We thank the careful review by both reviewers. Both of them mentioned that the paper was relevant to GMD, but needed to be edited more carefully. They offered suggestions on improving its readability. We have substantially edited the paper with all their comments addressed. Please see below our point-by-point response (in blue) to their general and specific comments (in black).

Response to Referee 1's comments

GENERAL COMMENTS
==========
The main goal of the manuscript by Huang et al. is to quantify the effect of different initial/boundary conditions on meteorological variables that drive isoprene emissions. The authors show convincingly how choice of land initial conditions (and in particular the time used for initialisation) have a much larger impact on atmospheric temperatures, and therefore isoprene emissions, the atmospheric conditions.

This is important work that should be of interest to both the meteorological and atmospheric chemistry modelling communities. The work is well within the scope of GMD. My main concerns are around presentation – I find the text very hard to read, the title uninformative. As a result, the important and interesting messages get lost amongst the details. Substantial editing is required before publication in GMD. More details below.

Title – The title is long and detailed, but to me it doesn't accurately represent what is to come in the text. As someone who doesn't use the NASA-Unified WRF model, I wouldn't read this paper, but having now read it I realise it is relevant to my work after all! I encourage the authors to revise it to inform the reader that the focus is on the impacts of different choices of land and atmosphere initial conditions on ability to simulate biogenic emissions. I don't think the choice of SEAC4RS and DISCOVER-AQ is relevant enough to include in the title – these just happened to be the most relevant evaluation observations available.

The revised title is: "Biogenic isoprene emissions driven by regional weather predictions using different initialization methods: Case studies during the SEAC^{4}RS and DISCOVER-AQ airborne campaigns".

We noticed that GMD has a requirement for "model description papers" of *"The main paper must give the model name and version number (or other unique identifier) in the title"* (http://www.geoscientific-model-development.net/about/manuscript_types.html), and to be safe we included "NASA-Unified WRF v7" in the previous title although this is a "model evaluation paper". We have removed the model name and version because we do agree with the reviewer that the methodology in this paper may interest other regional modelers. The model name and version can still be found in the abstract.

The revised primary title has been shortened with the campaign names moved to the secondary title. The campaign names in the title serve as identifiers of the study region/period as well as the datasets used. This usage is often seen in publications, and has been recognized to be informative.

Readability – The text throughout is dense, often hard to follow, and full of grammatical and spelling errors. I highlight some but not all of these below. I suggest a very careful read-through and edit (if not by the authors, then perhaps by a professional editor). There are a very large number of acronyms used throughout, and since this work should appeal to multiple modelling communities, it may be useful where possible to use a description rather than an acronym. An appendix listing all acronyms (and what they refer to, since in some cases model names are not intuitive) would also be helpful/

The readers would take extra efforts to digest this paper, in part due to its cross-disciplinary nature, which actually makes this study unique. The paper has been substantially edited, with all of the language-related specific and technical comments addressed. We took your suggestion to add a categorized list of acronyms in the appendix-For each model related item, a short description is included; For each instrument/product related item, the measured variables we used for this study are included. We also deleted unnecessary definitions of acronyms (i.e., those only used once throughout the paper).

Abstract – There are too many acronyms in the abstract, making it hard to follow. Most of these are not needed until the main text. For example, lines 19-20, at this stage the reader doesn't need to know the specific land surface model or reanalysis data set used. Just stating ". . . demonstrating that initialising the input land surface model using a coarser resolution dataset led to significant positive biases. . ." would be much clearer. Same applies elsewhere. Also the sentence "This study emphasizes. . . chemical data assimilation" (lines 29-32) is very long and could be split into two for clarity.

These are good suggestions. We are now not specific about the land surface model (Noah) and datasets (NARR or NAM) used in the abstract. The sentence near L29-32 has been split and now reads as: "This study emphasizes the importance of proper land initialization to the coupled atmospheric weather modeling and the follow-on emission modeling. We anticipate it to be also critical to accurately representing other processes included in air quality modeling and chemical data assimilation."

Introduction – Much of the introduction doesn't seem well suited to the work presented in this particular manuscript. For example, the entire first paragraph that discusses isoprene impacts on ozone seems irrelevant as there is no further mention (or simulation) of ozone. As this is a GMD paper, the focus on O3 seems irrelevant (and in any case, the references seem spotty and cherry-picked to only discuss those that have shown large responses to isoprene – some others off the top of my head include Wu et al. 2008 (doi: 10.1029/2007JD008917), Millet et al., 2016 (doi:10.1021/acs.est.5b06367)). In the second paragraph, there is too much detail about MEGAN that can wait until the model description (for example, "on flexible scales", "MEGAN computes emissions based on. . .". I suggest reorganising to start with the current paragraph 3, then following with some of paragraph 2 (with emphasis on uncertainties/errors in MEGAN), then paragraph 4.

This study aims to improve the biogenic emission estimates as that may further improve ozone (a regulated pollutant) air quality modeling. There are indeed many publications on isoprene-ozone relationships, but we chose to cite in the opening paragraph several modeling studies over the similar focus regions to this study's. Several long sentences in this paragraph have been reworded. As evaluating isoprene emissions is a primary goal in this work, we introduce MEGAN first including its sensitivity to weather inputs, followed by the contents on weather modeling.

In the conclusion, the findings in this study are connected with air quality modeling: "We anticipate that improved weather fields using the better land initialization approach will also benefit the representation of the other processes (other weather-dependent emission calculations, transport, transformation, deposition) included in air quality modeling, and therefore can help reduce uncertainty in the simulated chemical fields."

"Usual" and "Control" – the choice of language to describe the simulations is very confusing. What is described here as "usual" is what most authors mean when they say "control" (i.e. the control simulation is the normal or base method, without modifications. So using the word "control" to describe the simulations where something new/different is done is very counterintuitive. It took me until the 2nd reading of the paper to actually work out which simulations used something new. Suggest "usual" becomes "control" and "control" becomes "sensitivity".
We added and modified some sentences in Section 2.1.3 to clarify the naming criteria. "Usual" means the method used "in default and many WRF applications". "Control" was chosen to be consistent with the usage in hydrological modeling community. It means that high quality forcing data were used in the offline LIS simulation. Such simulation is often compared with "open-loop" and "assimilated" LIS runs, in which low-quality precipitation data were used, to assess the usefulness of soil moisture data assimilation.

An example: Hain, C. R., W. T. Crow, M. C. Anderson, and J. R. Mecikalski (2012), An ensemble Kalman filter dual assimilation of thermal infrared and microwave satellite observations of soil moisture into the Noah land surface model, Water Resour. Res., 48, W11517, doi:10.1029/2011WR011268.

The definitions and descriptions of these various cases can be found at multiple places of this paper (Section 2.1.3, Table 1, Figure 1), so the readers won't miss them.

3.3 Uncertainty discussions – while I appreciate the attempt to discuss remaining uncertainties, I find the current version in Section 3.3 doesn't add much value. It needs to include some discussion (i.e. literature based) of what the expected impacts of each of these things would be (I'm mainly referring to the NUWRF-MEGAN section here). I also find (d) in this section completely meaningless ("other limitations") – if they aren't going to be stated or explained, then why bother mentioning them?
The list of error sources of NUWRF-MEGAN serves as a summary for the readers' quick references. Point a) has been expanded, and points b)-c) were quantitatively discussed in previous sections. It'd be challenging to quantify impact from points a) and d) for this specific case, but we wanted to explicitly point them out, and they can be good topics of future studies. Although we are not specific about "other limitations", point d) is included to help strengthen the point that having more confidence in the weather inputs would help better determine and quantify those impacts.

Fig. 1 – I like the idea of an overview figure, but find this one hard to understand. Suggest adding resolution to the figure, and separating (b) to show the different types of sensitivity simulations that were done (perhaps including names from the table).
This figure is designed to illustrate the different land and atmospheric initialization approaches we compared in this study. The squares, circles, and hexagons represent (NU)WRF model, its

atmospheric lateral boundary conditions, and initial conditions, respectively. We now include resolutions in the figure captions, and also have added a column in Table 1 to connect this figure with the different cases listed in the table.

SPECIFIC COMMENTS
==========

Pg 2, line 21-22: Also Zeng et al., 2015 (doi:10.5194/acp-15-7217-2015); Emmerson et al., 2016 (doi:10.5194/acp-16-6997-2016)
Thanks for the suggestions. We added Emmerson et al. (2016) which fits better into this context.

Pg 2, line 28-29: "Much less has been done. . ." – see Zheng et al., 2015 (doi:10.5194/acp-15-8559-2015), Bauwens et al., 2016 (doi:10.5194/acp-16-10133- 2016)
These are very good global-scale studies. We added "at multiple spatial-temporal scales" in this sentence. Smaller-scale analyses like this study are novel and would be more relevant to air quality modeling and management. The "less" word is relevant to the significant attention given to temperature and radiation. Both the direct (γ_{SM}) and indirect impacts (atmosphere-land coupling) are mentioned in the following sentences.

Pg 3, line 18: "key variables" – state them here
Added.

Pg 5, line 29 – Pg 6, line 18: I got very lost in this paragraph – which is an important one for understanding what was actually done! I think it needs to be rewritten to first highlight what the different types of simulations are (in simplified form), followed by one paragraph to describe the basic simulation and another (or two) to describe the modifications in the other simulations.
We broke this long paragraph into separate ones (points a-c). Long sentences were shortened. Figure 1 was also revised according to your suggestion, and its a-c panels now match the three points here.

Pg 7, first paragraph: reference Toon et al. (doi: 10.1002/2015JD024297)
Added.

Pg 7, line 25 and elsewhere: I think Figs. A1-A3 belong in a Supplement rather than an Appendix as there is no associated text and they are not related to one another.
We keep Figure A1, which is now included in Figure 2. Figures A2 and A3 were moved to the SI.

Pg 8, line 14: "Close to the estimated OH concentrations. . . 2006" – what are these? Please provide the value here to allow the reader to make the comparison
Reading from Figure 7 of Warneke et al. (2010), at the beginning and near the end of that flight (near Houston), the estimates fell within 2-6×10^6 molecule/cm^3, so we added "of approximately 2-6×10^6 molecule/cm^3..."

Pg 10, line 22: DIAL-HSRL derived PBLH is mentioned here but not show or discussed, besides saying the model is "closer to the reality" of it. More discussion or plots needed to justify that statement.

We added "Figure 2a" to be more specific. At the beginning of Section 3.1, we introduced that "..where mixed layer heights indicated by the DIAL-HSRL instrument were mostly below 2 km", so the ctrl run results in Table 2, which are described here, appear to be closer to the DIAL-HSRL data.

Pg 11, line 32: 30% seems to me like a big bias, seems like this requires some discussion
We compared this bias with literature in the following sentence, and some discussions were placed in Section 3.3, NUWRF point b. More sensitivity analyses and discussions on this would be divergent from the focus of this paper.

Pg 13-14, Section 3.2.2: I don't get anything out of this section and am unclear what I am meant to take away. I suggest moving it to a Supplement and referencing briefly in the main text.
This section is already very concise. The paper mainly focuses on a short period during two field campaigns, and decadal satellite data help contrast this period to conditions in the other years so that the readers would get a better sense about the general conditions and temporal variability over this region. The satellite based analysis also indicates the relationships between soil moisture, vegetation and HCHO/isoprene. Related sentences in the last paragraph of Section 1 have been modified.

Pg 14, line 27: "random" – is it really a random error? It seems like if there were even reasonably consistent wind directions in a given location, this would be a systematic error. . .
We removed "random".

Pg 15, line 9: "result in Wolfe et al. (2015) of 8+/1 mg/m2/h)". This comparison value is given in a completely different unit from the one shown in the figure (and it is also hard to compare when the figure and text are not co-located). Please provide the value in the text and make sure the units of the calculated value and the Wolfe et al. value are the same (doesn't matter which).
We now use the converted values from the results in Wolfe et al. (2015).

Pg 15, line 16: "LIS simulation" – explain what type of model LIS is for the reader that skips straight to the conclusion
Added: "a flexible land surface modeling and data assimilation framework"

Pg 16, line 8: "self spin-up method" – what is this? Needs some explanation
This is a method suggested by Angevine et al. (2014). Before this citation, we added: "i.e., running the model for a certain spinup period (e.g., a month) at least once, cycling its own soil variables, to allow the land variables to develop appropriate spatial variability".

Fig 3: Not obvious why (b) and (c) show different coverage in the observations, and the choice of >=80% is not justified. Either justify or (preferable) just show all data in (c)
Unlike model data, the daily CCI data do not always cover everywhere in September 2013, so in (c) we meant to compare the results over the region with better satellite data coverage. To improve the presentation, we now show data in (c) without screening by the data size. However, to help the readers understand these results, we added Figure S2 to include the usable data size and uncertainty (defined as the standard deviation of data from multiple sensors) of the 11 September daily combined soil moisture data.

Fig 8: Would be easier to understand/interpret is bars were grouped by AM, noon, and PM (then colored by different dates/campaigns)
Revised.

TECHNICAL COMMENTS
==========
As noted previously, there are many errors of grammar, spelling, or poor word choice. I point out some but not all of them here.
Pg 1 line 25: "modify" -> "reduce"
Since we are not specific about which atmospheric initialization (NARR or NAM) worked better than the other, "modify" would be more appropriate than "reduce".

Pg 1, line 29: "resulted" -> "resulting"
Done.

Pg 2, line 5: "50% of reduction" -> "50% reduction"
Done.

Pg 2, line 14: "over the NA" -> "over NA"
Done.

Pg 3, line 17: "experimented" not right word, perhaps "tested"
Done.

Pg 4, line 6: should "LAI" actually be "gamma_LAI" (with symbol)?
The original form is consistent with the usage in Guenther et al. (2012) and Sindelarova et al. (2014).

Pg 4, lines 7-8: "which needs to be better understood" – irrelevant to this work, delete
Done.

Pg 4, line 16: reference the actual data rather than a blog post
We now report the weekly data along with the data source from the Scripps.

Pg 4, line 27-28: this sentence references Fig 2 before Fig 1 – either need to rearrange the figure or (preferable) remove this sentence and put it in the section where Fig 2 is described
Thanks for catching this. This information is now first introduced in Section 3.1. We also split the previous Figure 2 to improve its readability.

Pg 5, line 13: "Same as in. . ." -> "As in. . ."
Done.

Pg 5, line 14: "four-soil layer" - should this be "four-layer soil"?
Pg 5, line 15: "widely used" – any references?

This sentence has been rewritten, with two references added on scientific and NOAA operational use of the Noah LSM.

Pg 5, line 26: "full clocks" – what does this mean? I've never seen this term.
Rewritten as: "..recorded hourly at 00:00 (minute:second).."

Pg 6, line 2: "Same as in. . ." -> "As in. . ."
Done.

Pg 6, line 5: "representing" -> "represent:
Done.

Pg 6, line 7: "focused" -> "focus"
Done.

Pg 6, line 26: "rate coefficients with OH" – does this mean "rate coefficient of isoprene with OH"?
Yes, and we corrected 12 to -12.

Pg 8, lines 27, 28: "were" "was" (twice)
The "were" in both sentences were changed to "was".

Pg 9, line 6-7: "Long-term soil moisture changes. . . p72-73)" – this sentence is irrelevant to the work presented here, delete.
Removed.

Pg 9, line 12: "at where" -> "where"
Changed to "over the regions where"

Pg 10, lines 10-13: "The magnitudes. . . daily minima" – long and confusing sentence, suggest splitting into two
It now reads as: "The magnitudes of the downtown aircraft isoprene measurements were slightly lower than most of the nearby surface measurements during the daytime (Figure 2f). Measured surface isoprene levels during the daytime were ~twice as high as during the nighttime (~0.2-0.3 ppbv) when biogenic emissions are at their daily minima."

Pg 10, lines 20-22: "RMSEs . . . uncertain) – long and confusing sentence, suggest splitting into two
It now reads as: "Root Mean Square Errors (RMSEs) of the modeled air temperature from the ctrl runs are ~1.5 °C lower than the 12 km usual run-produced. The PBLHs from the ctrl run were thinner (~0.6 km on average) and less spatially variable. They may be closer to the reality referring to the DIAL-HSRL data in Figure 2a, which can also be uncertain."

Pg 11, line 10: Fig 5 referenced before Fig 4 – reorder these figures
Figure 4 was first referenced in Pg 10, before this line.

Pg 12, line 3: “activate” -> “active”
Done.

Pg 12, line 26: “weakest” -> “weaker”
Done.

Pg 15, line 21: “resulted” -> “resulting”
Done.

Pg 16, line 10: “Experimenting simulations” -> “Experiments using simulations”
Done.

Fig 2a: colors are hard to see; how about plotting the color on top of the black line rather than below?
Fig 2d: colorbar missing label
Figure 2 has been split to improve its readability. A label “LAI (m^2/m^2)” was added near the colorbar of the previous Figure 2d (current 2c).

Response to Referee 2's comments

General Comments:

The manuscript is not carefully edited. There are many errors (e.g. page 6, line 27 exponential "12" should be "-12"), many misuses of punctuation (e.g. semicolons and colons in the sentence on page 6, lines 4-10), and often lengthy hard-to-follow sentences (e.g. page 2, lines 7-10; page 2, lines 11-14; page 6, lines 4-10). All of these issues stand in the way of communicating the science and leave the reader confused. Some examples of these are provided, but please read through the manuscript carefully to identify and address these issues.

The readers would take extra efforts to digest this paper, in part due to its cross-disciplinary nature, which actually makes this study unique. The paper has been substantially edited, with all of the comments by both reviewers addressed. We believe the readability of the paper has been significantly improved.

The exponential 12 was corrected to "-12" and we apologize for any confusion due to this typo.

The sentences near lines 4, page 6 have been rewritten per Referees' suggestion. The semicolon was used to separate point a) and point b), but now points a)–c) are written as separate paragraphs. Figure 1 was also revised according to Referee 1's suggestion, and its a-c panels match points a-c here.

