than the western flow, prevents the high CO_2 concentration in the West Coast Basin from propagating further east and dilutes into the Central Basin. Controlled by the orography, strong southerly flows occur between the Santa Monica and San Gabriel Mountains, keeping the contaminated air from propagating to the west. Driven by the same meteorology, the 1.3-km WRF-Vulcan run shows a more smeared out CO_2 distribution over the LA basin (Figure 11c) due to the coarser resolution of the original Vulcan emissions. High CO_2 plumes seen in the 1.3-km WRF-Hestia run from point sources are replaced by broad areas of elevated CO_2 concentration in the 1.3-km WRF-Vulcan. The large differences in the simulated surface CO_2 fields between the 1.3-km WRF-Hestia and WRF-Vulcan runs are found around LAX and north of the Palos Verdes Peninsula where strong point sources are located (dipole-like pattern in Figure 11d).

5 Sampling density of the 2015-era GHG measurement network

In this section, we present a forward network design framework, using the modelled CO_2 concentrations and their relationship with neighbouring grid cells. Note no actual observation data but only pseudo data are used in this section. Compared to previous studies using tower footprints (i.e. linearized adjoint models) as in Kort et al. (2013), we propose here a forward model assessment of the network using the high-resolution model results. We assume that each observation site can be associated with a specific CO_2 air mass at any given time. To define this CO_2 air mass, we estimate the spatial coherence in the modelled CO_2 concentration fields. We constrain the coverage of each LA GHG measurement site by calculating the simultaneous correlation of the site to the rest of the domain using the simulated CO_2 concentration time series. Figure 12 shows the correlation map (R) of each site for the 1.3-km WRF-Hestia run. Only areas meeting a significance level of 0.01 in the t-test ($|R| \geqslant 0.46$) are coloured. Based on the spatial patterns of the correlation maps, all of the observation sites can be grouped into (i) coastal/island sites, i.e., UCI, SCI, and Palos Verdes (right three panels in bottom row of Figure 12), (ii) western basin sites, i.e., GH, Pasadena, MWO, USC, and Compton (top row in Figure 12), (iii) eastern basin sites, (i.e., CSUF, Ontario, SB; middle row in Figure 12), and (iv) desert sites, i.e., Dryden and VV (left two panels in bottom row of Figure 12).

Not surprisingly, the coastal/island sites are mainly correlated with CO_2 concentration in upwind areas offshore where there is limited $FFCO_2$ contamination. The white channel from Catalina Island to the Huntington Beach area demonstrates the influence of terrain-induced flows and mountain blocking. The western basin sites are mainly correlated with CO_2 concentration throughout the western portion of the basin, and the eastern basin sites are mainly correlated with CO_2 concentrations throughout the eastern portion of the basin. The desert sites are anti-correlated with the basin. CSUF also shows anti-correlation with the desert. Two reasons can explain this anti-correlation. Firstly, CO_2 is trapped and accumulates in the basin due to the mountain barrier; the basin is contaminated, the desert is clean. Secondly, after CO_2 accumulates in the basin over a certain amount of time, episodic strong sea breezes may push this basin CO_2 over the mountains to the desert. As a result, the basin will be relatively clean while the desert is contaminated.

Based on the correlation maps, we can also see how the coverage of each site varies with the $FFCO_2$ emissions data products and with the model resolutions. Figure 13 shows the correlation maps across the runs for the Compton, Palos Verdes, and CSUF stations. All runs use the optimal physics we determined for the LA megacity, i.e., MYNN_UCM. The correlation maps for each site differ with the $FFCO_2$ emissions data product used, model resolution, or their combination (Figure 13). Given that the 1.3-km WRF-Hestia is the reference run, the difference of this to the 1.3-km WRF-Vulcan run reflects the errors induced by emissions resolution. The discrepancy between the 1.3-km WRF-Hestia run and the 4-km WRF-Hestia run reflects the model representation errors. The 4-km WRF-Vulcan run is subject to model representation errors and emission aggregation errors at the same time. For simplicity, we will not emphasize but only show the comparison of the 4-km WRF-Vulcan to the others.

Compton is isolated from the rest of the basin in the 1.3-km WRF-Hestia run but correlated with most of the basin in the 1.3-km WRF-Vulcan run. A similar discrepancy is seen for Palos Verdes. Additionally, Palos Verdes appears to be a clean site in the 1.3-km WRF-Hestia run but dramatically contaminated in the 1.3-km WRF-Vulcan run (even correlated with the LA downtown area). For CSUF, the anti-correlation between basin and desert noted above is not visible in the 1.3-km WRF-Vulcan run. Compared to the 1.3-km WRF-Hestia run, the 4-km WRF-Hestia run overall shows a somewhat larger region with significant correlation for each site.

To highlight the discrepancy in the spatial patterns caused by the model representation errors and emission aggregation errors in the view of the existing GHG measurement network, a composite map for each run is shown in Figure 14. These maps are constructed by determining the number of sites for which the absolute value of R is greater than 0.46 for each grid cell (i.e., colour-filled area in Figure 12 and 11). R=0.46 is the critical value for the t-test at the significance level of 0.01. In the 1.3-km WRF-Hestia run (reference), the West Coastal Basin and Orange County Coastal Plain are correlated with up to 6 measurement sites. A gap appears over the Central Basin correlated with up to 3 sites due to the wind pattern (Figure 11a and 11b). The San Gabriel Mountains and Peninsular Ranges are rarely correlated to any of the sites due to the elevated terrain. The 4-km WRF-Hestia run shows a similar pattern but with more sites covered over the Peninsular Ranges and the coast because of the failure to resolve topography by the 4-km model resolution.

In the 1.3-km WRF-Vulcan run, by contrast, a large area of the basin is correlated with most of the sites (nine out of 13). The Compton area is even correlated with 11 sites, which is only correlated with about two sites in the 1.3-km WRF-Hestia run. A similar contrast can be seen for the GH, USC, and Palos Verdes areas where the multiple strong point sources nearby in Hestia-LA have been aggregated into one 10 km by 10 km grid cell in Vulcan (Figure 1d vs.1c). Relatively coarser $FFCO_2$ emissions artificially increase the coverage of each site, which highlights the importance of using a high-resolution emission product, i.e., Hestia, for the CO_2 simulation for urban environment to represent the spatial variability in CO_2 and design the optimal network of surface GHG measurement.

6 Discussion

The results presented in this paper have shown that the choice of model resolution and emission products can strongly influence the interpretation of atmospheric CO_2 signals. Hestia quantifies $FFCO_2$ emissions down to individual buildings and roadways, such that strong point sources create large plumes that are extremely sensitive to atmospheric transport. Reproducing dynamics realistically by the atmospheric transport model is crucial around strong point sources, such as power plants, refineries, airports, etc. For instance, a considerable number of point sources are located in Long Beach harbour (Figure 1d), about 7 km away from the Palos Verdes site. In late spring and summer, Palos Verdes is a clean site, with little evidence of $FFCO_2$ emissions from the LA megacity most of the time. However, we can clearly see that Palos Verdes is often simulated to be contaminated by $FFCO_2$ in all of the runs, especially during early morning (Figure 9b) due to incorrectly simulated east marine flows advecting the strong $FFCO_2$ emissions, which cannot be seen in the observations. Biases in wind speed and direction become critical for such a location. Palos Verdes may be challenging for the atmospheric inversion if used as a background site.

Simulating CO_2 at locations with strong CO_2 fluxes gradients remains challenging. For a location like Compton with strong point sources nearby emitting CO_2 at 86.9 million kgC per year (recorded in Hestia-LA version 1.0), a fine resolution emission product becomes very important due to the strong $FFCO_2$ gradient. A relatively coarse emission product likely produces a spurious signal due to aggregating a strong point source into a large grid cell (Figure 11b and 9c). For instance, dipole-like CO_2 gradients were created in the difference between the 1.3-km WRF-Vulcan and WRF-Hestia runs (Figure 11d).

In this paper, we focus on the spatial distribution of the CO_2 concentration over the LA megacity. The choice of model resolution also significantly impacts the vertical gradients of the CO_2 concentration as a result of the terrain resolved. In the 1.3-km model grids, the elevation of MWO is 1129 m, while in the 4 km grids it is 753 m; the actual elevation is 1670 m. The representation errors in the 4-km model resolution are relatively large. When there is finer topographic resolution, more CO_2 is accumulated in the basin due to blocking by the mountains. Around noon, the model results show CO_2 enhancement of 10 ppm over MWO in both the 1.3-km WRF-Vulcan and WRF-Hestia runs but only up to 3 ppm in the 4-km model runs. Sampling strategies should be investigated for mountain sites like MWO (e.g., Law et al., 2008) as well as coastal sites where the topography resolved varies by model resolution. Meteorological evaluation at surface sites is not sufficient to show differences in vertical mixing.

Figure 12 presents the simultaneous correlation maps for each site in terms of the simulated CO_2 concentration time series. The coverage of the correlation maps is determined by two factors at the same time: atmospheric transport and surface fluxes. This method differs from the footprint method (Kort et al., 2013). The footprint method maps the influence of atmospheric transport only at the location of the observation; no emission pattern is considered. Here both transport and emissions play a role in the area covered by the observation site. Therefore, the correlation maps are subject to overestimation of the influence area versus the footprint method, due to the complicated nature of the atmospheric integrator. As an example, in Figure 12, the coloured grids of the correlation map are not necessarily *physically* related to the observation site. Those far from the site may lose the track of the initial sources. Conversely, there is definitely no *physical* influence from the uncorrelated areas to the observation site.

However, this new network design method has a unique strengths compared to the footprint method. First of all, this method is computationally economical relative to the footprint method. Secondly, the method does not require adjoint models, avoiding another complexity. Most importantly, it brings extreme flexibility without any complexity for evaluating the existing measurement network or designing the measurement network with various observation platforms (i.e., in-situ, satellite, etc.) and, especially, outpaces the analysis for dense sampling techniques, such as use of remote sensing datasets. Applying the footprint method to satellite data for regional scale modelling is extremely computationally time-consuming and complex.

Figure 15 shows the fraction of the total $FFCO_2$ emissions detected over the LA megacity as function of the number of the observation sites for all of the runs. Because the correlation maps have the possibility of overestimating the influence area, we focus on the uncorrelated areas only. Assuming that the coverage of the GHG measurement network is not sufficient if an area is correlated to no more than two sites, then ~28.9 % of $FFCO_2$ is potentially under-constrained by the current GHG measurement sites (Figure 15a: WRF-Hestia 1.3-km). These areas include most of the mountains, Santa Monica Bay and the upwind coast, and the south part of the Central Basin (Figure 13), about 21.1 % of total area. However, this analysis is a qualitative assessment of the observational constraint. Consideration of errors in the CO_2 emissions needs to be taken into account for a complete assessment of the network.

Figure 15 also reflects the impact of the $FFCO_2$ emissions used to simulate the CO_2 fields. In the 1.3-km WRF-Hestia run, there are no areas covered by more than six sites, while the 1.3-km WRF-Vulcan run shows 39.8 % of $FFCO_2$ emissions over the LA megacity to be covered by more than six sites. Additionally, the distribution appears nearly normal for the 1.3-km WRF-Vulcan run. A similar discrepancy is seen between the 4-km WRF-Hestia and WRF-Vulcan runs. These differences further highlight the importance of using the high-resolution $FFCO_2$ emissions product for the urban CO_2 simulation.

The LA climate has two typical local regimes. From April to September, LA is warm, dry, and stable. Steady alongshore wind flow predominates. In contrast, from October to March, moist onshore flows bring precipitation to LA (Conil and Hall, 2006). The period of interest for this study is from the middle of May to the middle of June 2010. The results of this study represent the model performance for the dry seasons. Studying anther time of a year may yield different results. A longer-term model evaluation is also desired, which, however, is computationally and observationally time-consuming. This one-month long high-resolution simulation took 11520 CPU hours (45 hours × 256 processors) on the petascale supercomputer Pleiades at the NASA Advanced Supercomputing (NAS) Division.

