RESPONSE TO REVIEWERS

Journal: Atmospheric Chemistry and Physics
MS No.: egusphere-2026-1935
Manuscript Title: Addressing systematic underestimation in global ship emission inventories through improved fleet representation and fuel compliance characterization

Dear reviewer,
Many thanks to you for the positive comment and interest in our work. All of the comments are very professional and valuable for the revision and improvement of our research. We have carefully addressed all the comments and revised the manuscript accordingly. The correction details are listed below point by point.

The revision notes are as follows:

Reviewers' comments are in red and in bold font.

Authors' responses are in indent and in black normal font.

Revisions on manuscript are in indent and in blue normal font.

Content

Response to Reviewer #1: 4
OVERALL 4
DETAILS 4
Q1. Contribution analysis 4
Q2. Implications for environment and health 8
Q3. Distinguish between VLSFO, MGO, and ULSFO 10
Q4. Emission factors 11
MINOR COMMENTS 14
Q5. Graphical abstract 14
Q6. Text revision 15
Q7. Text revision 15
Q8. Implications for missing new vessels 15
Q9. Emission calculation formula 17
Q10. Annual data preprocess 17
Q11. Implications for scrubber installation 19
Q12. Reference 20
Q13. Scrubber installation determination 20
Q14. Figure 4 22
Q15. Validation 24
Q16. Text revision 25
Q17. Share of global shipping emissions 26
Q18. Text revision 27
Q19. Particular number 27
Q20. Uncertainties 28
Response to Reviewer #2: 32
OVERALL 32
Major Specific Comments 33
Q21. Line 27 33
Q22. Line 31 33
Q23. Line 34-37 34
Q24. Specify ship emission models 34
Q25. Number of scrubber-equipped ships 35
Q26. Clarification of newly identified vessels and distinction from newly built ships 36
Q27. Update STSD with daily AIS 37
Q28. Daily AIS-based activity processing 38
Q29. Ship identification 38
Q30. Data source 39
Q31. Three-day AIS processing 39
Q32. Information gap between STSD and AIS 40
Q33. Figure2 41
Q34. AIS cleaning 42

Q35. Line142....42
Q36. Emission factor for BC....43
Q37. LNG option....44
Q38. Figure S3....44
Q39. FLSCCM methodology and sensitivity analysis....45
Q40. Line174....47
Q41. Line 179 and Fig. S5....48
Q42. Figure 4....49
Q43. Line215....50
Q44. Figure 6....51
Q45. CEDS....51
Q46. Daily update of STSD, Fig. 7, and Fig. 8....52
Q47. Line 281....54
Q48. Line 290....55
Q49. Line 294....55
Minor Comments....56
Q50. Line 10....56
Q51. Line 14....56
Q52. Line 20....57
Q53. Line 83....57
Q54. Line 93....57
Q55. Line 183....58
Q56. Text revision....58
Reference....59

Response to Reviewer #1:

OVERALL

This paper presents an updated version of a ship activity model to better estimate the atmospheric emissions of CO2, SOX, BC, NOX and PM2.5 where both number of vessels, distance sailed and compliance option with respect to the IMO global sulfur cap is included. While presenting important new data, showing the underestimation of all emissions of atmospheric pollutants, it lacks a further discussion.

Response:

Thank you for your insightful and constructive feedback on our manuscript. We are glad that you recognize the value of our study in underestimation in global shipping emissions. Your comments have enlightened us with valuable perspectives on the broader implications of our work, and we greatly appreciate the thoughtful consideration you've given to our work.

We will seriously address these questions by reorganizing a more comprehensive and detailed description of the methodology and clarifying the assumptions on which the model is based. Further analyses were also conducted on the uncertainties and limitations of the results. These efforts are expected to strengthen the rigor and robustness of the conclusions of this study.

Thanks again for your time and effort in providing such thoughtful input.

DETAILS

Q1. Contribution analysis

If authors think it is the adoption to daily resolution, the data curation and/or the addition of more vessels to the model that have the greatest influence on the results?

Response:

We appreciate the opportunity to further clarify the relative contributions of the methodological improvements introduced in this study.

As introduced in the METHOD *(page4, line95)*, SEIMv2.3 differs from previous

versions through three major improvements: (1) daily AIS collection and preprocessing, (2) daily updating of the Ship Technical Specification Database (STSD) by incorporating newly identified vessels, and (3) the Fleet Low Sulfur Compliance Choice Module (FLSCCM). Below we discuss the contribution of each improvement to the emission estimates.

