Response to Reviewer #2 comments

*For clarity, our response to the reviewer's comment is in **purple** font*

Review of the preprint "Measurement Report: The effects of SECA regulations on the atmospheric SO2 concentrations in the Baltic Sea, based on long-term observations at the Finnish Utö Island" by A. Maragkidou et al., submitted to Atmospheric Chemistry and Physics

The manuscript describes a long term observation of air pollutants on the Finnish Utö island. The analysis is focused on SO2 concentrations in order to demonstrate the effects of SECA regulations to reduce sulphur emissions from shipping. It is an interesting paper that deserves publication, but improvements and clarifications are needed.

Major general comments:

It isn't exactly clear how you use the STEAM data in your analysis. Please include it in the data interpretation or leave it out.
Reply: As requested, we have added a couple of lines to subsection 3.3, in which we describe how STEAM data was used in the analysis.

I am not convinced that we can learn something from the RoRo ship case study (section 4.5). Please explain this better in the paper or remove this part.
Reply: We thank the reviewer for his feedback. Therefore, after careful consideration and per his request, we decided to remove subsection 4.5.

Please include a paragraph/section about the limitations of your study and about the uncertainties. This should include a discussion about the representativity of the observations for a larger region. Can we really say something about the compliance to the SECA rules in the Baltic Sea when we have approx. 8 ships passing by per day?
Reply: Firstly, we would like to inform the reviewer that Figure 4 was revised to include more information on the number of ships that pass Utö on a daily and yearly basis, with many of these vessels passing multiple times per day. Moreover, the purpose of this study was not to assess the compliance of ships with SECA regulations, but to assess the effectiveness and effects of the impact of SECA regulations on SO_2 concentrations in the Baltic Sea. We recognize that this part wasn't clear in the introduction, therefore we revised the introduction to reflect that point.
Regarding uncertainties, given the nearly two-decade observation period, several factors could introduce uncertainties in the results. Instrumentation for all variables changed over time, and quality assurance methods improved. Additionally, there were changes in the personnel responsible for maintaining the instruments and ensuring data quality. Although standard protocols and measurement diaries were followed, these transitions may have affected the results.
While our observations are important for long–term, high time resolution air quality observations at remote marine research stations, in the vicinity of a heavily trafficked ship lane, as already mentioned in the manuscript, they do not directly represent the broader Baltic Sea region.
Per the reviewer's request, we have addressed the uncertainties of this study in a new subsection "4.4 Uncertainties" of the manuscript.
In terms of limitations, our dataset offers unique fine temporal resolution data of air quality observations at a remote marine station near a heavily trafficked shipping lane. However, the data collected from 2003-2005 were recorded at hourly intervals, whereas the data from 2006-2020 were gathered at 1-minute intervals.

Additionally, CO_2 data was not available for the whole period for the same location and therefore, we were unable to estimate fuel sulfur content. Moreover, no local scale dispersion modelling was employed as it was outside of this study scope, However, the dataset of this work could serve as a benchmark for local and regional scale dispersion modelling. Utö itself is a sparsely populated island with minimal land traffic and no significant local sources of NOx and SO2.

Major specific comments:

Line 135-137: Which data was used for the STEAM model (2006 – 2020)? Why was there a change in AIS data source for the RoRo ship in 2016?
Reply: For the case study, we used AIS data from the AIS receiver installed in Utö in 2015. Therefore, for the period 2016-2019 AIS data from AIS receiver at the Utö marine station was used. For earlier years (2013-2015), HELCOM data set was used. Furthermore, AIS data prior to 2013 was not accessible to us. However, per reviewer's request, subsection 4.5 and hence these lines were removed. For the STEAM model, the AIS data used was from HELCOM AIS for the Baltic Sea since 2006.

Line 151/152 and section 4.5: I do not see the purpose of the case study looking at one RoRo ship. What do you want to demonstrate? What can we learn from this?
Reply: We thank the reviewer for his feedback. Therefore, after careful consideration and per his request, we decided to remove subsection 4.5.

