

Review of “Contribution of shipping NO_x emissions to the marine nitrogen budget of the western Baltic Sea – A case study” by Neumann et al.

This study presents a model-based analysis of the impact of ship-borne atmospheric nitrogen (N) on the biogeochemistry in the western Baltic Sea, a region with significant shipping traffic and in reach of shipping emissions in the North Sea and Baltic Sea. The authors combine the physical-biogeochemical MOM-ERGOM model with an element tagging technique to track ship-borne atmospheric N through the environment. Forcing fields of (ship-borne) atmospheric N deposition are provided by the CMAQ atmospheric model.

The authors provide a sound model validation and conduct a detailed analysis for the year 2012. They find that, on a regional/basin scale, the contribution of ship-borne N to near-surface total nitrogen (TN) stays well below 5%. However, its contributions to dissolved inorganic (DIN) and organic nitrogen (DON) can exceed 10% at some offshore locations not affected by riverine N inputs during the seasonal cycle. Hence, the shipping sector plays a small, yet non-negligible role for the biogeochemistry in the western Baltic Sea.

The manuscript is generally well written, concise and easy to read. I do have one major concern, which is that the manuscript lacks a thorough discussion of the results, both with respect to the limitations of the study and in the context of existing literature. Another a little bit bigger point relates to the manuscript title and the presented results. From my point of view, this study only quantifies the ship-borne contribution to the N **inventory**, but not to the **budget**, because only N variables (or sums of such) and no processes are considered. Although the latter is not absolutely necessary for this study, a TN budget for overall and ship-borne N could be a very nice and interesting contribution.

Aside from that, I have mostly minor suggestions/technical corrections. Therefore, I recommend reconsideration for publication after moderate revisions.

General comments

The “Summarizing Discussion” is less of a discussion but more of a summary. Results should be discussed in the context of existing literature (e.g. comparison with the approach of Raudsepp et al. (2013) mentioned in the Introduction; is the tagging approach better/preferable? Why?). It needs to be very clear what the main insights and contributions of the present study are, and how they expand on previous knowledge. In addition, the limitations of the study need to be discussed. For instance, does the year 2012 represent an average year in terms of environmental conditions, inputs from different types of nutrient sources etc. or was it an exception with respect to some factors? If the latter is the case, what implications could that have on the generality of the results. Some limitations of the model are touched in the course of the Results section but those get a bit lost and they are not sufficient for an in-depth discussion. The authors further emphasize the role of the sediment in the Conclusions. So, what influence of the relatively simple sediment parameterization can be expected with respect to the presented results? (Note that this list of potential discussion points is not meant to be exhaustive.) The authors may want to consider separating Results from Discussion more clearly.

As mentioned above, an actual TN budget for overall and ship-borne N (e.g. for the basins defined in Fig. 4) could be a very nice addition to the manuscript. Or even a budget of the overall and ship-borne N fluxes between the different model state variables (similar to a model schematic with numbers for the different fluxes for overall and ship-borne N). Since the manuscript is quite concise in its current form, this would not make it overly lengthy. If it's not possible to calculate such budget, i.e. the required fluxes were not stored, the authors should consider changing the title by using “inventory” instead of “budget”.

The Results section is a bit of a laundry list without a clear transitions between the individual results describing why the specific (upcoming) results are shown. Including such transitions from one subsection/result to the next would help the reader to get what the key focus of each figure is. Also, as a general comment on the results and the discussion: please regularly include cross-references to figures and figure panels to allow for an easier link between text description and figures. You may want to add panel labels (a, b, c, ...) for easier in-text referencing.

I don't think all additional figures provided in the supplement are necessary, especially since they are only mentioned as existent in the main text. If they don't provide additional relevant information, I'd rather remove them (e.g. time series for the three additional stations listed in the lower part of Table 1). If some of them show very specific features worth discussing, include them in the main manuscript.

Except for the first paragraph, the Conclusions are rather an Outlook with suggestions for future studies. It would be nice to have one or two more actual Conclusion points (which may require additional analysis).

