Response to comments by Anonymous Referee #1

We would like to thank the Anonymous Referee #1 for the constructive comments on our manuscript to improve it. We have replied to each point in turn below and will consider these carefully into the revised manuscript. Note that the referee's comments are in bold.

General comment

1. I couldn't see the main points clearly in the abstract. My first suggestion is that the authors should make the main points clear in the abstract (e.g., is KORPFT is better than two others or KORPFT can be used one of the PFTs?).

The abstract will be updated in the revised manuscript with a clear point that KORPFT can be used one of the PFTs to support biogenic emission and ozone air quality modeling. We will keep this point clearly throughout the manuscript. We thank the referee for this suggestion.

2.1. Through the contents, the authors mentioned that the KORPFT is better than two other PFTs because O3 and isoprene mean differences are smaller than the other two (Figure 9), compared with model-averaged O3 and isoprene from three different CMAQ runs. I think that it is hard to say that the KORPFT is better than the others.

We agree with the referee's view. It is hard to say that KORPFT is more accurate (or reliable) than two other PFTs based on the results showing that O3 and isoprene mean differences are smaller than the other two scenarios, compared with model-averaged O3 and isoprene from three different CMAQ runs. The main aspect we wished to address through the contents is that the use of different PFT distribution data need much caution for O3 air quality modeling (or forecasting) in complicated urban atmospheric condition including meteorology and emissions since it can provide temporally and spatially far different O3 prediction results. To support this aspect, we have synthesized and addressed our results showing that the spatial difference of vegetation data of each PFT scenario (e.g., Figure 3 in the original manuscript) affect the results of the MEGAN BVOC emission modeling (e.g., Figure 4) and consequently affect the results of CMAQ O3 predictions (e.g., Figures 8 and 9 in the original manuscript).

To our knowledge, no study has been conducted in any Asian mega city areas with a focus on the impact of different sources of PFT data on biogenic emission and ozone predictions. Therefore, we believe that our results can provide implications to atmospheric research community members who have interest in serious air pollution problems in Asian mega cities.

2.2. Rather, the authors could say that the mean values from the CMAQ including KORPFT

are between those from two others. It is simply caused by that the O3 and isoprene concentrations from CMAQ with MODIS based PFT are larger than those with KORPFT and the concentrations from CDP based CMAQ are smaller than those from those with KORPFT.

We agree to the referee's comment. The mean values from the CMAQ including KORPFT are between those from two others. It is simply caused by that the O3 and isoprene concentrations from CMAQ with MODIS based PFT are larger than those with KORPFT and the concentrations from CDP based CMAQ are smaller than those from those with KORPFT (Table 4). We will add this description into Section 3.4 in the revision.

2.3. I think that Figure 7 suggests that isoprene from the MODIS based simulations are better than the other two regarding on similar R2 values and better slopes. Thus, I want to see how the authors conclude that the KORPFT is better than two others.

We agree with referee. While all of the performance metrics are similar, CMAQ with MODIS scenario produced slightly better predictions than other scenarios. We acknowledged this (line 28, page 24946 – line 1, page 24947).

As we responded to General comment1, we will clearly address the point that KORPFT can be used one of the PFTs to support biogenic emission and ozone air quality modeling. In addition, we will add some discussions about the benefit of using local detailed sources of PFT distribution data for biogenic emission and ozone air quality modeling.

It should be noted that the main objective of the performance evaluation section is not to determine which one of the three different PFT database should be preferred. The main objective of this performance evaluation is to check whether our CMAQ simulations (with different PFT scenarios) can provide us with reasonable O3 concentrations in our study domain although these are a limited number of comparisons. As the result, the CMAQ predictions follow the observations reasonably well (line 22 of page 24945- line 23 of page 24947).

We feel that the title of section 3.4 "Impact of PFT distribution differences on model performance" convey somewhat misleading implication to readers like "There must be great differences between the model performance metrics." Therefore, to avoid misleading implication, the title of the section 3.4 will be altered to "Performance check of CMAQ predictions." In addition, a paragraph in the concluding section (lines 14-17 of page 24955) "Based on the findings and conclusions presented here, we suggest that the use of representative PFT distribution data can provide less biased results in regional or local biogenic emission and photochemical O3 predictions and that other sources of PFT distribution data (e.g., MODIS) can serve as an alternative." will be removed.

