Answer to Referee #1

(General Comments)

It is not clear to me, why only choose those periods because it should be based on the overall general or a year-round period to secure the generality of characterization. In this sense, this conclusion needs to be very carefully characterized and the re-analysised according to the separate specific events and their differing conditions.

There are several choices how to select models and seasons for the discussion of discrepancies between models and observation in the marine O_3 under the framework of MICS-Asia III project. As for the selection of the month, according to Figure 3 of the overview paper for ozone in the MICS-Asia III (Li, J. et al., 2019), the substantial overestimate of surface O_3 in EA3 region covering the Northwest Pacific and the Sea of Japan, can be seen only in summer and early september by most of the models. This is the reason we selected only July as a representative month of summer for the discussion in this paper. The relative importance of long-range transport, in-situ photochemistry and dry deposition to ocean surface contributing to the overestimate would be different by months during the season, quantitative discussion of the cause of discrepancies among the models is out of the scope of this paper. A short notes relevant to the selection of only July have been added in Line 103-107.

(Major Comments)

(1) Under- or over-estimation by model is presumably caused by very complicated factors in the regional CTM models. Authors employed two models with different versions of CMAQ. I thought it may make sense to compare it to a completely different and diverse models, such as the CAMx or WRF-Chem or GEOS-Chem model. This is because, as we do not know the "true" values of dry deposition velocities, and thus authors should open to the different possibilities of other controlling factors. For example, as depicted in Fig. 8, NAQM showed considerably lower ozone mixing rations over the whole domain than CMAQ, which in turn could derive the relatively lower ozone than CMAQ over Hedo and Ogasawara as well (it would be quite natural to me), rather than pointing out relatively higher dry deposition velocities than CMAQ. As authors are aware, models employ their different dry deposition parameterizations, mostly generating different values.

We agree that inclusion of other models would be more persuasive and useful for the discussion of this paper. However, the reason we selected only the two CMAQ and NAQM is as follows. As seen in Figure 7 of the overview paper for ozone (Li, J. et al., 2019), there are large variability of surface O_3 spatial distribution in the targeted marine region in summer

among the 13 models submitted to MICS-Asia III. Even among the six WRF-CMAQ, and the three WRF-Chem models, the variability is substantial and it is not possible to select representative WRF-CMAQ or representative WRF-Chem. Since the purpose of our papers in this special issue is to elucidate the main element causing the discrepancies, rather than to make quantitative discussion of the cause of whole variability among the models, we selected only limited number of models for discussion. More practical reason that we selected two CMAQs and NAQM is that we could use the hourly data of concentrations of O_3 and NO_x, and process analysis of net chemical O_3 production only for these models. Since the submission of hourly data and process data were not specifically requested by the MICS-Asia III project, such data were not submitted from other models.

Although NAQM gives relatively lower O_3 values than CMAQ over the whole region in East Asia as shown in Fig. 7, we think the cause of lower O_3 is different between over land with high photochemical O_3 production and over ocean where such activity is minimal. The cause of lower O_3 in the heavily polluted region giving better agreement with the observation have already been discussed by our previous paper (Akimoto et al., 2019) in this special issue. We think the lower O_3 in the oceanic region is caused by quite different reason as we discussed in this paper.

(2) As authors described in page 7 (Long range transport of O3), Oki is influenced by O3 inland area (and thus mostly urban O3). Probably CMAQ which usually simulate "higher-than observation" ozone mixing ratios over inland, and CMAQ model improvement over "land" area may automatically remove ozone biases in "marine" area where authors claim that dry deposition velocity should be higher.

Yes, we think overestimate of O_3 over land by CMAQ definitely contribute to the overestimate of the transported peak on July 6 and 23. The reduction of the overestimate would remove the substantial part of the overestimate on these days. A sentence has been added in Line 234-236.

(3) (Line 355-375) It is also confusing that, among three models, NAQM reproduce ozone well over Hedo and Ogasawara primarily due to the higher dry deposition velocities than those of CMAQ. However, in the case that dry deposition velocities of Bohai Bay and Yellow sea have to be raised, then Oki will be expected to be also down, but ozone will be also down simultaneously in Hedo and Ogasawara where originally there were no biases of ozone mixing ratios. I guess increasing dry deposition velocities in model NAQM over Bohai Bay and Yellow sea will not satisfy both.