General note on figures/figure captions: I could only print the manuscript in grey-scale. Many of the described features are barely or not at all visible when printed on grey-scale, i.e. not visible for color-blind people either. Please try some other color scales and also avoid references to color in the text/figure captions. In most cases, the color references can just be removed. Figures that caused me trouble in grey-scale are: Fig 1a,b (basically no spatial differences visible); Fig. 3 (colors of transects not distinguishable); Fig. 4 (colored regions not at all distinguishable from each other and from land; transect colors not distinguishable; remove color references in caption); Figs. 5 and 6 (just remove color references); Fig. 7a (gradient in darker colors not visible); Figs. 8-10 (color references need to be removed, perhaps different line styles do the trick?); Fig. 11 (gradients in darker colors barely visible: TN looks almost all the same throughout the year; locations of station hard to see)

Specific comments

Title: I think the speciality of the applied tracing approach is its quantitative nature. Therefore, I'd suggest a slight rephrasing of the title: "Quantifying the contribution of shipping NOx emissions to the marine nitrogen inventory – A case study for the western Baltic Sea".

Line 11: Please state why it is reduced during cyanobacteria blooms.

Lines 27/28: Is the difference in deposition velocity really the most important factor for land-sea differences? Isn't the spatial distribution of sources (over land >> over water) more important? You also mention it in the next sentence.

Lines 61-64: I would rephrase this part and get rid of the bullet point with the question, especially since it is only one question. E.g.: "Here, we combine such model with the nutrient tagging to quantify the contribution of shipping related nitrogen deposition to the total nitrogen (TN) and the different inorganic and organic nitrogen fractions."

Lines 69-72: I would move the part on Raudsepp et al (2013) to the discussion and discuss why one approach is better than the other depending on the research question. The tagging approach is preferable if the current state of a system is to be described as it doesn't change the balance

between sources (“non-disruptive approach”; see Menesguen and Lacroix (2018), doi: [10.1016/j.scitotenv.2018.04.183](https://doi.org/10.1016/j.scitotenv.2018.04.183)). However, if the effect of nutrient reductions should be determined, the actual removal of the considered source is required.

Lines 72-75: This should go right after the first mention of the tagging method, i.e. after the sentence ending on line 61.

Line 88: In the Discussion, you should include if 2012 is an average or exceptional year in terms of environmental conditions, nutrient inputs from different (types of) sources etc., and how this may affect your results (in case it is somehow exceptional).

Figs. 1 and 2: Change their order or adapt in-text cross-referencing. Currently, Fig. 2 is referred to first (line 104).

Lines 1007-111: Based on the introduction, do I understand correctly that shipping emissions only contribute to NO_x? If so, maybe you could explicitly state that here. It could be worthwhile to provide a number or even a figure panel (in Fig. 1) for how much of the total NO_x deposition is from ships.

Lines 121-124: Should this go into section 2.2 “Marine modelling”?

Lines 126-131: I suggest to move this to the discussion and add a statement on if and how this may affect the study results.

Lines 137/138: Is it really “ice cover thickness” and “extent”? Do you mean “ice cover” (as a fraction of the grid cell area) and “ice thickness”?

Lines 140-156: This needs a little bit of reordering. The sentence on the river loads should go to the end of the paragraph and the first sentence should be merged with the one on ERGOM’s development at the IOW.

Lines 146/147: Has ERGOM atmospheric P deposition included? I am not sure after reading “phosphate” and “atmospheric deposition” in the same sentence. Please clarify.

Lines 160-162 (and thereafter throughout the manuscript): I suggest to simply distinguish between N (without “all”) and ship-borne N (“N_{ship}”). For ratios, I would then simply write “TN_{ship}/TN” etc. It makes it more legible and text descriptions less cumbersome. Further suggest to rephrase this sentence to: “... another variable containing only shipping-related nitrogen (subscript “ship”). Process rates for the latter ones are equal to the process rates for the original state variables scaled by the relative contribution of shipping N to the educts.”

Please also add on what frequency model output has been stored.

Lines 190-195: I suggest removing the part on oxygen. It is no relevant for the study, and it seems to me that persistent hypoxia/anoxia in the deep basins of the Baltic Sea is mixed up with seasonal hypoxia in parts of the coastal zone.

Fig. 4: Could be merged with Fig. 3 or at least put as Fig. 3b (ideally with the map inset shown in Fig. 3a/nor Fig. 3). There is a horizontal line a few pixels above the horizontal border between

“Belt Sea” and “Arkona Basin”, which has the same color as the “Oresund”. Please double-check that there is no error with the region mask in your analyses.

Lines 208/209: Please state why you picked the three stations that are shown. Do they represent specific regimes? E.g. coastal vs. offshore? Re-evaluate whether you really need the three stations shown in the supplement. If you want to make the validation more sound, then you should point out somewhere in the validation section that model-data agreement is good also at the other stations presented in the supplement. Remove the sentence on the stations that were ignored.