In the section 3.4, we described that most monitoring sites (121 sites in 148 sites) were located at urban sites (lines 19-20 of page 24946). Thus, we can expect that the little difference in those performance metrics does not necessarily mean the same impacts of the different PFT distribution scenarios on CMAQ O3 predictions. To support this, we presented noticeable deviations of hourly

CMAQ O3 predictions with different PFT scenarios in our modeling domain (Figure 8 in the original manuscript). We further indicated that different PFT distribution scenario make large differences with increasing ambient temperature (Figure 9 in the original manuscript).

3.1. In the CMAQ system, the NOx concentrations are significantly overpredicted and O3 and isoprene are underpredited compared with the in-situ measurement. With the large overpredictions of NOx concentrations, it is risky to evaluate the O3 sensitivity to the changes in the isoprene emissions.

We agree with the referee.

3.2. As the authors indicated, O3 chemistry is complicated and the ratio of NOx/VOC really matters in the O3 formation. The authors also addressed that the simulated high NOx biases are caused by high NOx emissions.

We thank the referee for this comment and agree with the referee's view.

3.3. Ideally, in the system, in order to see the sensitivities of O3 and isoprene to the different PFTs, the simulated NOx should be fixed by changing the NOx emissions (e.g., using an inverse method or using the ratios of in-situ measured NOx and corresponding simulated NOx concentrations).

Following referee's comment, we have done some O3 sensitivity runs after fixing CMAQ NOx. We fixed CMAQ NOx concentrations using the ratios of in-situ measured NOx and corresponding CMAQ NOx concentrations. These sensitivity works were conducted for June 2008. The results and discussion will be added into the revised manuscript.

Specific comments

1. In abstract, the authors say, "Multiple regression analyses with the different PFT data (delta O3 vs. delta PFTs) suggest that KORPFT can provide reasonable information to the framework of MEGAN biogenic emissions modeling and CTM O3 predictions". What does this mean? Can KORPFT be used like other PFTs or is KORPFT better than two others. Make it clear.

The abstract will be updated in the revised manuscript. We will clearly point that KORPFT can be used like other PFTs. As we responded to the referee #2, we have re-conducted spatial regression analysis with a consideration of the spatial dependence of dependent variables (O3 concentration changes) and the spatial collinearity between explanatory variables (PFT area changes). The results of this new analysis will be added into the revised manuscript.

2.1. In abstract, the authors says, "Exponentially diverging hourly BVOC emissions and O3 concentrations with increasing ambient temperature suggest that the use of representative PFT distributions becomes more critical for O3 air quality modeling (or forecasting) in support of air quality decision-making and human health study". As the author addressed, three different PFT make large differences with increasing ambient temperature. Can the author make another scatter plot like Figure 7 for O3.

As the referee requested, a scatter plot for O3 is shown in Figure 1. O3 derived from CMAQ with three different PFT scenarios show similar R^2 values and slopes.

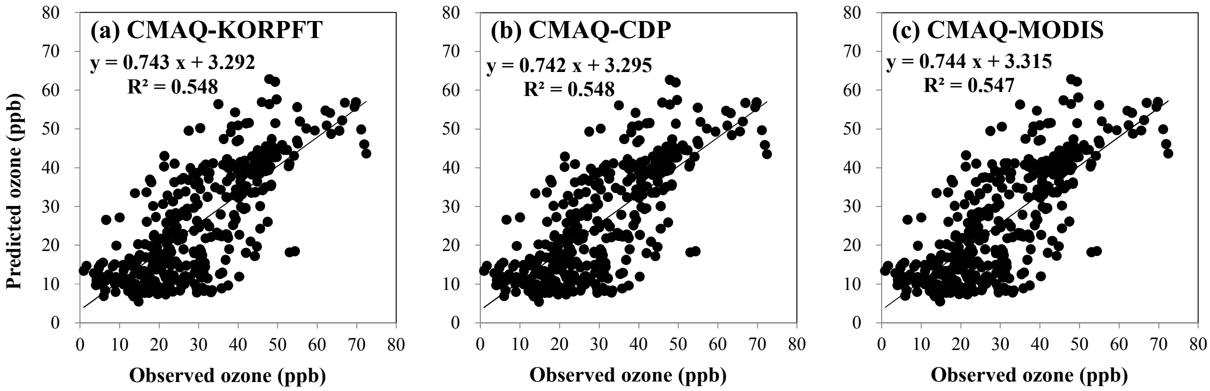

Fig 1. Scatter plots of observed values versus CMAQ predicted ozone for each BVOC emission scenario for the period of May–June 2008.

2.2. Again, to me, from Figure 7, the MODIS-based PFT looks better than two others for simulating surface isoprene.

We agree with the referee. Please refer to our responses to General comment 2.3 and Specific comment 4.