We think the dry deposition velocity over Bohai Bay/Yellow Sea region is much different

from the open Pacific south of Japan, since the former is expected to be much larger than the latter due to the presence of much higher amount of organic materials deposited from the polluted atmosphere to cover the ocean surface. Underestimate of dry deposition velocity of NAQM is expected only over Bohai Bay/Yellow Sea region. Therefore, region-specific dry deposition velocity should be given to the model simulation although the necessary quantitative data is not available at the moment. Therefore, in the Section of Future Research Recommendations, we recommended the measurement of deposition velocity of O_3 in the Northwestern part of the Pacific Ocean particularly over Bohai Bay and Yellow Sea.

(4) Finally I recommend to secure some more simulation cases. Authors used at least global model for initial condition such as GEOS-chem and CHASER: dry deposition velocities from two global models vs. over- (or under-) estimation of ozone mixing ratios can be also useful to justify authors' conclusion from multiple cases. Or in some cases, authors may reach a different conclusion (i.e., model internal errors instead of different dry velocities).

The dry deposition velocity used in the GEOS-Chem and CHASER has been given in the text, and a comment has been added in Line 364-369.

(Specific comments)

1. *Specify the dry deposition parameterizations (with references) for all two (or three) models.*

   Description on the parameterizations of Wesely (1989) and M3DRY (Pleim et al., 2001) used in the NAQM and CMAQ was added in Line 131-132.

2. *During July 23-26 in Fig. 3(c), observation of ozone in Ogasawara showed strong diurnal variations with big differences between max. in daytime and min. in night time. Should there be a photo-chemical reactions? because Ogasawara is thought to be a real background site: it has nothing to do with both local photo-chemical formation and transport of NOx from other areas I guess.*

   Although Ogasawara is generally thought to be a real background site based on its remote location over the open Pacific ocean, the surroundings of the monitoring station is actually surrounded by trees and the data may be affected by the very local emissions of biogenic VOCs and soil NO_x, which may possibly cause some photochemical activity to form in-situ O_3 under certain meteorological conditions. Such a note has been added in Line 207-213.

3. *Line 230. --- three models reproduce observation reasonably well.. It is confusing because, in the previous sentences, only NAQM's results matches well with observations in Ogasawara.*

Here the "increment" of transported O_3 is reproduced well in all the three models. Overestimate of by the CMAQ is mostly caused by the overestimate of "background" O_3 and resulted the overall better matching of NAQM at Ogasawara (Line 248).

4. Please indicate the locations of Bohai Bay and Yellow sea areas.

The location of Bohai Bay and Yellow Sea has been added in Fig.1.

5. is it possible to analyze the case different periods like May or June where there are high ozone mixing ratios with (or without Long-range transport processes) over East Asia?, because just one single month test might have a possibility of sometimes misleading the conclusions.

As noted in the answer to the general comments at the top of this document, substantial overestimate of surface O_3 in EA3 region covering the Northwest Pacific and the Sea of Japan, can be seen only from July to September. In May there is excellent agreement of monthly averaged mixing ratio of O_3 and there is much less overestimate by the model in June than July. We agree that analysis of these months may yield different contribution of the transport, in-situ photochemistry and dry deposition to each site, we wish to limit the purpose of this paper to identify these three elements as possible cause of discrepancies between the model simulations and observation.

Over

Answer to Referee #2

(Major Comment)

1. *What're the major differences between CMAQ 5.02 and 4.7.1? If they're very similar in lots of aspects, what's the point to compare these "two" models? It would be much more interesting to compare it with another "independent" model.*

   As a result, the difference between CMAQ 4.7.1 and 5.0.2 was not very much for the targeted Marine region although substantial differences were found in the simulated NO_x mixing ratios and net O_3 production at Oki and Hedo as shown in Figs. 5and 6. Since we found that the different version of two CMAQ models gave substantially different results for heavily polluted land area over Beijing and Tokyo (Akimoto et al., 2019 in this special issue), we are interested in how much difference can be seen between the two different versions of CMAQ over oceanic region. Practical reason that we selected only NAQM and two CMAQ models in this paper is that we could use the submitted hourly data of concentrations of O_3 and NO_x, and process analysis of net chemical O_3 production only for these models. Since the submission of hourly data and process data were not specifically requested by the MICS-Asia III project, such data were not submitted from other models for comparison.

2. *I would suggest a detailed analysis of O3 budgets, including chemical production/ loss terms and physical removal terms of O3, at the sites, which might be also helpful for understanding the model overestimates. For example, underestimates of chemical loss of O3 due to the halogen chemistry in the models would also contribute to the overestimates of O3 at these oceanic sites.*

   Unfortunately, the detailed terms consisting of net chemical production, production and loss terms, were not available under this project, and the direct comparison between the chemical and physical loss terms could not be made. We agree that the additional discussion of the contribution of the gas-phase halogen chemistry is worthwhile to be included in the paper. A description was added at the end of **3.4** in Lines 410-412.