First, the adoption of daily AIS data itself does not change the emission calculation results compared with annual AIS data, but changes the temporal resolution of the inventory from annual to daily and enables near-real-time emission estimation. To accommodate the characteristics of daily AIS data, we developed a dedicated preprocessing, or data curation framework which includes temporal anomaly removal, spatial anomaly removal, interpolation across long signal gaps, and route restoration. The data curation influenced the emission results in many ways. For example, filtering out spatial anomaly signals decreases the emissions while route restoration increases emissions.

Data curation provides the foundation for emission calculations, rather than constituting an improvement to the underlying emission calculation methodology. Therefore, we did not analyze the influence of data curation on the emission results, while we did evaluate its effectiveness by comparing the spatial distribution of AIS signals before and after preprocessing. As shown in Figure 1 for representative regions and Figure S1 for the global dataset, the preprocessing framework effectively removes abnormal signals while preserving realistic shipping activities.

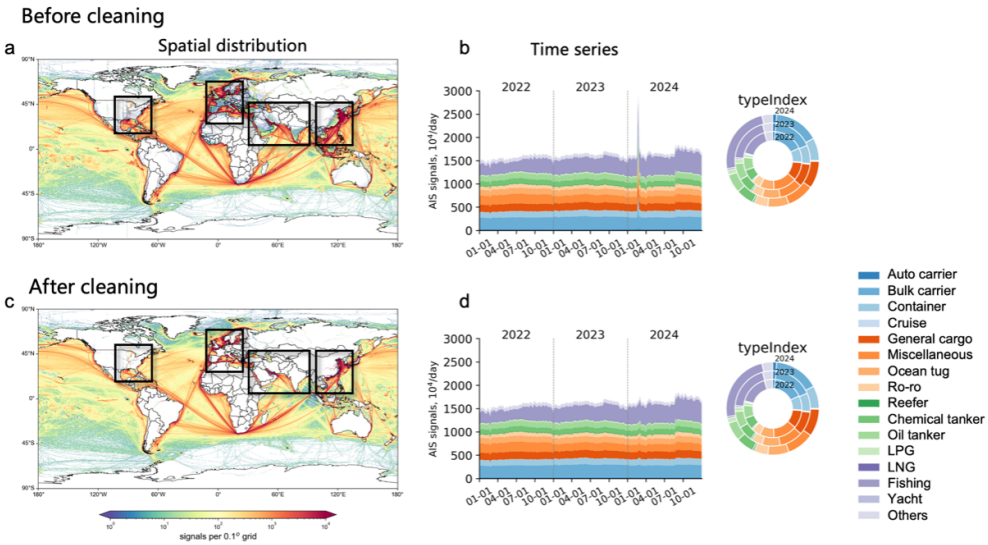


Figure S1. The spatial distribution and time series of AIS before and after preprocessing

Second, to quantitatively distinguish the contributions of the remaining two methodological improvements, namely the daily STSD update and the FLSCCM, we designed three numerical experiments for the 2024 emission inventory based on fully curated AIS dataset:

- We first calculated emissions using the previous SEIMv2.2 framework, which neither incorporates newly identified vessels nor includes the FLSCCM, yielding emission estimate E1.
- We then applied the complete SEIMv2.3 framework, including both the daily STSD update and the FLSCCM, yielding E2.
- Finally, we ran SEIMv2.3 with the FLSCCM disabled while retaining the updated STSD, yielding E3.

Accordingly, the difference E2 − E3 represents the contribution of the FLSCCM (labelled as FLSCCM-2024 in Figure 6), whereas E3 − E1 represents the contribution of the daily STSD update (labelled as STSD-2024 in Figure 6).

