***We thank the anonymous reviewer for their detailed comments that helped improve the manuscript.***

Reviewer #2:

In this study, the authors performed eight WRF model simulations (at 12, 4, and 1.3 km horizontal resolutions) to assess the impact of different parameterization schemes, land/lake surface initialization approaches, and analysis nudging methods on the simulated surface energy fluxes and near-surface atmospheric conditions over the Lake Michigan region during the 1-month LMOS field campaign period in 2017. Model evaluation presented in this work helped the same group to select meteorological inputs of the CMAQ simulations described in a companion paper by Pierce et al., which is also currently under review for the same journal.

My major comments include:

**Novelty:** Comparing Pleim-Xiu with indirect soil nudging and Noah is not a new idea-more than 7 years ago, a TCEQ funded project on such a topic was conducted. Many modeling communities including at NASA and NOAA plan to (or are already actively working on) migrating from Noah to Noah-MP land surface model due to known limitations in Noah. Initialization WRF using output from LIS or similar frameworks to benefit weather and air quality studies is not new neither, which has been recognized by the authors themselves. Running at very high-resolution over regions with complex surface types (e.g., land vs water) is also broadly appreciated by modeling communities. Although this is a companion paper of Pierce et al., it should also be able to stand alone with its own highlights. Thus the author are encouraged to clearly underscore the novel aspects of this study, and this may lead to adding some additional modeling experiments and more rigorous evaluation. And perhaps the expected paper would fit better into GMD. Otherwise, shortening the paper and merging the key information into Pierce et al. is suggested.

***The novel contribution of this paper is its detailed assessment of a multi-resolution (12 km to 1.3 km) and multi-constraint modeling framework for the Lake Michigan area. As discussed in the introduction, given the important role that boundary layer meteorology and the land-lake breeze circulation have on ozone production and transport in this region, it is critical to explore the ability of different parameterization schemes and surface datasets to improve the accuracy of near-surface meteorological and air quality simulations. Regarding the Pierce et al. (2023) companion paper, we have added some text to better tie these two papers together. This includes describing the importance of surface meteorology on ozone production in the last paragraph of the introduction and discussing how differences in meteorology impact biogenic emissions in the chemistry simulations. We have also added a sentence to the last paragraph of this paper to serve as a segue to the Pierce et al. (2023) paper. It is not feasible to combine these two papers into a single paper given the amount of material being presented; however, we are willing to transfer this paper to GMD, if necessary, as long as it does not require another set of reviews.***

**Methods and presentation:** While a lot of information is given, clarifications on the methods are still necessary. Specifically:

1. The authors stated at L222-223 that "*Direct insertion into the WRF model was possible because of the similarly configured Noah LSM used in both the LIS and WRF simulations*". Please provide the version of Noah in LIS that was used in this study as well as evidence showing it's similar to Noah embedded in WRF3.8.1. Also, this statement cannot be agreed if the static inputs (land use/land cover, soil type, terrain, etc) of the land model are consistent in LIS/Noah and WRF3.8.1. LIS and WRF3.8.1 static inputs are generated from LDT and WPS tools, respectively. Please clarify what exactly has been done. This study area has complex surface type (not only land vs water, but also for land, urban vs non-urban categories), so some discussions on how these surface characteristics are represented by the model would be very informative.

***Our use of "direct insertion" in lines 222-223 when describing how we used the SPoRT LIS soil moisture and soil temperature data was imprecise. It was simply meant to convey that the use of the same vertical layers in the Noah LSM used in both SPoRT LIS and the WRF model made it easier to use the LIS output in our WRF model sensitivity experiments since there would have been no reason to interpolate between vertical layers. We have deleted this sentence from the revised paper because it was unnecessary. Version 3.6 of the Noah LSM was used in the SPoRT LIS, which is very similar to that used in version 3.8.1 of the WRF model. All of our simulations with the Noah LSM used the default settings for the vegetation and soil properties, with the exception of the YNT_GVF and YNT_SSNG simulations where the climatological vegetation fraction was replaced by the high-resolution daily VIIRS green vegetation fraction data. To address your request for more information, we have expanded the first paragraph of the methods section describing the WRF model configurations to provide additional detail about how the simulations were performed and to mention that the WRF Preprocessing System (WPS) was used to prepare the surface datasets. In particular, we have added these sentences: "Except for the two baseline simulations described below, all of the simulations were performed in daily increments using the standard WRF model restart files to allow for daily updates of high-resolution surface datasets using the WPS. A sentence has also been added to Section 2.1 stating that: "The 40-category National Land Cover Dataset (NLCD) 2011 land use dataset (Jin et al. 2013) was used to determine the vegetation type and soil properties for each model grid point."***