Lengthy sentences pointed out by both reviewers were reworded.

There are many instances where "the" is used when it is not necessary, e.g. the is not needed in "the sunset" or "the sunrise" (page 7, lines 27-28). "The" is only needed when the noun is specific or particular, for example there is only one NUWRF-MEGAN model (page 8, line 2), one North America (page 2, line 14) and one sunrise (page 7, line 27), so "the" is not needed.

We agree and removed those unnecessary "the"s.

Figures are not presented in order in the text. The authors first introduce Figure 2(d) then goes on to mention Figure 1. Reorder figures to reflect the order in which they appear in the text.

This issue was also brought up by Referee 1. The figures have been rearranged in order. Also, Figure 2 has been split to improve its readability. Figure A1 is now Figure 2e, and Figures A2-A3 are now in the SI.

There is unnecessary repetition, in particular in Section 3.1. That SEAC4RS has no/minimal biomass burning interference (page 10, lines 2-3) is already stated in Section 2.2.1, as is the limited contribution of anthropogenic VOC interference to measurements in Conroe (page 10, lines 7-10). No need to state all this again in Section 3.1.

The sentence in Section 2.2.1 of "Furan is found in significant concentrations only in biomass burning plumes and no enhancement in the biomass burning tracer acetonitrile was observed in our case studies (details in Section 3.1)" is rather general. The data used to support this statement were not shown and discussed until Section 3.1. We believe the related figures and descriptions in Section 3.1 are still very important. Some descriptions on DISCOVER-AQ data in the following paragraph were merged into Section 3.1 to avoid repetition.

Possessive is not necessary when describing data from a model or measurement platform. For example, replace "NUWRF's day time surface air temperature" with "NUWRF day time surface air temperature". There are many other instances where apostrophes are used, but aren't required.

Please identify and correct these.
We agree and modified the related sentences.

The equation used to infer isoprene emissions relies on OH concentration and boundary layer height as input. I would like to see some discussion and evaluation of the diurnal variability of these parameters, as these are used as input in Eq. (2) to estimate isoprene emissions from the isoprene concentration observations.
A rough evaluation of NAQFC OH is in Section 2.2.2, with the descriptions on its diurnal variability added: "Close to the estimated OH concentrations of approximately 2-6×10^6 molecule/cm^3 near Houston on 16 September, 2006 (Warneke et al., 2010), the simulated PBL OH on 11 September, 2013 range from ~1.8×10^6 to ~4.0×10^6 molecule/cm^3 along the P-3B around Conroe. The OH levels are higher in late morning and around noon (>3.1×10^6 molecule/cm^3) than in the afternoon (~1.8×10^6 molecule/cm^3), qualitatively consistent with the observations in downtown Houston in May 2009 (Czader et al., 2013). The averaged OH on all P-3B flight days is within the same range, following similar diurnal variability."

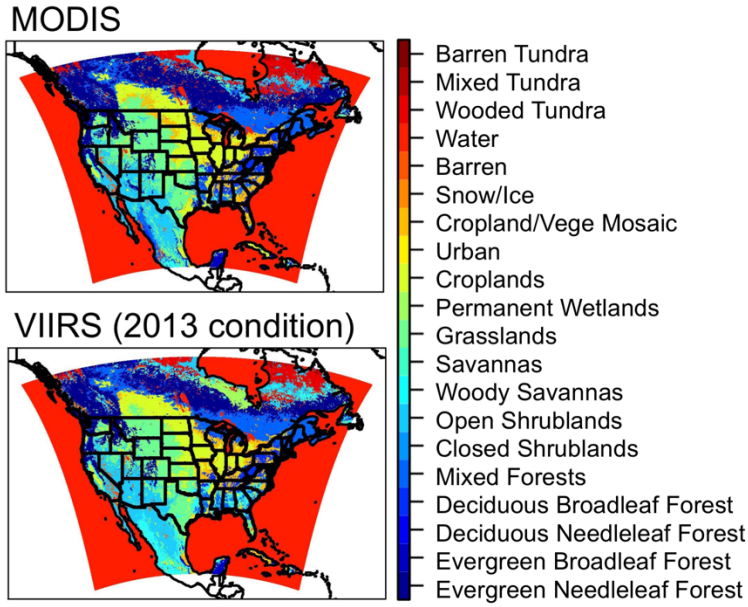
The evaluation of model PBLH is included in Section 3.1.1 and Table 3 for 11 September. The text briefly mentioned that the ctrl run PBLH "… are likely more realistic referring to the observed isoprene vertical profiles." The multi-day mean numbers are not significantly different: ~0.6/1.6/1.8km (AM/~Noon/PM).

Specific Comments:
Page 4, line 4: What emissions come from the atmosphere to the canopy? Do the authors mean the emissions consumed/deposited within the canopy (term ρ in Eq. (1))? If this is not considered in the emission model used in this study, then please clarify that this value is set to 1 to avoid confusion.
We clarified that ρ is assumed to be 1.0 in MEGAN v2.1.

Page 5, lines 15-18: What is the difference in land cover between the IGBP-derived land cover that is used in this study and the default used in MEGAN to justify using an updated land cover map? Some discussion of how using this updated land cover impacts isoprene emission modelling is needed.
The only two land use/land cover options currently available in WRF are USGS (24-category, Advanced very-high-resolution radiometer (AVHRR)-based) and IGBP-modified MODIS (20-category):
http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/users_guide_chap3.htm#_Land_Use_and
A third option is available in NUWRF called "University of Maryland", which is also AVHRR-based (Tao et al., 2013).
The MODIS option was chosen for this study because it was relatively more up-to-date than the other available options, and it is expected to lead to better meteorological output which was used to drive MEGAN. See the language "..which reflect more recent conditions than the other available options (Tao et al., 2013; Yu et al., 2012)."

None of these options represents the 2013 conditions. It is understood that many factors modified the US landscape such as biomass burning during SEAC^{4}RS in some regions. We are currently in

the process of testing and comparing LIS/NUWRF simulations with surface type inputs of IGBP-modified MODIS and VIIRS (Visible Infrared Imaging Radiometer Suite) (Zhang, R. et al., 2016). See the figure below comparing these two inputs, which is also included in the SI.



In Section 3.3, we recommended to use up-to-date land use/land cover input data in both (NU)WRF and MEGAN in future studies. Considering MEGAN needs a different format of land cover input, this would require mapping recent satellite information to the Community Land Model PFT categories, as in some previous studies (e.g., Lawrence and Chase, 2007).

Zhang, R., Huang, C., Zhan, X., Dai, Q., and Song, K. (2016), Development and validation of the global surface type data product from S-NPP VIIRS, Remote Sens. Lett., 7, 51-60, doi: 10.1080/2150704X.2015.1101649.

Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res., 112, G01023, doi:10.1029/2006JG000168.

Page 5, lines 23-24: What is "the urban surface option"?
We meant to activate the WRF urban canopy model (Chen et al., 2011) for the Noah LSM. However, since NUWRF couples the land surface model from LIS in which this does not seem to be activated. Therefore, this line has been removed.

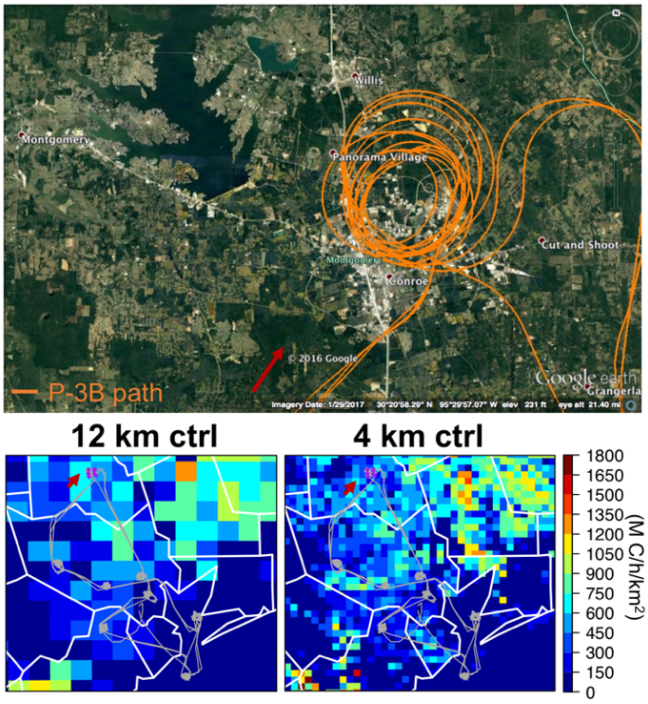
Page 5, line 26; Table 3 footnote: What is "full clocks"? This isn't standard terminology. Rather describe this in a way that can be understood.
Rewritten as: "..recorded hourly at 00:00 (minute:second)".

Page 7, lines 4-5: Why mention the August western US flights? Seems it has no relevance to this study and so can be removed.
This sentence was removed and the nearby sentences were reworded.

Page 11, line 32: Is "vastly similar" correct? In the next clause, the authors state that the difference is >30%.
"Vastly similar" refers to the comparisons among model runs. ">30%" refers to the model-observation differences. So it is correct. We modified this sentence slightly to avoid confusion.

Page 12, lines 29-31: Why does resolution induce a difference in isoprene emissions in the

atmospheric initialization sensitivity test in Figure 7c) and not in Figure 4c)?



This is a very good question. There can be three reasons for the discrepancies among 12 km, 4 km and aircraft-derived emissions. Two of them were already included in the earlier paragraph: 1) "…This is in part due to the coolest temperature from this NUWRF run, especially in the afternoon, as well as its weakest photosynthetically active radiation than the 12 km simulations' (i.e., by ~10 W/m^2 on average during the daytime)." 2) "This may also be resulting from some limitations of MEGAN's parameterization and uncertainty in its other inputs (e.g., PFT and LAI) on grid scale."

A third reason has just been identified and added, defined as the "representation error" due to discrepancies in different data resolutions as well as neglecting horizontal transport in deriving emissions from aircraft data. Multiple P-3B aircraft data points correspond to several NUWRF model grids, and the averaged NUWRF-MEGAN emissions were used in the comparisons. The bottom panels in the left figure (adapted from Figure 6b of the paper) show that for the grids collocated with the west side of the aircraft spiral (built-up land as indicated in the upper panel Google image), 4 km emissions are significantly lower than 12 km emissions. In the forestal/agricultural areas southwest to the spiral (highlighted by red arrows in the figure), 4 km and 12 km emissions are similar. It is very likely that higher emissions are from regions outside the spiral as the 4 km results indicate, and they were transported to the aircraft-sampled regions. On the 12 km grid, spatial variability of the emissions is not well distinguished, and the impact of horizontal transport did not appear to be evident. As a result, we see smaller discrepancies between the modeled and observation-derived.

Page 16, lines 1-2: Please provide references to back up the statement that many model comparison studies don't adequately assess the impact of model inputs.

A sentence in the introduction also supports this point. A few references have been provided there: "e.g., Millet et al., 2008; Warneke et al., 2010; Canty et al., 2015; Hogrefe et al., 2011". The Carlton and Baker (2011) includes some discussions on the impact of meteorological input. We added "cited in Section 1" to this sentence.

Figure 3: Please increase the size of the points in the Figure 3a) Observations panel so that the reader can easily compare the observations and model or instead show a scatterplot of the model versus the observations and include regression statistics.

The original observational data were regridded to the 12 km model grid for this plot, as mentioned in the figure caption, and therefore the data shown were in grid-size. We agree that this panel was a bit hard to read and regridding is unnecessary for making this plot. We now just show the original observations at their actual locations, in bigger size. Bulk statistics would not be quite relevant to

the point we made from this figure.

Figure 3 caption: Please say where the temperature observations are from. Are these NCEP?
Added.
We also added in the caption that "Note that for (b) and (c), warm (cool) colors indicate high (low) soil moisture values, opposite to the commonly used color scheme in hydrological studies." In most hydrological studies, warmer colors indicate drier conditions. We flipped the color scheme so that it is consistent with the usage in emission and air quality studies.

Biogenic isoprene emissions driven by regional weather predictions using different initialization methods: Case studies during the SEAC^{4}RS and DISCOVER-AQ airborne campaigns

Min Huang[1,2], Gregory R. Carmichael[3], James H. Crawford[4], Armin Wisthaler[5,6], Xiwu Zhan[7], Christopher R. Hain[2,a], Pius Lee[8], Alex B. Guenther[9]

[1]George Mason University, Fairfax, VA, USA
[2]University of Maryland, College Park, MD, USA
[3]University of Iowa, Iowa City, IA, USA
[4]NASA Langley Research Center, Hampton, VA, USA
[5]University of Oslo, Oslo, Norway
[6]University of Innsbruck, Innsbruck, Austria
[7]NOAA National Environmental Satellite, Data, and Information Service, College Park, MD, USA
[8]NOAA Air Resources Laboratory, College Park, MD, USA
[9]University of California, Irvine, CA, USA
[a]Now at: NASA Marshall Space Flight Center, Huntsville, AL, USA

Correspondence to: Min Huang (mhuang10@gmu.edu)

Abstract. Land and atmospheric initial conditions of the Weather Research and Forecasting (WRF) model are often interpolated from a different model output. We perform case studies during NASA's SEAC^{4}RS and DISCOVER-AQ Houston airborne campaigns, demonstrating that using land initial conditions directly downscaled from a coarser resolution dataset led to significant positive biases in the coupled NASA-Unified WRF (NUWRF, version 7) (near-) surface air temperature and planetary boundary layer height (PBLH) around the Missouri Ozarks and Houston, Texas, as well as poorly partitioned latent and sensible heat fluxes. Replacing land initial conditions with the output from a long-term offline Land Information System (LIS) simulation can effectively reduce the positive biases in NUWRF surface air temperature by ~2°C. We also show that the LIS land initialization can modify surface air temperature errors almost ten times as effectively as applying a different atmospheric initialization method. The LIS-NUWRF based isoprene emission calculations by the Model of Emissions of Gases and Aerosols from Nature (MEGAN, version 2.1) are at least 20% lower than those computed using the coarser resolution data-initialized NUWRF run, and are closer to aircraft observation-derived emissions. Higher resolution MEGAN calculations are prone to amplified discrepancies with aircraft observation-derived emissions on small scales. This is possibly resulting from some limitations of MEGAN's parameterization, uncertainty in its inputs on small scale, as well as the representation error and neglecting horizontal transport in deriving emissions from aircraft data. This study emphasizes the importance of proper land initialization to the coupled atmospheric weather modeling and the follow-on emission modeling. We anticipate it to be also critical to accurately representing other processes included in air quality modeling and chemical data assimilation. Having more confidence in the weather inputs is also beneficial for determining and quantifying the other sources of uncertainties (e.g., parameterization, other input data) of the models that they drive.

Deleted: Linkages between land initialization of the NASA-Unified WRF v7 and biogenic isoprene emission estimates

Deleted: initializing the Noah
Deleted: surface model
Deleted: using
Deleted: North American Regional Reanalysis (NARR)
Deleted: 's
Deleted: the
Deleted: 's
Deleted: fields
Deleted: the

Deleted: NARR
Deleted: the
Deleted: errors
Deleted: ,
Deleted: ed
Deleted: and its inputs'
Deleted: , which we

1 Introduction

The weather-dependent emissions of biogenic volatile organic compounds (BVOCs), including the highly reactive species isoprene (C_5H_8), contribute to the formation of secondary short-lived climate pollutants such as ozone (O_3) and secondary organic aerosol. Therefore, these emissions affect air quality on local, regional, and global scales, which feed back to the climate. For example, a modeling study by Li et al. (2007) showed that a 50% reduction in Houston isoprene emissions led to 5-25 ppbv summertime afternoon O_3 reductions at its urban areas, and the transport of isoprene from north of the urban Houston area had non-negligible impact on its isoprene budget within several days. Summertime isoprene emitted from the Missouri Ozarks (also known as "isoprene volcano", where a high density of oak trees efficiently emit isoprene), along with its oxidation product formaldehyde (HCHO), can be transported to urban areas (e.g., Chicago and St. Louis) and affect their O_3 burdens (Wiedinmyer et al., 2005). Ozone and peroxyacetyl nitrate produced from isoprene and other O_3 precursors can also affect air quality on hemispheric scale. Fiore et al. (2011) compared the O_3 sensitivity to a 20% change in North American (NA) isoprene emissions with the sensitivity to a 20% change in NA anthropogenic emissions. Over NA, the former was ~1/3 of the latter, and over Europe and North Africa, the former was more than half of the latter in summer and fall. Therefore, possible increases in future isoprene emissions due to land cover and climate change may offset the surface O_3 decreases due to controlling anthropogenic emissions in NA and its downwind continents.