Line 164/165: What is the data source for this graph?
Reply: The data source for Figure 2 is EMEP (European Monitoring and Evaluation Programme). The following clarification was also added to the Figure caption "*The source of SOx emissions data for the different sectors is EMEP (European Monitoring and Evaluation Programme)*".

Line 179/180 and Fig. 3: Why are PM2.5 emissions only available with one decimal place?
Reply: We did not provide any numerical values for $PM_{2.5}$ emissions in lines 179-180. The $PM_{2.5}$ emissions shown on the right y-axis of Figure 3 are displayed with one decimal place as the default setting.

Line 180 and Fig. 3: What happens with the NOx emissions in 2017? After a steady decrease until end of 2017 something looks different in the data with an increase in 2018. It think you cannot say that the emissions remained stable. It would also be interesting to see the CO2 emissions from STEAM in order to see the effect of increased ship traffic vs. more efficient fuel use. Besides the sulphur content in ship fuels, the total amount of fuel burned has an influence on SO2 concentrations. This is not much discussed in the paper.
Reply: Thank you for your comment. Regarding the NO_x emissions in 2017, we observed a significant reduction—approximately 95%—in the amount of AIS data received starting from 6 November 6 2017 until 31 December 31 2017. This reduction appears to be an issue with AIS data reception rather than with data storage, as the decline is evident in both the HELCOM AIS and global AIS datasets. Therefore, it is unlikely that this discrepancy is related to HELCOM's data storage system. The following statement "*except for between 6 November 2017 and 31 December 2017, when a 95% decrease in AIS data was observed due to an issue with AIS data reception*" was added to the end of this sentence in subsection 4.1: "*The predicted annual NOx emissions during the period between 2006 and 2020 have remained relatively stable*" to explain why the significant reduction that was observed during that period.

Fig. 5, b): Why does the median for PM2.5 reach 0 between 2016 and 2018? It looks like the median and the mean are lower in 2015-2017 and then go up again in 2018. Is this connected to a change in instruments?

Reply: Figures 5a-d were revised by filtering out extreme peaks and negative values identified as invalid data, likely caused by changes in instrumentation and improvements in quality assurance methods over time. $PM_{2.5}$ specifically had some negative values that were removed in the revised version of Fig. 5b, resulting in gaps in the time series. While all changes followed standard protocols and were documented in measurement diaries, shifts in equipment and personnel may have introduced some variability. We have carefully accounted for these factors and performed additional quality checks to ensure the integrity of the data presented. Moreover, the following sentence was added to subsection 4.2, below Table 2: "*$PM_{2.5}$ concentrations also had some negative values, which were removed, leading to gaps in the time series (Fig. 5b)*." to justify the gaps observed during some periods.

Fig. 5, c): It looks like NO is lower since 2014, this holds in particular for the median. This is also visible in the data in the appendix and it is in contrast to what you say in the text. Can you comment on it? Is it because of the change in instruments? Could you please elaborate on the effects of instrument changes in general?
Reply: As previously mentioned, Figures 5a-d were revised by filtering out extreme peaks and negative values, which were identified as invalid data, likely due to changes in instrumentation and the changes in quality assurance protocols over time.
Specifically for NO (and NO_x), data from the period 22.5.2010 to 15.6.2011 were excluded in the revised version of Fig. 5c (as well as of Fig. 5d) due to abnormally low values, potentially caused by overly strict data processing during that time. Additionally, since the medians for hourly data (2003-2005) are not directly comparable to the medians for 1-minute data (2006-2020), we have removed the medians for the 2003-2005 period from Tables 2 and Tables A1-A2 and A4 to ensure consistency and accuracy in the data presented.
The following sentence: "*However, for NO and NO_x, a period of data from 22 May 2010 to 15 June 2011 was removed (Figs. 5c and 5d) due to abnormally low values, likely caused by overly strict data processing*." was added to subsection 4.2, below Table 2, to justify the gaps observed during this period.