2. The evaluation of the model runs are not rigorous and not well connected with Pierce et al. The uncertainty of VIIRS GVF is not introduced in the paper - this product sounds to have short latency but for retrospective analysis like it'd be important to tell its quality on a daily time scale for this region relevant to air pollution events studied. Similar comment on GLSEA. The SPoRT LIS product is not discussed clearly (it is not very clear whether land data assimilation is enabled in the LIS system and if so, some data assimilation diagnostics could be shown) - my understanding is that SPoRT hosts documentations and visualizations of these routinely generated products elsewhere which may be cited in the paper. In terms of WRF model evaluation, some statistics and maps are presented but only for a limited number of variables, and as the authors noted at L232-233, "*these surface observations were also used to perform surface nudging during the EPA simulation, which will impact the results presented in Section 3 because surface nudging was not used during any of the YNT simulations*". As the model outputs served as meteorological input of CMAQ, a list of variables central to pollutants to be studied should be selected with

justifications, followed by model performance of them. The performance could be discussed in connection with the air pollution events and time series presented in Pierce et al., and additional evaluation metrics such as correlations between modeled and observed time series may be added. Furthermore, are there really no in-situ flux measurements/PBL info across the entire three WRF domains as stated at L402?

***To address your comment about the VIIRS and GLSEA datasets, we have added sentences to Sections 2.2.1 and 2.2.2 stating that: "Only satellite observations are used to produce the daily lake surface temperature analyses, which Schwab et al. (1992) showed have small bias and root mean square error (1-1.5 °C) when compared to buoys." and "Ding and Zhu (2018) have shown that the VIIRS GVF product has smaller errors and bias than other satellite derived GVF datasets because of reduced atmospheric influences, improved observing capabilities in high biomass regions, better representation of vegetation canopies, and reduced bidirectional reflection distribution function effects." Regarding the NASA SPoRT LIS run, we now state in Section 2.2.3 of the revised manuscript that no observations were assimilated during the LIS runs. The model variables that we chose to evaluate in this paper (2-m temperature, 2-m water vapor mixing ratio, 10-m wind speed, and PBL height) are all important for air quality modeling applications. They are typically used in model verification studies given their relevance and availability over large spatial domains. As for flux tower measurements, there are a few stations across the domain; however, we have chosen not to include them in this analysis because of their sparse distribution and difficulties handling representativeness issues due to differences in spatial scale. Surface fluxes can vary greatly over short distances, which makes it difficult to use them for model verification. Finally, as requested, we computed the correlations between the simulated and observed meteorological variables using hourly data over the 7-week study period. The correlations are shown in the right column in the figure below. It is evident that the correlations are very similar among all of the simulations for 2-m temperature, 2-m mixing ratio, and 10-m wind speed on the 12-km and 4-km domains. On the 1.3-km domain, however, the correlations are less for the AP-XM simulation whereas they remain similarly high among the various YNT simulations. These results are consistent with what was shown in the RMSE percentage statistics shown in the original manuscript and therefore we decided not to include the correlations in the revised manuscript.***