The Model of Emissions of Gases and Aerosols from Nature (MEGAN, Guenther et al., 2006, 2012) has been frequently used to generate BVOC emissions on flexible scales for air quality modeling. MEGAN computes emissions based on emission source types and their densities, ambient carbon dioxide (CO_2) concentrations, and meteorological conditions (e.g., temperature, solar radiation, and moisture). It has been found that the MEGAN emissions are often higher than those calculated using other emission models, and are possibly associated with positive biases (e.g., Millet et al., 2008; Warneke et al., 2010; Carlton and Baker, 2011; Canty et al., 2015; Hogrefe et al., 2011; Emmerson et al., 2016). These biases, which still need careful validation with observation-based emission fluxes, can pose significant difficulties to accurately simulating isoprene and secondary air pollutants by chemical transport models. In addition to the impact of MEGAN parameterization, outdated/unrealistic land cover input data and uncertainties of the meteorological inputs are important causes of these biases (e.g., Guenther et al., 2006, 2012; Carlton and Baker, 2011). The positive biases in surface air temperature and radiation fields from meteorological models have been identified as major sources of uncertainty, and certain solutions have been established to reduce the biases such as substituting the modeled radiation with satellite radiation products. Much less has been done at multiple spatial-temporal scales to explore the biases imported from the modeled soil moisture fields although satellite observations have suggested negative correlations between BVOC emissions and soil moisture (e.g., Duncan et al., 2009). MEGAN sensitivity calculations by Sindelarova et al. (2014) showed weak direct impact of soil moisture on the isoprene emissions over vegetated and moist surfaces. However, the variability in soil moisture can indirectly impact BVOCs emissions and their atmospheric distributions through affecting air/canopy temperature and planetary boundary

Deleted: (SOA).
Deleted: of
Deleted: Wiedinmyer et al. (2005) found that summertime
Deleted: eastward and northward
Deleted: to
Deleted: .
Deleted: (PAN)
Deleted: reported that for summer and fall,
Deleted: over Europe and North Africa
Deleted: was more than half of
Deleted: , greater than
Deleted: O_3 response over
Deleted: NA.
Deleted: the
Deleted: the
Deleted: the
Deleted: the
Deleted:).
Deleted: the
Deleted: the
Formatted: English (US)
Deleted: the
Deleted: emission
Deleted: .,
Deleted: '
Deleted: the

layer height (PBLH), the key factors controlling isoprene emissions and satellite column measurements (e.g., Palmer et al., 2003; Duncan et al., 2009), especially over the US transitional climate zones including the Great Plains and some east Asian regions (e.g., Miralles et al., 2012; Zeng et al., 2014; Lee et al., 2016; Zaitchik et al., 2012 and the references therein). Therefore, accurately simulating land states and correct representation of land-atmosphere interactions by meteorological models can provide improved inputs for the MEGAN emission calculations.

The performance of coupled land-atmospheric modeling relies on numerous factors such as the choice of land surface model (LSM), nudging methods, and land use/land cover input data (e.g., Jin et al., 2010; Byun et al., 2011; Huang et al., 2016). The initialization of soil moisture and other land fields have also been shown important to the modeled atmospheric weather states (e.g., air temperature, humidity, winds, precipitation, and PBLH) and latent/sensible heat fluxes. Suitable and sufficient LSM spin-up as well as land data assimilation can benefit land surface modeling and the coupled atmospheric weather prediction (e.g., Rodell et al., 2005; Case et al., 2008; 2011; Zeng et al., 2014; Angevine et al., 2014; Collow et al., 2014; Lin and Cheng, 2015; Santanello et al., 2013, 2016). However, potential benefit of appropriate land initialization of numerical weather models to emission and air quality related studies needs to be better understood.

In this study, we performed a number of NASA-Unified Weather Research and Forecasting (NUWRF) sensitivity simulations, in which different land and atmospheric initialization methods and model grid resolutions were tested. The simulated weather states, especially the key variables impacting isoprene emissions such as surface air temperature and radiation, as well as heat fluxes were evaluated against in-situ and remote sensing observations. Isoprene emissions were then calculated by MEGAN, driven by these various NUWRF simulations, and these emission estimates were compared with the aircraft observation-derived during two NASA airborne campaigns in September 2013. The paper is structured as follows: We will first introduce the isoprene emissions calculated by MEGAN (Section 2.1) and observations (Section 2.2), followed by model evaluation datasets (Section 2.3). The NUWRF performance and its impact on isoprene emission estimates will be shown on a specific day in September 2013 (Section 3.1), as well as for multiple days in that month when research flights were executed for an airborne campaign. We will also show extended analyses on interannual variability of drought and vegetation conditions in relation to isoprene emissions during 2005-2014 (Section 3.2). The sources of uncertainty of the emissions will be summarized (Section 3.3) before the conclusions and suggestions on future directions are given in Section 4.

2 Methods and data

Deleted: the
Deleted: the
Deleted: the
Deleted: the
Deleted: the
Deleted: the
Deleted: the
Deleted: experimented
Deleted: the
Deleted: , which
Deleted: then be further compared with the decadal mean
Deleted: discussed in detail

2.1 Bottom-up emission calculations

2.1.1 MEGAN model version 2.1

The most recent version of MEGAN (version 2.1, Guenther et al., 2012) generates the net primary biogenic emissions that escape into the atmosphere, i.e., these are only emissions from the canopy to the atmosphere and do not include the chemical fluxes from the atmosphere into the canopy, which on average can be a few percent of the net primary emissions (Guenther et al., 2012). The emissions are estimated based on Equation (1):

$$\text{Emission} = [\varepsilon][\gamma][\rho] \quad (1)$$

where [ε] stands for the emission factor at standard conditions, [ρ] accounts for the production and loss within the plant canopies, assumed to be 1.0, [γ] is a unitless emission activity factor, a product of multiple factors that account for the emission response to light (γ_P), temperature (γ_T), soil moisture (γ_{SM}), leaf age (γ_A), leaf area index (LAI), as well as CO_2 inhibition (γ_{CO2}), the process that reduces isoprene emissions when ambient CO_2 concentration increases above the level of 400 ppmv. Among the meteorological variables, MEGAN emissions are strongly sensitive to radiation and air temperature (Guenther et al., 2012, and the references therein), but less sensitive to soil moisture over vegetated and moist surfaces including the central/southeastern US (Sindelarova et al., 2014), where root zone soil moisture is usually larger than a threshold (the sum of a small empirical value and the soil type-dependent wilting point) above which γ_{SM}=1.0.

The stand-alone version of MEGAN 2.1 was used in this study, which requires the users to provide meteorological and land cover inputs. The land cover and meteorological inputs we used in this study will be introduced in detail in Sections 2.1.2 and 2.1.3, respectively. We ignored the CO_2 impacts on the emissions (i.e., γ_{CO2} = 1.0), as September 2013 CO_2 in-situ measurements at the Mauna Loa, Hawaii Observatory are nearly 400 ppmv (i.e., weekly averages within 393.22-393.53 ppmv, http://scrippsco2.ucsd.edu/assets/data/atmospheric/stations/in_situ_co2/weekly/weekly_in_situ_co2_mlo.csv). Sensitivity calculation by Sindelarova et al. (2014) showed that for the year of 2003, the inclusion of CO_2 impact resulted in a 2.7% increase in MEGAN emissions globally under the 373.1237 ppmv CO_2 environment, corresponding to a γ_{CO2} of 1.0277. Therefore, omitting the CO_2 impacts in this study would not introduce large biases. The γ_{SM} value was also 1.0, as the root zone soil moisture from our meteorological input exceeded the sum of the empirical value and the wilting point (from Chen and Dudhia, 2001) over the regions of interest.

2.1.2 Plant functional type (PFT) and LAI input data

The recommended high-resolution 30 arc-second PFT input files for the year of 2008 (Guenther et al., 2012; http://lar.wsu.edu/megan/docs/NorthAmericaPlantFunction/), based on the Community Land Model 16 PFT classification system, were interpolated to the NUWRF model grids for use in this study. The LAI input was based on the Terra-Moderate

Deleted: .

Deleted: which needs to be better understood.

Deleted: the

Deleted: the

Deleted: the

Deleted: the monthly mean CO_2 concentration in

Deleted: was measured to be nearly 400 ppmv (i.e., 393.31 ppmv)

Deleted: (https://weatherdem.wordpress.com/2013/10/10/september-2013-co2-concentrations

Deleted: -31ppm). The sensitivity

Resolution Imaging Spectroradiometer (MODIS) 8-day product, and the grids with missing data were filled with the monthly-mean MODIS product.

Deleted: Figure 2d shows the LAI input for MEGAN on 11 September, 2013 over the Greater Houston area, indicating denser vegetation north and northeast to the downtown Houston area.

2.1.3 NUWRF meteorological simulations using different land and atmospheric initialization methods

Deleted: The NASA-Unified WRF

The MEGAN emission calculations in this study were driven by the meteorological fields simulated by the NUWRF (Peters-Lidard et al., 2015) modeling system version 7. The WRF component within this version of NUWRF was modified from the core WRF version 3.5.1, and it simulates atmospheric processes on a terrain-following mass vertical coordinate system over flexible spatial and temporal scales (Skamarock et al., 2008). NUWRF supports coupling between WRF and NASA's Land Information System (LIS), a flexible land surface modeling and data assimilation framework developed to integrate satellite and ground observations and advanced land surface modeling techniques to produce optimal fields of land surface states and fluxes (Kumar et al., 2006, 2008). This coupled system enables the investigation of land-air interactions including evaluating the impact of land initialization and land data assimilation on atmospheric states (e.g., Santanello et al., 2016).

Deleted: The

Deleted: /

Deleted: the

A number of NUWRF meteorological simulations (Table 1) were performed over the contiguous US (CONUS) for September 2013 on 12 km (479×399 grids) and 4 km (1248×900 grids) horizontal resolution Lambert conformal grids that are both centered at 40°N/95°W. As in Huang et al. (2016), the vertical grid spacing recommended by the Texas Commission on Environmental Quality (TCEQ) was implemented. We applied the Noah LSM (Chen and Dudhia, 2001), which contains four soil layers of 0.1, 0.3, 0.6, 1.0 m thicknesses, and is an option widely used in scientific and operational applications (e.g., Ek et al., 2003; Santanello et al., 2016). The Noah LSM is based on grid-dominant land use/land cover types and we chose to use the recommended International Geosphere-Biosphere Programme modified MODIS 20-category land use/land cover (Figure S1, upper), which reflect more recent conditions than the other available options (Tao et al., 2013; Yu et al., 2012). The commonly-used Mellor-Yamada-Janjic PBL scheme (Janjic, 2002) and the matching Monin-Obukhov (Janjic Eta) surface layer scheme (Monin and Obukhov, 1954) were chosen, although these might lead to shallower, cooler PBL and less vertical mixing than other available schemes in WRF (e.g., Saide et al., 2011; Angevine et al., 2012; Huang et al., 2013; Zhang, Y. et al., 2016). Other key physics options include: the Eta microphysics (Rogers et al., 2001), the Rapid Radiative Transfer Model short- and long-wave radiation (Iacono et al., 2008), and the Betts-Miller-Janjic cumulus parameterization (Janjic, 2000). These simulations were started at 06 UTC (00 Central US Standard Time) of each day. The 4 km and 12 km calculations used 4s and 24s time steps, respectively, and they were recorded hourly at 00:00 (minute:second) for 24 and 48 hours, respectively. The effect of simulation length (i.e., day 1- and day 2- forecasts, defined as the simulations 00-24h and 25-48h since the initial time, respectively) on the 12 km NUWRF performance will be included in the discussion.

Deleted: Same as

Deleted: The four-soil layer (with thicknesses of 0.1, 0.3, 0.6, 1.0 m)

Deleted: , was applied.

Deleted: (IGBP)-

Deleted: ,

Deleted: (RRTM)

Deleted: The urban surface option (Chen et al., 2011) was turned on only in the 4 km simulations.

Deleted: full clocks

As illustrated in Figure 1, we performed three types of NUWRF simulations to evaluate the impact of two land initialization methods (points a and b below, and Figures 1a-b, focus of this study), and two atmospheric initialization methods (points b and c below, and Figures 1b-c):

a) A usual method applied to the 12 km NUWRF grid, in which the atmospheric and land initial conditions (IC) were downscaled from the output of a coarser model North American Regional Reanalysis (NARR, at 32 km horizontal resolution with a 3-hourly time interval, Mesinger et al., 2006). NUWRF atmospheric lateral boundary conditions (LBC) were also downscaled from NARR. NARR is known to be generally drier and warmer than the observations (e.g., Royer and Poirier, 2010; Kennedy et al., 2011). As in default and many WRF simulations, the green vegetation fraction (GVF) input data in this case were based on climatological monthly mean satellite normalized difference vegetation index (NDVI). In previous studies, realistic vegetation density in LSMs has been shown important to accurately represent the partitioning of soil evaporation and canopy transpiration (e.g., Bell et al., 2012). However, in this study, the model did not show considerable sensitivities in response to replacing the climatological monthly GVF with satellite near real-time GVF over this study's focus regions, and these will be briefly discussed in Section 3.1.1;

b) The 12 km and 4 km "control (ctrl) simulations": Same as a), except that NUWRF land IC were instead from the output of long-term (i.e., cold-started from 01 January, 2001, cycled twenty times from 01 January, 2001 to 31 December, 2001 before running all the way through September 2013) offline LIS simulation that allowed land conditions to reach thermodynamical/water equilibrium. The LIS offline spin-up was completed on the same horizontal resolutions as NUWRF, forced by highly resolved atmospheric fields from the Global Data Assimilation System (GDAS) and precipitation data from the Global Land Data Assimilation System (GLDAS). The daily near real-time satellite GVF were used within LIS and NUWRF. "Control" was chosen to name these simulations, consistent with the usage in hydrological modeling and data assimilation communities;

c) Two additional 12 km and 4 km "ctrl" simulations: Same as b), except that NUWRF IC and LBC in these simulations were taken from the atmospheric fields of the North American Mesoscale Forecast System (NAM, at 12 km horizontal resolution with a 6-hourly time interval, Janjic, 2003; Janjic et al., 2004), which is known to usually have positive biases in temperature, moisture, and wind speed in the CONUS (e.g., McQueen et al., 2015a, b).

2.2 Emissions derived from in-situ isoprene measurements

The mixed-PBL approach introduced by Warneke et al. (2010) was adopted to derive isoprene emissions during two NASA airborne campaigns in September 2013, which were compared with NUWRF-MEGAN bottom-up emissions. The mixed-PBL approach accounts for isoprene's atmospheric lifetime but neglects the impact of horizontal advection, and it estimates isoprene emissions based on Equation (2):

$$\text{Emission}_{\text{isoprene}} - F_e = [\text{isoprene}] \times \text{boundary layer height} \times k_{OH} \times [OH] \quad (2)$$

Deleted: Our
Deleted: include three sensitivity simulations that
Deleted: , as illustrated in Figure 1:
Deleted: 's
Deleted: Same as
Deleted: Although
Deleted: land surface models is
Deleted: ing
Deleted:),
Deleted: the
Deleted: ed
Deleted: 's
Deleted: the
Deleted: NUWRF "control" simulations were conducted at
Deleted: resolutions to assess the impact of atmospheric IC and LBC on NUWRF's performance. NUWRF's
Deleted: the
Deleted: the

where [isoprene] and [OH] are the concentrations of isoprene and hydroxyl radical (OH), respectively, and the data used in our calculations will be introduced in Sections 2.2.1-2.2.2; k_{OH} refers to the rate coefficients of isoprene with OH which was set to be 101×10^{-12} cm^3/molecule/s; and F_e represents the entrainment flux from the boundary layer to the free troposphere, set constantly to be 30% of the emission flux, based on aircraft isoprene flux measurements over the Amazonian rain forest (Karl et al., 2007). Our NUWRF modeled PBLHs, after qualitatively evaluated with the aircraft measurements, were used in the emission calculations. The uncertainty of the isoprene emissions derived by this approach will be further discussed in Section 3.3.

2.2.1 Isoprene measurements

NASA's Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC^{4}RS, Toon et al., 2016, https://espo.nasa.gov/home/seac4rs/content/SEAC4RS) was conducted in August-September 2013. More than 20 research flights were executed, including ten September DC-8 flights over the southeastern US primarily focusing on the attribution and the quantification of trace gas pollution and their distributions as a result of deep convection. In-situ isoprene data measured by the Proton Transfer Reaction-Mass Spectrometry (PTR-MS, de Gouw and Warneke, 2007) on board the DC-8 aircraft were used in the emission calculations.

We used data obtained in the Missouri Ozarks ("isoprene volcano") region where biogenic isoprene emissions were high and the potential measurement interferences from furan and 2,3,2-methylbutenol (232-MBO) were negligible: Furan is found in significant concentrations only in biomass burning plumes and no enhancement in the biomass burning tracer acetonitrile was observed in our case studies (details in Section 3.1). 232-MBO is only emitted from the coniferous ecosystems in the western US. These isoprene observation data have an accuracy of ±5%.

Nine research flights were executed in September 2013 (i.e., on 4, 6, 11, 12, 13, 14, 24, 25, 26 of this month) to support the Houston portion of the NASA DISCOVER-AQ (Deriving Information on Surface conditions from Column and Vertically Resolved Observations Relevant to Air Quality, Crawford et al., 2014, https://discover-aq.larc.nasa.gov) field experiment. High-resolution in-situ isoprene measurements during DISCOVER-AQ were made using a proton-transfer-reaction time-of-flight mass spectrometry instrument (PTR-ToF-MS, with an accuracy of ±10%, Müller et al., 2014) on board the P-3B aircraft over selected locations in the Greater Houston area three times a day (i.e., morning, noon-early afternoon, and mid-afternoon) to explore their spatial and diurnal variability. Isoprene measurements from the PTR-ToF-MS are possibly interfered by other VOCs from anthropogenic sources in Houston (e.g., from oil and gas industries). Therefore, we focus on deriving biogenic emissions at the Conroe site, a region north to downtown Houston area with medium vegetation coverage and less strongly influenced by urban transportation/industrial sources and biomass burning plumes (details in Section 3.1). Additionally, we investigated the hourly surface isoprene measurements available at eight TCEQ Automated Gas Chromatograph (AutoGC) monitoring stations, mostly located in downtown Houston area. The data before sunrise and after

Deleted: 10^{12}

Deleted: -

Deleted: , when more

Deleted: . The August deployment over the western US focused on the influences of biomass burning pollution, their temporal evolution, and their impacts on meteorological processes and air quality. The

Deleted: deployment

Deleted: ed

Deleted: the

Deleted: the

Deleted: The isoprene

Deleted: the

Deleted: the

Deleted: (Figure 2d)

Deleted: , as indicated by the isoprene-carbon monoxide (CO) relationships

Deleted: Figure A1.

Deleted: the

Deleted: (marked as triangles in Figure 2d).

Deleted: the

Deleted: the

sunset, when biogenic emissions are at their daily minima, are particularly useful for determining the regional background and non-biogenic isoprene levels, and therefore they helped quantify the uncertainty in the observation-derived emissions. The Limit of Detection applied to all AutoGC target compounds is currently 0.4 ppbC (0.08 ppbv for isoprene).