7 Conclusion

A set of WRF configurations varying by PBL scheme, urban surface scheme, and model resolution has been evaluated by comparing the PBL height determined by aircraft profiles and ceilometer, wind speed and wind direction measured by radar wind profiler, and surface atmospheric states measured by NWS stations. The results suggest that there is no significant difference between the 4-km and 1.3-km resolution simulations in terms of atmospheric model performances at the surface, but the 1.3-km model runs resolve the vertical gradients of wind fields and PBL height somewhat better. The model inter-comparisons show the model using the WRF configured MYNN_UCM PBL and urban surface schemes has overall better performance than others. Coupled to $FFCO_2$ emissions products (Hestia-LA and Vulcan 2.2), a land-atmosphere modelling system was built with MYNN_UCM for studying the heterogeneity of urban CO_2 emissions over the LA megacity.

The Vulcan and Hestia-LA $FFCO_2$ emission products were used to investigate the impact of the model representation errors and emission aggregation errors in the modelled CO_2 concentration. Compared to in-situ measurements during CalNex-LA, the 1.3-km modelled CO_2 concentrations clearly outperform the results at 4-km resolution for capturing both the spatial distribution and the temporal variability of the urban CO_2 signals due to strong $FFCO_2$ emission gradients across the LA megacity, even though no clear improvement in the meteorological evaluation was observed across the basin. The inter-comparison of the WRF-Hestia and WRF-Vulcan runs reinforces the importance of using high-resolution emission products to represent correct, large spatial gradients in atmospheric CO_2 concentrations for urban environments.

Based on the 1.3-km WRF-Hestia run, the coverage of the current GHG measurement site over the LA megacity was evaluated using the modelled spatial correlations. Kort et al. (2013) concluded a network of eight surface observation sites provided the minimum sampling required for accurate monitoring of $FFCO_2$ emissions in LA using Vulcan at 4-km model resolution. In this study, however, using Vulcan $FFCO_2$ emissions tend to overestimate the observational constraint spatially, suggesting that the information lies in multiple fine-scale plumes rather than a single urban dome over the Los Angeles basin. Thanks to the much finer-resolution model and $FFCO_2$ emission product Hestia-LA, the coverage of each observation site seems constrained to a more limited area. Using a high-resolution emission data product and a high-resolution model configuration is necessary for accurately assessing the urban measurement network.

8 Author contributions

S. Feng and T. Lauvaux designed the model experiments, evaluated the model performance, and developed the assessment of the measuring network; S. Newman provided the calibrated CO_2 measurements and support for the model evaluations. P. Rao, R. Patarasuk, D. O'Keeffe, J. Huang, Y. Song, and K.R. Gurney developed and prepared the Vulcan and Hestia emission products; R. Ahmadov contributed to the development of the WRF-VPRM model and relevant guidelines; A. Deng provided quality control for the observations from the National Weather Stations; L.I. Díaz-Isaac tested PBL algorithms; S. Jeong and M.L. Fischer provided the background CO_2 concentration for the LA megacity (region); R.M. Duren, C. Gerbig, Z. Li, C. E. Miller, S. Sander, K.W. Wong, and Y. L. Yung provided comments and discussion on the results of the study.

Acknowledgements

A portion of this work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. The Megacities Carbon Project is sponsored in part by the National Institute of Standards and Technology (NIST). S. Newman acknowledges funding from the Caltech/JPL President & Director's Research and Development Fund. K. R. Gurney thanks NIST grant 70NANB14H321.R. Ahmadov was supported by the US Weather Research Program within the NOAA/OAR Office of Weather and Air Quality. S. Jeong and M.L. Fischer acknowledge the support by the Laboratory Directed Research and Development Program, Office of Science, of the US Department of Energy under Contract No. DE-AC02-05CH11231. Thanks to W. Angevine at NOAA for radar wind profiler data, K. Aikin at NOAA for Aircraft WP-3D data, and B. Lefer at University of Houston for ceilometer data.

References

Ahmadov, R., Gerbig, C., Kretschmer, R., Koerner, S., Neininger, B., Dolman, A. J., and Sarrat, C.: Mesoscale covariance of transport and CO2 fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model, Journal of Geophysical Research: Atmospheres, 112, D22107, 10.1029/2007JD008552, 2007.

Ahmadov, R., Gerbig, C., Kretschmer, R., Körner, S., Rödenbeck, C., Bousquet, P., and Ramonet, M.: Comparing high resolution WRF-VPRM simulations and two global CO2 transport models with coastal tower measurements of CO2, Biogeosciences, 6, 807-817, 10.5194/bg-6-807-2009, 2009.

Angevine, W. M., Eddington, L., Durkee, K., Fairall, C., Bianco, L., and Brioude, J.: Meteorological Model Evaluation for CalNex 2010, Monthly Weather Review, 140, 3885-3906, 10.1175/MWR-D-12-00042.1, 2012.

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K.: A multiyear, global gridded fossil fuel CO2 emission data product: Evaluation and analysis of results, Journal of Geophysical Research: Atmospheres, 119, 10,213-210,231, 10.1002/2013JD021296, 2014.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y. H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO2 fluxes, 1988–2003, Global Biogeochemical Cycles, 20, n/a-n/a, 10.1029/2004GB002439, 2006.

Baker, K. R., Misenis, C., Obland, M. D., Ferrare, R. A., Scarino, A. J., and Kelly, J. T.: Evaluation of surface and upper air fine scale WRF meteorological modeling of the May and June 2010 CalNex period in California, Atmospheric Environment, 80, 299-309, http://dx.doi.org/10.1016/j.atmosenv.2013.08.006, 2013.

Bosart, L. F.: Analysis of a California Catalina Eddy Event, Monthly Weather Review, 111, 1619-1633, 10.1175/1520-0493(1983)111<1619:AOACCE>2.0.CO;2, 1983.

Bougeault, P., and Lacarrere, P.: Parameterization of Orography-Induced Turbulence in a Mesobeta--Scale Model, Monthly Weather Review, 117, 1872-1890, 10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2, 1989.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707-1724, 10.5194/acp-15-1707-2015, 2015.

Brioude, J., Angevine, W. M., Ahmadov, R., Kim, S. W., Evan, S., McKeen, S. A., Hsie, E. Y., Frost, G. J., Neuman, J. A., Pollack, I. B., Peischl, J., Ryerson, T. B., Holloway, J., Brown, S. S., Nowak, J. B., Roberts, J. M., Wofsy, S. C., Santoni, G. W., Oda, T., and Trainer, M.: Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NOx and CO2 and their impacts, Atmos. Chem. Phys., 13, 3661-3677, 10.5194/acp-13-3661-2013, 2013.

C40: Climate 40 Group, http://live.c40cities.org/, 2012.

Chen, D., Li, Q., Stutz, J., Mao, Y., Zhang, L., Pikelnaya, O., Tsai, J. Y., Haman, C., Lefer, B., Rappenglück, B., Alvarez, S. L., Neuman, J. A., Flynn, J., Roberts, J. M., Nowak, J. B., de Gouw, J., Holloway, J., Wagner, N. L., Veres, P., Brown, S. S., Ryerson, T. B., Warneke, C., and Pollack, I. B.: WRF-Chem simulation of NOx and O3 in the L.A. basin during CalNex-2010, Atmospheric Environment, 81, 421-432, http://dx.doi.org/10.1016/j.atmosenv.2013.08.064, 2013.

Chen, F., and Dudhia, J.: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity, Monthly Weather Review, 129, 569-585, 10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2, 2001.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., and Zhang, C.: The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems, International Journal of Climatology, 31, 273-288, 10.1002/joc.2158, 2011.

Cohen, A. E., Cavallo, S. M., Coniglio, M. C., and Brooks, H. E.: A Review of Planetary Boundary Layer Parameterization Schemes and Their Sensitivity in Simulating Southeastern U.S. Cold Season Severe Weather Environments, Weather and Forecasting, 30, 591-612, doi:10.1175/WAF-D-14-00105.1, 2015.

Conil, S., and Hall, A.: Local Regimes of Atmospheric Variability: A Case Study of Southern California, Journal of Climate, 19, 4308-4325, 10.1175/JCLI3837.1, 2006.

Davis, C., Low-Nam, S., and Mass, C.: Dynamics of a Catalina Eddy Revealed by Numerical Simulation, Monthly Weather Review, 128, 2885-2904, 10.1175/1520-0493(2000)128<2885:DOACER>2.0.CO;2, 2000.

Deng, A., Stauffer, D. R., Gaudet, B. J., Dudhia, J., Hacker, J., Bruyere, C., Wu, W., Vandenberghe, F., Liu, Y., and Bourgeois, A.: Update on WRF-ARW End-to-End Multi-scale FDDA System, 10th Annual WRF Users' Workshop, Boulder, CO, June 23, 2009.

Díaz Isaac, L. I., Lauvaux, T., Davis, K. J., Miles, N. L., Richardson, S. J., Jacobson, A. R., and Andrews, A. E.: Model-data comparison of MCI field campaign atmospheric CO2 mole fractions, Journal of Geophysical Research: Atmospheres, 119, 2014JD021593, 10.1002/2014JD021593, 2014.

Djuricin, S., Pataki, D. E., and Xu, X.: A comparison of tracer methods for quantifying CO2 sources in an urban region, Journal of Geophysical Research: Atmospheres, 115, n/a-n/a, 10.1029/2009JD012236, 2010.

Engelen, R. J., Denning, A. S., and Gurney, K. R.: On error estimation in atmospheric CO2 inversions, Journal of Geophysical Research: Atmospheres, 107, 4635, 10.1029/2002JD002195, 2002.

Enting, I. G., Heimann, M., Wigley, T. M. L., Commonwealth, S., and Industrial Research, O.: Future emissions and concentrations of carbon dioxide: key ocean/atmosphere/land analyses, Division of Atmospheric Research technical paper ;no. 31, 120 p., CSIRO, Australia, 120 p. pp., 1994.

Etheridge, D. M., Steele, L. P., Langenfelds, R. L., Francey, R. J., Barnola, J. M., and Morgan, V. I.: Natural and anthropogenic changes in atmospheric CO2 over the last 1000 years from air in Antarctic ice and firn, Journal of Geophysical Research: Atmospheres, 101, 4115-4128, 10.1029/95JD03410, 1996.

Gerbig, C., Körner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmos. Chem. Phys., 8, 591-602, 10.5194/acp-8-591-2008, 2008.

Grell, G. A., and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophysical Research Letters, 29, 38-31-38-34, 10.1029/2002GL015311, 2002.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, Nature, 415, 626-630, http://www.nature.com/nature/journal/v415/n6872/suppinfo/415626a_S1.html, 2002.

Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakumar, S., and de la Rue du Can, S.: High Resolution Fossil Fuel Combustion CO2 Emission Fluxes for the United States, Environmental Science & Technology, 43, 5535-5541, 10.1021/es900806c, 2009.

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., and Abdul-Massih, M.: Quantification of Fossil Fuel CO2 Emissions on the Building/Street Scale for a Large U.S. City, Environmental Science & Technology, 46, 12194-12202, 10.1021/es3011282, 2012.

Gurney, K. R., Romero-Lankao, P., Seto, K. C., Hutyra, L. R., Duren, R., Kennedy, C., Grimm, N. B., Ehleringer, J. R., Marcutuillio, P., Hughes, S., Pincetl, S., Chester, M. V., Runfola, D. M., Feddema, J. J., and Sperling, J.: Climate change: Track urban emissions on a human scale citation, Nature, 525, 179–181, 10.1038/525179a, 2015.

Haman, C. L., Lefer, B., and Morris, G. A.: Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment, Journal of Atmospheric and Oceanic Technology, 29, 697-710, 10.1175/JTECH-D-11-00114.1, 2012.

Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation, Monthly Weather Review, 132, 103-120, 10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2, 2004.

Hong, S.-Y., Noh, Y., and Dudhia, J.: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes, Monthly Weather Review, 134, 2318-2341, 10.1175/MWR3199.1, 2006.

Houghton, R. A.: The annual net flux of carbon to the atmosphere from changes in land use 1850–1990*, Tellus B, 51, 298-313, 10.1034/j.1600-0889.1999.00013.x, 1999.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, n/a-n/a, 10.1029/2008JD009944, 2008.