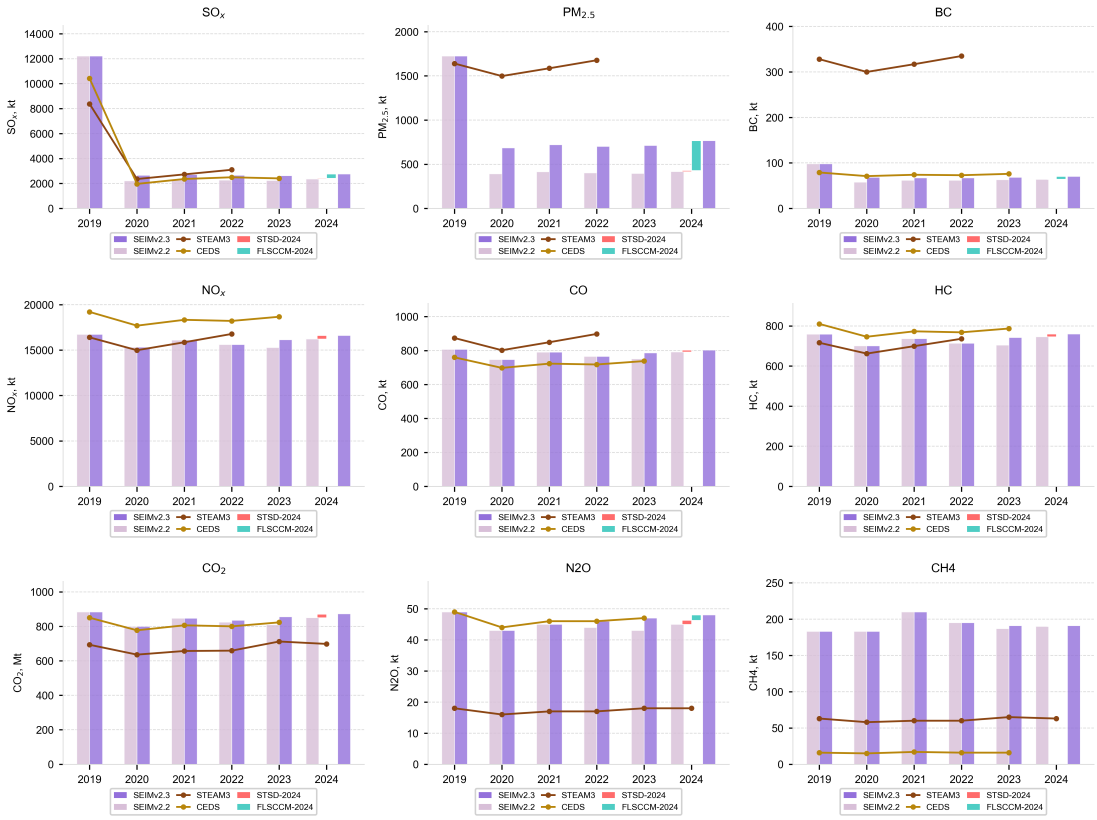


Figure 6: Comparison of annual global ship emissions results from SEIMv2.3, SEIMv2.2, and other emission inventories from 2019 to 2024. SEIMv2.2 results were sourced from (Yi et al., 2024a). SEIMv2.3 represents the method presented in this paper. STEAM3 represents the emission results from EDGAR

(Crippa M., 2025), whose shipping emission results are based on STEAM3. CEDS represents the emission results from Community Earth atmospheric Data Systems, sourced from (Hoesly, 2025). STSD-2024 and FLSCCM-2024 denote configurations incorporating updated fleet information and reconstructed sulfur compliance pathways, respectively.

The results demonstrate that the relative importance of these two improvements depends on the pollutant considered. For CO_2, NO_x, CO, HC, N_2O, CH_4, and most gaseous pollutants, the daily STSD update contributes the majority of the mitigated underestimation because it increases the completeness of the active fleet represented in the model. For example, failure to identify and characterize newly active ships would have resulted in an underestimation of global CO_2 emissions by 5.4% in 2023 and 2.6% in 2024. In contrast, the FLSCCM plays the dominant role for $PM_{2.5}$ and BC because their emission factors are highly sensitive to the fuel-compliance pathway adopted after implementation of the IMO global sulfur limit. Neglecting continued HFO use after 2020 would underestimate $PM_{2.5}$ and BC emissions by approximately 43% and 15%, respectively, of which approximately 97% is mitigated by the FLSCCM. For SO_2, both compliance pathways achieve substantial sulfur reductions, and the difference between their contributions is therefore less pronounced than for $PM_{2.5}$ and BC. These quantitative results are presented in Figure 6 and discussed in Sect. 3.2.

Based on your suggestion, we have further clarified the respective roles of daily AIS processing, the daily STSD update, and the FLSCCM in both the Methods and Results sections.

Revisions:

(1) In RESULTS *(page12, para2)*, we presented the analysis of the contributions of different factors on total emissions in a separate paragraph at the end of Sect.3.2 and added a description of the analytical methods and results:

To distinguish the contributions of daily STSD updates and the FLSCCM, numerical experiments were conducted for the 2024 inventory. Experiment (i) applied SEIMv2.2 without daily STSD updates or FLSCCM; experiment (ii) applied the updated STSD without FLSCCM; and experiment (iii) applied the complete SEIMv2.3 framework. The difference between experiments (iii) and (ii) represents the contribution of FLSCCM, whereas the difference between experiments (ii) and (i) represents the contribution of daily STSD updates. The results indicate that emission underestimation sources differ among pollutants:

STSD updates mainly improve CO_2 and gaseous pollutant estimates by increasing fleet completeness, while FLSCCM remains particularly important for $PM_{2.5}$ and BC because these pollutants are directly affected by fuel sulfur content and scrubber-related removal efficiencies.