The ground speed of the DC-8 and P-3B aircraft was around 8-9 km/minute near the "isoprene volcano" areas during SEAC^{4}RS and 9-14 km/minute at around Conroe during DISCOVER-AQ within the focused time period. Therefore, the aircraft data averaged in 1-minute interval (released on 10 February, 2016 and 23 July, 2015 for SEAC^{4}RS and DISCOVER-AQ, respectively) were used to estimate the emissions, as they represent isoprene concentrations on similar spatial scales to NUWRF-MEGAN. At around Conroe, multiple P-3B aircraft data points correspond to several NUWRF model grids, and the averaged emissions based on NUWRF-MEGAN and the median PBL observations were used in the comparisons.

2.2.2 OH from the NOAA National Air Quality Forecasting Capability (NAQFC)

Due to the lack of aircraft OH measurements in September 2013, the OH concentrations simulated by the NOAA NAQFC 12 km Community Modeling and Analysis System (CMAQ, Byun and Schere, 2006; Pan et al., 2014) were used to derive isoprene emissions. The NAQFC CMAQ is driven by the NAM meteorological fields, and biogenic emissions are computed online from the Biogenic Emission Inventory System (BEIS) version 3.14, that often produces much lower emissions than MEGAN at the "isoprene volcano" and in the eastern Texas (e.g., Warneke et al., 2010; Carlton and Baker, 2011). The NAQFC CMAQ OH performance near the "isoprene volcano" was generally satisfactory for the studied period: i.e., the mean±standard deviation of the predicted OH of $(1.8\pm0.8)\times10^6$ on 11 September and $(1.5\pm0.3)\times10^6$ molecule/cm^3 on 06 September along the Missouri flight paths (to be shown in Sections 3.1 and 3.3) are of the close magnitudes to the observationally constrained OH concentrations in that area during SEAC^{4}RS (e.g., $(1.3\pm0.3)\times10^6$ molecule/cm^3 by Wolfe et al., 2015). Close to the estimated OH concentrations of approximately 2-6×10^6 molecule/cm^3 near Houston on 16 September, 2006 (Warneke et al., 2010), the simulated PBL OH on 11 September, 2013 range from ~1.8×10^6 to ~4.0×10^6 molecule/cm^3 along the P-3B around Conroe. The OH levels are higher in late morning and around noon (>3.1×10^6 molecule/cm^3) than in the afternoon (~1.8×10^6 molecule/cm^3), qualitatively consistent with the observations in downtown Houston in May 2009 (Czader et al., 2013). The averaged OH for all P-3B flight days is within the range of 11 September, following similar diurnal variability. Little prior knowledge exists on CMAQ OH performance in the Greater Houston area, except the moderate negative biases (with observed-to-modeled ratios of 1.15-1.36) reported by Czader et al. (2013) for May 2009. As their modeling system was configured differently from the NAQFC, the biases of the modeled OH fields from the NAQFC CMAQ system need to be investigated further in the future.

2.3 Evaluation datasets

2.3.1 Ground and aircraft measurements of air temperature, solar radiation, and PBLH

Deleted: (LOD)

Deleted: the

Deleted: 's

Deleted: the

Deleted: s

Deleted: 7

Deleted: ,

Deleted: levels on

Deleted: are also

Deleted: is

Deleted: .

Deleted: the

We focus on evaluating the sensitivities of NUWRF air temperature, solar radiation, and PBLH to initialization methods, as they are the most important weather variables to the estimated isoprene emissions. The NUWRF modeled (near-) surface air temperature fields were compared with the National Centers for Environmental Prediction (NCEP) Global Surface Observational Weather Data (also used in Huang et al., 2016), the DC-8 aircraft air temperature measurements, and the 5-minute TCEQ special observations at the Conroe site taken in support of the airborne campaigns. The NUWRF modeled solar radiation was briefly compared with the measurements by pyranometers on board the DC-8 and at Conroe. The NUWRF-simulated PBLH was also roughly compared with the estimated PBLH by: a) the Differential Absorption Lidar (DIAL)-High Spectral Resolution Lidar (HSRL) measurements on board the DC-8 aircraft during the SEAC^{4}RS campaign (Figure 2a); and b) the vertical gradients of the in-situ isoprene observations measured on board the P-3B aircraft during DISCOVER-AQ at around the Conroe site at different times of the day (Figure 2d).

Deleted: 's
Deleted: the
Deleted: -
Deleted: -
Deleted: were
Deleted: were
Deleted: 2c

2.3.2 Satellite soil moisture and heat flux products

The European Space Agency (ESA) soil moisture Climate Change Initiative (CCI, http://www.esa-soilmoisture-cci.org) project produces daily surface soil moisture data at 0.25°×0.25° horizontal resolution, based on multiple passive and active sensors, as well as by merging both passive and active products. Fang et al. (2016) reported that the merged CCI product exhibited higher anomaly correlation (than the individual active/passive CCI products) with both Noah LSM simulations and in-situ measurements during 2000-2013. Version 02.2 of this merged product, which was released in 2015 and covers the period of 1978-2014, has enhanced spatial and temporal coverage and intercalibration between different instruments. We used this version to evaluate the modeled soil moisture fields and the normalized soil moisture anomalies (as defined in Equation (3)), over the regions where the CCI data quality flag equals zero:

Deleted: during 1978-2014
Deleted: Long-term soil moisture changes in the US based on the CCI product contributed to the US National Climate Assessment report (Melillo et al., 2014, p72-73).
Deleted: Therefore, the most recent version
Deleted: (version 02.2,
Deleted:) with
Deleted: was
Deleted: at

$$\text{Normalized soil moisture (SM) anomaly} = \frac{daily\ SM - monthly\ mean\ SM}{monthly\ SM\ standard\ deviation} \quad (3)$$

Soil moisture controls the partitioning of energy into latent (the energy related to changes in phase) and sensible heat (the energy related to temperature changes) fluxes. To evaluate the appropriateness of NUWRF land initialization, we compared NUWRF modeled absolute heat fluxes and their partitioning (i.e. evaporative ratio, defined as latent heat/(latent heat+sensible heat)) with the Atmosphere-Land Exchange Inversion (ALEXI, Anderson et al., 2007; Hain et al., 2011) retrievals. The ALEXI heat flux product using the NOAA Geostationary Operational Environmental Satellite (GOES) thermal-infrared (TIR) land surface temperature, along with its soil moisture proxy retrievals, is a part of the NOAA operational GOES Evapotranspiration and Drought Product System (http://www.ospo.noaa.gov/Products/land/getd). Although limited to clear-sky conditions, ALEXI provides retrievals over a wide range of vegetation cover on horizontal resolution close to that of NUWRF (i.e., 0.08°×0.08° for this study).

Deleted: 's
Deleted: the
Deleted: -
Deleted: the
Deleted: IR
Deleted: GET-D;
Deleted: our NUWRF's

3 Results and discussion

3.1 Case study of 11 September, 2013

We first show a case study on a specific day of 11 September, 2013, when aircraft measurements were available from both the SEAC^{4}RS and DISCOVER-AQ campaigns. For SEAC^{4}RS, the DC-8 aircraft sampled over broad areas in the central/southeastern US on this day (Figure 2a), passing the "isoprene volcano" region in Missouri at the early afternoon time (18:30-19:30 UTC/12:30-13:30 local standard time), where mixed layer heights indicated by the DIAL-HSRL instrument were mostly below 2 km. Elevated isoprene concentrations (up to ~10.4 ppbv) were observed by the PTR-MS near the surface (<1 km, a.g.l.). Biomass burning plumes had little interference to these isoprene measurements, as determined by the low acetonitrile concentrations (Figure 2b) in the sampled airmasses. For DISCOVER-AQ, the P-3B repeatedly took measurements at different times of the day at around the Conroe site in Houston, an area with slightly denser vegetation than downtown Houston (i.e., ~1m^2/m^2 larger LAI, Figure 2c). The observed isoprene vertical profiles at Conroe indicate the growth of PBLH from the morning (a few hundred meters, a.g.l.) to the afternoon (~2 km, a.g.l.), and ~50% higher near-surface isoprene concentrations in the afternoon (~2.5 ppbv) than in the morning (~1.7 ppbv) (Figure 2d). The CO concentrations in the sampled airmasses were below 200 ppbv (Figure 2e), indicating negligible biomass burning source impacts. Anthropogenic emission sources are mainly located at downtown Houston, where the daytime P-3B aircraft isoprene concentrations (i.e., at the Moody Tower, Deer Park, Channelview spirals) did not exceed ~0.6 ppbv and the isoprene-CO enhancement ratio differed from that in Conroe (Figure 2e). The magnitudes of the downtown aircraft isoprene measurements were slightly lower than most of the nearby surface measurements during the daytime (Figure 2f; locations of these surface sites are shown as triangles in Figure 2c). Measured surface isoprene levels during the daytime were ~twice as high as during the nighttime (~0.2-0.3 ppbv) when biogenic emissions are at their daily minima. Therefore, we expect that non-biogenic emissions contributed to no more than 0.3 ppbv of the P-3B observed isoprene over that region.

3.1.1 Evaluation of NUWRF surface air temperature, PBLH, soil moisture, and heat fluxes

Figure 3a compares the NUWRF modeled surface air temperature in the central/eastern US with ground observations in the early afternoon where the DC-8 flew past Missouri. The 12 km usual run shows 2-4 °C positive biases in Missouri, which are of the similar magnitudes to the findings by Carlton and Baker (2011) for these regions. These positive biases were dramatically reduced in NUWRF ctrl runs. Evaluation of the modeled air temperature and PBLH along the DC-8 flight paths in Missouri was summarized in Table 2. Root Mean Square Errors (RMSEs) of the modeled air temperature from the ctrl runs are ~1.5 °C lower than the 12 km usual run-produced. The PBLHs from the ctrl runs are thinner (~0.6 km on average) and less spatially variable. They may be closer to the reality referring to the DIAL-HSRL data in Figure 2a, which can also be uncertain. The higher resolution 4 km ctrl run generated slightly (~0.04 °C) better air temperature and ~0.02 km thinner mean PBLH than the 12 km ctrl run.

Deleted: (
Deleted: 2d), where the
Deleted: 2c
Deleted: A1
Deleted: the
Deleted: presented a different
Deleted: Conroe's
Deleted: A1
Deleted: 2e) that
Deleted: their isoprene levels
Deleted: the
Deleted: the
Deleted: the
Deleted: , corresponding to
Deleted: PBLH which
Deleted: (that
Deleted: quite
Deleted:).

Figure 4a compares the NUWRF modeled daytime surface air temperature at the Conroe site against the TCEQ special measurements, and Table 3 summarizes the statistical evaluation of NUWRF PBLH and surface air temperature performance in Conroe. Similar to the conditions in Missouri, temperatures from the 12 km usual run are positively biased by 1.8-3.0°C during the daytime, and the ctrl runs significantly better captured the observed magnitudes (i.e., with 1.6-1.8 °C lower RMSEs than the 12 km usual run), corresponding to at least ~300 m lower PBLH, which are likely more realistic referring to the observed isoprene vertical profiles. The 4 km ctrl simulation produced noticeably lower air temperature and PBLH at the morning (by up to ~0.5 °C/~170 m) and afternoon (by up to ~2.6 °C/~290 m) times than the 12 km ctrl run.

Figure 3b-c compare the CCI daily surface soil moisture fields with the NARR and LIS modeled at the NUWRF initialization time on this day. The NARR soil moisture fields are at least 0.1 m^3/m^3 drier than the LIS-NUWRF systems at the beginning of the simulation (Figure 3b), causing the spurious NUWRF temperature/PBLH fields as described earlier. The impact of initial soil moisture states on simulated temperature at later times is similar to the results in Collow et al. (2014) for the Great Plains in May 2010. Figure 3c shows that the normalized soil moisture anomalies from NARR and LIS overall demonstrate similar spatial patterns, which was difficult to be validated with the CCI product due to small sizes of usable data in September 2013 (Figure S2). This suggests that when downscaling land fields to a different modeling system, adjusting the large-scale dataset based on the climatology (preferably for a much longer record) of both systems would be helpful. This adjustment, sometimes also called "bias-correction", is indeed useful in satellite land data assimilation (e.g., scaling satellite soil moisture before assimilation based on the climatology of the model and the satellite). Figure 5 compares the modeled heat fluxes with ALEXI retrievals, indicating that the usual land initialization method resulted in significantly underpredicted latent heat and overpredicted sensible heat, and the partitioning between these heat fluxes were poorly represented. These evaluation results confirm that the usual land initialization method is inappropriate for this case.

It's worth pointing out that by replacing the WRF-default monthly-mean climatological GVF input with the daily near real-time GVF in the 12 km usual run, we did not find significantly changes in the modeled temperature (i.e., <±0.5°C, as shown in Figure S3 right) and PBLH (not shown in figures) fields near the Missouri Ozarks and Conroe, where the GVF differences are within ±0.1 (Figure S3 left). Therefore, soil states at the initialization were the major causes to the different temperature and PBLH fields from the usual and ctrl runs over these regions. In contrast, weather fields over some other central/southeastern US regions, particularly in the eastern Arkansas, are shown very sensitive to this GVF update, with negative (positive) GVF differences resulting in positive (negative) temperature differences. Over these regions, the different weather fields in usual and ctrl runs indicate the net effect of GVF and soil initialization.

Figure 4b evaluates the impact of simulation length on the modeled surface air temperature at Conroe, and in this case higher temperature biases were shown in the longer simulation (day 2-forecast) regardless of the land initialization method, especially during the morning and early afternoon times. The RMSEs of daytime air temperatures from the day 2-forecasts are ~0.3 °C higher than the day 1-forecasts. Figure 4c shows the impact of atmospheric IC/LBC on the modeled air

Deleted: the

Deleted: too many missing

Deleted: .

Deleted: fields

Deleted: model's and the satellite's

Deleted: before they are assimilated).

Deleted: the

Deleted: Such

Deleted: confirmed

Deleted: A2

Deleted: A2

Deleted: the

temperature. In both 12 km and 4 km grids, replacing the NARR IC/LBC with NAM's resulted in larger temperature amplitude, associated with greater negative biases in the morning and positive biases at around the mid-afternoon. The RMSEs of daytime air temperatures from the NAM-related cases are ~0.2 °C higher than the NARR-related cases. Figures 4a and 4c together also suggest that an inappropriate land initialization for a regional simulation can result in almost ten times larger model errors than using an alternative atmospheric IC/LBC.

The NUWRF modeled solar radiation fields were briefly evaluated. It was found that regional NUWRF solar radiation fields in Missouri from these various runs are vastly similar in the early afternoon local time, and they are >30% (a couple of hundred of W/m^2) larger than the DC-8 measurements. These biases are close to what has been reported by Carlton and Baker (2011), and the WRF-satellite differences in Guenther et al. (2012). The daytime NUWRF solar radiations at Conroe had time-varying biases but on average are a few percent different from the observations, and the photosynthetically active radiation at Conroe differed by up to ~12W/m^2 among these simulations.

3.1.2 NUWRF-MEGAN and observation-derived isoprene emissions in Missouri and Houston

Figure 6a shows the spatial distributions of the MEGAN isoprene emissions driven by these multiple NUWRF simulations, compared with the observation-derived emissions at the early afternoon time, when the DC-8 aircraft sampled at the "isoprene volcano" and isoprene emissions approached their daily maxima. Similar spatial patterns of the MEGAN emissions were produced when different NUWRF runs were used. The emissions based on the 12 km NUWRF usual run are at least 20% larger than those driven by NUWRF ctrl runs, corresponding to a ~2 °C larger positive bias in NUWRF temperature. Such emission sensitivities to the air temperature are close to the magnitudes reported in literature for other regions (Guenther et al., 2006, 2012; Wang et al., 2011). NUWRF-MEGAN emissions are 22-49% higher than the observation-derived emissions along the DC-8 flight tracks, with the 4 km NUWRF ctrl run-based MEGAN emissions the closest to the observation-derived.

Figure 6b shows the spatial distributions of the MEGAN isoprene emissions driven by these different NUWRF runs over Houston near the local standard noon time, the second time P-3B sampled over Conroe on that day, when isoprene emissions almost reached their daily maxima. Similar to the Missouri conditions, the MEGAN emissions driven by the 12 km NUWRF usual run are >20% larger than the cases driven by NUWRF ctrl runs. Figure 7a compares NUWRF-MEGAN daytime isoprene emissions at Conroe. The 12 km NUWRF usual run-based daily peak emissions during local noon/early afternoon times are ~20% higher than the 12 km NUWRF ctrl run-based, the latter of which is closer to the observation-derived. The daytime-integrated emissions derived using the 12 km NUWRF usual run are ~21% higher than the 12 km NUWRF ctrl run-based. Again this discrepancy corresponds to a ~2 °C temperature differences on this day (Figure 4a; Table 3). The emissions driven by the 4 km NUWRF ctrl run are the lowest, with the daytime integrated and the peak emissions ~40% lower than the 12 km NUWRF ctrl run-based, and they substantially deviate from the observation-derived. This is in part

Deleted: from the multiple NUWRF runs
Deleted: the
Deleted: at
Deleted: the different
Deleted: the
Deleted: the
Deleted: The
Deleted: calculated
Deleted: um
Deleted: the
Deleted: the

due to the coolest temperature from this NUWRF run, especially in the afternoon, as well as its weaker photosynthetically active radiation than the 12 km simulated (i.e., by ~10 W/m^2 on average during the daytime). This may also be resulting from some limitations of MEGAN's parameterization and uncertainty in its other inputs (e.g., PFT and LAI) on small scale. As illustrated in Figure S4, representation error (i.e., due to different data resolutions) along with neglecting horizontal transport in deriving emissions from aircraft data, also contributed to the discrepancies among the 12 km and 4 km NUWRF- and aircraft-derived emissions.