3.1. Page 24946, The authors say, "The over-prediction of NOx concentrations is primarily due to the overestimation of anthropogenic NOx emissions, and the under-prediction of isoprene concentrations is due to the combined effects of the overestimations in NOx and underestimations VOC and ambient temperature (i.e., under-predictions)". Probably, the authors could make better NOx concentrations by changing NOx emissions using the ratio of in-situ measured and corresponding NOx concentrations.

We thank the referee for this comment and agree with the referee's view. Please see our response to General comments 3.3.

3.2. After the NOx simulations are fixed then, the authors could do some sensitivity of isoprene and O3 to the changes in the PFT. I think that it is critical for this study, but it depends upon how the authors feel about this suggestion.

We accepted the referee's suggestion (Please see the response to General comments 3.3).

4. Page 24948, the authors say, "Among the three CMAQ isoprene results, the CMAQ provided values closer to the observations with MODIS (MB = -0.02 ppb and NMB =-7.78 %) than with the others (MB = -0.05 and NMB = -22.62% with KORPFT; MB =-0.08 and NMB = -32.82% with CDP). However, the CMAQ shows noticeably better performance for isoprene time variation with KORPFT (r = 0.622) than with the others (r = 0.598 with MODIS and r = 0.591 with CDP) (Fig. 7)." How do you conclude that the KORPFT CMAQ shows noticeably better performance for isoprene? Is it because the r value from the KORPFT has higher than others?

We agree to the Referee's point and will reword this sentence as below:

"Among the three CMAQ isoprene results, the CMAQ provided values closer to the observations with MODIS (MB = -0.02 ppb and NMB =-7.78 %) than with the others (MB = -0.05 and NMB = -22.62% with KORPFT; MB =-0.08 and NMB = -32.82% with CDP). Meanwhile, in terms of time change, the CMAQ provided slightly higher r values with KORPFT (r = 0.622) than with the others (r = 0.598 with MODIS and r = 0.591 with CDP)."

5.1. Page 24929, the authors say, "Thirdly, we changed only PFT datasets without changing any other model input configurations, such as LAI, meteorological or chemical variables, in order to isolate the impacts of the different PFTs on atmospheric chemistry (or O3 concentrations)." The MEGAN computes emissions for plant functional types as a function of temperature, solar radiation, leaf area index (LAI), and leaf age, which means that PFTs are dependent upon the LAI data. For example, Pfister et al. (2008) used three different sets of LAI and PFT input data.

We agree with the referee.

5.2. The authors need to justify how they use three different PFT values with the same MODIS LAIs.

We thank the referee for this comment. To justify how we use three different PFT values with the same MODIS LAIs, we will add more description to the section 2.2.2 in the revised manuscript as followed:

“MEGAN uses an approach that divides the surface of each grid cell into different PFTs and non-vegetated surface (Guenther et al., 2006). MEGAN use LAIv to simulate the seasonal variations in leaf biomass and age distribution rather than use LAI values (Guenther and Sakulyanontvittaya, 2011; Guenther et al., 2006). MEGAN assumes that plant leaves cover only that part of the grid cell containing vegetation (Guenther et al., 2006). Thus, the LAIv calculations can be performed with LAI at only the grid cells with PFT values. LAIv is different from LAI in that it is estimated by dividing the grid average LAI by the vegetation covered fraction. The upper limit of LAIv is set up 6 to eliminate the very high values that can be estimated for grids with very little vegetation (Guenther et al., 2006). Guenther and Sakulyanontvittaya (2011) suggested the two main reasons of using LAIv in MEGAN. First, LAIv is the actual LAI of the canopy, thus more appropriate input for a canopy environment model. Second, lower- and upper- bounds can be placed on LAIv since vegetation covered areas rarely have a maximum LAI of less than 0.1 or more than 10.

In our study, we adopted three different sources of PFT data (i.e., KORPFT, CDP, and MODIS) and a single source of LAI data (i.e., MODIS LAI (1km × 1km resolution)). Raw MODIS LAI values go 0 to 10. Following the definition of LAIv in MEGAN literatures (e.g., Guenther and Sakulyanontvittaya, 2011; Guenther et al., 2006), the MODIS LAI values were divided by PFT covered fraction at each grid cell (3km×3km size) and converted to LAIv values (e.g., 3km×3km zonal averaged LAI values / 3km×3km zonal averaged PFT fractional values). The LAIv calculations were performed for three different vegetation datasets (i.e., MODIS LAI-KORPFT, MODIS LAI-CDP, and MODIS LAI-MODIS PFT) independently. It is important to note that the LAIv are not computed unless a grid cell includes both PFT and LAI values concurrently.”