Fig. 6: Why is there no data in 2011? And could you give some statistical information? How large is the reduction in 2015-2020 compared to the earlier years? And how does this compare to the expected reduction because of more stringent SECA rules.
Reply: As mentioned earlier, there was no NO data during the period 22.5.2010 to 15.6.2011 due to abnormally low values, potentially caused by overly strict data processing during that time. In addition, we didn't employ a chemical dispersion model in our study; therefore, we couldn't calculate and compare the expected reduction resulting from the implementation of stricter SECA regulations.

Line 155 and Fig. 7: Please enhance the figures with some statistical information or use a separate table for this. How does the data compare to that from all wind sectors?
Reply: As requested, we have revised Figure 7 by adding the time series of moving mean and moving median. The figure's caption was accordingly revised, as well.

Line 267: It is not clear why you select 2019 as the year with lower sulphur emissions from shipping. Why not 2016/17/18? Having possible changes in emissions from other sources in mind, 2019 seems to be not the preferred choice.
Reply: In terms of SO_x emissions from shipping, the years 2016, 2017, 2018 and 2019 are not substantially different, as can be seen based on Fig. 2. The reviewer probably refers here to the impacts of COVID-19, which is of course a relevant point. However, many studies have shown that the impacts of COVID-19 on emissions in Europe were not felt yet in 2019; these only affected the subsequent years, since 2020. For instance, the first case of COVID-19 documented in the UK was on 31 January 2020.
We have revised the manuscript to clarify why we selected the year 2019. We have added the following rationale to subsection 4.3, below Figure 7: "*The years 2016, 2017, 2018 and 2019 were not substantially different in terms of SO_x emissions from shipping (cf. Fig. 2) or the number of ships (Fig 4). In this regard,*

any of these years could have been selected as an example year for the post-SECA 2015 analysis. The selected year 2019 was prior to the COVID-19 pandemic and selected for comparison. The pandemic did not yet affect the emissions in Europe in 2019; these effects were felt only during the subsequent years."

Line 273/274 and Fig. 8 a) and b): It seems that SO2 mean and/or median does not depend very much on wind direction. Can you comment on this?

Reply: Thank you for your observation. However, we would like to point out that Figure 8 depicts histograms of SO_2 concentrations, not mean or median values. The radial axis in Fig. 8 represents the number of measured cases for each wind direction sector.

These histograms represent the measured concentrations of SO_2 at the island of Utö, with a time resolution of 1–minute, plotted as a function of wind direction. From Figure 8, it is clear that the dominant wind directions are southwesterly, with relatively high contributions also from northwesterly directions (Figure 8a, Figure 8b and Figure 8c).

Figures 8c-d present only the occurrence of the highest concentrations. It is very clear based on panel c that a vast majority of the highest measured concentrations in 2014 originated from the direction of the shipping lane (which was west of the island).

Moreover, a significant difference is observed between the years 2014 and 2019, suggesting that the implementation of SECA regulations has effectively reduced the number of high concentration peaks (Figure 8c and Figure 8d).

Fig.8 a) and b): The reader has the impression that violet values are 0 – 0.5 µg/m³, yellow values are 0.5 – 1.5 µg/m³, and red values are >1.5 µg/m³. This is in contrast to the legend and the caption.