The impacts of simulation length and atmospheric initialization on NUWRF-MEGAN isoprene emissions at Conroe are generally much smaller than the impact of land initialization (Figures 7b-c), mainly due to the smaller temperature sensitivities (Figures 4b-c): The day-2 forecast derived emissions are higher than the day 1 forecast-based emissions by ~10% in Conroe at the local standard noontime, but their daytime-integrated isoprene emissions differ much less (~1.5%). Daytime maximum emissions disagree by only <±2% in both 12 km and 4 km grids. Noontime isoprene emissions related to NAM and NARR IC/LBC differ by less than 2% in both resolutions, and the daytime-integrated emissions related to NAM IC/LBC are higher than the NARR related by 0.8% and 5.2% in 12 km and 4 km grids, respectively.

3.2 Conditions on extended time periods

3.2.1 Conditions on multiple flight days during DISCOVER-AQ in September 2013

As the 12 km NUWRF ctrl run-based MEGAN isoprene emissions showed the best agreement with the observation-derived emissions at Conroe on 11 September (Section 3.1), we calculated MEGAN isoprene emissions using this set of NUWRF simulation also for the other eight DISCOVER-AQ flight days when variable meteorological conditions were present (details are in the flight reports at: https://discover-aq.larc.nasa.gov/planning-reports_TX2013.php), and the multi-flight day averaged MEGAN calculations were compared with the P-3B aircraft observation-derived at the Conroe site (Figure 8). The multi-day averaged MEGAN and observation-based emissions are higher than the estimates for 11 September, except in the morning. The multi-day mean morning emissions from MEGAN are ~44% higher than the observation-derived, a larger discrepancy than on 11 September. A possible reason for this morning-time overestimation is that MEGAN does not account for the circadian control that can lower the isoprene emissions from some canopies (Hewitt et al., 2011). At local noontime and afternoon times, unlike the 11 September condition, the multi-day averaged MEGAN emissions were slightly (by <5%) lower than the observation-derived.

3.2.2 September 2013 comparing with decadal mean conditions

We extend the analyses to the interannual variability of drought and vegetation conditions in relation to the isoprene emissions in Conroe. The monthly anomalies were calculated for HCHO column (which is often used to derive biogenic emissions) from the Ozone Monitoring Instrument (OMI, De Smedt et al., 2015), Terra-MODIS LAI, ESA CCI microwave

Deleted: st
Deleted: at
Deleted: ions'
Deleted: Uncertainty
Deleted: ed
Deleted: grid
Deleted: the
Deleted: The noontime

soil moisture and ALEXI TIR soil moisture proxy in September 2013, related to the decadal (2005-2014) September means. Eastern Texas was under extreme drought conditions in September 2011 as indicated by the Palmer Drought Severity Index (http://www.ncdc.noaa.gov/temp-and-precip/drought/historical-palmers.php), which was excluded from the decadal mean calculations as severe drought can reduce or terminate isoprene emissions (Pegoraro et al., 2004) and complicate the anomalies. At Conroe, close-to-1 anomalies are found in September 2013 for the ALEXI and CCI data (~0.99 and ~0.98), and vegetation was slightly thinner than the decadal mean conditions (the LAI anomaly of ~0.96). A much lower than average HCHO column (the anomaly of ~0.77) was observed by OMI in this month. A higher OMI anomaly (~0.99) was found in September 2006 studied by Warneke et al. (2010), under drier conditions (ALEXI and CCI anomalies of ~0.77 and ~0.91, respectively) with denser-than-average vegetation (the LAI anomaly of ~1.07). Note that these interannual differences can be complicated by the uncertainty in these satellite data, and also reflect the possible influences by the temporal changes in non-biogenic VOC emissions, local/regional chemistry, and plant types in this area.

Deleted: The eastern

Deleted: less dense

Deleted: higher

Deleted: coverage

Deleted: s

Deleted: the

Deleted: the

3.3 Uncertainty discussions

In addition to the biases in NUWRF surface air temperature, a number of other factors can affect NUWRF-MEGAN isoprene emission calculations. These include:

Deleted: 's

Deleted: the

a) The outdated PFT data that represent year 2008 conditions and the uncertainty in the MODIS LAI input. Future studies should consider implementing in both (NU)WRF and MEGAN the up-to-date land cover input data, e.g., a recently developed product from the Visible Infrared Imaging Radiometer Suite (Zhang, R., et al., 2016), which is compared with the MODIS input in Figure S1 (lower). It would be also worth performing sensitivity calculations using LAI from (NU)WRF, which is either prescribed to its GVF input or computed by some LSMs.

Deleted: the

b) The known positive biases in NUWRF solar radiation fields partially due to the lack of aerosol impacts and the misplaced/missing clouds. It has been shown that implementing certain satellite solar radiation product can reduce the biases in MEGAN emissions for other time periods (Carlton and Baker, 2011; Guenther et al., 2012). Identifying suitable satellite radiation products for this case will be included in future work.

Deleted: 's

c) As described in Section 2.1, due to the omission of deposition, MEGAN version 2.1 net primary emissions are higher than the net emission flux by a few percent on average, and this bias may be larger at a specific location. Adding that contribution in future emission calculations is important.

Deleted: the

Deleted: always larger

Deleted: (

Deleted: but

Deleted:) due to the omission of deposition

d) Other limitations in MEGAN's parameterization which good input data can help better diagnose.

The uncertainties of aircraft observation-derived isoprene emissions are expected to come from:

a) The PTR-MS and PTR-ToF-MS measurements have accuracies of ±5% and ±10%, respectively, which can be propagated to the emission calculations. These were smaller than the ±15% from the Warneke et al. (2010) study.

Deleted: that

b) The biases introduced from the NAQFC CMAQ OH fields as mentioned in Section 2.2.2 which will need to be investigated further on grid-scale (e.g., by comparing them with other modeling products covering our studied period).

c) As discussed in Warneke et al. (2010), the mixed-PBL approach neglects horizontal transport, which may attribute transported isoprene to the incorrect grid boxes. For this study, observed wind speed along the SEAC^{4}RS DC-8 flight path ranged from 0.27 to 5.47 m/s, with the mean value of ~1.68 m/s. The TCEQ 5-minute surface wind speed observations were no larger than <3.5 m/s on 11 September. Assuming isoprene lifetime in this study is ~an hour, the aircraft observed isoprene may be actually emitted from the nearby 1-2 model grids on the 12 km scale. Therefore, this approach introduces an error which may not significantly affect the magnitude of regional emission calculations in Missouri but may have a larger impact on the Conroe case especially on a single day (See the illustration in Figure S4). Developing and applying top-down methods that also account for atmospheric transport should be strongly encouraged.

d) As discussed in Warneke et al. (2010), the constant 30% entrainment flux may not be realistic for the regions/times we studied, which needs further validation.

e) Regional non-biogenic emission sources may contribute to 10-20% of the aircraft observed isoprene at Conroe, as estimated by the ground in-situ data (Sections 2.2.1 and 3.1).

f) The mixed-PBL approach assumes complete vertical mixing which may not be true in practice. Additionally, the control run based NUWRF modeled PBLHs (Tables 2-3) were used, possibly associated with uncertainty on a magnitude of a few hundred meters (~20%).

Warneke et al. (2010) estimated the uncertainty of their aircraft observation-derived emissions over Texas to be a factor of 2 (-50%, +100%). We anticipate the uncertainty of ours to be of the similar magnitude for the single-day Conroe case, but smaller in the multi-day averaged emissions in Conroe. The regional-averaged aircraft observation-derived emissions over the "isoprene volcano" region (Figure 6a for 11 September, and Figure S5 for 06 September with more descriptions in the figure caption) from this study are close to the result in Wolfe et al. (2015) of 587±73 M C/km^2/h, derived using a different method for similar regions during SEAC^{4}RS.

4 Conclusions and suggestions on future direction

We performed case studies during the SEAC^{4}RS and DISCOVER-AQ Houston field campaigns, showing that a usual method to initialize the Noah LSM (i.e., directly downscaling the land fields from the coarser resolution NARR) led to significant positive biases in the coupled NUWRF (near-) surface air temperature and PBLH around the Missouri Ozarks and Houston, Texas, as well as poorly partitioned latent and sensible heat fluxes. Replacing the land initial conditions with the output from a long-term offline LIS (a flexible land surface modeling and data assimilation framework) simulation effectively reduced the positive biases in NUWRF surface air temperature fields. We also showed that using proper land initialization modified NUWRF surface air temperature errors almost ten times as effectively as applying a different

Deleted: (
Deleted:),
Deleted: the
Deleted: a random
Deleted: .
Deleted: the
Deleted: A3
Deleted: 8±1 mg/m^2
Deleted: the
Deleted: 's
Deleted: 's

atmospheric initialization method. The LIS-NUWRF based MEGAN version 2.1 isoprene emission calculations were at least 20% lower than those computed from the NARR-initialized NUWRF run, closer to the aircraft observation-derived emissions. Higher resolution MEGAN calculations were prone to amplified discrepancies with the aircraft observation-derived emissions on small scales. This was possibly resulting from some limitations of its parameterization uncertainty in its inputs on small scale, as well as the representation error and neglecting horizontal transport in deriving emissions from aircraft data. This study emphasizes the importance of proper land initialization to the coupled atmospheric weather modeling and the follow-on biogenic emission modeling. We anticipate that improved weather fields using the better land initialization approach will also benefit the representation of the other processes (other weather-dependent emission calculations, transport, transformation, deposition) included in air quality modeling, and therefore can help reduce uncertainty in the simulated chemical fields. The study is limited to selected locations and times considering the availability of aircraft data, and the observation-derived emissions may also be associated with large uncertainty. In future, developing methods to combine satellite land and atmospheric chemical data assimilation should be encouraged to further improve air quality modeling and top-down emission estimation over broader regions/extended time periods to help interpret the trends and variability of the atmospheric composition. Improved chemistry output from regional models can also help evaluate the current "a priori" used in satellite retrievals, and may serve as an alternative.

It should be noted that many published model comparison studies cited in Section 1 did not adequately assess the impacts of model inputs versus their parameterization. Having more confidence in the weather inputs is beneficial for quantifying the other sources of uncertainties (e.g., parameterization, other input data) of the models that they drive. In future, the impact of atmospheric weather input on emissions computed using other biogenic emission models (e.g., BEIS, future versions of MEGAN) will be explored. Efforts will be made to improve the other inputs data (e.g., radiation, land cover).

Although we recommend initializing WRF or NUWRF with the LIS land fields, when long-term atmospheric forcing data are not available to facilitate the offline LIS spin-up, we suggest: 1) "bias-correcting" the land fields from the initial condition model, based on the climatology of the initial condition model and the target model; or 2) adopting the self spin-up method, i.e., running the model for a certain spinup period (e.g., a month) at least once, cycling its own soil variables, to allow the land variables to develop appropriate spatial variability (Angevine et al., 2014). Experiments using simulations with different LSMs along with suitable nudging methods can also be helpful.

5 Model and data availability

Instructions for obtaining and running the used models can be found at: LIS (lis.gsfc.nasa.gov/documentation/lis); NUWRF (nuwrf.gsfc.nasa.gov/doc); MEGAN (lar.wsu.edu/megan/docs). The satellite land products, and the NUWRF/MEGAN output can be made available upon request. The open access to the used aircraft and ground observations is acknowledged:

Deleted: The higher
Deleted: NUWRF-
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Deleted: errors
Formatted: Font color: Text 1
Deleted: ,
Deleted: ed
Deleted: and the inputs'
Deleted: the
Deleted: via improved
Deleted: improve
Deleted: es
Deleted: The improved
Deleted: , and efforts
Deleted: adopting the self spin-up method (Angivine et al., 2014) or
Deleted: model's land fields
Deleted: . Experimenting

Aircraft data were obtained from: http://www-air.larc.nasa.gov/index.html

SEAC^{4}RS: doi: 10.5067/Aircraft/SEAC4RS/Aerosol-TraceGas-Cloud

DISCOVER-AQ: doi: 10.5067/Aircraft/DISCOVER-AQ/Aerosol-TraceGas

TCEQ AutoGC data: https://www.tceq.texas.gov/cgi-bin/compliance/monops/agc_daily_summary.pl

NCEP Global Surface Observational Weather Data (DS461): http://rda.ucar.edu/datasets/ds461.0/

OMI HCHO column data: http://h2co.aeronomie.be

Appendix: List of acronyms by category

A.1 Model related (with short descriptions in the parentheses)

Models

BEIS (emission model): Biogenic Emission Inventory System

NAQFC CMAQ (regional air quality model): National Air Quality Forecasting Capability Community Modeling and Analysis System

LIS (land surface modeling and data assimilation framework): Land Information System

MEGAN (emission model): Model of Emissions of Gases and Aerosols from Nature

Noah LSM (land surface model): NCAR, OSU, AirForce, office of Hydrology Land Surface Model

NUWRF (observation-driven regional earth system modeling and assimilation system): NASA-Unified Weather Research and Forecasting model

Model forcing datasets

GDAS (LIS forcing): Global Data Assimilation System

GLDAS (LIS forcing): Global Land Data Assimilation System

IC/LBC: Initial Conditions/Lateral Boundary Conditions

NAM (NUWRF forcing): North American Mesoscale Forecast System

NARR (NUWRF forcing): North American Regional Reanalysis

A.2 Observation related (measurements used in this study are indicated in the parentheses)

Field experiments and aircraft instruments

DC-8 and P-3B: Aircraft used in SEAC^{4}RS and DISCOVER-AQ, respectively

DIAL-HSRL (PBLH): Differential Absorption Lidar-High Spectral Resolution Lidar

DISCOVER-AQ: Deriving Information on Surface Conditions from COlumn and VERtically Resolved Observations Relevant to Air Quality

SEAC^{4}RS: Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys

PTR-MS (isoprene): Proton Transfer Reaction-Mass Spectrometry

PTR-ToF-MS (isoprene): Proton-Transfer-Reaction Time-of-Flight Mass Spectrometry

Satellite related

ALEXI (heat flux, soil moisture proxy): Atmosphere-Land Exchange Inversion

CCI (soil moisture): Climate Change Initiative

GOES: Geostationary Operational Environmental Satellite

MODIS (land use/land cover, LAI, GVF): Moderate Resolution Imaging Spectroradiometer

OMI (HCHO): Ozone Monitoring Instrument

TIR: Thermal Infrared

VIIRS (land use/land cover): Visible Infrared Imaging Radiometer Suite

Surface Measurements

AutoGC (isoprene): Automated Gas Chromatograph

A.3 Terminology

BVOCs: biogenic volatile organic compounds

CH_4: methane

CO: carbon monoxide

CO_2: carbon dioxide

GVF: green vegetation fraction

HCHO: formaldehyde

LAI: leaf area index

NDVI: normalized difference vegetation index

N_2O: nitrous oxide

OH: hydroxyl radical

O_3: ozone

PBLH: planetary boundary layer height

PFT: plant functional type

RMSE: root mean square error

SM: soil moisture

232-MBO: 2,3,2-methylbutenol

A.4 Organizations/Agencies/Geographical regions

CONUS: contiguous United States

ESA: European Space Agency

NA: North America

NASA: National Aeronautics and Space Administration
NCEP: National Centers for Environmental Prediction
NOAA: National Oceanic and Atmospheric Administration
TCEQ: Texas Commission on Environmental Quality
TX: State of Texas

Acknowledgements

We thank Heather Stewart and Mark Estes (TCEQ) for providing the ground special measurements at Conroe. The DIAL-HSRL mixed layer height data during SEAC^{4}RS were produced by Richard Ferrare, Johnathan Hair, and Amy Jo Scarino (NASA LaRC). The aircraft CO measurements were made by Glenn Diskin (NASA LaRC), and the DC-8 solar radiation measurements during SEAC^{4}RS were made by Anthony Bucholtz (NRL). Isoprene measurements on board the NASA aircraft during SEAC^{4}RS and DISCOVER-AQ were supported by the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit) through the Austrian Space Applications Programme (ASAP) of the Austrian Research Promotion Agency (FFG). Markus Müller and Tomas Mikoviny are acknowledged for data acquisition and analysis. Min Huang is grateful for the financial support from a NASA grant (NNX16AN39G) and NOAA GOES-R Risk Reduction, as well as the technical support from Li Fang, Jifu Yin (U Maryland), and the LIS/NUWRF teams at NASA GSFC.

Deleted: the

Deleted: the

Deleted: MH

References

Anderson, M.C., J.M. Norman, J.R. Mecikalski, J.P. Otkin, and W.P. Kustas (2007), A climatological study of surface fluxes and moisture stress across the continental United States based on thermal infrared remote sensing, Part I: model formulation, J. Geophys. Res., 112, D11112.

Angevine, W. M., Eddington, L., Durkee, K., Fairall, C., Bianco, L., and Brioude, J. (2012), Meteorological model evaulation for CalNex 2010, Mon. Weather Rev., 140, 3885–3906, doi:10.1175/MWRD-12-00042.1.

Angevine, W. M., Bazile, E., Legain, D., and Pino, D. (2014), Land surface spinup for episodic modeling, Atmos. Chem. Phys., 14, 8165-8172, doi:10.5194/acp-14-8165-2014.

Bell, J. R., Case, J. L., LaFontaine, F. J., Kumar, S. V. (2012), Evaluating the Impacts of NASA/SPoRT Daily Greenness Vegetation Fraction on Land Surface Model and Numerical Weather Forecasts, the 16th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, New Orleans, LA, 22−26 January 2012, http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120004024.pdf (last access: January 2017).

Deleted: Bell et al.

Byun, D. and Schere, K. L. (2006), Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, Appl. Mech. Rev., 59, 51–77, doi: doi:10.1115/1.2128636.

Byun, D., Kim, H.-C., Ngan, F. (2011), Final Report: Improvement of Meteorological Modeling by Accurate Prediction of Soil Moisture in the Weather Research and Forecasting (WRF) Model, available at: https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/582886246FY1009-NOAA_WRF_Soil_Moisture_20110331.pdf, reported by NOAA ARL to Texas Commission on Environmental Quality (last access: January 2017).