Line 213: Could you specify what bathymetry criteria are used. Perhaps you can add relevant isobaths to Fig. 4?

Line 219 and Figs. 5 and 6: I would suggest switching the order of the T and S columns as you first talk about T and then S.

Lines 222/223: The vertical mixing might be a point for the discussion.

Lines 225-227: These two sentences contradict each other. First “all is good”, then nitrate is not.

Lines 239-248: Should this go into the discussion?

Lines 244-246: I don't understand this part. Please clarify what is supposed to happen but does not in the model, and how this affects the deposition.

Fig. 8, caption: The last sentence explains how percentiles (10%, 50%, 90%) were calculated for the ratios. However, following this approach, all ratios get the same weight, which may change the results especially during periods of strong short-term changes (e.g. during the spring bloom). I would suggest calculating daily (assuming daily model output), spatial integrals of mass (i.e. concentration times volume) of TN and TN_{ship}. You can then calculate TN_{ship}/TN ratios and use the TN mass as weights to calculate weighted percentiles. This way you ensure that short-term changes in concentrations and volume (ERGOM uses a free surface, right?) are accounted for correctly. Analogously, you can calculate the weighted percentiles for TN by dividing the daily time series of regions' TN mass by the daily time series of the region' volumes, which gives the spatially weighted-average TN concentration, and using the daily time series of the regions' volumes as weights. Analogously for the other variables.

Fig. 10: What is the cause for the strong peaks in the 90% percentiles of PON (and TN) in fall and winter at DMU547?

Line 292: 6% is not “very high”

Line 297: An introductory sentence why you show the vertically resolved plots would be nice.

Lines 303-313: I would suggest rewriting this text part as a “normal” paragraph. The bullets make it appear like a list of very important information but they are mostly a detailed description of the dynamics of DIN, DON and PON. If you prefer keeping it as a bullets, please correct the numbering (“3” occurs twice).

Fig. 11: Would it make sense to add a row for PON since it's referred to a lot on lines 303-313? Perhaps you could label the two transects in Figs. 3/4 (e.g. T1/T2 or S1/S2) and use the label in the figure caption. I was a bit confused about the term "profiles": do you have profile data available? If so, why didn't you use them for validation?

Lines 329-332: Please explain the cause of the difference between Oresund and Arkona Basin.

Section 3.5: Include references to figures, it's hard to remember what result derives from which figure. Discussion should be in the same order as the results, i.e. annual cycle (lines 329-332) should be discussed before the vertical. And as mentioned in my general comments, please provide an actual discussion.

Technical corrections

Line 8: "the atmospheric input"

Line 9: "in shallow coastal regions"

Line 19: Include reference to MSFD (EU, 2008) for the GES

Line 24: remove "approximately"

Line 25: "atmospheric nitrogen deposition"

Line 32: "region where high amounts"

Line 43: move HELCOM reference to end of sentence

Line 45: "That means"

Line 46: "ships built after 2021"

Lines 56/57: "in the system and their deposition sites"

Line 60: "key variables in biogeochemical models"

Lines 63/64: You also provide analyses for organic N, not just TN and DIN

Lines 67/68: remove the sentence "We used an ..."

Line 68: "Within ERGOM, the tagging" instead of "Previously, this tagging"

Line 71: "and another one without"

Line 79: "were not coupled on-line."

Line 98: "deposited"; use full name of SMOKE

Line 99: use full name of STEAM

Line 100: use full term for IMO

Caption of Fig. 1: remove long name of CMAQ and the note

Line 109: Please add in parentheses what HONO is

Line 138: remove "past"

Line 146: "Inorganic" instead of "Basic"?

Line 155/156: remove the name of the supplement; it can all be found in the readme.

Line 157: "by the method"

Line 158: "and used"

Line 173: "in the eastern"

Lines 178-180: change order of the two sentences

Line 181: "emergence of stratification"

Line 183: remove "read"

Line 185: "N:P" => N and P have not been introduced

Lines 189/190: "algal bloom period ends in autumn when stratification"

Line 193/194: remove "- denoted as oxygen minimum zones -"

Lines 205/206: "were aggregated into a monthly 'climatology'" ('climatology' with apostrophes as seven years do not really warrant a proper climatology)

Table 1, caption: remove "and model evaluation"; remove color references to figures

Table 1 should come before Fig. 4 as it is referred to first.