Reply: The figure caption was slightly revised to clarify what each color represents and how it was plotted. It now reads as follows: *"Figure 8: The measured concentrations of SO_2 at the island of Utö as a function of the wind direction, as polar histograms with 1-minute resolution. The panels a) and c) correspond to the data in 2014 and the panels b) and d) to those in in 2019. These two years were selected to represent the situation before and after the SECA regulation in 2015. The upper panels (a and b) include all datapoints, whereas the lower panels (c and d) include only the data, for which the concentrations were higher than a selected threshold value (5 µg m^{-3}). In panels a) and b), the red colour in the legend represents all data points with values larger than 0 µg m^{-3}. Overlaid on the red, the yellow colour shows data points with values below 1.5 µg m^{-3}, and the violet colour, which overlays the yellow, indicates data points with values below 0.5 µg m^{-3}. The radial axis represents the number of measured cases for each wind direction sector*."

Lines 286-329, Section 4.5: As said before, this section does not provide new insights. You may completely skip it unless you describe better, what is new and what can be learned from it.

Reply: We thank the reviewer for his feedback. Therefore, after careful consideration and per his request, we decided to remove subsection 4.5.

Line 335: The STEAM results are not used very much. You should improve this.

Reply: The purpose of this paper was to present the long-term time series of SO_x measurements collected at a marine monitoring station located near a major shipping lane. While we acknowledge the importance of modeling, our focus was on measurements, with modeling serving to complement and enhance the interpretation of the time series. Specifically, the STEAM data was employed to support the analysis by providing information on ship traffic, such as the number of vessels, and to examine the temporal trends in the measured SO_x concentrations.

Line 339/340: You would underpin this statement with a trend analysis including statistical significance of the trend. It is also in contrast to steady emission reductions in many emission sectors in Europe (Fig. 2). Therefore, you may a few words on this.

Reply: Regarding the concentrations of SO_2, if one would make a statistical trend analysis, it should consider separately the 3 periods, separated by the SECA changes. One should address annual average concentrations, due to the changing meteorological conditions for each year. However, there are far too few annual data points in each of the 3 periods for conducting such a statistical analysis.

This study focused on SO_2 concentrations. A detailed statistical study of the other considered pollutants is outside the scope of this manuscript. An examination of the data shows that there were no substantial trends for the other pollutants. Please note also that Fig. 2 in the manuscript presents only the emissions of SO_x and Fig.3 presents emissions of SO_x, $PM_{2.5}$ and NO_x, but only those from shipping.

Line 340/341: This could also be caused by changes or variations in emissions

Reply: We agree with this point and have improved the text in the Conclusions section to read: *"The year–to–year variations of the concentrations were substantial for all pollutants; these were attributed partly to the variations in regional meteorology, partly to the variations of emissions*".

Line 350/351: I would like to read some words about the possibilities to check compliance to the NECA since 2021. What are the prospects for the future for these observations related to air pollution from shipping?

Reply: We thank the reviewer for the suggestion. However, checking compliance with the NECA regulations since 2021 could be the focus of a separate, new study. As highlighted in our paper, the primary objective of this work was to assess the effects and effectiveness of SECA regulations on atmospheric SO_2 concentrations in the Baltic Sea, as well as on $PM_{2.5}$, NO_x, NO, and O_3 concentrations, using long-term air quality data measured at the Utö Island, and not assessing the rate of compliance of ships with the SECAs regulations implemented. In addition, monitoring NO_x compliance is more challenging due to the regulations being based on engine power output, which set emission limits based on engine performance.

The dataset of this work is particularly unique because there are no other similar long-term air quality datasets from remote locations in the Baltic Sea, and it has not been previously published or analyzed.

Moreover, Utö Island's location in the middle of the Baltic Sea makes it an ideal site for studying long-term pollution trends from shipping, with minimal influence from other pollution sources. Our work underscores the significance of long-term, high-resolution air quality monitoring at remote marine research stations, especially those near heavily trafficked shipping lanes. Such observations are crucial for both quantitative and qualitative analyses of the impacts of regulatory environmental regulations.

Minor comments:

Line 98: Is there really no local wood burning that may have an effect on SO2 and PM2.5 concentrations? Or were these events removed from the data?