| | Bias | | | % RMSE Change vs. AP_XM | | | % RMSE Change vs. YNT | | | Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **a)** 2-m Temperature [K] | | | **b)** 2-m Temperature [K] | | | **c)** 2-m Temperature [K] | | | **d)** 2-m Temperature [K] | | |
| Simulation | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km |
| AP_XM | -0.66 | -0.83 | -0.14 | 2.27 | 2.36 | 3.03 | | | | 0.93 | 0.92 | 0.85 |
| YNT | 0.16 | 0.47 | 0.55 | 1.37 | -5.12 | -25.83 | 2.30 | 2.24 | 2.25 | 0.93 | 0.93 | 0.93 |
| YNT_SST | 0.17 | 0.48 | 0.56 | 0.79 | -5.67 | -26.22 | -0.57 | -0.58 | -0.53 | 0.93 | 0.93 | 0.93 |
| YNT_SOIL | -0.39 | -0.19 | -0.22 | 0.35 | -9.91 | -31.01 | -1.00 | -5.04 | -6.99 | 0.93 | 0.93 | 0.93 |
| YNT_N2KM | 0.25 | 0.58 | 0.67 | 0.79 | -5.72 | -25.33 | -0.57 | -0.62 | 0.67 | 0.93 | 0.93 | 0.93 |
| YNT_GVF | -0.28 | -0.02 | -0.03 | 0.88 | -7.32 | -28.53 | -0.48 | -2.32 | -3.65 | 0.93 | 0.93 | 0.93 |
| YNT_SSNG | -0.56 | -0.32 | -0.38 | -0.84 | -10.46 | -30.32 | -2.17 | -5.62 | -6.06 | 0.93 | 0.93 | 0.93 |
| YNT_SSN | -0.29 | -0.07 | -0.09 | -2.29 | -12.57 | -32.50 | -3.61 | -7.85 | -8.99 | 0.93 | 0.94 | 0.94 |

| | **e)** 2-m Mixing Ratio [g/kg] | | | **f)** 2-m Mixing Ratio [g/kg] | | | **g)** 2-m Mixing Ratio [g/kg] | | | **h)** 2-m Mixing Ratio [g/kg] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simulation | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km |
| AP_XM | 0.38 | 0.64 | 0.60 | 1.67 | 1.80 | 1.70 | | | | 0.90 | 0.90 | 0.89 |
| YNT | 0.19 | 0.00 | -0.20 | -10.98 | -19.87 | -14.86 | 1.48 | 1.44 | 1.45 | 0.90 | 0.90 | 0.90 |
| YNT_SST | 0.20 | 0.00 | -0.20 | -11.76 | -20.42 | -15.62 | -0.88 | -0.69 | -0.90 | 0.90 | 0.90 | 0.90 |
| YNT_SOIL | 0.24 | 0.10 | -0.02 | -11.16 | -20.37 | -16.74 | -0.20 | -0.62 | -2.21 | 0.90 | 0.90 | 0.90 |
| YNT_N2KM | 0.22 | 0.05 | -0.14 | -10.68 | -19.20 | -13.68 | 0.34 | 0.83 | 1.38 | 0.90 | 0.90 | 0.90 |
| YNT_GVF | 0.30 | 0.17 | 0.02 | -11.16 | -19.87 | -15.97 | -0.20 | 0.00 | -1.31 | 0.91 | 0.91 | 0.90 |
| YNT_SSNG | 0.36 | 0.28 | 0.24 | -13.62 | -21.14 | -16.91 | -2.96 | -1.59 | -2.41 | 0.91 | 0.91 | 0.91 |
| YNT_SSN | 0.27 | 0.14 | 0.04 | -12.36 | -21.14 | -17.50 | -1.55 | -1.59 | -3.10 | 0.90 | 0.90 | 0.90 |

| | **i)** 10-m Wind Speed [m/s] | | | **j)** 10-m Wind Speed [m/s] | | | **k)** 10-m Wind Speed [m/s] | | | **l)** 10-m Wind Speed [m/s] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simulation | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km | 12 km | 4 km | 1.3 km |
| AP_XM | -0.02 | -0.22 | -0.23 | 1.51 | 1.50 | 1.62 | | | | 0.71 | 0.71 | 0.64 |
| YNT | 0.45 | 0.34 | 0.36 | 7.10 | 2.46 | -3.26 | 1.61 | 1.54 | 1.57 | 0.69 | 0.70 | 0.69 |
| YNT_SST | 0.46 | 0.34 | 0.36 | 7.37 | 2.80 | -2.34 | 0.25 | 0.32 | 0.95 | 0.69 | 0.70 | 0.69 |
| YNT_SOIL | 0.38 | 0.24 | 0.23 | 5.91 | 1.53 | -4.43 | -1.12 | -0.91 | -1.21 | 0.69 | 0.70 | 0.69 |
| YNT_N2KM | 0.42 | 0.32 | 0.34 | 5.44 | 0.87 | -4.99 | -1.55 | -1.56 | -1.78 | 0.70 | 0.71 | 0.70 |
| YNT_GVF | 0.60 | 0.54 | 0.60 | 11.75 | 8.26 | 4.13 | 4.34 | 5.65 | 7.64 | 0.69 | 0.70 | 0.69 |
| YNT_SSNG | 0.53 | 0.47 | 0.49 | 8.90 | 5.53 | -0.18 | 1.67 | 2.99 | 3.18 | 0.70 | 0.71 | 0.70 |
| YNT_SSN | 0.36 | 0.23 | 0.22 | 4.65 | 0.07 | -6.47 | -2.29 | -2.34 | -3.31 | 0.70 | 0.71 | 0.70 |