Q2. Implications for environment and health

Given the systematic underestimation (and 20% is a lot), what are the implications of the shipping fleet's impact on the environment and human health and what can be proposed as measures to improve reporting, data sharing and inclusion of global shipping in future climate/impact models?

Response:

We agree that the approximately 20% systematic underestimation in shipping emission can propagate into biases in environmental, climatic, and health-impact assessments based on bottom-up ship emission inventories. The implications differ across environmental pathways and geographical regions.

Ship-emitted primary PM and secondary particulate matter formed from SO_2 can alter marine aerosol concentrations and cloud condensation nuclei, thereby affecting cloud microphysical properties, cloud formation, and aerosol radiative forcing (Sofiev et al., 2018; Shi et al., 2023; Yi et al., 2026). Underestimated $PM_{2.5}$ and SO_2 emissions may therefore bias assessments of the climatic effects of shipping. Sulfur and nitrogen compounds deposited onto the ocean may also affect surface-water acidity and marine biogeochemistry, particularly along intensively trafficked shipping routes and in coastal or semi-enclosed waters, although the magnitude of these effects depends on seawater buffering capacity, circulation, and the spatial scale considered (Shi et al., 2023).

The implications for human health should be interpreted more cautiously. As described in the METHOD section, most of the additional emissions identified in this study originate from open-ocean regions outside Emission Control Areas, rather than near densely populated coastal areas. Therefore, the effect of the inventory correction on population exposure and associated health burdens is likely to be limited. Nevertheless, long-range transport may still influence coastal air quality. The spatially and temporally resolved inventory developed in this study can provide a basis for future modelling studies to quantify the resulting effects on

climate, air quality, and human health.

Our results also highlight the need for improved data accessibility and observational validation. We therefore propose the following measures:

(i) Establishing more timely and accessible global databases for vessel technical specifications, new-build records, engine characteristics, and exhaust-gas cleaning system installations. At present, much of this information is available primarily through costly commercial providers, such as Clarksons Research, limiting its use in independent scientific research. Future data-sharing mechanisms should better balance commercial interests and confidentiality with scientific needs, for example by providing standardized, aggregated, or research-access versions of essential vessel data.

(ii) Conducting additional onboard, plume-based, remote-sensing, and port-based measurement campaigns covering different ship types, engine loads, operating conditions, fuels, and emission-control technologies. Such observations would support the validation and refinement of emission factors, fuel-compliance assumptions, and emission estimates at both individual-vessel and aggregated regional scales.

Together, these measures would improve the completeness, transparency, and reliability of global shipping emission inventories and strengthen their application in future assessments of climate, air quality, human health, and marine environmental impacts.

Revisions:

In DISCUSSION *(page16, para2)*, we expanded the discussion on the environmental and health impacts of the underestimated emissions:

Near-real-time emission inventories enable tracking of short-term emission changes and provide a data foundation for emission reduction policies (Xu et al., 2023). The underestimation addressed in this study may affect assessments of the environmental and health impacts of global shipping. Underestimated $PM_{2.5}$ and SO_2 emissions may bias estimates of shipping-induced aerosol–cloud interactions and radiative forcing (Chen et al., 2017; Goldsworthy et al., 2019; Yuan et al.,

2024). Such underestimation may also affect assessments of atmospheric deposition impacts in coastal and marine environments (Shi et al., 2023). The additional contribution to population exposure and health burdens may be limited, because the additional emissions identified in this study were mainly in open-ocean regions, distant from populated coastal areas.

In DISCUSSION *(page 18, para1)*, we combine our idea of how to improve data reporting with your last comment about uncertainties:

Future reductions in uncertainty require improved ship-specific technical databases, broader sharing of scrubber installation records and fuel-use data, and additional validation datasets such as onboard measurements, EU MRV, or ship bunkering records. Direct single-ship validation remains challenging because such data are rarely publicly available. Nevertheless, comparisons with independent inventories and sensitivity analyses indicate that these uncertainties are unlikely to alter the main conclusions regarding the magnitude and characteristics of global shipping emissions.