IPCC: Climate Change 2013. The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535pp., 2013.

Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B., Thompson, A. M., Shaw, G. E., McCauley, E., Pederson, J. R., and Fisher, J. A.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results, Atmos. Chem. Phys., 10, 5191-5212, 10.5194/acp-10-5191-2010, 2010.

Janjić, Z. I.: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes, Monthly Weather Review, 122, 927-945, 10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2, 1994.

Jeong, S., Hsu, Y.-K., Andrews, A. E., Bianco, L., Vaca, P., Wilczak, J. M., and Fischer, M. L.: A multitower measurement network estimate of California's methane emissions, Journal of Geophysical Research: Atmospheres, 118, 11,339-311,351, 10.1002/jgrd.50854, 2013.

Kort, E. A., Frankenberg, C., Miller, C. E., and Oda, T.: Space-based observations of megacity carbon dioxide, Geophysical Research Letters, 39, L17806, 10.1029/2012GL052738, 2012.

Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO2 emissions: Minimum site location requirements for the Los Angeles megacity, Journal of Geophysical Research: Atmospheres, 118, 1577-1584, 10.1002/jgrd.50135, 2013.

Kretschmer, R., Gerbig, C., Karstens, U., and Koch, F. T.: Error characterization of CO2 vertical mixing in the atmospheric transport model WRF-VPRM, Atmos. Chem. Phys., 12, 2441-2458, 10.5194/acp-12-2441-2012, 2012.

Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, A., Vogel, F., Hammer, S., and Totsche, K. U.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO2, Atmos. Chem. Phys., 14, 7149-7172, 10.5194/acp-14-7149-2014, 2014.

Kusaka, H., Kondo, H., Kikegawa, Y., and Kimura, F.: A Simple Single-Layer Urban Canopy Model For Atmospheric Models: Comparison With Multi-Layer And Slab

Models, Boundary-Layer Meteorol, 101, 329-358, 10.1023/A:1019207923078, 2001.

Kusaka, H., and Kimura, F.: Thermal Effects of Urban Canyon Structure on the Nocturnal Heat Island: Numerical Experiment Using a Mesoscale Model Coupled with an Urban Canopy Model, Journal of Applied Meteorology, 43, 1899-1910, 10.1175/JAM2169.1, 2004a.

Kusaka, H., and Kimura, F.: Coupling a Single-Layer Urban Canopy Model with a Simple Atmospheric Model: Impact on Urban Heat Island Simulation for an Idealized Case, Journal of the Meteorological Society of Japan. Ser. II, 82, 67-80, 10.2151/jmsj.82.67, 2004b.

Lac, C., Bonnardot, F., Connan, O., Camail, C., Maro, D., Hebert, D., Rozet, M., and Pergaud, J.: Evaluation of a mesoscale dispersion modelling tool during the CAPITOUL experiment, Meteorology and Atmospheric Physics, 102, 263-287, 10.1007/s00703-008-0343-2, 2008.

Lac, C., Donnelly, R. P., Masson, V., Pal, S., Riette, S., Donier, S., Queguiner, S., Tanguy, G., Ammoura, L., and Xueref-Remy, I.: CO2 dispersion modelling over Paris region within the CO2-MEGAPARIS project, Atmos. Chem. Phys., 13, 4941-4961, 10.5194/acp-13-4941-2013, 2013.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. Phys., 8, 3459-3471, 10.5194/acp-8-3459-2008, 2008.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089-1102, 10.5194/bg-6-1089-2009, 2009.

Lauvaux, T., Schuh, A. E., Bocquet, M., Wu, L., Richardson, S., Miles, N., and Davis, K. J.: Network design for mesoscale inversions of CO2 sources and sinks, 2012, 64, 10.3402/tellusb.v64i0.17980, 2012.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P. B., Sweeney, C., Turnbull, J. C., and Wu, K.: High resolution atmospheric inversion of urban CO2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), 2015.

Law, R. M., Rayner, P. J., Steele, L. P., and Enting, I. G.: Data and modelling requirements for CO2 inversions using high-frequency data, Tellus B, 55, 512-521, 10.1034/j.1600-0889.2003.00029.x, 2003.

Law, R. M., Peters, W., Rödenbeck, C., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S. J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P. K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO2: Experimental overview and diurnal cycle results for 2002, Global Biogeochemical Cycles, 22, n/a-n/a, 10.1029/2007GB003050, 2008.

Le Quéré, C., Peters, G. P., Andres, R. J., Andrew, R. M., Boden, T. A., Ciais, P., Friedlingstein, P., Houghton, R. A., Marland, G., Moriarty, R., Sitch, S., Tans, P., Arneth, A., Arvanitis, A., Bakker, D. C. E., Bopp, L., Canadell, J. G., Chini, L. P., Doney, S. C., Harper, A., Harris, I., House, J. I., Jain, A. K., Jones, S. D., Kato, E., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Koven, C., Lefèvre, N., Maignan, F., Omar, A., Ono, T., Park, G. H., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Schwinger, J., Segschneider, J., Stocker, B. D., Takahashi, T., Tilbrook, B., van Heuven, S., Viovy, N., Wanninkhof, R., Wiltshire, A., and Zaehle, S.: Global carbon budget 2013, Earth Syst. Sci. Data, 6, 235-263, 10.5194/essd-6-235-2014, 2014.

Levin, I., Kromer, B., Schmidt, M., and Sartorius, H.: A novel approach for independent budgeting of fossil fuel CO2 over Europe by 14CO2 observations, Geophysical Research Letters, 30, n/a-n/a, 10.1029/2003GL018477, 2003.

Lu, R., and Turco, R. P.: Air pollutant transport in a coastal environment—II. Three-dimensional simulations over Los Angeles basin, Atmospheric Environment, 29, 1499-1518, http://dx.doi.org/10.1016/1352-2310(95)00015-Q, 1995.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochemical Cycles, 22, GB2005, 10.1029/2006GB002735, 2008.

Description of the modifications made in WRF.3.1 and short user's manual of BEP, 2009.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American Regional Reanalysis, Bulletin of the American Meteorological Society, 87, 343-360, 10.1175/BAMS-87-3-343, 2006.

Miller, J. B., Lehman, S. J., Montzka, S. A., Sweeney, C., Miller, B. R., Karion, A., Wolak, C., Dlugokencky, E. J., Southon, J., Turnbull, J. C., and Tans, P. P.: Linking emissions of fossil fuel CO2 and other anthropogenic trace gases using atmospheric 14CO2, Journal of Geophysical Research: Atmospheres, 117, n/a-n/a, 10.1029/2011JD017048, 2012.

Mu, L., Mu, L., Stammerjohn, S. E., Lowry, K. E., and Yager, P. L.: Spatial variability of surface pCO2 and air-sea CO2 flux in the Amundsen Sea Polynya, Antarctica, Elementa (Washington, D.C.), 2, 000036, 10.12952/journal.elementa.000036, 2014.

Nakanishi, M., and Niino, H.: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog, Boundary-Layer Meteorol, 119, 397-407, 10.1007/s10546-005-9030-8, 2006.

Nehrkorn, T., Henderson, J., Leidner, M., Mountain, M., Eluszkiewicz, J., McKain, K., and Wofsy, S.: WRF Simulations of the Urban Circulation in the Salt Lake City

Area for CO2 Modeling, Journal of Applied Meteorology and Climatology, 52, 323-340, 10.1175/JAMC-D-12-061.1, 2012.

Newman, S., Xu, X., Affek, H. P., Stolper, E., and Epstein, S.: Changes in mixing ratio and isotopic composition of CO2 in urban air from the Los Angeles basin, California, between 1972 and 2003, Journal of Geophysical Research: Atmospheres, 113, n/a-n/a, 10.1029/2008JD009999, 2008.

Newman, S., Jeong, S., Fischer, M. L., Xu, X., Haman, C. L., Lefer, B., Alvarez, S., Rappenglueck, B., Kort, E. A., Andrews, A. E., Peischl, J., Gurney, K. R., Miller, C. E., and Yung, Y. L.: Diurnal tracking of anthropogenic CO2 emissions in the Los Angeles basin megacity during spring 2010, Atmos. Chem. Phys., 13, 4359-4372, 10.5194/acp-13-4359-2013, 2013.

Newman, S., Xu, X., Gurney, K. R., Hsu, Y. K., Li, K. F., Jiang, X., Keeling, R. F., Feng, S., O'Keefe, D., Patarasuk, R., Wong, K. W., Rao, P., Fisher, M. L., and Yung, Y. L.: Toward consistency between bottom-up CO2 emissions trends and top-down atmospheric measurements in the Los Angeles megacity, Atmos Chem Phys, *16*(6), 3843–3863, doi:10.5194/acp-16-3843-2016, 2016.

Pataki, D. E., Alig, R. J., Fung, A. S., Golubiewski, N. E., Kennedy, C. A., McPherson, E. G., Nowak, D. J., Pouyat, R. V., and Romero Lankao, P.: Urban ecosystems and the North American carbon cycle, Global Change Biology, 12, 2092-2102, 10.1111/j.1365-2486.2006.01242.x, 2006.

Pataki, D. E., Xu, T., Luo, Y. Q., and Ehleringer, J. R.: Inferring biogenic and anthropogenic carbon dioxide sources across an urban to rural gradient, Oecologia, 152, 307-322, 10.1007/s00442-006-0656-0, 2007.

Pillai, D., Gerbig, C., Marshall, J., Ahmadov, R., Kretschmer, R., Koch, T., and Karstens, U.: High resolution modeling of CO2 over Europe: implications for representation errors of satellite retrievals, Atmos. Chem. Phys., 10, 83-94, 10.5194/acp-10-83-2010, 2010.

Pillai, D., Gerbig, C., Ahmadov, R., Rödenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B., and Lavrié, J. V.: High-resolution simulations of atmospheric CO2 over complex terrain – representing the Ochsenkopf mountain tall tower, Atmos. Chem. Phys., 11, 7445-7464, 10.5194/acp-11-7445-2011, 2011.

Rao, P., Gurney, K. R., Patarasuk, R., Song, Y., Miller, C. E., Duren, R. M., and Eldering, A.: Spatio-temporal Variations in Onroad Vehicle Fossil Fuel CO2 Emissions in the Los Angeles Megacity, Atmospheric Environment, under revivew, 2015.

Riette, S., and Lac, C.: A New Framework to Compare Mass-Flux Schemes Within the AROME Numerical Weather Prediction Model, Boundary-Layer Meteorol, 1-29, 10.1007/s10546-016-0146-9, 2016.

Riley, W. J., Hsueh, D. Y., Randerson, J. T., Fischer, M. L., Hatch, J. G., Pataki, D. E., Wang, W., and Goulden, M. L.: Where do fossil fuel carbon dioxide emissions from California go? An analysis based on radiocarbon observations and an atmospheric transport model, Journal of Geophysical Research: Biogeosciences, 113, n/a-n/a, 10.1029/2007JG000625, 2008.

Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. Phys., 9, 5331-5342, 10.5194/acp-9-5331-2009, 2009.

Rogers, R. E., Deng, A., Stauffer, D. R., Gaudet, B. J., Jia, Y., Soong, S.-T., and Tanrikulu, S.: Application of the Weather Research and Forecasting Model for Air Quality Modeling in the San Francisco Bay Area, Journal of Applied Meteorology and Climatology, 52, 1953-1973, 10.1175/JAMC-D-12-0280.1, 2013.

Ryerson, T. B., Andrews, A. E., Angevine, W. M., Bates, T. S., Brock, C. A., Cairns, B., Cohen, R. C., Cooper, O. R., de Gouw, J. A., Fehsenfeld, F. C., Ferrare, R. A., Fischer, M. L., Flagan, R. C., Goldstein, A. H., Hair, J. W., Hardesty, R. M., Hostetler, C. A., Jimenez, J. L., Langford, A. O., McCauley, E., McKeen, S. A., Molina, L. T., Nenes, A., Oltmans, S. J., Parrish, D. D., Pederson, J. R., Pierce, R. B., Prather, K., Quinn, P. K., Seinfeld, J. H., Senff, C. J., Sorooshian, A., Stutz, J., Surratt, J. D., Trainer, M., Volkamer, R., Williams, E. J., and Wofsy, S. C.: The 2010 California Research at the Nexus of Air Quality and Climate Change

(CalNex) field study, Journal of Geophysical Research: Atmospheres, 118, 5830-5866, 10.1002/jgrd.50331, 2013.