Fig. 5: remove “var [unit] (see header)” on the left; Why are some individual ticks on the y axes not labelled (e.g. 0 in bottom left panel or 15 in the one next to it)? Use “mmol N m⁻³” and “mmol P m⁻³” as units for NO₃ and PO₄, respectively; caption: remove color references; “Each column presents one state variable”; “Vertical lines and shaded areas show the monthly”; remove everything after the first “Supplement”

Line 212: “Basin definitions by Omstedt”

Line 215: “the model’s open boundary”

Line 219: “Sea surface temperature [...] but sea surface salinity”

Fig. 6: remove “var [unit] (see header)” on the left; Why are some individual ticks on the y axes not labelled (e.g. 15 in bottom left panel)? Use “mmol N m⁻³” and “mmol P m⁻³” as units for NO₃ and PO₄, respectively; Why are the depth ranges at the first 2 stations not equivalent to 10m as stated on line 198? Caption: “Same as Fig. 5 but for the bottom 10 m.”

Line 222: Here and throughout the manuscript, do you mean “seasonal” with “intra-annual”? Since you show monthly data, any shorter-term variability is averaged out.

Lines 227/228: replace “The modeled water column is stronger stratified than the real water column” with “Simulated salinity suggests that stratification is overestimated by the model”

Fig. 7: Should the unit be “mmol m⁻³”, not “μmol m⁻³”? Remove the “all” subscript; stations/transects hard to see in grey-scale; caption: remove color references; “White areas are on land in the MOM-ERGOM domain.”

Lines 239-241: “origin and caused by a lower horizontal grid resolution of the CMAQ compared to MOM-ERGOM and interpolation over the land-sea interface.”; remove the last sentence

Line 247: remove “to the ground/sea”

Line 249: here and later: “The contribution of shipping-related nitrogen to TN (TN_{ship}/TN) ...”

Fig. 8: I would suggest removing the absolute shipping TN, DIN, etc. since it is pretty much zero compared to the overall quantities and hardly visible especially in grey-scale. The y axes labels could then be “overall” for the even rows and “shipping”; I suggest putting TN in the first column as you previously analyzed TN, so now you increase the level of detail from TN to its inorganic and organic compartments; add to the caption that it’s 2012 only and remove “in 2012” from the x axes labels; caption: “thick” instead of “think”; remove “in the odd rows”; “For the ratios, ...”; remove color references

Line 255: remove “with all nitrogen” and “all” subscript

Lines 259/260: “which are the sum of DIN, DON and PON”

Line 261: “In contrast, the DIN concentrations are elevated (~5 mmol N m⁻³) throughout the year in the Oresund .”

Lines 262/263: “by riverine nutrient loads”; remove the names of the rivers listed in parentheses

Lines 264/265: “The relative contributions of shipping N to DIN, DON and PON are very small.”

Lines 267/268: “from 1.5-2% in January to about 1% in July”

Fig. 9: These are no daily values (see caption); put only one station/depth label per station (like in Fig. 8: change caption to: “Same as Fig. 8 but for specific stations (see Fig. 1).”

Line 274: “However” instead of “But”

Line 275: remove “as presented in Sect. 3.2”; “in the open ocean”

Line 278: These are no means and no daily data

Fig. 10: put only one station/depth label per station (like in Fig. 8); again why I sthe depth range not equal to 10 m for the first two stations? Change caption to: “Same as Fig. 9 but for the bottom 10 m.”

Line 286: remove “data”

Line 291: “at the surface due to vertical stratification.”

Line 297: State what quantities are shown.

Line 299: "causing the low values in the south."

Line 300: "later" instead of "delayed"

Line 303: remove "reaches ~12.5%"

Fig. 11: add "latitude (N)" to x axis on left and right; station lines are barely visible in grey scale, same for the temporal development; caption: remove "is plotted", remove last sentence

Line 315: Please rewrite the sentence such that models and data are only mentioned once

Line 317: "The concentration of shipping-related TN ..."

Line 320: "... the contribution of shipping-related N to DIN was highest ..."

Line 334: "to TN"

Line 335: "the NO_x emissions"

Line 337: "to DIN"

Line 349: "." after ")"

Line 368: add first name of co-author (although I would suggest to only use initials in the whole author contribution section)

Line 383: "program whose intense"