6. Does using the same MODIS LAI affect the analysis that the authors have performed on?

We assume that the MODIS LAI mentioned here is just raw MODIS LAI (not LAIv) (Please refer to our response to Specific comment 5.2). We haven’t seen any study for the case using the raw MODIS LAI for MEGAN biogenic emission estimation and subsequent CMAQ modeling. It would be another important issue that requires further research. In the future work, we could look into this issue by comparing with the current study.

7.1. In Section 3.1, the authors say, "This artifact (i.e., PFT area missing) can occur in the process of the LAIv calculations due to the geo-locational disparity between the PFT and the LAI distributions". As the author indicated, the artifact such as PFT area missing affect the LAI calculations. Again, the PFT and LAI are closely associated when they are estimated in the preparation of the data.

The referee is right to point out the association between the PFT and LAI in the preparation of LAIv input for MEGAN. We feel that the lines 13-16 of page 24940 "The zones missing biogenic emissions (the squared zones shown with the red-dotted line in Fig. 4a–d) occurred due to missed

PFT area. This artifact (i.e., PFT area missing) can occur in the process of the LAIv calculations due to the geo-locational disparity between the PFT and the LAI distributions "are misleading and should be reworded in the revision.

For the clarity, the terms “PFT area missing” will be altered to just “data missing.” In this study, there are two sources of the data missing:

The first one is PFT data omission (Figure 2 (a)). KORPFT (Figure 2 (a1)) finely well provide the PFT distribution information though it has a little bit of data omission at some grid cells, whereas CDP (Figure 2 (a2)) and MODIS have notable omissions of the PFT distributions at some islands, some costal city areas off the Incheon Metropolitan Area, and some Seoul Metropolitan Areas. While the KORPFT has the PFT omissions at only 11 grid cells, CDP and MODIS have the PFT omissions at 120 and 118 grid cells, respectively. These omissions of PFT distributions can inhibit to compute LAIv.

The second one is LAI data omission (Figure 2 (b)). The MODIS LAI has the LAI value omissions at some islands, some costal city areas, and some Seoul Metropolitan Areas although these areas are covered by vegetation (Figure 2 (a)). The MODIS LAI has the LAI omissions at 285 grid cells. These LAI omissions can inhibit to compute LAIv.

As described above, either PFT or LAI omission can inhibit to compute LAIv. This LAIv computation inhibition (i.e., LAIv missing) causes the biogenic emission missing implying the BVOC reactivity missing. The cause of LAIv missing was different according to the sources of PFT. For the KORPFT, most (about 96%) of LAIv missing were caused by the omitted MODIS LAI data. For the CDP and MODIS, about 60 % of the LAIv missing was caused by the omitted MODIS LAI data and about 40 % of the LAIv missing was by the omitted PFT data.

Therefore, we will reword the lines 13-16 of page 24940 in the revision as below:

"The zones missing biogenic emissions (the squared zones with the red color in Fig. 4a–d) occurred due to data missing problem. In this study, there are two sources of the data missing: PFT and LAI data omissions. Either PFT or LAI data omission can inhibit to compute LAIv and then inhibit to compute biogenic emissions. The cause of the LAIv computation inhibition (i.e., LAIv missing) was different according to the sources of PFT data. For the KORPFT, most (about 96%) of LAIv missing were caused by the omitted MODIS LAI data. For the CDP and MODIS, about 60 % of the LAIv missing was caused by the omitted MODIS LAI data and about 40 % of the LAIv missing was by the omitted PFT data.”

In the original manuscript, we tried to present the effect of the data missing problem (i.e., BVOC emission missing due to LAIv missing) on the CMAQ O3 predictions by applying a regression equation (line 16 of page 24948 – line 9 of page 24950). We feel that this is not sound approach. For the revision, we have carried out CMAQ simulations with new biogenic emissions estimated by MEGAN with new LAIv input data and investigated the effect of data missing on CMAQ O3 predictions. We made the new LAIv input data by filling in each nearest neighbor LAIv value for each missing grid cell. This new result will be added in the revised manuscript.