3. *One major conclusion in this paper is that the overestimate in O3 over the study region in CMAQ is due to too small O3 dry depositions whereas the better agreement between NAQM and observation might have resulted from a relatively larger O3 dry dep. It would be interesting to conduct a few sensitivity model simulations, e.g., applying higher O3 dry deposition rates in CMAQ or implementing the dry deposition that used in NAQM into CMAQ, to support the authors' major conclusions.*

Although we agree that such sensitivity analysis is very useful to strengthen our conclusion, it was beyond the activity of MICS-Asia III project and we could not perform this activity (Line 407-409).

(Minor comments)

1. *Line 69: I don't think the present paper aims to "solve" the discrepancies between models and observations. It may provide some useful information on advancing the current understandings of discrepancies between the two models and between modeled and observed O3 at the sites.*

   We agreed and modified the sentence in Line 56-58.

2. *Line 102-105: It would be good to elaborate a little bit here to illustrate how important the role of air-sea exchange plays instead of just citing the papers.*

   According to the suggestion, a paragraph was revised as in Line 93-102.

3. *Line 106: It's not clear to me that a "more detailed comparison" is compared to which study?*

   "more detailed" has been deleted.

4. *Does JST mean Japanese local time? Need to clarify.*

   "Japan Standard Time" has been added.

5. *unit of O3 needs to be consistent: in some places, it's ppb but others as ppbv.*

   All have been unified to "ppbv"

6. *Fig 1: Overlapping monthly winds in the figure may help readers better understand the transport patterns of O3 in the region.*

   We added the wind vectors to Fig.1.

7. *Line 197: I would suggest the authors to avoid stating the model/models "excellently" capture observations. It seems that NAQM still overestimates observed O3 by ~30-40% during nighttime.*

   "excellently" has been deleted.

8. Line 229: "… along with the edge of the Pacific High": it would be good to add isopiestic lines or wind patterns in Fig 4. It would be better to add the spatial pattern of O3 on this day from NAQM as well.

"along with the edge of the Pacific High" has been deleted. Spatial distribution of surface O_3 by NAQM has been added in Fig. 4 as 4(c).

9. Line 281: I'm wondering if this soil NOx from sugar cane fields had been considering any of these models?

MIX emission inventory used for the modeling considers NO_x emission from agricultural sources without specifying crop species. The sentence has been added in Line 288-290.

10. Line 282: It's not clear to me what the term "diurnal O3 formation" means.

"diurnal O_3 formation" has been changed to "daytime O_3 peaks".

11. Fig 5: I would suggest the authors plot the comparisons of observed and simulated NO2 in one plot and the comparisons of NO in the other. That would be more consistent with the analyses in other similar figures and clearer for NO values.

Figures for NO_2 and NO have been separated in Fig. 5.

12. Fig 6: I would suggest plotting the comparisons of modeled chemical production between three models in the same figure.

Fig 6 has been modified according to the suggestion.

13. Fig 7: What're the simulated values from the other two models like at these two sites?

Fig. 7 has been modified including the data for CMAQ 5.0.2 and NAQM according to the suggestion and combined to Fig. 6 as Fig. 6(b) and 6(c).

14. Line 419: ": : :identified as the cause of overestimate": I would say the three factors are possible causes of models overestimates in O3 since there could be other possible reasons.

"cause" has been changed to "possible cause".

15. Line 422: May need to clarify the regions or sites that were suffered from overestimates in transported O3 from continents. I'm also curious about whether the transport patterns of O3

over the study region are significantly different in June or August from that in July.

"at Oki in July" has been specified in the sentence (Line 456). From our former experience of study (Pochanart et al., Atmos. Environ, 36, 4235–4250, 2002), frequency of the transport of continental outflow is higher in June, followed by August and July is the least at Oki and Ogasawara during summer.

over

Answer to Referee #3

(General Remarks)

The authors proposed couple of hypotheses, but did not show solid evidence to support them. First, the current manuscript used the surface observations (O3, NO2, etc.) and modeled surface air pollutants and dry deposition velocity for LRT effects. The LRT is usually dominated by the transport in the free troposphere, and the PBL dynamics (the buildup of PBL in the early morning and the downward mixing of ozone and its precursors to the surface). So it is hard to make this conclusion only using the surface data and model results. The revised manuscript may investigate the PBL simulations if the authors can find aircraft observations or sounding data.*

It is true that the amount of long-range transport is affected by the mixing ratios of both in the boundary layer and free troposphere. It is generally analyzed by using three-dimensional trajectory analysis (Akimoto et al., 1996; Pochanart et al., 2002). However, the assessment of quantitative contribution of each is beyond the scope of this study, and no aircraft campaign data of vertical distribution of O_3 is available in the targeted period of summer in 2010 in this area. A note has been added in the text in Line 384-389.