Canty, T. P., Hembeck, L., Vinciguerra, T. P., Anderson, D. C., Goldberg, D. L., Carpenter, S. F., Allen, D. J., Loughner, C. P., Salawitch, R. J., and Dickerson, R. R. (2015), Ozone and NO_x chemistry in the eastern US: evaluation of CMAQ/CB05 with satellite (OMI) data, Atmos. Chem. Phys., 15, 10965-10982, doi:10.5194/acp-15-10965-2015.

Carlton, A. G. and K. R. Baker (2011), Photochemical Modeling of the Ozark Isoprene Volcano: MEGAN, BEIS, and Their Impacts on Air Quality Predictions, Environ. Sci. Technol., 45 (10), 4438-4445, doi: 10.1021/es200050x.

Case, J., W. Crosson, S. V. Kumar, W. Lapenta, and C. Peters-Lidard (2008), Impacts of High-Resolution Land Surface Initialization on Regional Sensible Weather Forecasts from the WRF Model, J. Hydrometeorology, 9 (6), 1249-1266, doi: 10.1175/2008JHM990.1.

Case, J. L., S. V. Kumar, J. Srikishen, and G. J. Jedlovec (2011), Improving numerical weather predictions of summertime precipitation over the southeastern United States through a high-resolution initialization of the surface state, Weather Forecasting, 26, 785–807, doi:10.1175/ 2011WAF2222455.1.

Chen, F., and J. Dudhia (2001), Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Mon. Wea. Rev., 129, 569-585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Collow, T. W., A. Robock, and W. Wu (2014), Influences of soil moisture and vegetation on convective precipitation forecasts over the United States Great Plains, J. Geophys. Res. Atmos., 119, 9338–9358, doi:10.1002/2014JD021454.

Deleted: . et al.

Deleted: Chen, F., H. Kusaka, R. Bornstain, J. Ching, C.S.B. Grimmond, S. Grossman-Clarke, T. Loridan, K. Manning, A. Martilli, S. Miao, D. Sailor, F. Salamanca, H. Taha, M. Tewari, X. Wang, A. Wyszogrodzki, and C. Zhang (2011), The integrated WRF/urban modeling system: development, evaluation, and applications to urban environmental problems. International Journal of Climatology, 31, 273-288, doi: 10.1002/joc.2158. ... [1]

Crawford, J. H., R. R. Dickerson, and J. C. Hains (2014), DISCOVER-AQ: Observations and early results, Environ. Managers, September 2014, 8-15.

Czader, B. H., X. Li, and B. Rappenglueck (2013), CMAQ modeling and analysis of radicals, radical precursors, and chemical transformations, J. Geophys. Res. Atmos., 118, 11,376–11,387, doi:10.1002/jgrd.50807.

De Gouw, J. and Warneke, C. (2007), Measurements of volatile organic compounds in the earth's atmosphere using proton-transferreaction mass spectrometry, Mass Spectrom. Rev., 26, 223–257, doi:10.1002/mas.20119.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendael, M. (2015), Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, Atmos. Chem. Phys., 15, 12519-12545, doi:10.5194/acp-15-12519-2015.

Duncan, B. N., Y. Yoshida, M. R. Damon, A. R. Douglass, and J. C. Witte (2009), Temperature dependence of factors controlling isoprene emissions, Geophys. Res. Lett., 36, L05813, doi:10.1029/2008GL037090.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108, 8851, doi:10.1029/2002JD003296, D22.

Emmerson, K. M., Galbally, I. E., Guenther, A. B., Paton-Walsh, C., Guerette, E.-A., Cope, M. E., Keywood, M. D., Lawson, S. J., Molloy, S. B., Dunne, E., Thatcher, M., Karl, T., and Maleknia, S. D. (2016), Current estimates of biogenic emissions from eucalypts uncertain for southeast Australia, Atmos. Chem. Phys., 16, 6997-7011, doi:10.5194/acp-16-6997-2016.

Fang, L., C. R. Hain, X. Zhan, and M. C. Anderson (2016), An inter-comparison of soil moisture data products from satellite remote sensing and a land surface model, International Journal of Applied Earth Observation and Geoinformation, 48, 37-50, doi: 10.1016/j.jag.2015.10.006.

Fiore, A. M., Levy II, H., and Jaffe, D. A. (2011), North American isoprene influence on intercontinental ozone pollution, Atmos. Chem. Phys., 11, 1697-1710, doi:10.5194/acp-11-1697-2011.

Moved (insertion) [1]

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C. (2006), Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181-3210, doi:10.5194/acp-6-3181-2006.

Guenther, A. B., X. Jiang, C. L. Heald, T. Sakulyanontvittaya, T. Duhl, L. K. Emmons, and X. Wang (2012), The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5 (6), 1471-1492, doi:10.5194/gmd-5-1471-2012.

Hain, C. R., W. T. Crow, J. R. Mecikalski, M. C. Anderson, and T. Holmes (2011), An intercomparison of available soil moisture estimates from thermal infrared and passive microwave remote sensing and land surface modeling, J. Geophys. Res., 116, D15107, doi:10.1029/2011JD015633.

Hewitt, C.N., K. Ashworth, A. Boynard, A. Guenther, B. Langford, A. R. MacKenzie, P. K. Misztal, E. Nemitz, S. M. Owen, M. Possell, T. A. M. Pugh, A. C. Ryan, and O. Wild (2011), Ground-level ozone influenced by circadian control of isoprene emissions, Nature Geoscience, 4, 671-674, doi:10.1038/ngeo1271.

Hogrefe, C., Isukapalli, S., Tang, X., Georgopoulos, P., He, S., Zalewsky, E., Hao, W., Ku, J., Key, T., and Sistla, G. (2011), Impact of biogenic emission uncertainties on the simulated response of ozone and fine Particulate Matter to anthropogenic emission reductions, J. Air Waste Manage., 61, 92-108, doi: 10.3155/1047-3289.61.1.92.

Huang, M., Carmichael, G. R., Chai, T., Pierce, R. B., Oltmans, S. J., Jaffe, D. A., Bowman, K. W., Kaduwela, A., Cai, C., Spak, S. N., Weinheimer, A. J., Huey, L. G., and Diskin, G. S. (2013), Impacts of transported background pollutants on summertime western US air quality: model evaluation, sensitivity analysis and data assimilation, Atmos. Chem. Phys., 13, 359-391, doi:10.5194/acp-13-359-2013.

Huang, M., P. Lee, R. McNider, J. Crawford, E. Buzay, J. Barrick, Y. Liu, and P. Krishnan (2016), Temporal and spatial variability of daytime land surface temperature in Houston: Comparing DISCOVER-AQ aircraft observations with the WRF model and satellites, J. Geophys. Res. Atmos., 121, 185–195, doi:10.1002/2015JD023996.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res., 113, D13103, doi:10.1029/2008JD009944.

Janjic, Z. I. (2000), Comments on "Development and Evaluation of a Convection Scheme for Use in Climate Models", J. Atmos. Sci., 57, 3686, doi:10.1175/1520-0469(2000)057<3686:CODAEO>2.0.CO;2.

Janjic, Z. I. (2002), Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, NCEP Office Note, No. 437, 61.

Janjic, Z. I. (2003), A nonhydrostatic model based on a new approach, Meteorol. Atmos. Phys., 82, 271–285, doi:10.1007/s00703-001- 0587-6.

Janjic, Z., Black, T., Pyle, M., Chuang, H., Rogers, E., and DiMego, G. (2004), An Evolutionary Approach to Nonhydrostatic Modeling, Symposium on the 50th Aniversary of Operational Numerical Weather Prediction, College Park, MD, Amer. Meteor. Soc, available at: http://www.wrf-model.org/wrfadmin/publications/Chuang_Janjic_NWP50yearsfinalshort.pdf (last access: January 2017).

Jin, J., Miller, N. L., and Schlegel, N. (2010), Sensitivity Study of Four Land Surface Schemes in the WRF model, Advances in Meteorology, vol. 2010, Article ID 167436, 11 pages, 2010. doi:10.1155/2010/167436.

Karl, T. G., A. Guenther, R. J. Yokelson, J. Greenberg, M. Potosnak, D. R. Blake, and P. Artaxo (2007), The tropical forest and fire emissions experiment: Emission, chemistry, and transport of biogenic volatile organic compounds in the lower atmosphere over Amazonia, J. Geophys. Res., 112, D18302, doi:10.1029/2007JD008539.

Kennedy, A. D., Dong, X., Xi, B., Xie, S., Zhang, Y., and Chen, J. (2011), A Comparison of MERRA and NARR Reanalyses with the DOE ARM SGP Data, Journal of Climate, 24, 4541-4557, doi: http://dx.doi.org/10.1175/2011JCLI3978.1.

Kumar, S. V., Y. Tian, C. Peters-Lidard, et al., (2006), Land information system: An interoperable framework for high resolution land surface modeling, Environ Modelling Software, 21, 1402-1415, doi: 10.1016/j.envsoft.2005.07.004.

Kumar, S. V., C. Peters-Lidard, Y. Tian, et al. (2008), An Integrated Hydrologic Modeling and Data Assimilation Framework, Computer, 41 (12): 52-59, doi: 10.1109/MC.2008.475.

Lee, E., R. Bieda, J. Shanmugasundaram, and H. Basara Richter (2016), Land surface and atmospheric conditions associated with heat waves over the Chickasaw Nation in the South Central United States, J. Geophys. Res. Atmos., 121, 6284–6298, doi:10.1002/2015JD024659.

Li, G., R. Zhang, J. Fan, and X. Tie (2007), Impacts of biogenic emissions on photochemical ozone production in Houston, Texas, J. Geophys. Res., 112, D10309, doi:10.1029/2006JD007924.

Lin, T.-S. and Cheng, F.-Y. (2015), Impact of soil moisture initialization and soil texture on simulated land-atmosphere interaction in Taiwan, Journal of Hydrometeorology, doi: http://dx.doi.org/10.1175/JHM-D-15-0024.1.

McQueen, J., Huang, J., Shafran, P., Rogers, E., Pondeca, M., DiMego, G., and Stajner, I. (2015a), Evaluation of NCEP Atmospheric Models for Driving Air Quality Prediction, 27th Conference on Weather Analysis and Forecasting, Chicago, IL, available at: https://ams.confex.com/ams/27WAF23NWP/webprogram/Paper273598.html (last access: December 2016).

McQueen, J., Lee, P., Huang, J., Huang, H.-C., Shafran, P., Rogers, E., Pondeca, M., DiMego, G., and Stajner, I. (2015b), NWS NWP models and their Potential Impact for Air Quality Prediction, 7th International workshop on air quality research, College Park, MD, available at: http://www.arl.noaa.gov/documents/IWAQFR/Presentations2015/S4_McQueen_IWAQFR_2015.pdf (last access: December 2016).

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovic, 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. (2006), North American Regional Reanalysis, B. Am. Meteor. Soc., 87, 343–360, doi:10.1175/BAMS-87-3-343.

Millet, D. B., D. J. Jacob, K. F. Boersma, T. M. Fu, T. P. Kurosu, K. Chance, C. L. Heald, and A. Guenther (2008), Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, J. Geophys. Res., 113, D02307, doi:10.1029/2007JD008950.

Miralles, D. G., M. J. van den Berg, A. J. Teuling, and R. A. M. de Jeu (2012), Soil moisture-temperature coupling: A multiscale observational analysis, Geophys. Res. Lett., 39, L21707, doi:10.1029/2012GL053703.

Monin, A. S., and A. M. Obukhov (1954), Basic laws of turbulent mixing in the surface layer of the atmosphere, Tr. Akad. Nauk. SSSR Geophiz. Inst., 24 (151), 163–187.

Deleted: Melillo, J. M., Richmond, T.

Moved up [1]: C.,

Deleted: and Yohe, G. W. (Eds.) (2014), Climate Change Impacts in the United States: The Third National Climate Assessment, US Global Change Research Program, 841 pp., doi:10.7930/J0Z3 ... [2]

Müller, M., Mikoviny, T., Feil, S., Haidacher, S., Hanel, G., Hartungen, E., Jordan, A., Märk, L., Mutschlechner, P., Schottkowsky, R., Sulzer, P., Crawford, J. H., and Wisthaler, A. (2014), A compact PTR-ToF-MS instrument for airborne measurements of volatile organic compounds at high spatiotemporal resolution, Atmos. Meas. Tech., 7, 3763-3772, doi:10.5194/amt-7-3763-2014.

Palmer, P. I., D. J. Jacob, A. M. Fiore, R. V. Martin, K. Chance, and T. P. Kurosu (2003), Mapping isoprene emissions over North America using formaldehyde column observations from space, J. Geophys. Res., 108(D6), 4180, doi:10.1029/2002JD002153.

Pan, L., Tong, D. Q., Lee, P., Kim, H., and Chai, T. (2014), Assessment of NO_x and O_3 forecasting performances in the U.S. National Air Quality Forecasting Capability before and after the 2012 major emissions updates, Atmos. Environ., 95, 610–619, doi: 10.1016/j.atmosenv.2014.06.020.

Pegoraro, E., Rey, A., Greenberg, J., Harley, P., Grace, J., Malhi, Y., and Guenther, A. (2004), Effect of drought on isoprene emission rates from leaves of Quercus virginiana Mill, Atmos. Environ., 38, 6149–6156, doi:10.1016/j.atmosenv.2004.07.028.

Peters-Lidard, C.D., E. M. Kemp, T. Matsui, J. A. Santanello, Jr., S. V., Kumar, J. P. Jacob, T. Clune, W.-K. Tao, M. Chin, A. Hou, J. L. Case, D. Kim, K.-M. Kim, W. Lau, Y. Liu, J.-J. Shi, D. Starr, Q. Tan, Z. Tao, B. F. Zaitchik, B. Zavodsky, S. Q. Zhang, and M. Zupanski (2015), Integrated modeling of aerosol, cloud, precipitation and land processes at satellite-resolved scales, Environmental Modelling & Software, 67, 149–159, doi: 10.1016/j.envsoft.2015.01.007.

Rodell, M., Houset, P. R., Berg, A. A., and Famiglietti, J.S. (2005), Evaluation of 10 Methods for Initializing a Land Surface Model, Journal of Hydrometeorology, 6, 146-155, doi: http://dx.doi.org/10.1175/JHM414.1.

Rogers, E., T. Black, B. Ferrier, Y. Lin, D. Parrish, and G. DiMego (2001), Changes to the NCEP Meso Eta analysis and forecast system: Increase in resolution, new cloud microphysics, modified precipitation assimilation, modified 3DVAR analysis, NWS Technical Procedures Bulletin, available at: http://www.emc.ncep.noaa.gov/mmb/mmbpll/eta12tpb (last access: December 2016).

Royer, A., and S. Poirier (2010), Surface temperature spatial and temporal variations in North America from homogenized satellite SMMR-SSM/I microwave measurements and reanalysis for 1979–2008, J. Geophys. Res., 115, D08110, doi:10.1029/2009JD012760.

Saide, P. E., Carmichael, G. R., Spak, S. N., Gallardo, L., Osses, A. E., Mena-Carrasco, M. A., and Pagowski, M. (2011), Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRFChem CO tracer model, Atmos. Environ., 45, 2769–2780, doi:10.1016/j.atmosenv.2011.02.001.

Santanello, J. A., S. V. Kumar, C. D. Peters-Lidard, K. W. Harrison, and S. Zhou (2013), Impact of Land Model Calibration on Coupled Land–Atmosphere Prediction, J. Hydrometeorology, 14 (5), 1373-1400, doi: 10.1175/JHM-D-12-0127.1.

Santanello, J. A., S. V. Kumar, C. D. Peters-Lidard, and P. M. Lawston (2016), Impact of Soil Moisture Assimilation on Land Surface Model Spinup and Coupled Land-Atmosphere Prediction, J. Hydrometeorology, doi:10.1175/jhm-d-15-0072.1.

Skamarock, W., J.B. Klemp, J. Dudhia, D.O. Gill, D. Barker, M.G. Duda, X.-Y. Huang, and W. Wang (2008), A Description of the Advanced Research WRF Version 3, NCAR Technical Note NCAR/TN-475+STR, doi:10.5065/D68S4MVH.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W. (2014), Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317-9341, doi:10.5194/acp-14-9317-2014.

Tao, Z., J. A. Santanello, M. Chin, S. Zhou, Q. Tan, E. M. Kemp, and C. D. Peters-Lidard (2013), Effect of land cover on atmospheric processes and air quality over the continental United States – a NASA Unified WRF (NU-WRF) model study. Atmospheric Chemistry and Physics, 13, 6207-6226, doi:10.5194/acp-13-6207-2013.

Toon, O. B., H. Maring, J. Dibb, R. Ferrare, D. J. Jacob, E. J. Jensen, Z. J. Luo, G. G. Mace, L. L. Pan, L. Pfister, K. H. Rosenlof, J. Redemann, J. S. Reid, H. B. Singh, A. M. Thompson, R. Yokelson, P. Minnis, G. Chen, K. W. Jucks and A. Pszenny (2016), Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC^{4}RS) field mission, J. Geophys. Res. Atmos., 121, 4967–5009, doi:10.1002/2015JD024297.

Wang, X., Situ, S., Guenther, A. B., Chen, F., Wu, Z., Xia, B., and Wang, T. (2011), Spatiotemporal variability of biogenic terpenoid emissions in Pearl River Delta, China, with high-resolution land-cover and meteorological data, Tellus B, 63, 241–254, doi: 10.1111/j.1600-0889.2010.00523.x.

Warneke, C., J. A. de Gouw, L. Del Negro, J. Brioude, S. McKeen, H. Stark, W. C. Kuster, P. D. Goldan, M. Trainer, F. C. Fehsenfeld, C. Wiedinmyer, A. B. Guenther, A. Hansel, A. Wisthaler, E. Atlas, J. S. Holloway, T. B. Ryerson, J. Peischl, L. G. Huey, and A. T. Case Hanks, (2010), Biogenic emission measurement and inventories determination of biogenic emissions in the eastern United States and Texas and comparison with biogenic emission inventories, J. Geophys. Res., 115, D00F18, doi:10.1029/2009JD012445.