| | | | |
|---|---|---|---|
| -1 0 1 | -40 0 40 | -10 0 10 | 0.5 0.75 1.0 |

3. More justifications on the design of model configurations should be added: although the focus of the study is on land initialization/model and nudging, the selections of all the other physics, ICs/BCs, and the distributions of vertical layers (40 layers, is this fine enough?) to study this area should be justified, particularly, are the Noah-related setups based on literature or recommendations from any of their partnering local agency? Also, some extended discussions on ACM2 PBL scheme vs YSU scheme and how they affect the different model runs and conclusions would be very helpful.

***Thank you for your comments. We mention in the revised paper that seven of the 40 vertical layers are located below 2 km. Though the number of vertical layers is relatively small, this choice was made to reduce computational expense due to the large number of simulations, domain size, and simulation length used during this project (including CMAQ simulations described in the Pierce et al. companion paper). The model configurations used during this study, including the number of vertical layers, were determined based on feedback from our partners. We also noted in the original manuscript that: "This particular set of schemes was chosen based on our previous studies showing that they performed well during the warm season across the United States (e.g., Harkey and Holloway 2013; Cintineo et al. 2014; Greenwald et al. 2016; Griffin et al. 2021; Henderson et al. 2021). Because there are dozens of parameterization schemes to choose from in the WRF model, we do not aim to find necessarily the best physics suite but instead to assess the potential of using other schemes***

***to improve upon the performance of the baseline AP-XM configuration." Finally, because multiple parameterization schemes were changed when switching from the AP-XM to YNT baseline simulations, it is impossible to determine how the ACM2 and YSU PBL schemes by themselves affected the simulations. However, to address your comment, we have added a sentence to the final paragraph in Section 3.2.5 (surface energy budget constraints) in the revised manuscript stating: "Though it is not the focus of this research, differences in PBL height between the AP-XM and YNT simulations could be due to differences in vertical mixing strength and entrainment flux in the AMC2 and YSU PBL schemes (Hu et al. 2010)."***

Minor comments:

*Pleim-Xu* should be Pleim-Xiu throughout the paper

***Thank you for noticing this spelling mistake. We have revised the spelling throughout the paper.***

Table 1: *IC/LC* should be IC/BC

***We have revised this as suggested.***

Using SST as the short form of lake surface temperature is a little confusing

***We agree that this can be confusing, however, this is the naming convention that is used for this dataset ([https://coastwatch.glerl.noaa.gov/erddap/griddap/GLSEA_GCS.html](https://coastwatch.glerl.noaa.gov/erddap/griddap/GLSEA_GCS.html)).***

The authors defined soil moisture/soil temperature as SOIL but still use soil moisture and (/) soil temperature in multiple places

***We defined this as "SOIL" in the context of the shortened simulation name (YNT_SOIL) and in the abstract. We prefer to explicitly refer to the soil moisture and soil temperature variables in the rest of the manuscript.***

I think using "evaluation" instead of analysis in many places of this paper would be less confusing

***Thank you for the suggestion. We have changed "analysis" to "evaluation" in various locations (such as the titles of the subsections) to avoid confusion with "analysis nudging" and the input data analyses.***

Abstract is very descriptive and specific to this modeling experiment, rather than delivering messages that could impact a broader audience.

***A sentence was added to the end of the abstract stating that: "These results demonstrate the value of using high-resolution satellite-derived surface datasets in model simulations."***

L208: spell out NLDAS-2

***This acronym was already spelled out in the original text.***

Units of Figure 2 differences plot are missing. Text in Figures 7-9 are small.

***Thank you for noticing that the units were missing from the figure caption for Fig. 2. They have been added to the figure in the revised manuscript. We have also increased the font size for this figure, as well as for Figs. 7-9.***