Second, the paper did not state how the net in-site photochemical production rate was calculated. From the observation at Oki, the peak NO2 or NO2 on July 2 is not associated with high ozone levels, which did not support the hypothesis of local photochemical production.*

In the model, net photochemical production of O_3 was calculated by the difference between model-calculated O_3 formation terms and O_3 loss terms by chemical reactions. As for the NO_2^* peak on July 2 at Oki appeared in the observation, it is apparently due to accidental local perturbation near the monitoring site or due to malfunction of the instrument, since the peak appeared only in the one-hour averaged data at 7 am (Japan Standard Time). Such a note has been added in the figure caption.

Third, the NAQS model did show good results simulating the ozone concentrations at these marine observation sites. However, NAQS tends to have consistent low bias as compared with two CMAQ simulations. If NAQS has significant underestimation of ozone levels in the source regions, we cannot conclude that NAQS has better model performance in simulating marine ozone concentrations.

NAQM gave substantially lower O_3 than CMAQ agreeing reasonably well with observation in the polluted source region over land according to our previous paper (Akimoto et al., 2019). We think, however, the lower continental O_3 does not contribute much to the lower O_3 at Hedo

and Okinawa in July, since the frequency of long-range transport of continental airmass to these sites is rare in this month. Meanwhile, as noted just below, inclusion of the gas-phase halogen reactions in the model would decrease surface O_3, and NAQM would underestimate the mixing ratio of monthly averaged O_3 in this region (see Lines 392-409).

Lastly, the CMAQ model was developed by EPA to regulate the air pollution mainly over the land. So for the two versions used in this study, complex air-sea interactions and halogen chemistry are not included. So it is not surprised to see CMAQ has poor performance here. In summary, the current manuscript shows some results but lacks further discussion or analysis.

We added the discussion of the possible reduction of marine O_3 by inclusion of the gas-phase chemistry of halogens into the models (Lines 392-409).

(Detailed Remarks/Suggestions for Revision)

Line 82: Please define 'NOz' here

Here, 'NO_z' has been changed to 'NO_y' (Line 71).

Line 125: This paragraph discussed the set-up of these 3 CTMs, and WRF was used to generate meteorological fields. But there is no information about the WRF simulations, such as the physical options and if observation/analysis nudging was used, which are important for ozone transportation and deposition. I also cannot find these details in Akimoto et al., 2019 ACP paper. Different configuration of WRF could influence the regional CTMs simulations. The authors need to add explanations in the revised manuscript.

Some explanation about WRF has been added in Line 128-131.

Line 195: I believe the 'transport amplitude' means the ozone enhancements due to the LRT. If yes, please revise this sentence to make it clear.

We changed the 'transport amplitude' to 'amount of O_3 increase due to the transport' (Line 206).

Line 211-214: As raised above, did these WRF runs generate consistent circulation patterns for these two episodes? Figure 4 shows the results from CMAQ v4.7.1, how about the ozone contours in CMAQ v5.0.2 and NAQS? Please include figures similar as Figure 4 in the supplementary material.

We added the similar figures for CMAQ5.0.2 and NAQM as Figs 4(b) and (c).

Line 239: This statement needs further analysis to support it. I agree that if the in-situ photochemical production is important at Oki, the observations should show a similar diurnal cycle which did not exist. Another possible explanation is that the LRT occurred usually in the free troposphere, and the downward mixing due to the PBL build-up can cause the same diurnal cycle I suggest the authors to examine the vertical profiles of ozone from CTMs over Oki to rule out this possibility.

We confirmed that since O_3 does not accumulate in the upper MBL during nighttime being different from over land where near surface O_3 is depleted by NO titration and high deposition to ground during nighttime, the downward mixing after sunrise doesn't cause O_3 buildup in the morning. Therefore, we think the observed diurnal variation of O_3 is mainly due to in-situ photochemistry rather than vertical mixing.

