General comment

The paper reports a modelling study of the contribution of shipping to NO2 and O3 concentrations in the Mediterranean area. In addition, deposition is included in the analysis. Results of different models are compared among themselves and with measurements in some specific stations. The topic is interesting and the paper generally well written. It has elements of novelty and I believe that it could be published after a revision step necessary to clarify some aspects and put results in a better perspective, see my specific comments.

Specific comments

Lines 31-32. It is not clear this sentence. The other two models use different meteorological input?

- ➔ These two models used the exact same input since they were run by the same institution. We added a sentence saying that the other models use different meteorological input. (*p2l33*)

Lines 41-43. I suggest to mention the recent work of Contini et al (Atmosphere 2021, 12, 92.) that gives a global overview of the effects of shipping on air quality and health.

- ➔ Thanks for this, we included it (*p2l46*)

Line 140. Please remove etc. if authors want to add something it is better to do it explicitly.

- ➔ Removed

Line 243. Actually, looking at the map in Fig. 1 it seems that it is included also the major part of eastern Mediterranean.

- ➔ The description is adjusted (*p10l275*)

Section 2.5. It should be mentioned how these stations have been chosen and if a certain threshold of distance from the coast or from the main routes of ships. This because it is known that the impact of the emissions from ships to air quality is strongly depending from the distance from the harbours/routes and I see some stations that are quite inland, especially in Northern Italy. A discussion on this should be provided even because I believe that the impact of shipping on such stations would be really small.

- ➔ No detailed distance, but we included stations with a distance of < 30 km from the coast (*p14l347*)
- ➔ Yes, the impact is small at inland stations, but nevertheless it makes sense to also include some stations inland to check the model performance (*p14l348*)

Another aspect that should be clarified and it is partially correlated to the previous point is if the emission dataset used include emissions of ships at berth. Several studies indicated that in EU harbours the emission at berth lead to the majority of the impact on local air quality in port cities, see for example Merico et al. (Atmospheric Environment 139 (2016) 1e10). Considering the use of low sulphur fuels at berth, this phase is particularly relevant for nitrogen oxides and could also lead to local exceedances of air quality standards. If neglected it could be present an underestimation of the impacts.

- ➔ The STEAM dataset also includes the ships at berth
- ➔ But there was no separate evaluation of harbor cities in the present study, it will be nevertheless be a part of the project to investigate harbors in city scale models

Line 336. A correlation with R=0.06 is not weak, rather it is a total absence of correlation.

- ➔ It is changed to “no to weak correlation” in the description
- ➔ It is a slightly higher value now due to new calculations which were done based on adjusted domains.
  (*p15l379*)

Lines 410-415. I believe that the results here are also comparable with those obtained with CAMx in the central/eastern part of Mediterranean area reported in Merico et al (Transportation Research Part D 50 (2017) 431–445).

- ➔ Thanks for this hint, this study is good for comparing and we now make reference of it
  (*p20l472*)

Lines 420-421. I would add or near the harbours.

- ➔ This additional information is added to the sentence
  (*p20l481*)

Figure 6. Please use the apex for m3 as in the other figures.

- ➔ The figures are adjusted
  (*Figure 3; Figure 7*)

Line 446. The absence of negative values in the tagging method is a consequence of how the method is formulated rather than a relevant result. Could this lead to problems in evaluation titration of O3?

- ➔ Tagging is no longer considered in the present study: we decided to leave this part out due to the length of the manuscript plus because it is a separate method. Furthermore, it was not goal of this study to compare the different methods.

Section 3.1.5. Regarding O3. There are several experimental evidences, some of them also in the papers that I already mentioned in my previous points, that emission of NO from ships could lead to a local reduction of O3 concentrations, especially in the spring/summer period. At larger distances instead there could be an increase. There are also some hypothesis that models could catch this behaviour more or less efficiently according to the spatial resolution of simulations. Could this be an issue in your results considering that outcomes ranging from negative to positive impacts for O3 were observed? Also Figure 16 shows that relevant differences are observed especially in the negative part.

- ➔ Yes, that explains the differences. I added these information and explained in more detail the local reduction of O_3
  (*p28l589*)

Line 656. I would not say air pollution considering that only NO2 and O3 are considered in this work.

- ➔ Right, this can be misleading. It is changed to “to air pollution by NO_2 and O_3”
  (*p42l741*)

Lines 714-719. This part is a little vague. It could be useful to understand if there is any possibility to understand if one of the model performs better than the other. In addition, it should be mentioned how to use results from the different models with different resolution, averaging the results?

- ➔ Goal of the present study was not to show if some model would perform better than another. We tried to show 1) that it is more reliable to include several model in data evaluation and 2) the range of potential impacts of ships
- ➔ Yes, we averaged the results in form of the boxplots (*Figure 3; Figure 7*). The resolution of 12 km and 0.1° is very similar.