Reply: As we mentioned in the manuscript, Utö is a small island in the Baltic Sea, with less than 1 km². It is located 70 km off the coast of mainland Finland, and it is surrounded by open sea and a few smaller islands. During the winter, less than 40 inhabitants live in Utö. Because Utö is so sparsely populated and has very little land traffic, there aren't any significant local sources of NO_x or SO_2. Although some firewood is used for heating, it has to be imported, and its contribution to SO_2 emissions is minimal. While local wood burning might have a slight effect on $PM_{2.5}$, it's unlikely to have a major impact on the long-term trends we've observed in our study.

Line 131: These vessel categories are not well defined. What is "small" and what is "large". Please give more details. And why is there no further distinction of cargo ships into e.g. container ships, tankers, bulk cargo, …? I would assume that STEAM considers more categories than six.

Reply: Vessels with an IMO registry number are classified as "large," while smaller vessels transmit only an MMSI code and no IMO number. The primary aim of this paper is to demonstrate the reduction in SO_x concentration trends, regardless of vessel type.

Line 158: Correct: "is therefore originating"
Reply: The sentence was corrected to "*Part of the SO_2 observed at Utö therefore originates from long-range transport of regional background pollution, while another portion is attributed to shipping traffic in the vicinity of Utö.*", since grammatically is more correct.

Line 173: Effects of the pandemic on emissions were not visible before 2020.
Reply: According to HELCOM (2021), air pollutant emissions from the Baltic Sea ship fleet have declined since 2019, largely due to the COVID-19 pandemic, which significantly decreased vessel activity in the region.

Line 190/191: "The amount of ship traffic has been fairly constant during this period." repeats what was said before.
Reply: After careful consideration we decided to delete this sentence, as it repeats what was said before.

Line 205, Table 2: Please explain STD, N, TBA. Are there no units for STD?
Reply: STD= Standard deviation in [µg m^{-3}].
N (%) represents the fraction of the year for which high-quality data is available, expressed as a percentage. High-quality data, used in our analysis, are defined as valid data recorded, excluding those compromised by factors such as instrument malfunctions, environmental interference, or calibration issues.
We have added all the necessary clarifications to the respective Tables and revised them.

Line 232/233: please make clear that these changes always refer to the 2003-2005 values (and not to the previous period).
Reply: This sentence was changed to "*Our findings revealed that three-year average SO_2 concentrations from the pre-SECA period (2003–2005) decreased by 38 %, 39 %, and 67 % in comparison to the post-SECA periods (2007–2009, 2011–2013, and 2016–2018), respectively*".

Line 240/241: Perhaps you want to introduce abbreviations like SWECA2006, SECA2011, SECA2015 to make clear which phase of the regulation you talk about. This might also help at other places (e.g. line 256, but there may be more).
Reply: We would like to thank the reviewer for this suggestion. We have incorporated the suggested abbreviations (i.e., SECA2006, SECA2010, SECA2015) throughout the text where applicable.

Line 242: replace "during" with "within"
Reply: We thank the reviewer for this correction. We will replace "during" with "within" per the reviewer's recommendation.

Line 250-258: These paragraphs need improvements of the English language (e.g. articles).
Reply: We would like to thank the reviewer for his suggestion. We have revised the text to improve clarity and correctness, especially in terms the English language. The updated paragraphs now read: **"***To study the impact of ships passing by Utö, we selected SO_2 concentration data based on wind direction. First, we*

separated the data points measured when the wind was blowing from the shipping lane (covering wind directions from 185° to 315°) towards the measurement site from the rest of the data.
All data, along with data from when the wind was blowing from the direction of the shipping lane and data from the background sector (wind directions excluding the shipping lane sector), are shown in Figs. 7a–c. Similar to the previously presented results, there is an evident decrease in SO_2 concentrations after the SECA2015, and a slight decrease after the SECA2010. This decreasing trend is visible in all three plots, but the most pronounced decrease occurs after 2015 when only wind directions from the shipping lane were considered."