Line 258-260: The observations in Fig 5a shows high NO2 concentrations around 07/02. However, I didn't see significant enhancement in ozone at Oki in Fig. 3a on the same day. The net photochemical production of ozone should be anticipated if the NO2 levels are higher. Need some explanation or discussion here.*

As stated in the answer to the second comments of General Remarks, we think the observational NO_2^* peak in the morning of July 2 at Oki, is apparently due to artifact, either by accidental local perturbation near the monitoring site or malfunction of the instrument. Such a note has been added in the figure caption.

Line 270: How these hourly net chemical ozone production rates are calculated?

As answered to the second comments in the General Remarks, the net photochemical production of O_3 was calculated in the model by the difference between model-calculated O_3 formation terms and O_3 loss terms by chemical reactions.

Line 276: Should be 'in-situ photochemical ozone production in the CTMs, which contributes to the overestimate : : :'

The sentence has been changed as suggested (Line 298-300).

Line 290: Better to show similar figures such as Fig 5 for Ogasawara site in the supplementary material to support this statement.

We added the similar figures for Ogasawara in Fig. 5S-2 in the supplementary material.

Line 298: I am surprised that NAQS predicted much lower ozone concentrations compared with two CMAQ simulations. Especially for the Pearl River Delta (PRD) region, NAQS predicted extremely low monthly mean ozone, as low as 10-20 ppbv, for July. Same phenomenons are found in the Yangtz River Delta (YRD), Wuhan, Seol, and Tokyo. So the good performance of NAQS simulating marine ozone at these 3 marine sites could not be that the model successfully captured the nature, but NAQS has a systematic low bias for surface ozone. If that is the case, why select this model run? The Akimoto et al., 2019 ACP paper listed 12 regional model simulations for the MICS-Asia III project, and some WRF-Chem simulations should be introduced here.

We selected NAQM not as a model reproducing well the observational data, but from the practical reason that we could use the submitted hourly data of concentrations of O_3 and NO_x, and process analysis of net chemical O_3 production along with two CMAQ models. Since the submission of hourly data and process data were not specifically requested by the MICS-Asia III project, such data were not submitted from other models. Although we agree that comparison including WRF-Chem and other models would strengthen our discussion, we limited ourselves within the framework of MICS-Asia III.

It is true that Fig. 8 shows NAQM gives lower O_3 in most of the region in East Asia, However, as shown in our previous paper (Akimoto et al., 2019), NAQM reproduced well the O_3 and NO_x in megacity areas in Beijing and Tokyo in contrast to the overestimate of CMAQ 4.7.1 and 5.0.2. This paper has suggested that the lower mixing ratios of marine O_3 in Northwestern Pacific would be due to the higher dry deposition velocity of O_3 over oceanic water. Thus, the cause of lower O_3 of NAQM would be different by region. It would be worthwhile to elucidate the cause of lower O_3 by NAQM in PRD and YRD region, but it is beyond the scope of this paper.

Line 362-364: Not sure about NAQS. CMAQ models did not include halogen chemistry until version 5.2. So I am not surprised that the halogen chemistry did not impact the dry deposition of ozone over Bohai Bay and the Yellow Sea.

As responded to the forth comments in the General Remarks and also to the comment of Reviewer #2, we agree that the additional discussion of the contribution of the gas-phase halogen chemistry is worthwhile to be included in the paper. A description was added at the end of **3.4** in Lines 392-409.

Line 368: Any observations support that the statement that the ozone concentrations in the Bohai Bay and Yellow Sea are overestimated?

Unfortunately, as far as we know, there is no reported data of the mixing ratios of O_3 over Bohai Bay and Yellow Sea. We recommended such observation in Future Research Recommendation (Line 421, 429-430)

Line 370: Which one? Do NAQS and CMAQ have similar sensitivity to water surface resistance?

Here we described general statement not for the specific model. We omitted the sentence and replaced by a new the statement (Line 389-391).

Line 426: I disagree with this argument. The LRT of ozone in the free troposphere should be more important than the transport near the water body. The KORUS-AQ campaign results in 2016 support this hypothesis. In my opinion, the underestimate of ozone deposition could only impact the surface ozone levels.

Since we didn't make any quantitative analysis of the contribution of marine boundary layer and free tropospheric O_3 to the transported O_3 in the marine region of Northeast Asia, the statement has been modified taking into account the reviewer's comment. (Line 384-389; 459-462)

Figures Figure 4: Consider using different shapes to represent these 3 sites, for the readers who are not familiar with the names.

We changed the symbols of the three sites and gave legend in the figure caption.

Over