Sarrat, C., Noilhan, J., Dolman, A. J., Gerbig, C., Ahmadov, R., Tolk, L. F., Meesters, A. G. C. A., Hutjes, R. W. A., Ter Maat, H. W., Pérez-Landa, G., and Donier, S.: Atmospheric CO2 modeling at the regional scale: an intercomparison of 5 meso-scale atmospheric models, Biogeosciences, 4, 1115-1126, 10.5194/bg-4-1115-2007, 2007.

Scarino, A. J., Obland, M. D., Fast, J. D., Burton, S. P., Ferrare, R. A., Hostetler, C. A., Berg, L. K., Lefer, B., Haman, C., Hair, J. W., Rogers, R. R., Butler, C., Cook, A. L., and Harper, D. B.: Comparison of mixed layer heights from airborne high spectral resolution lidar, ground-based measurements, and the WRF-Chem model during CalNex and CARES, Atmos. Chem. Phys. Discuss., 13, 13721-13772, 10.5194/acpd-13-13721-2013, 2013.

Strong, C., Stwertka, C., Bowling, D. R., Stephens, B. B., and Ehleringer, J. R.: Urban carbon dioxide cycles within the Salt Lake Valley: A multiple-box model validated by observations, Journal of Geophysical Research: Atmospheres, 116, n/a-n/a, 10.1029/2011JD015693, 2011.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Book, Whole, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2005.

Torres, R., Pantoja, S., Harada, N., González, H. E., Daneri, G., Frangopulos, M., Rutllant, J. A., Duarte, C. M., Rúiz-Halpern, S., Mayol, E., and Fukasawa, M.: Air-sea CO2 fluxes along the coast of Chile: From CO2 outgassing in central northern upwelling waters to CO2 uptake in southern Patagonian fjords, Journal of Geophysical Research: Oceans, 116, n/a-n/a, 10.1029/2010JC006344, 2011.

Turnbull, J., Rayner, P., Miller, J., Naegler, T., Ciais, P., and Cozic, A.: On the use of 14CO2 as a tracer for fossil fuel CO2: Quantifying uncertainties using an atmospheric transport model, Journal of Geophysical Research: Atmospheres, 114, n/a-n/a, 10.1029/2009JD012308, 2009.

Turnbull, J. C., Miller, J. B., Lehman, S. J., Tans, P. P., Sparks, R. J., and Southon, J.: Comparison of 14CO2, CO, and SF6 as tracers for recently added fossil fuel CO2 in the atmosphere and implications for biological CO2 exchange, Geophysical Research Letters, 33, n/a-n/a, 10.1029/2005GL024213, 2006.

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

Ulrickson, B. L., and Mass, C. F.: Numerical Investigation of Mesoscale Circulations over the Los Angeles Basin. Part II: Synoptic Influences and Pollutant Transport, Monthly Weather Review, 118, 2162-2184, 10.1175/1520-0493(1990)118<2162:NIOMCO>2.0.CO;2, 1990.

UN: World Urbanization Prospects e Revision 2005, Factsheet 7: Mega-cities, 2006. United Nations, Department of Economic and Social Affairs, Population Division. World Urbanization Prospects: The 2005 Revision. Working Paper No. ESA/P/WP/200, 2006.

UN: World Urbanization Prospects: The 2009 Revision, 2010.

Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L.: On the Sources of Methane to the Los Angeles Atmosphere, Environmental Science & Technology, 46, 9282-9289, 10.1021/es301138y, 2012.

Wong, K. W., Fu, D., Pongetti, T. J., Newman, S., Kort, E. A., Duren, R., Hsu, Y. K., Miller, C. E., Yung, Y. L., and Sander, S. P.: Mapping CH4 : CO2 ratios in Los Angeles with CLARS-FTS from Mount Wilson, California, Atmos. Chem. Phys., 15, 241-252, 10.5194/acp-15-241-2015, 2015.

Wu, L., Bocquet, M., Lauvaux, T., Chevallier, F., Rayner, P., and Davis, K.: Optimal representation of source-sink fluxes for mesoscale carbon dioxide inversion with synthetic data, Journal of Geophysical Research: Atmospheres, 116, n/a-n/a, 10.1029/2011JD016198, 2011.

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, Geophysical Research Letters, 36, L15810, 10.1029/2009GL039825, 2009.

Xiao, X., Hollinger, D., Aber, J., Goltz, M., Davidson, E. A., Zhang, Q., and Moore Iii, B.: Satellite-based modeling of gross primary production in an evergreen needleleaf forest, Remote Sensing of Environment, 89, 519-534, http://dx.doi.org/10.1016/j.rse.2003.11.008, 2004.

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

Zhou, Y., and Gurney, K.: A new methodology for quantifying on-site residential and commercial fossil fuel CO2 emissions at the building spatial scale and hourly time scale, Carbon Management, 1, 45-56, 10.4155/cmt.10.7, 2010.

Table 1. Common elements of the WRF-Chem configuration used in all runs.

Option	Description
Microphysics	WSM5 (Hong et al., 2004)
Longwave radiation	RRTMG (Iacono et al., 2008)
Shortwave radiation	RRTMG (Iacono et al., 2008)
Land surface	Noah land surface model (Chen and Dudhia, 2001)
Cumulus scheme	Grell-3 (Grell and Dévényi, 2002) applied to 12-km domain (d01) only
Advection	5^{th} and 3^{rd} order differencing for horizontal and vertical advection respectively
Time step	3^{rd} order Runge-Kutta; 45, 24, and 5 s for outermost, middle, innermost domains, respectively

Table 2. WRF configurations used for the sensitivity runs.

Configuration	PBL scheme	Urban surface scheme	Grid spacing (km)
BouLac_BEP_d02	BouLac	BEP	4
BouLac_BEP_d03	BouLac	BEP	1.3
BouLac_UCM_d02	BouLac	UCM	4
BouLac_UCM_d03	BouLac	UCM	1.3
MYJ_d02	MYJ	None	4
MYN_d03	MYJ	None	1.3
MYJ_UCM_d02	MYJ	UCM	4
MYJ_UCM_d03	MYJ	UCM	1.3
MYNN_d02	MYNN	None	4
MYNN_d03	MYNN	None	1.3
MYNN_UCM_d02	MYNN	UCM	4
MYNN_UCM_d03	MYNN	UCM	1.3

Table 3. Comparison Statistics of model performance on PBL height (unit: m AGL) relative to the ceilometer data over 1100 – 1700 PST at Caltech

	Mean	Bias	Stdv*	RMSE
OBS	835.7	-	223.8	-
MYNN_UCM_d03	828.8	-6.9	82.7	89.7
MYNN_UCM_d02	820.4	-15.3	66.1	94.5
MYNN_d03	1055.6	219.9	205.8	278.2
MYNN_d02	1029.4	193.7	200.0	254.3
MYJ_UCM_d03	961.4	125.8	154.9	168.8
MYJ_UCM_d02	971.4	135.7	109.3	157.7
MYJ_d03	1115.3	279.7	174.4	308.7
MYJ_d02	1105.1	269.5	150.9	291.6
BouLac_UCM_d03	936.1	100.5	147.3	149.9
BouLac_UCM_d02	958.7	123.1	104.8	148.7
BouLac_BEP_d03	1233.9	398.3	239.0	442.2
BouLac_BEP_d02	1244.3	408.6	219.5	446.0

*Stdv = standard deviation

Table 4. Statistics of hourly modelled CO_2 (unit: ppm) with different configurations relative to in-situ CO_2 between 1300 – 1700 PST

	Pasadena		Palos Verdes	
	bias	*RMSE*	bias	*RMSE*
WRF-Hestia 1.3-km	8.91	18.43	2.57	17.00
WRF-Hestia 4 km	7.03	14.50	8.09	19.64
WRF-Vulcan 1.3 km	1.20	11.10	5.03	10.62
WRF-Vulcan 4 km	-1.38	9.13	4.20	9.40

Table 5. Statistics of daily afternoon averaged modelled CO_2 (unit: ppm) with different configurations relative to in-situ CO_2*

	Pasadena		Palos Verdes	
	bias	*RMSE*	bias	*RMSE*
WRF-Hestia 1.3 km	-1.39	6.21	-0.75	4.71
WRF-Hestia 4 km	0.58	4.38	-1.77	4.59
WRF-Vulcan 1.3 km	-3.43	5.51	1.37	5.21
WRF-Vulcan 4 km	-4.41	6.12	0.58	4.38

*Averaged over 1300 – 1700 PST

Table 6. Locations of the 2015-era GHG measurement sites in the model domain❖

Code*	Name	Type	Lat. (° N)	Lon. (° E)
GH	Granada Hills	Tower	34.28	-118.47
Pasadena	Pasadena	Building top	34.14	-118.13
MWO	Mt. Wilson	Mountain top	34.22	-118.06
USC	University of South California	Building top	34.02	-118.29
Compton	Compton	Tower	33.87	-118.28
CSUF	California State University, Fullerton	Building top	33.88	-117.88
Ontario	Ontario	Tower	34.06	-117.58
SB	San Bernardino	Tower	34.09	-118.35
Dryden✚	Dryden	TCCON	34.95	-117.89
VV	Victorville	Tower	34.61	-117.29
UCI	University of California, Irvine	Building top	33.64	-117.84
SCI	San Clemente Island	Tower	32.92	-118.49
PV	Palos Verdes	In-situ non-standard	33.74	-118.35

❖La Jolla site is operating but not included in this paper

*Codes used in this paper

✚ In the analysis, we assume Dryden site is a near-surface point measurement like other sites rather than a column observation for simplicity. TCCON is the Total Carbon Column Observing Network (Wunch et al., 2011).