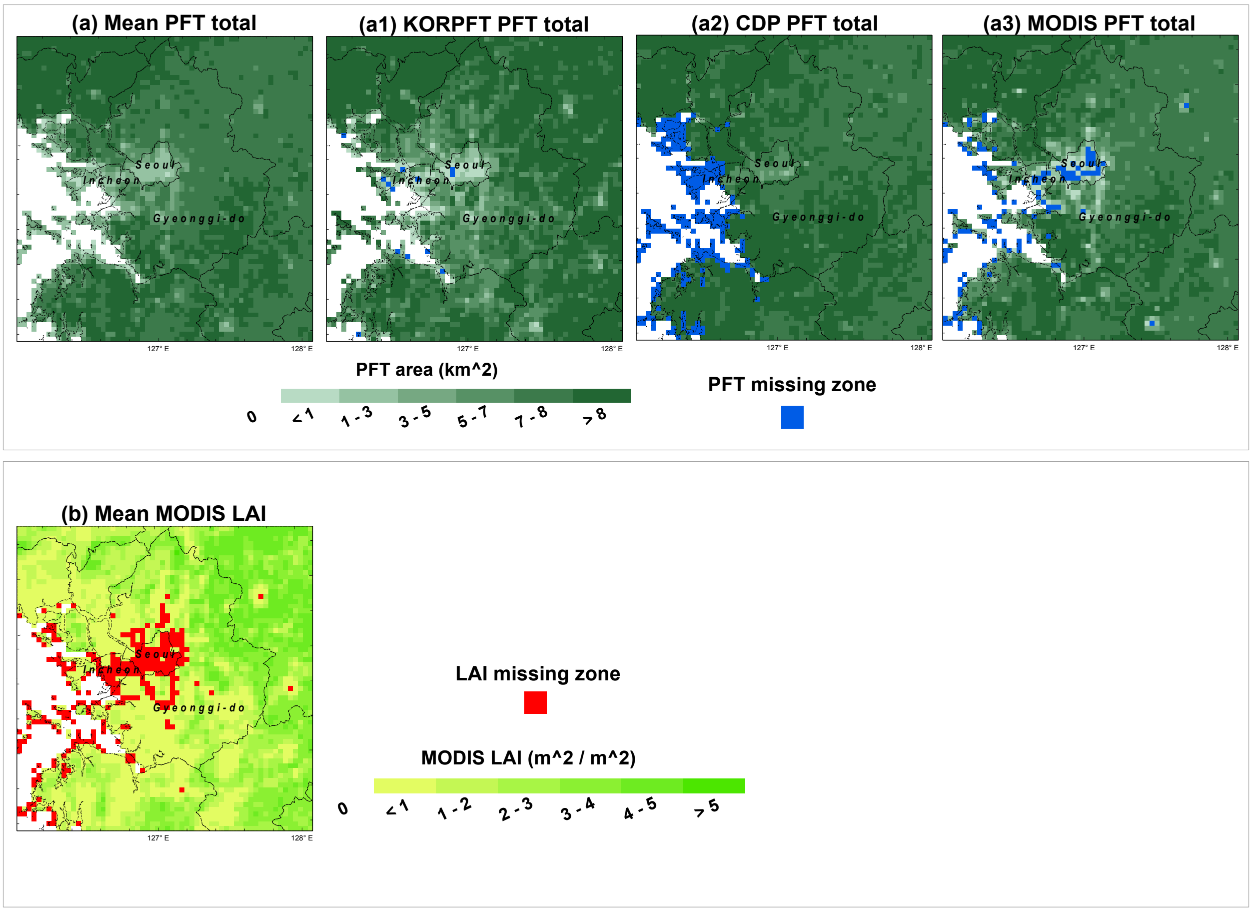

Fig 2. Spatial distribution of the PFT area for each PFT scenario and MODIS LAI. The mean spatial distributions for PFT total (i.e. sum of BT, NT, SB and HB areas) were derived by averaging the three different distribution data sources (i.e., KORPFT, CDP and MODIS). The mean MODIS LAI spatial distributions were derived by averaging the raw MODIS LAI values for the consecutive period of April to July 2008. It should be noted that the MODIS LAI values shown at Fig 2(b) is not the LAIv values.

7.2. The authors need to give some justifications on their approach to change only PFT.

Please refer to our response to Specific comment 5.2

Technical comments

The referee has provided a very helpful list of technical corrections required in the manuscript to improve readability and clarity of our text. Many thanks for the detailed list. We will address these in the revised manuscript.

1. Page 25495, typo ....are very strogn

It will be corrected in the revised manuscript.

2. Page of 24953, typo ....Figure 8 (Figure 9?)

It will be corrected in the revised manuscript.

3. From Figure 4, the red-dotted regions were not clearly shown.

The figure will be fixed for clarity.

4. From Figure 5, the reactive missing region were not clearly shown.

The figure will be fixed for clarity.

5. Figure 9 was never called in the content.

It will be called in the revised manuscript.

References

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

Guenther, A. and Sakulyanontvittaya, T.: Improved Biogenic Emission Inventories across the West Technical Analysis Report, 2011.