Wiedinmyer, C., J. Greenberg, A. Guenther, B. Hopkins, K. Baker, C. Geron, P. I. Palmer, B. P. Long, J. R. Turner, G. Pétron1, P. Harley, T. E. Pierce, B. Lamb, H. Westberg, W. Baugh, M. Koerber, and M. Janssen, (2005), Ozarks Isoprene Experiment (OZIE): Measurements and modeling of the "isoprene volcano", J. Geophys. Res., 110, D18307, doi:10.1029/2005JD005800.

Wolfe, G. M., T. F. Hanisco, H. L. Arkinson, T. P. Bui, J. D. Crounse, J. Dean-Day, A. Goldstein, A. Guenther, S. R. Hall, G. Huey, D. J. Jacob, T. Karl, P. S. Kim, X. Liu, M. R. Marvin, T. Mikoviny, P. K. Misztal, T. B. Nguyen, J. Peischl, I. Pollack, T. Ryerson, J. M. St. Clair, A. Teng, K. R. Travis, K. Ullmann, P. O. Wennberg, and A. Wisthaler, (2015), Quantifying sources and sinks of reactive gases in the lower atmosphere using airborne flux observations, Geophys. Res. Lett., 42, 8231–8240, doi:10.1002/2015GL065839.

Yu, M., Carmichael, G. R., Zhu, T., Cheng, Y. (2012), Sensitivity of predicted pollutant levels to urbanization in China, Atmos. Environ., 60, 544-554, 1352-2310, doi: 10.1016/j.atmosenv.2012.06.075.

Zaitchik, B. F., J. A. Santanello, S. V. Kumar, and C. D. Peters-Lidard (2012), Representation of Soil Moisture Feedbacks during Drought in NASA Unified WRF (NU-WRF), J. Hydrometeor, 14 (1), 360-367, doi: 10.1175/JHM-D-12-069.1.

Zeng, X.-M., Wang, B., Zhang, Y., Song, S., Huang, X., Zheng, Y., Chen, C., and Wang, G. (2014), Sensitivity of high-temperature weather to initial soil moisture: a case study using the WRF model, Atmos. Chem. Phys., 14, 9623-9639, doi:10.5194/acp-14-9623-2014.

Zhang, R., Huang, C., Zhan, X., Dai, Q., and Song, K. (2016), Development and validation of the global surface type data product from S-NPP VIIRS, Remote Sens. Lett., 7, 51-60, doi: 10.1080/2150704X.2015.1101649.

Zhang, Y., Y. Wang, G. Chen, C. Smeltzer, J. Crawford, J. Olson, J. Szykman, A. J. Weinheimer, D. J. Knapp, D. D. Montzka, et al. (2016), Large vertical gradient of reactive nitrogen oxides in the boundary layer: Modeling analysis of DISCOVER-AQ 2011 observations, J. Geophys. Res. Atmos., 121, 1922–1934, doi:10.1002/2015JD024203.

Deleted: et al.

Deleted: et al.

Deleted: et al.

Figures

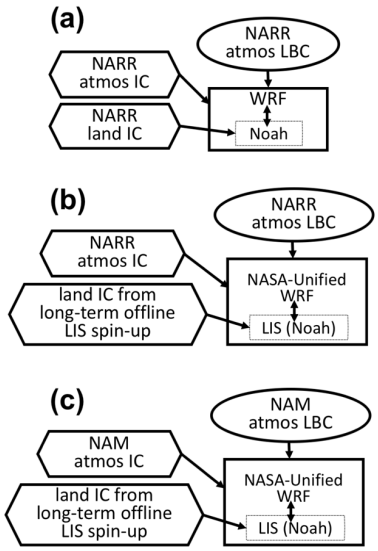


Figure 1: Illustration of the different (NU)WRF initialization methods compared in this study: (a) A usual method, applied only on a 12 km resolution grid, in which the land and atmospheric (atmos) initial conditions (IC) and lateral boundary conditions (LBC) are downscaled from a coarser model output North American Regional Reanalysis (NARR). (b) The ctrl runs on 12 km and 4 km resolution grids, in which the land IC are from long-term offline LIS spin-up at the same grid resolutions as NUWRF, forced by highly resolved atmospheric forcing and precipitation data; atmospheric IC and LBC are from NARR. (c) Same as (b), except that atmospheric IC and LBC are from the North American Mesoscale Forecast System (NAM). Table 1 summarizes all NUWRF simulations in this study including the corresponding initialization methods illustrated here.

Deleted: , ... [3]

Deleted: , ... [4]

Deleted: two

Deleted:);

Deleted: the

Deleted: both

Deleted:) are used

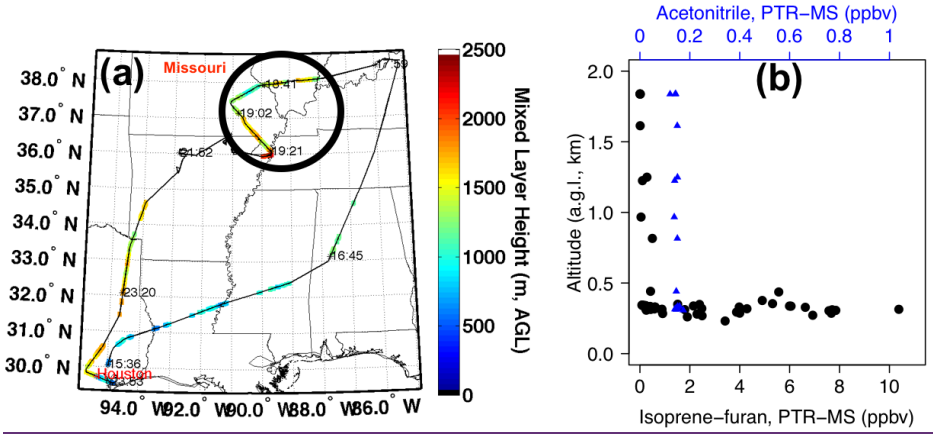


Figure 2: Measurements during the SEAC[4]RS campaign on 11 September, 2013: (a) the DC-8 flight path colored by mixed layer height from the DIAL-HSRL instrument. Areas within the black circle indicate the Missouri "isoprene volcano" regions sampled at ~19 UTC (~13 local standard time). (b) Vertical profiles of the PTR-MS measured isoprene (black dots) and acetonitrile (blue triangles) at around the "isoprene volcano" regions.

Deleted: . ... [5]

Deleted: and DISCOVER-AQ Houston campaigns

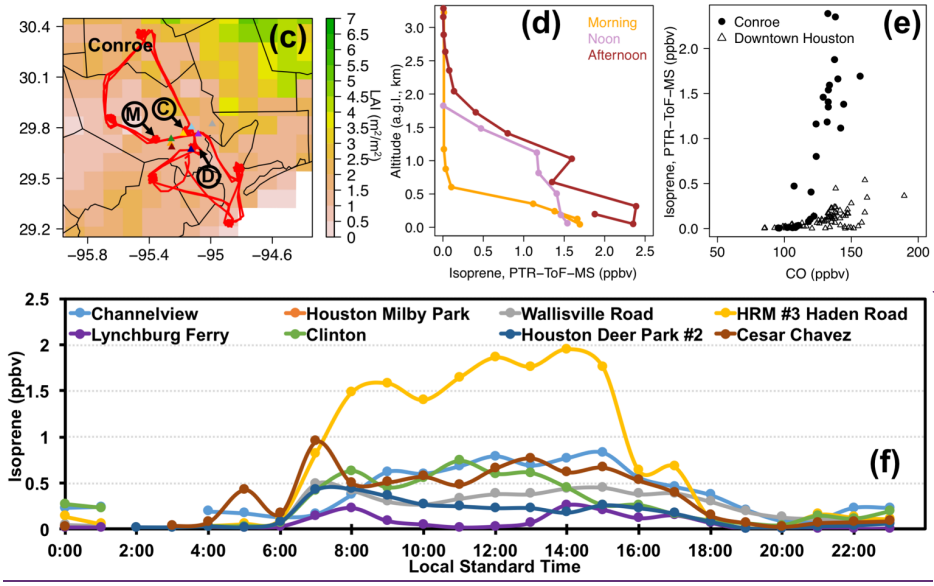


Figure 2 (cont.): Measurements during the DISCOVER-AQ Houston campaign on 11 September, 2013: (c) MODIS leaf area index (8-day mean with missing values filled with the monthly mean) over the Greater Houston area in the 12 km NUWRF grid. The red solid line indicates the DISCOVER-AQ P-3B flight path. (d) Vertical profiles of the PTR-ToF-MS measured isoprene at the Conroe spirals in the Greater Houston area at different times of the day. (e) Scatterplot of the P-3B measured isoprene-CO at the Conroe spirals (filled circles) and at three downtown Houston/Ship Channel sites of Moody Tower, Deer Park and Channelview (open triangles), on 11 September, 2013. Locations of Conroe, Moody Tower (M), Deer Park (D) and Channelview (C) are defined in Figure 2c. The CO measurements were taken using a Diode laser spectrometer measurements of CO, CH_4, N_2O instrument with uncertainty of 2% or 2 ppbv. (f) AutoGC isoprene measurements at multiple surface sites in Houston urban and ship channel areas. Locations of these sites are shown in Figure 2c as colored triangles. Note that only one AutoGC data point is available at the Milby Park site at 00 local standard time.

Formatted: Space After: 12 pt

Deleted: (c

Moved (insertion) [2]

Deleted: (d) MODIS leaf area index (8-day mean with missing values filled with the monthly-mean) over the Greater Houston area in the 12 km NUWRF grid.

Moved up [2]: The red solid line indicates the DISCOVER-AQ P-3B flight path.

Deleted: Colored triangles in Houston urban and ship channel areas denote the locations of surface sites that had hourly AutoGC isoprene measurements shown in (e).

Deleted: Locations of the Conroe and urban/ship channel spirals (M: Moody Tower; C: Channelview; D: Deer Park, cited in Figure A1) are marked in black in (d).

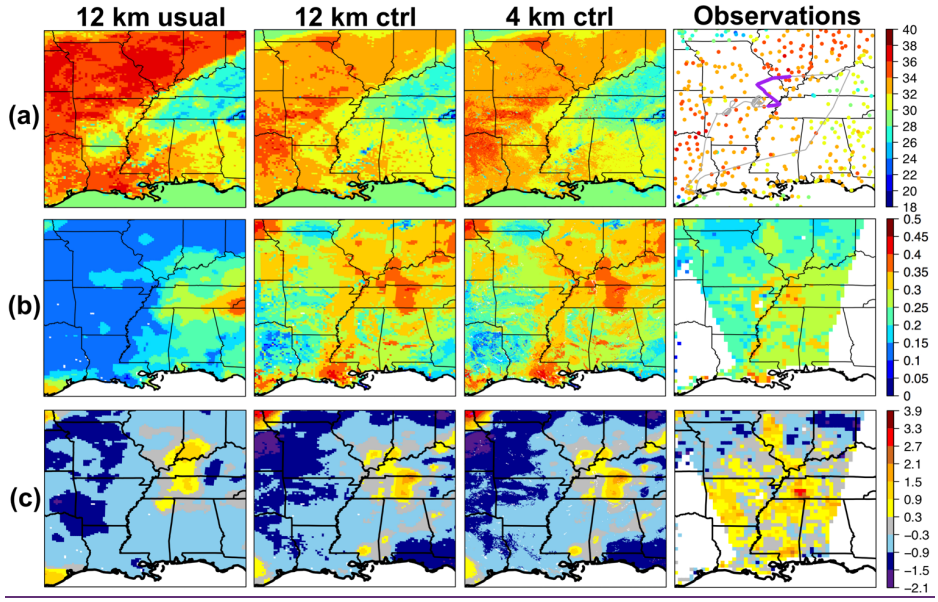


Figure 3: Evaluation of NUWRF (a) surface air temperature in °C at ~13 local standard time (b) surface soil moisture in m^3/m^3 and (c) normalized surface soil moisture at ~00 local standard time (NUWRF initial time) over the central/southeastern US on 11 September, 2013. NUWRF simulations are shown in the first three columns; NCEP surface in-situ temperature observations and the ESA CCI combined daily soil moisture product (0.25°×0.25°, only showing data with quality flag = 0) are shown in the right column. The uncertainty of the CCI product on 11 September, 2013 is shown in Figure S2 (left). The normalized anomaly is defined in Eq. (3) in the text, and Figure S2 (right) shows the size of usable CCI data in September 2013. Note that for (b) and (c), warm (cool) colors indicate high (low) soil moisture values, opposite to the commonly used color scheme in hydrological studies.

Deleted: .

Deleted: 's

Deleted: and

Deleted: the

Deleted: The

Deleted: Surface

Deleted: (in the 12 km NUWRF grid)

Deleted: the

Deleted: panel in (c) only

Deleted: results at where

Deleted: are available on >=80% of the days

Deleted: .

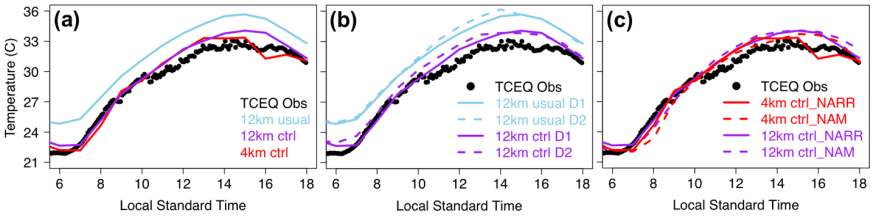


Figure 4: Comparing NUWRF daytime surface air temperature with the 5-minute TCEQ surface in-situ measurements (black dots) at Conroe, TX, on 11 September, 2013: (a) Evaluating the impact of NUWRF land initialization and grid resolution; (b) Evaluating the impact of NUWRF simulation length, including 1-day (D1) and 2-day (D2) forecasts, on the 12 km grid; (c) Evaluating the impact of NUWRF atmospheric initialization (using NARR and NAM) on 12 km and 4 km grids.

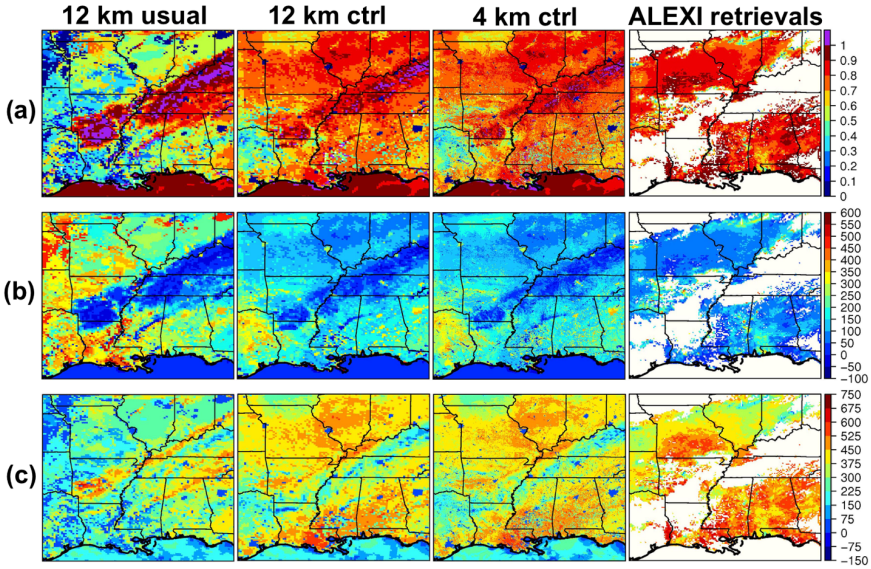


Figure 5: Comparing (a) evaporative ratio, unitless, defined as latent heat/(latent+sensible heat); (b) sensible heat in W/m^2, and (c) latent heat in W/m^2 from three NUWRF simulations (first three columns) with the NOAA operational 0.08°×0.08° ALEXI retrievals (right), on 11 September, 2013.

Deleted: 's
Deleted: the
Deleted: in
Deleted: 's
Deleted: in

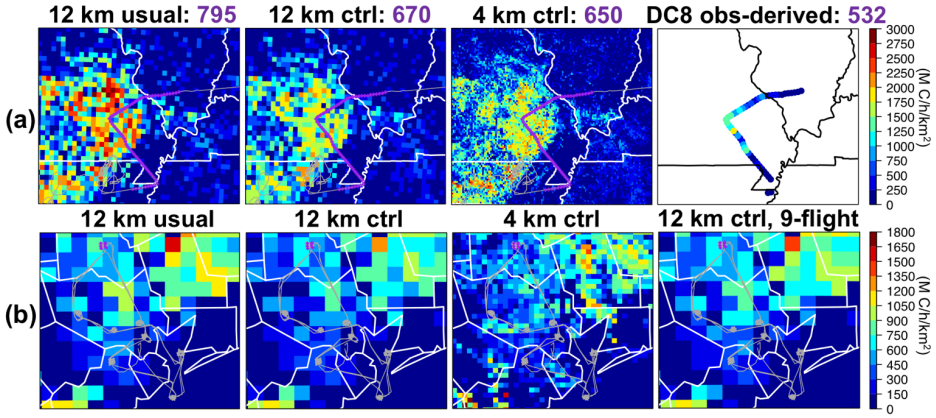


Figure 6: (a) Isoprene emissions around the "isoprene volcano" areas in Missouri based on NUWRF-MEGAN and aircraft observations at ~13 local standard time on 11 September, 2013. The mean values along the DC-8 flight path during 12:30-13:30 local standard time (in purple) are indicated in the figure captions. Open purple dots along the flight path refer to where isoprene data are missing or above the PBL. (b) Isoprene emissions in Houston, TX from NUWRF-MEGAN on 11 September, 2013 and on 9 DISCOVER-AQ flight days in September 2013, at local noontime. The DISCOVER-AQ P-3B flight path on 11 September, 2013 is overlaid in grey and the Conroe samples at around the noontime are highlighted in purple. Note that the flight paths for all other flight days are similar to but not exactly the same as that of 11 September.