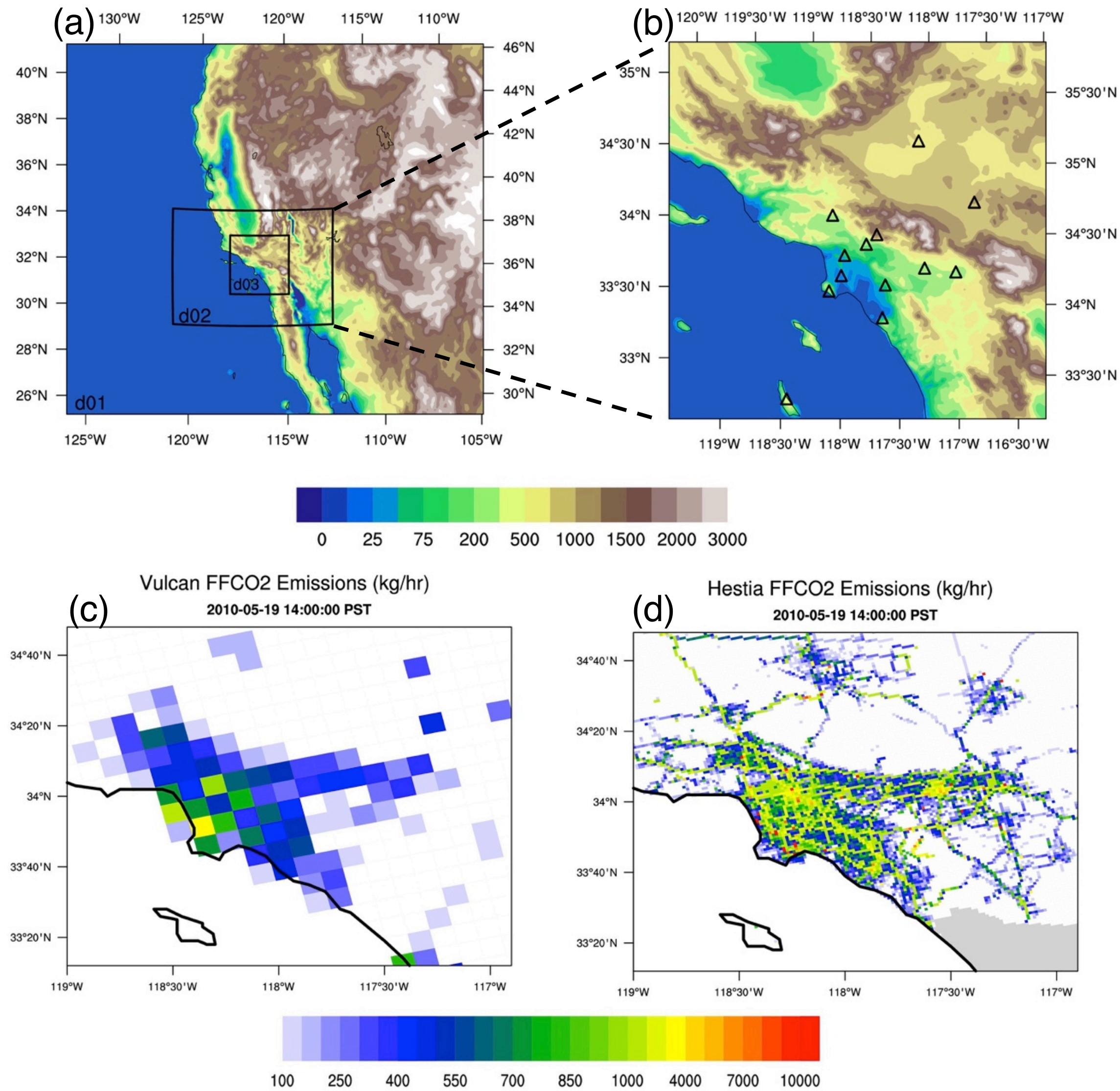

Figure 1. (a) Model domains. Contours are terrain height (unit: m). (b) The 1.3-km model domain (d03) and terrain height (unit: m). Triangles represent the locations of the GHG measurement sites. (c and d) Snapshots of the Vulcan and Hestia $FFCO_2$ emissions (unit: kg/hr) over the LA megacity at 14:00 PST on 15 May 2010.

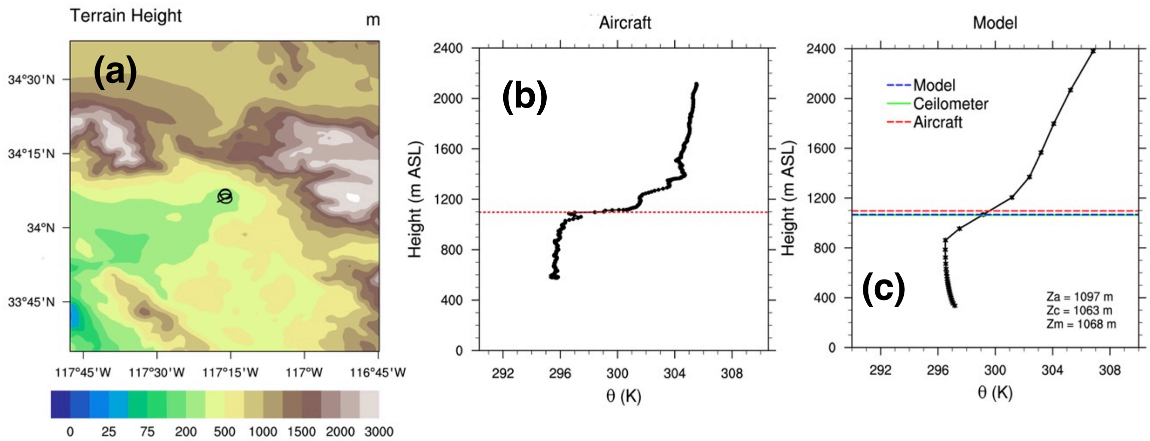

Figure 2. A case selected on 19 May 2010 at 12:25 (PST) (a) Location of the vertical profile flown by the CalNex aircraft and the neighbouring terrain heights (units: m). (b) In-situ potential temperature profile measured by the aircraft. The red dashed line at ~1100 m is the PBL height calculated based on the vertical gradient of potential temperature Θ(K). (c) Modelled potential temperature profile from the MYNN_UCM_d02 configuration. The red dashed line is the aircraft-determined PBL height (Z_a in masl). The solid green line is the PBL height measured by the Caltech ceilometer (Z_c in masl). The blue dashed line is the modelled PBL height (Z_m in m), almost identical to the green line.

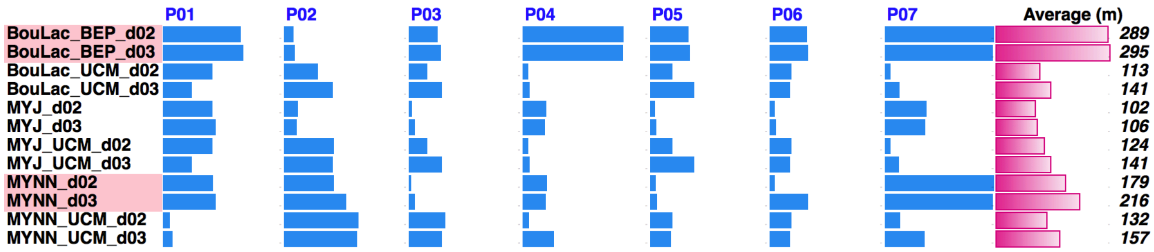

Figure 3. Absolute difference between the aircraft-determined and modelled PBL height for each profile: P01, P02, …, and P07 (blue bars). The pink bars in the last column represent the averaged bias over all of the profiles for each configuration. Note that the shorter the bar, the better agreement of the model with the observations.

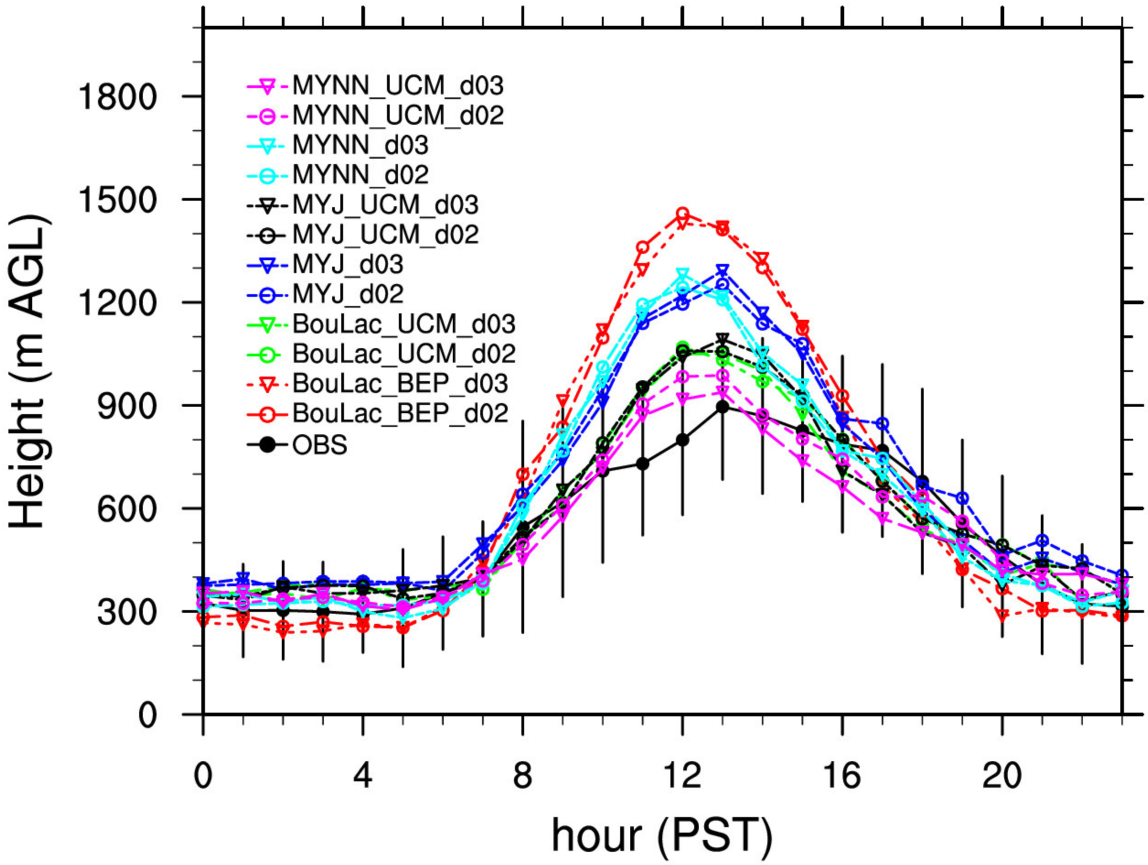

Figure 4. Average diurnal variation of the ceilometer-measured and modelled PBL heights at California Institute of Technology (Caltech) in Pasadena, CA during 15 May through 15 June 2010. Error bars indicate standard deviations of the means of the ceilometer measurement.

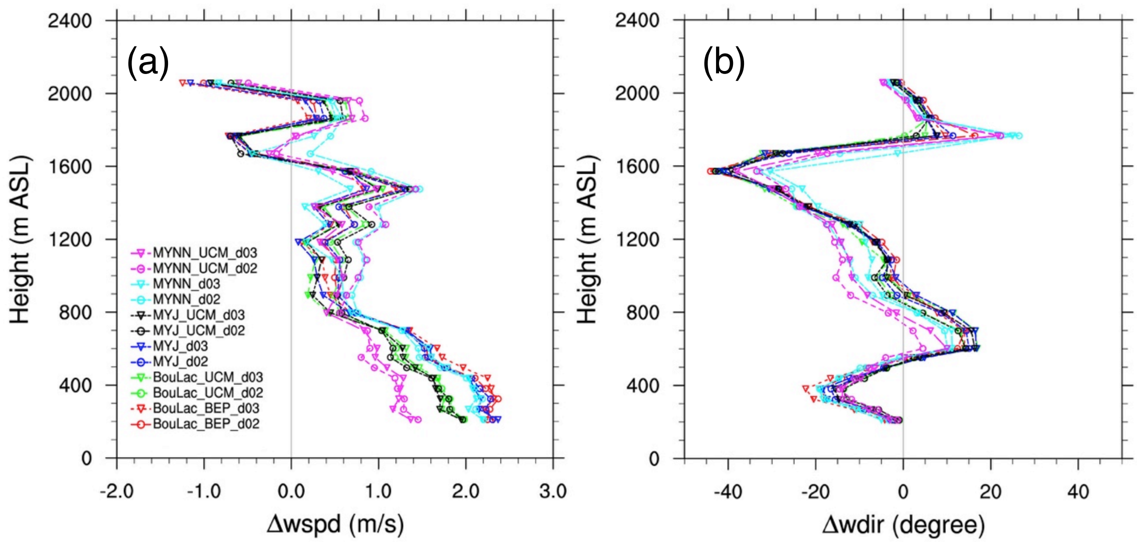

Figure 5. Average differences of wind profiles between the simulations and observations (model – wind radar profiler) at the Los Angeles International Airport (LAX). (a) The difference for wind speed (unit: m/s); (b) for wind direction (unit: degree). Note that these results are for daytime 1100 – 1700 PST only.

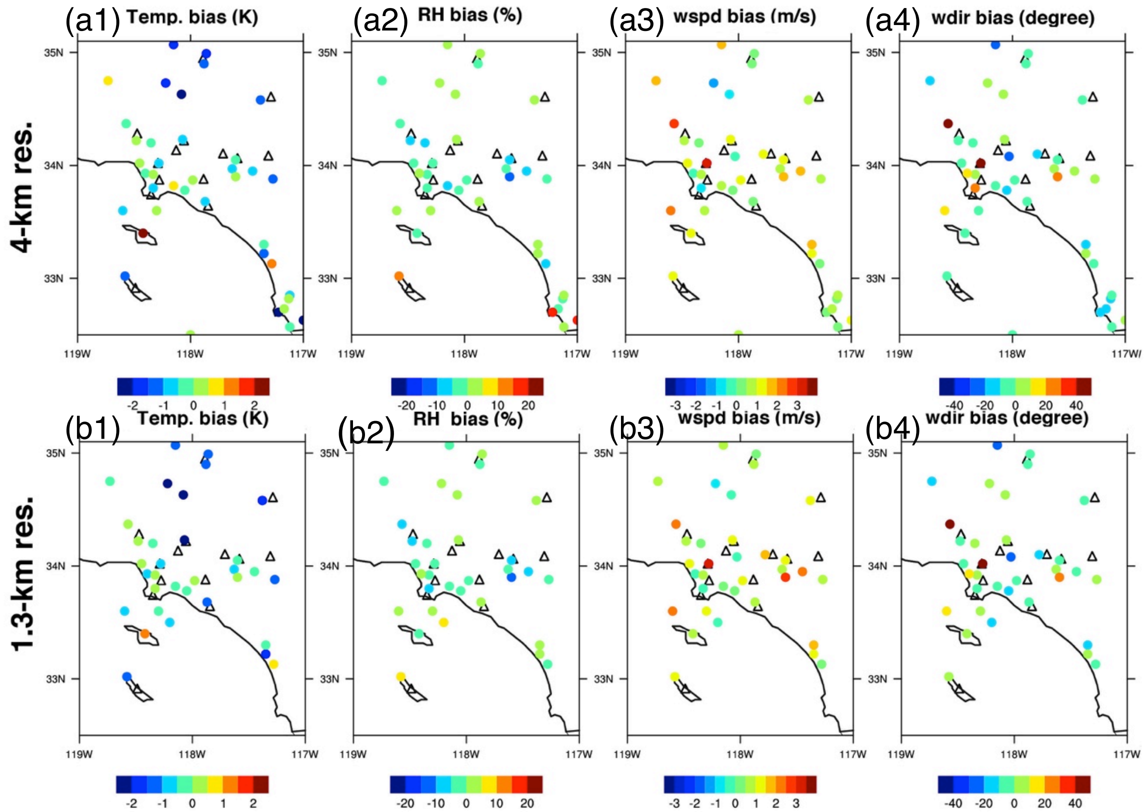

Figure 6. Bias maps of atmospheric state variables from the MYNN_UCM runs versus National Weather Stations (NWS) over the LA megacity (Model – NWS): (a1-a4) 4-km run; (b1 – b4) 1.3-km run. Black triangles indicate the locations of the GHG measurement sites. Note daytime 1100 – 1700 PST only.

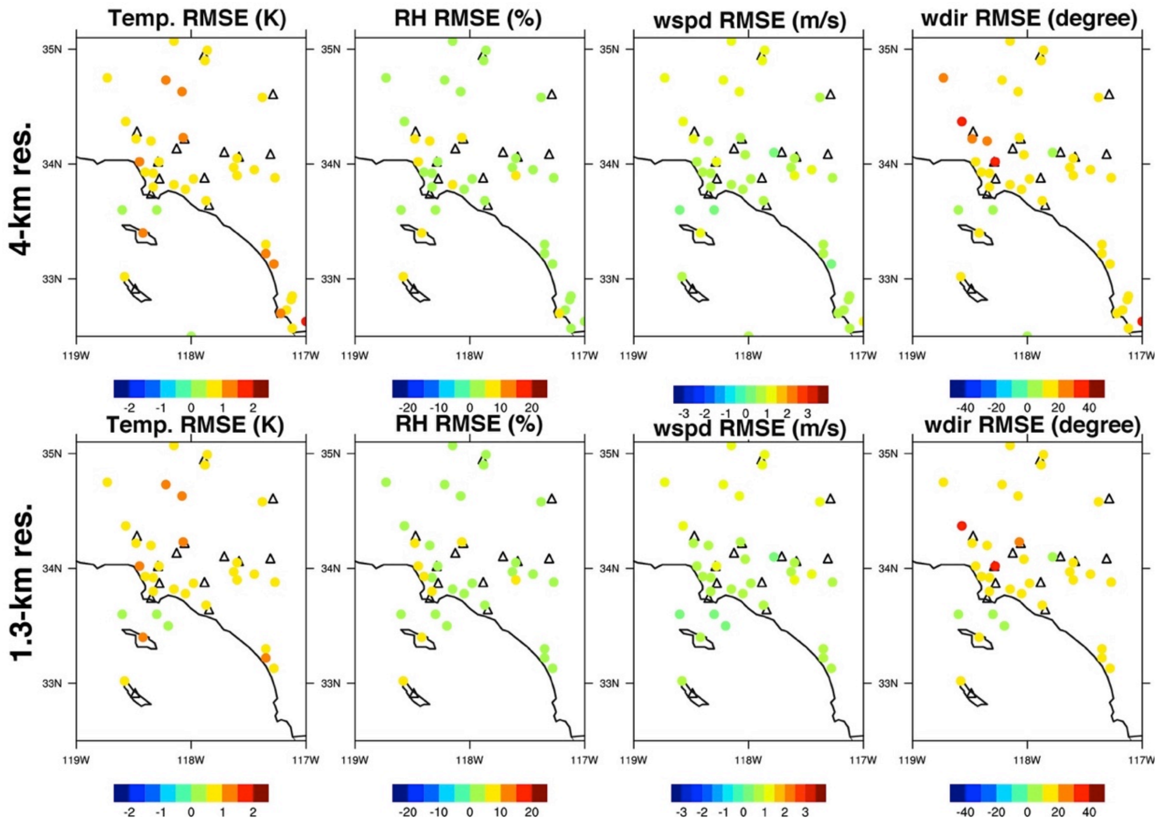

Figure 7. *RMSE* maps of atmospheric state variables from the MYNN_UCM runs versus National Weather Stations (NWS) over the LA megacity: (a1-a4) 4-km run; (b1 – b4) 1.3-km run. Black triangles indicate the locations of the GHG measurement sites. Note daytime 1100 – 1700 PST only.

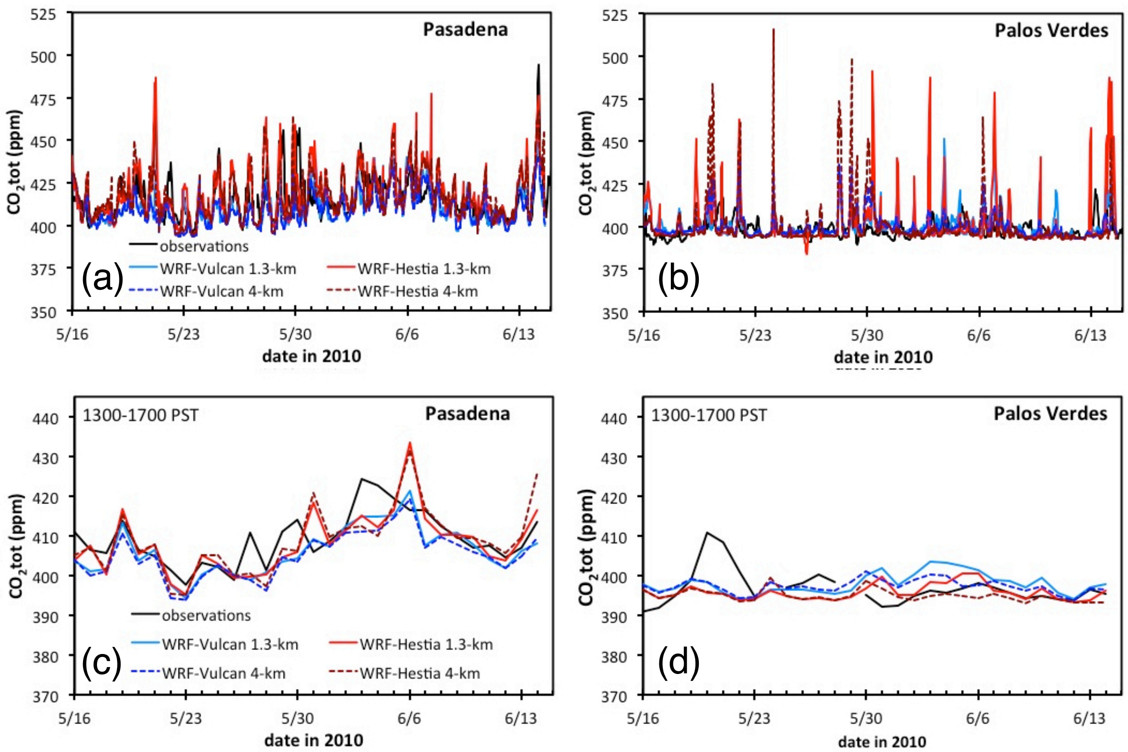

Figure 8. Comparison of the observed and modelled CO_2 concentrations at the (a and c) Pasadena and (b and d) Palos Verdes sites: (a and b) hourly time series, (c and d) daily afternoon averages for 1300 – 1700 PST.

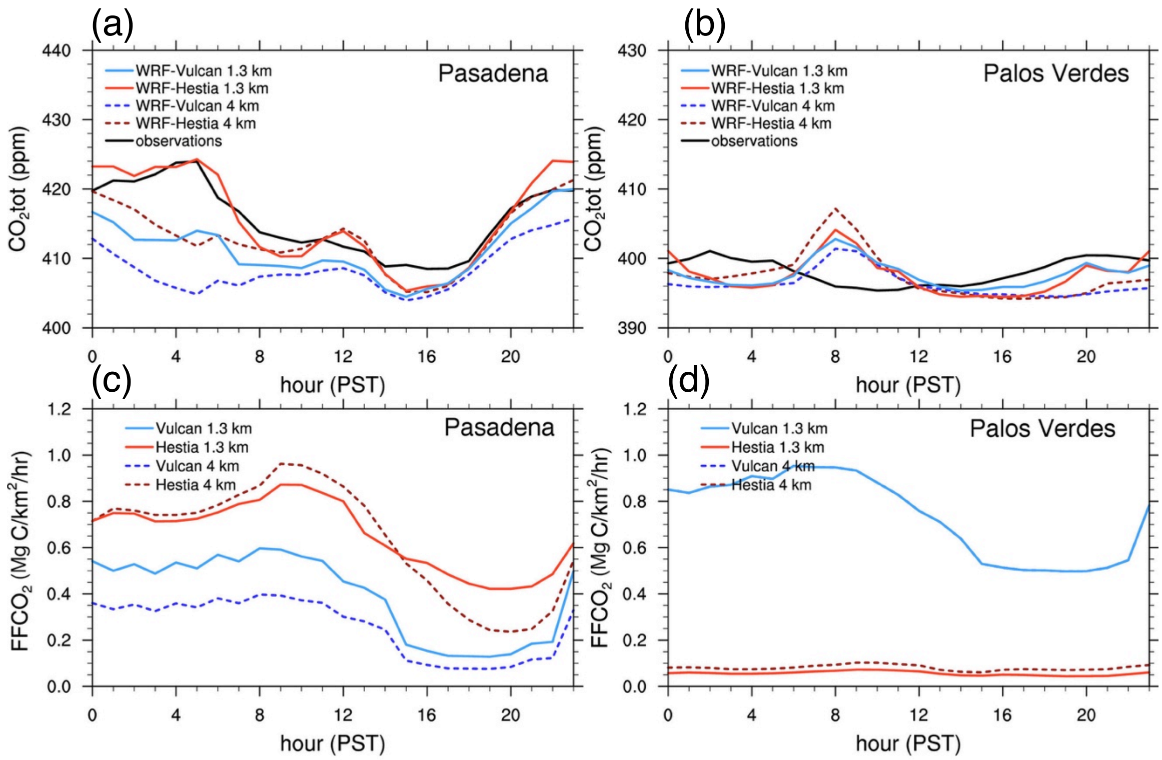

Figure 9. Averaged diurnal variation of observed and modelled CO_2 concentration and $FFCO_2$ emissions for the (a and c) Pasadena and (b and d) Palos Verdes sites during CalNex-LA. Note that Vulcan 4-km overlaps with Vulcan 1.3-km in Figure 9d.