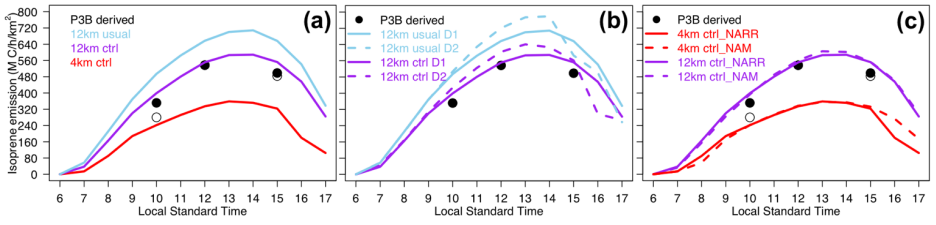


Figure 7: Comparing NUWRF-MEGAN daytime isoprene emissions in Conroe on 11 September, 2013 with the observation-derived emissions. The filled and open black circles indicate the calculations based on PBLHs from the 12 km and 4 km NARR IC/LBC ctrl runs, respectively: (a) Evaluating the impact of NUWRF land initialization and grid resolution; (b) Evaluating the impact of NUWRF simulation length, including 1-day and 2-day forecasts, on the 12 km grid; (c) Evaluating the impact of NUWRF atmospheric initialization (using NARR and NAM) on 12 km and 4 km grids.

Deleted: The open

Deleted: (note

Deleted: the

Deleted: September's) is overlaid in grey and the Conroe samples at around the noontime are highlighted in purple

Deleted: the

Deleted: black

Deleted: the

Deleted: in

Deleted: 's

Deleted: in

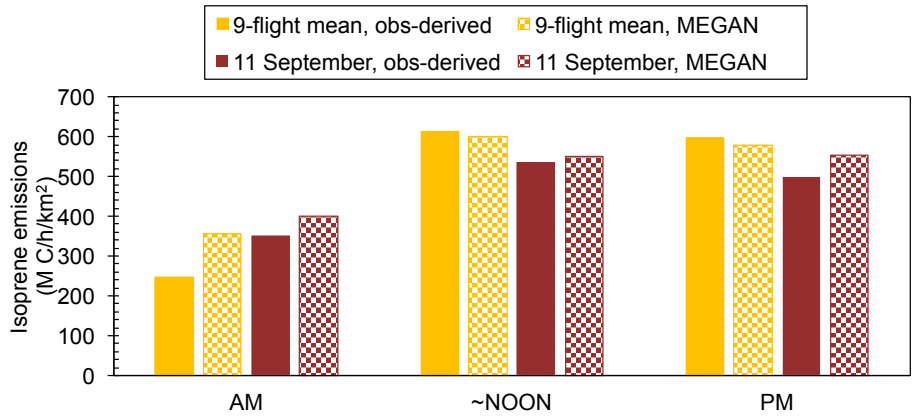
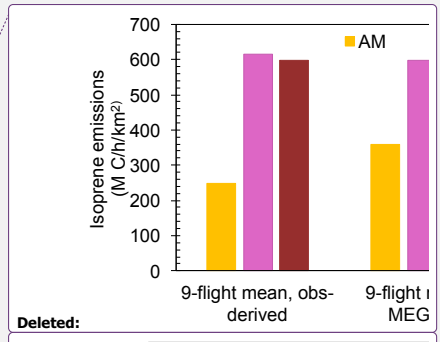
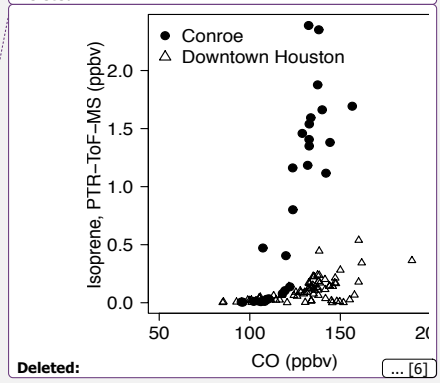


Figure 8: Temporal variability (AM: 15-16 UTC; ~noon: 18-19 UTC; PM: 20-21 UTC) of isoprene emissions in Conroe averaged on multiple DISCOVER-AQ flight days in September 2013 from NUWRF-MEGAN and aircraft observations, comparing with the 11 September, 2013 conditions. The multi-flight day mean P-3B isoprene, NAQFC CMAQ OH and NUWRF PBLH were used to derive the multi-day mean observation-based emissions.



Deleted:



Deleted:

... [6]

Tables

Table 1: Summary of all NUWRF simulations in this study. More descriptions on these simulations can be found in Figure 1 and Section 2.1.3.

Case Name	Horizontal resolution	Land initialization	Atmospheric initialization/ Lateral boundary conditions	Approach illustrated in Figure 1
12 km usual(_NARR)[a]	12 km	NARR	NARR	(a)
12 km usual_veg(_NARR)[a,b]	12 km	NARR	NARR	(a)
12 km ctrl(_NARR)[a]	12 km	LIS	NARR	(b)
4 km ctrl(_NARR)[a]	4 km	LIS	NARR	(b)
12 km ctrl_NAM	12 km	LIS	NAM	(c)
4 km ctrl_NAM	4 km	LIS	NAM	(c)

[a]The *_NARR simulations are the focus of this study and the "(*_NARR)" part in the case names was often omitted in figures and texts. [b]Results are only shown in Figure S3 and briefly discussed in Section 3.1.1.

Table 2: NUWRF air temperature and PBLH performance along the DC-8 flight path (at where the isoprene data are available) around the Missouri "isoprene volcano" region on 11 September, 2013. A 1.6 km PBLH, which approximates to the 4 km/12 km simulated mean values (in italic/bold), was used to derive the emissions.

NUWRF Case Name	PBLH (km), mean ± standard deviation	Air temperature RMSE (°C)
4 km ctrl	***1.551***±0.304	0.735
12 km ctrl	***1.569****±0.369*	0.771
12 km usual	2.190±0.630	2.241

Table 3: NUWRF-simulated median PBLH (km) and daytime (6-17 local standard time) surface air temperature performance at different times of 11 September, 2013, at around Conroe, TX. The bold italic numbers were used to derive emissions.

NUWRF Case Name	PBLH (km)			Surface air temperature (°C)[a]	
	Morning	~Noon	Afternoon	Mean bias	RMSE
4 km ctrl	***0.666***	***1.467***	***1.752***	0.159	0.747
12 km ctrl	***0.839***	***1.465***	***2.038***	0.640	0.864
12 km usual	1.195	1.746	2.337	2.473	2.507

Formatted: Right: 1.2"

Formatted: Right: 1.01"

Deleted: [b]Only

Deleted: A2

Deleted: in Missouri

Deleted: was close

[a]The two data points nearest to 00:00 (minute:second) from the TCEQ 5-minute special ground observations were averaged and compared with the NUWRF output hourly-recorded at 00:00.

Deleted: the full clocks

Deleted: NUWRF output

Page 20: [1] Deleted **Min Huang** **6/23/17 9:46 PM**

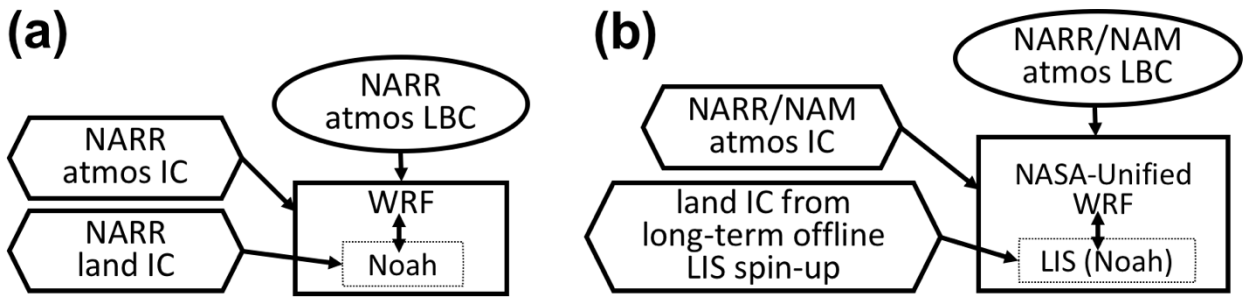
Chen, F., H. Kusaka, R. Bornstain, J. Ching, C.S.B. Grimmond, S. Grossman-Clarke, T. Loridan, K. Manning, A. Martilli, S. Miao, D. Sailor, F. Salamanca, H. Taha, M. Tewari, X. Wang, A. Wyszogrodzki, and C. Zhang (2011), The integrated WRF/urban modeling system: development, evaluation, and applications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

Page 24: [2] Deleted **Min Huang** **6/23/17 9:46 PM**

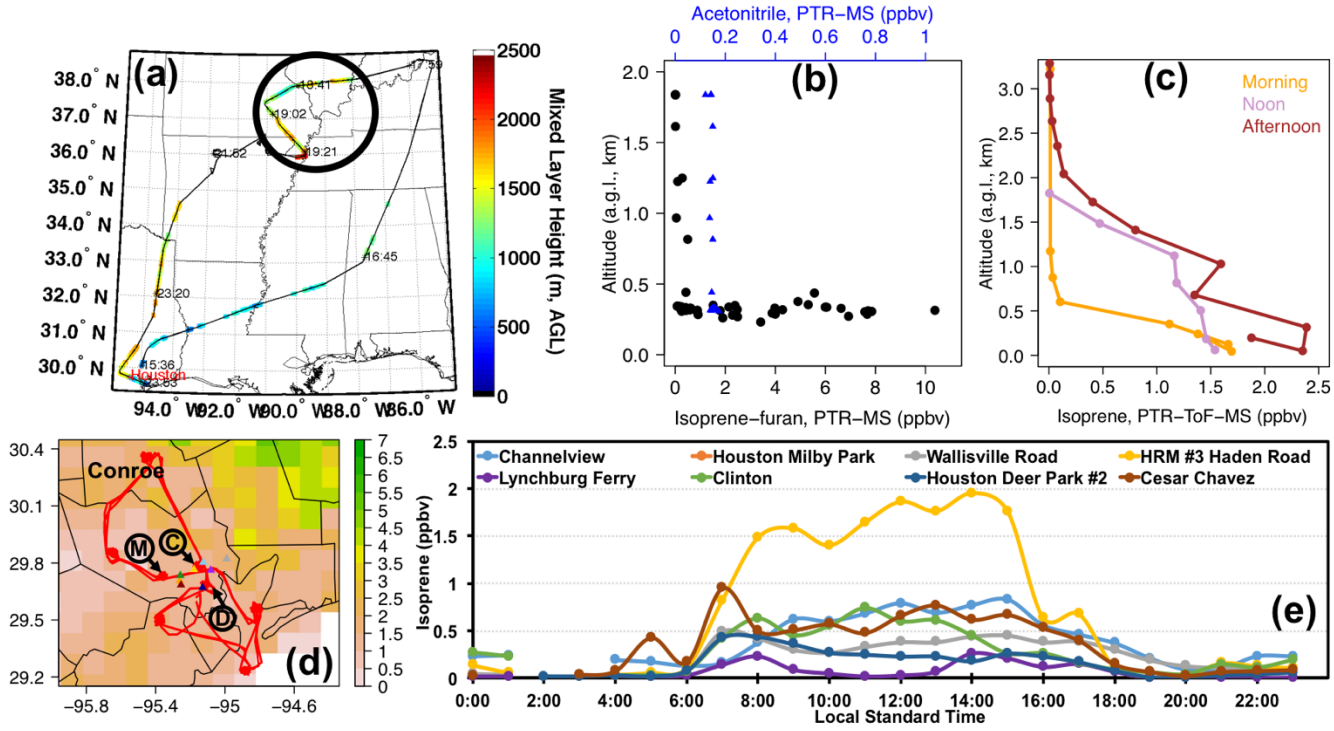
and Yohe, G. W. (Eds.) (2014), Climate Change Impacts in the United States: The Third National Climate Assessment, US Global Change Research Program, 841 pp., doi:10.7930/J0Z31WJ2.

Page 28: [3] Deleted **Min Huang** **6/23/17 9:46 PM**

Page 28: [4] Deleted **Min Huang** **6/23/17 9:46 PM**



Page 29: [5] Deleted **Min Huang** **6/23/17 9:46 PM**



Page 34: [6] Deleted	Min Huang	6/23/17 9:46 PM

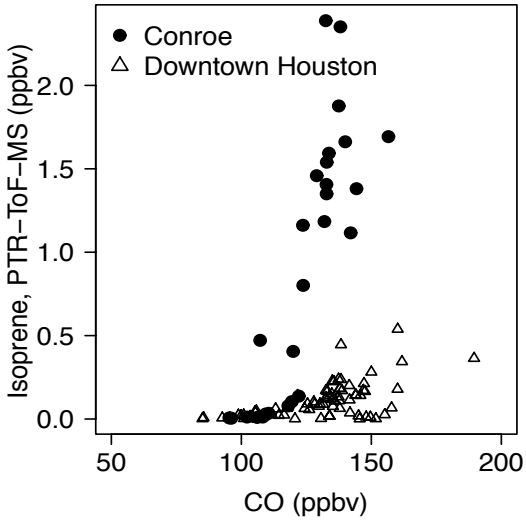


Figure A1: Scatterplot of the P-3B measured isoprene-carbon monoxide (CO) at the Conroe spirals (filled circles) and at three Downtown Houston/Ship Channel sites of Moody Tower, Deer Park and Channelview (open triangles), on 11 September, 2013. Locations of Moody Tower, Deer Park and Channelview are defined in Figure 2d. The CO measurements were taken using a Diode laser spectrometer measurements of CO, CH_4, N_2O (DACOM) instrument with uncertainty of 2% or 2 ppbv.

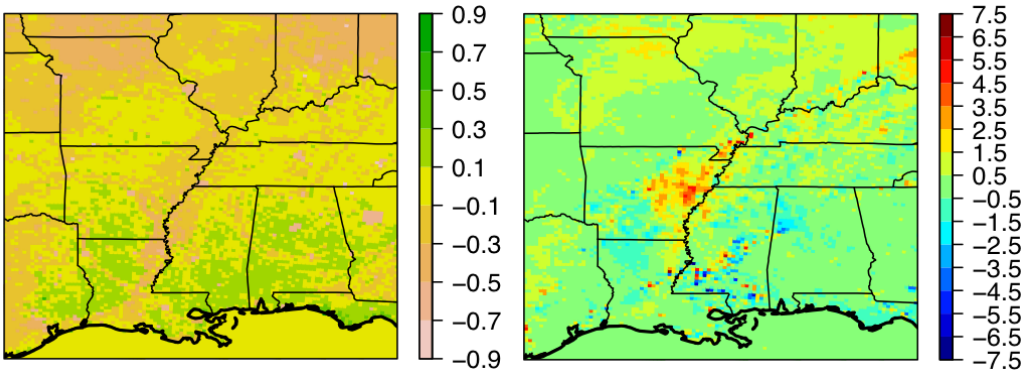


Figure A2: (left) The difference between daily near real-time GVF and the climatological monthly-mean GVF on 11 September, shown on the 12 km grid, and (right) the resulting differences in NUWRF usual runs' (12 km usual_veg-12 km usual, as defined in Table 1) surface air temperature in °C at 13 local standard time on this day.

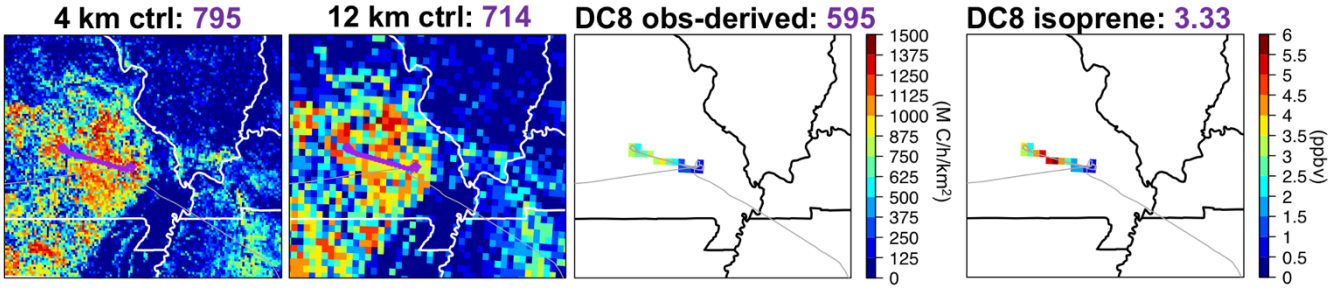


Figure A3: (Left three columns) Isoprene emissions around the "isoprene volcano" areas in Missouri from NUWRF-MEGAN at 14 local standard time and aircraft observations on 06 September, 2013. The mean values along the DC-8 flight path during 13:20-14:30 local standard time (when aircraft observations were made along four West-East transects) are indicated in purple in the figure captions. (Right column) Observed isoprene concentrations in ppbv along the DC-8 flight path during 13:20-14:30 local standard time, with the mean values shown in the figure captions. The observations and observation-derived isoprene emissions were averaged to the 12 km NUWRF grid for the plots. The 12 km ctrl NUWRF-MEGAN isoprene emissions overall have a slightly lower positive mean bias (~20%) from the observation-derived emissions than the 11 September result (~26%), whereas the positive biases of 4 km ctrl NUWRF-MEGAN emissions (~34%) are larger than the 11 September result (~22%). The largest overprediction occurs near the east side of the transects, where the biases are higher in the 4 km case than in the 12 km case. These biases are also shown by Wolfe et al. (2015) in their MEGAN emissions computed using a different meteorological input.