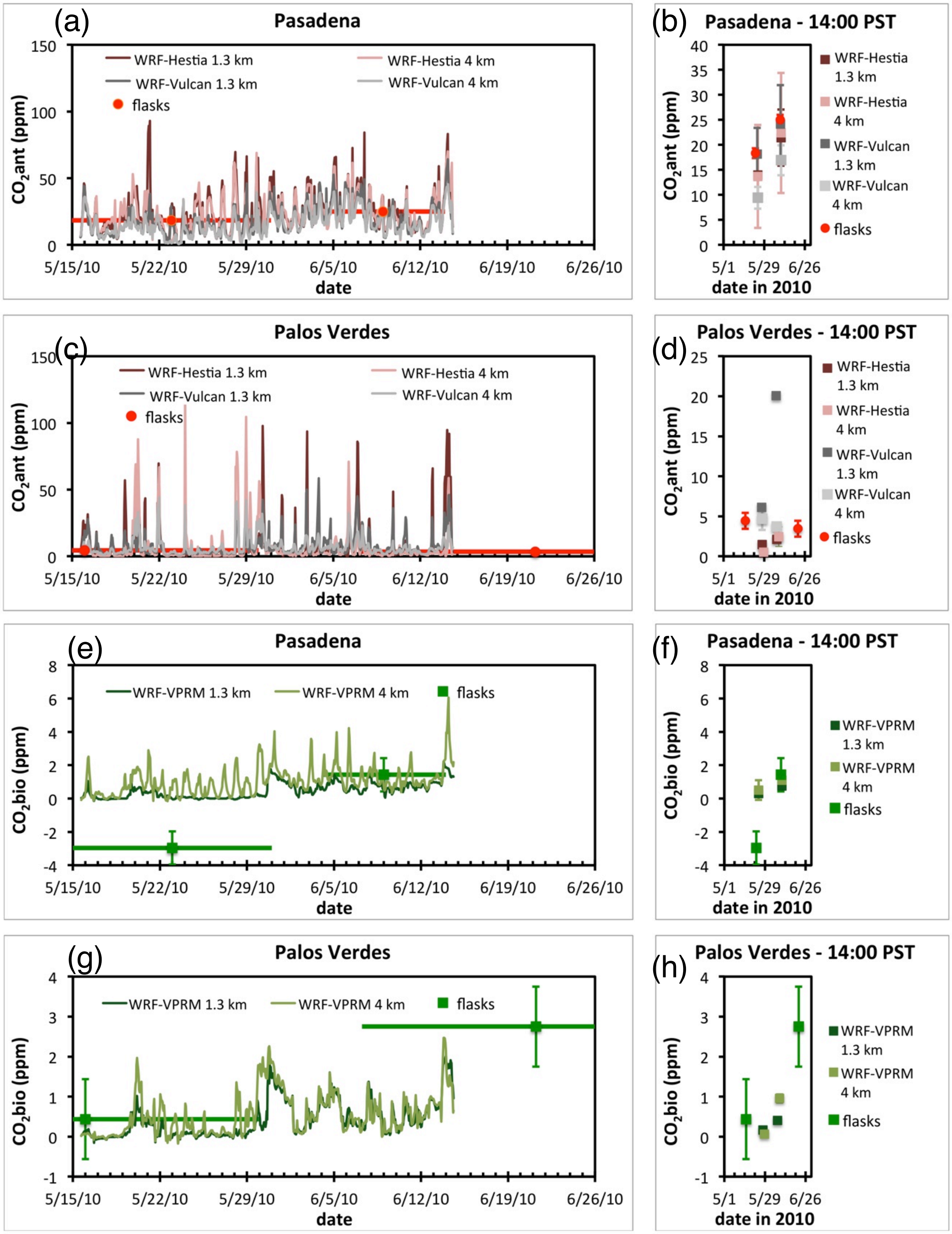

Figure 10. Comparisons of flask-sampled and modelled (a-d) anthropogenic fossil fuel and (e-h) biogenic CO_2 concentration. Left column: hourly time series. The horizontal error bars on the flask-sampled data points indicate the range of dates combined in each sample. Note that much of the time period for the $\Delta^{14}C$ samples at the Palos Verdes site

is before or after our modelling period. Right column: Averages at 1400 PST during CalNex-LA. See Newman et al. (2016) for details about the sites and sampling information.

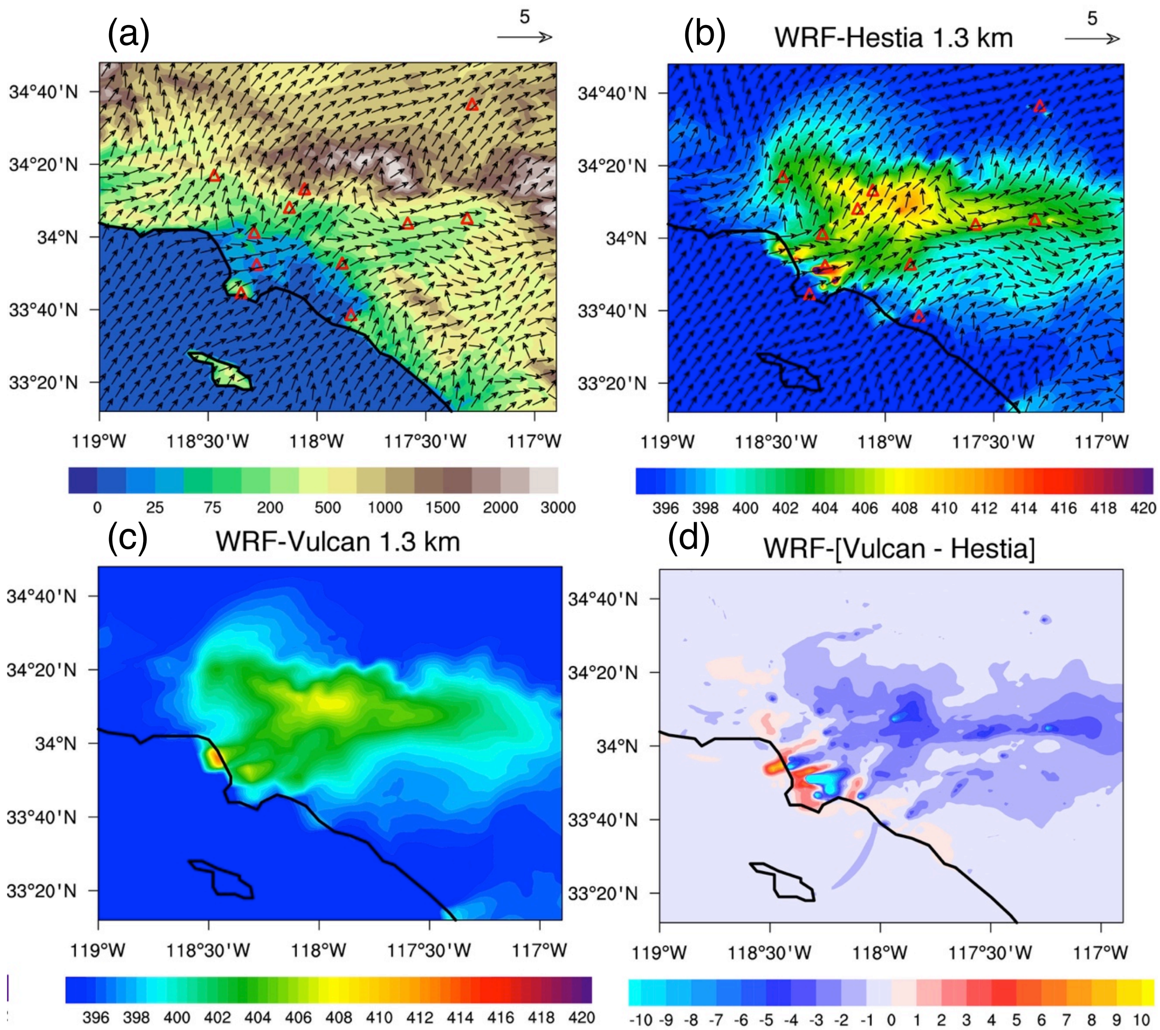

Figure 11. (a and b) The first empirical orthogonal function (EOF 1) for the surface wind pattern (black arrows) simulated by MYNN_UCM_d03 at 1400 PST during CalNex-LA. EOF 1 accounts for 48.1 % of the variance in the average winds. Contours: (a) terrain height (unit: m); (b) the modelled surface CO_2 concentration (unit: ppm) from the 1.3-km WRF-Hestia run. The red triangles indicate the locations of the GHG measurement sites. (c) The modelled CO_2 concentrations from the 1.3-km WRF-Vulcan run (unit: ppm). (d) The difference in the modelled CO_2 concentrations between the 1.3-km WRF-Vulcan and WRF-Hestia runs (unit: ppm).

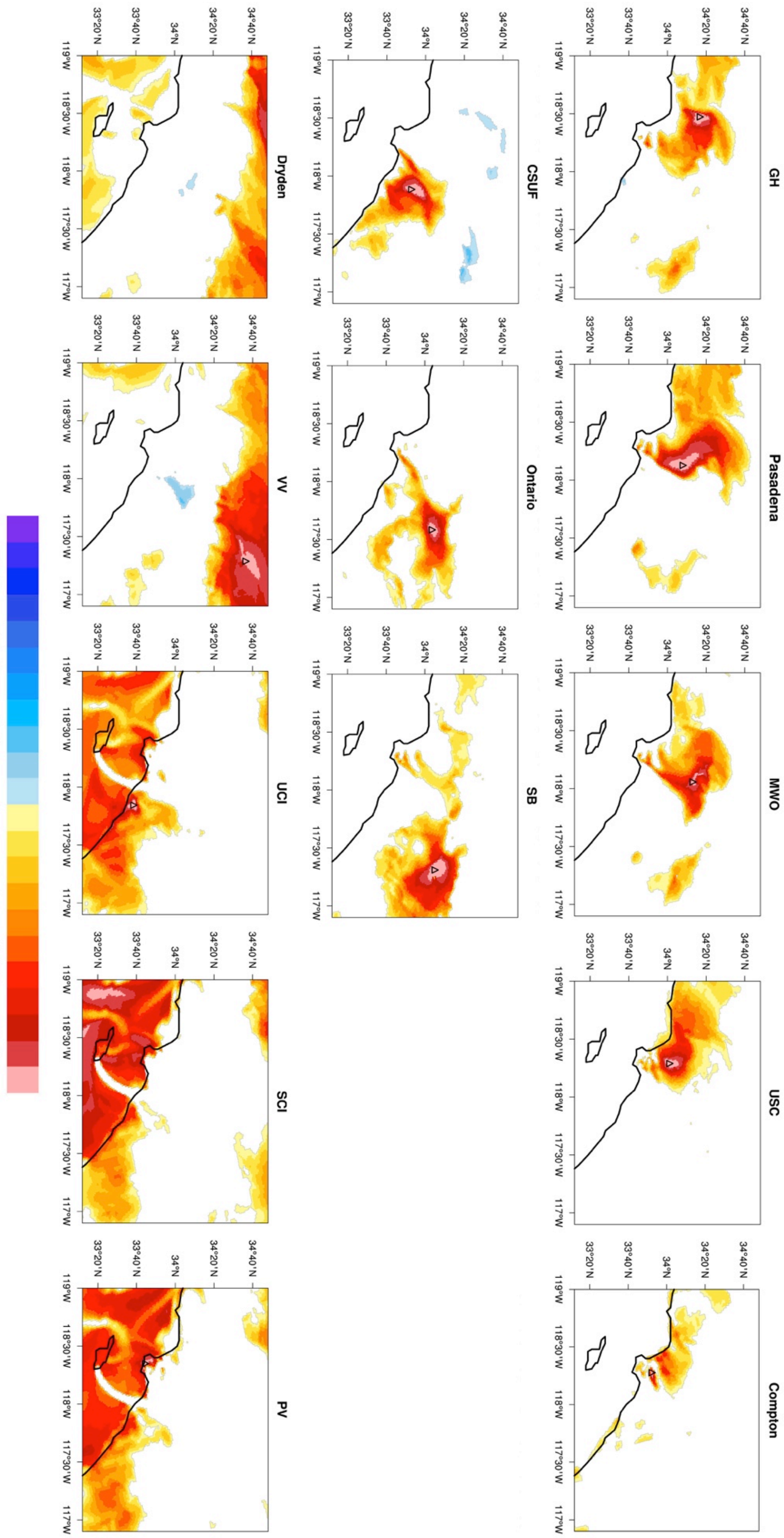
GH
Pasadena
MWO
USC
Compton
CSUF
Ontario
SB
Dryden
VV
UCI
SCI
PV
34°40'N
34°20'N
34°N
33°40'N
33°20'N
119°W
118°30'W
118°W
117°30'W
117°W
-0.9
-0.75
-0.6
0
0.6
0.75
0.9

Figure 12. The spatial correlation map (R) of the 1.3-km WRF-Hestia simulated CO_2 concentration between each site and the remainder of the domain at 1400 PST during the CalNex-LA campaign. The correlation map was constructed by calculating the simultaneous correlation of the site CO_2 to the CO_2 over rest of the LA megacity. Note that only those pixels that pass the *t*-test at the significance level of 0.01 ($|R| \geqslant 0.46$) are coloured.

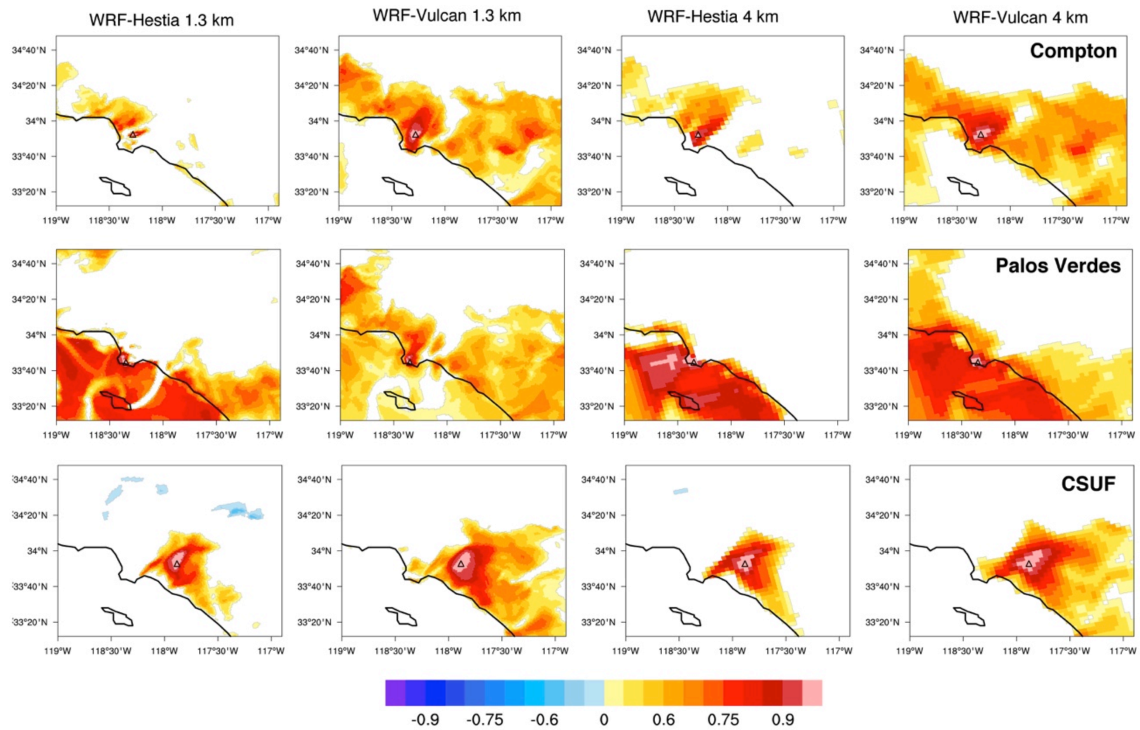

Figure 13. Same as Figure 12 but for the Compton (top row), Palos Verdes (middle row), and CSUF (bottom row) sites only. Shown are the correlation maps of these three measurement sites for the 1.3-km WRF-Hestia (first column), 1.3-km WRF-Vulcan (second column), 4-km WRF-Hestia (third column), and 4-km WRF-Vulcan runs (fourth column). Note that only those pixels that pass the *t*-test at the significance level of 0.01 (|R| ⩾ 0.46) are coloured.

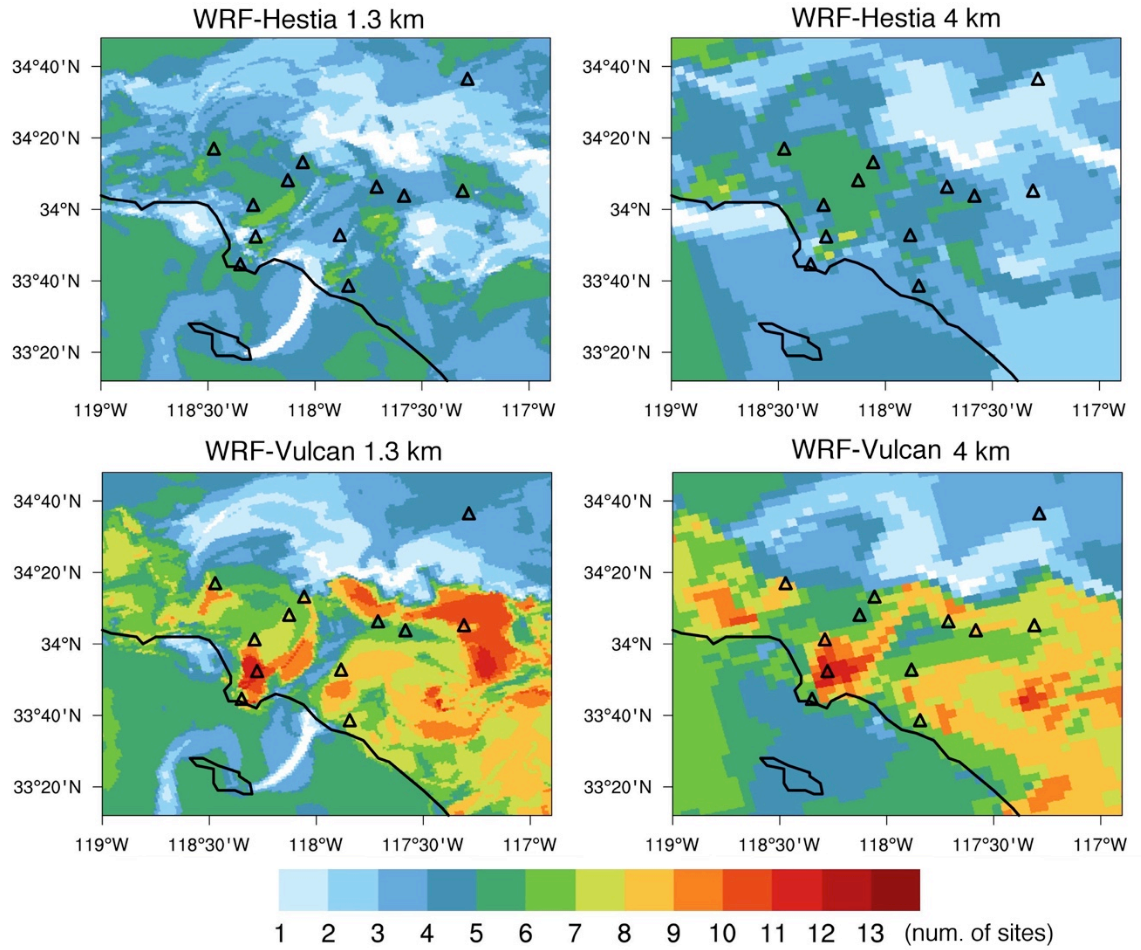

Figure 14. Composite maps of spatial correlation (R in Figure 12 and 13) for the 1.3-km WRF-Hestia, 1.3-km WRF-Vulcan, 4-km WRF-Hestia, and 4-km WRF-Vulcan runs. Each composite map was constructed by determining the number of the observation sites for which |R| is greater than 0.46 at each grid cell. |R| = 0.46 is the critical value at the significance level of 0.01 of *t*-test. Specifically, white cells indicate that no sites are correlated well at the location; dark red cells indicate that over 13 sites have good correlation at the location. The SCI and Dryden sites are not shown on these maps.

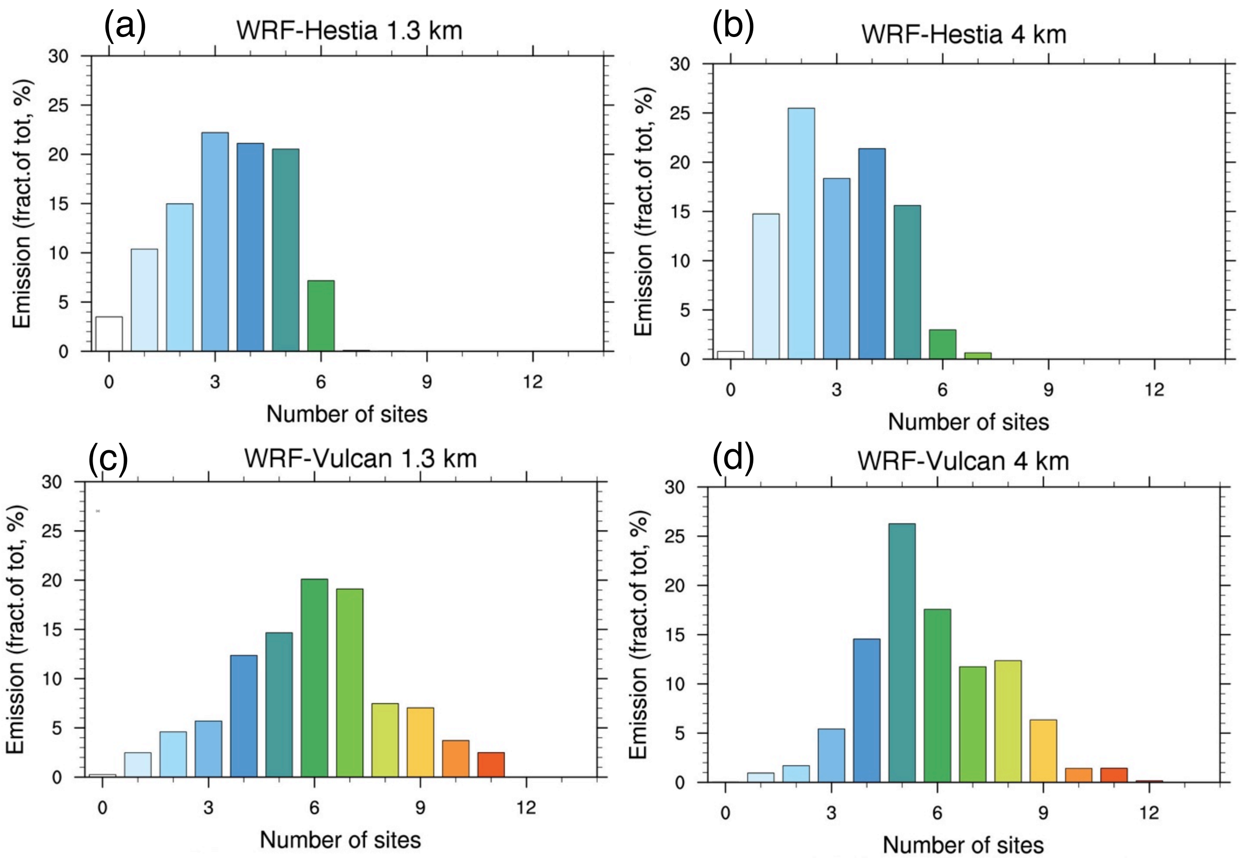

Figure 15. The fraction of the $FFCO_2$ emission over the LA megacity as function of the number of the GHG measurement sites that covers the area (see Figure 14) for (a) 1.3-km WRF-Hestia, (b) 4-km WRF-Hestia, (c) 1.3-km WRF-Vulcan, and (d) 4-km WRF-Vulcan runs during CalNex-LA. Colour scale is the same as in Figure 14.

Figure S1. Google Earth map showing the location of the 14 GHG measurement sites, only 13 of which are within the innermost model domain, the exception being the La Jolla site.