Q3. Distinguish between VLSFO, MGO, and ULSFO

The authors focus on two compliance options (low sulfur fuels and HFO + scrubber) but given the high temporal resolution of ship activity data, would it be possible to divide low sulfur options into VLSFO (in 0.5% areas) vs MGO (and ULSFO) (in 0.1% areas)? My guess is that emission factors are different but the question is if data is available to differentiate.

Response:

Thank you for this important comment. We agree that the low-sulfur fuel pathway should distinguish fuels used under the 0.5% and 0.1% sulfur limits because their emission factors differ. In the revised model description, VLSFO and MGO are represented as separate fuel categories, and all fuel categories and corresponding emission factors used in SEIMv2.3 are provided in Tables S2 and S3.

Specifically, high-resolution AIS data are used to determine the sulfur limit applicable to each vessel according to its operating time and location. For vessels following the fuel-switching pathway, fuel used in waters subject to the IMO 0.5% global sulfur limit is represented using VLSFO emission factors, whereas fuel used in ECAs and other waters (such as China DECA) subject to the stricter 0.1% sulfur limit is represented using MGO emission factors.

However, the model cannot reliably distinguish ULSFO from MGO within the 0.1% sulfur fuel category. AIS data identify vessel activity and location but not the actual fuel product consumed onboard, and globally consistent ship- or voyage-level data distinguishing ULSFO from MGO are unavailable. Although both fuels comply with the 0.1% sulfur limit and therefore have broadly similar SO_2 emission factors when their actual sulfur contents are comparable, residual-type ULSFO may have higher $PM_{2.5}$ and black carbon emission factors than distillate MGO. Representing all fuels used in 0.1% sulfur areas using MGO emission factors may therefore cause some underestimation of $PM_{2.5}$ and black carbon where ULSFO is used. We have clarified this assignment procedure and associated uncertainty in Sect. 2.3, the Discussion, and the Supplement.

Revisions：

In DISCUSSION *(page16, para4),* we have clarified the limitation regarding ULSFO and MGO:

Third, FLSCCM does not distinguish ULSFO and MGO, and fuels of 0.1% sulfur are represented using MGO emission factors. This simplification is expected to have limited influence on SO_2, but it may lead to some underestimation of $PM_{2.5}$ and BC emissions where residual-type ULSFO is used. Therefore, uncertainties may arise from incomplete scrubber records, assumed scrubber operation, and simplified representation of ULSFO using MGO emission factors. These uncertainties mainly affect $PM_{2.5}$ and BC estimates, while their influence on SO_2 remains limited.

Q4. Emission factors

On emission factors, would it be possible to present these as ranges instead of absolute numbers to reflect the variability? Is engine load accounted for in the modelling, if not the authors could discuss the implications of including this? For many pollutants, emission factors are tested at different loads so it should be possible to do some estimations (present ranges as proposed in previous point).

Response:

Thank you for this helpful suggestion, which allows us to clarify how engine-load variability is represented in SEIMv2.3. Engine load and low-load adjustment factors are already explicitly considered in the emission calculations. Table S2 reports the baseline energy-based emission factors assigned according to fuel type, engine type, engine-speed category, and NO_x Tier. For each AIS record, the

instantaneous main-engine load is estimated from vessel speed using the propeller law. When the estimated load is below 20%, the corresponding baseline emission factor is multiplied by a pollutant-specific low-load adjustment factor; at loads of 20% or higher, the baseline value is retained. Therefore, the effective emission factor applied in the calculation varies with the operating load rather than remaining constant.

We agree that presenting only the baseline values in the original Supplement did not adequately convey this load-dependent variability. However, we did not replace the baseline factors with a single minimum–maximum range. Instead, we retained the baseline factors in revised Table S2 and added the pollutant-specific LLAF curves and their applicable load range in Fig. S6. This allows the effective emission factor at any load below 20% to be reproduced directly, while preserving the signal-by-signal calculation framework.

We also reorganized the Supplementary Information so that revised Table S2 summarizes the baseline energy-based emission factors, whereas revised Table S3 summarizes the fuel properties, sulfur contents, and fuel-based emission factors used in the fuel-switching module. This reorganization, together with Fig. S6, more clearly shows both the baseline values and their load-dependent adjustment in SEIMv2.3.

Revision:

(1) In METHOD *(page4, para3)*, we added a sentence to clarify that engine load was considered in the calculation process:

Engine-load effects were explicitly considered by estimating the main-engine load for each AIS record and applying pollutant-specific low-load adjustment factors when the load was below 20%. Detailed factors are provided in Fig. S6.

(2) In Supplementary Information, we added Fig. S4. and a note.

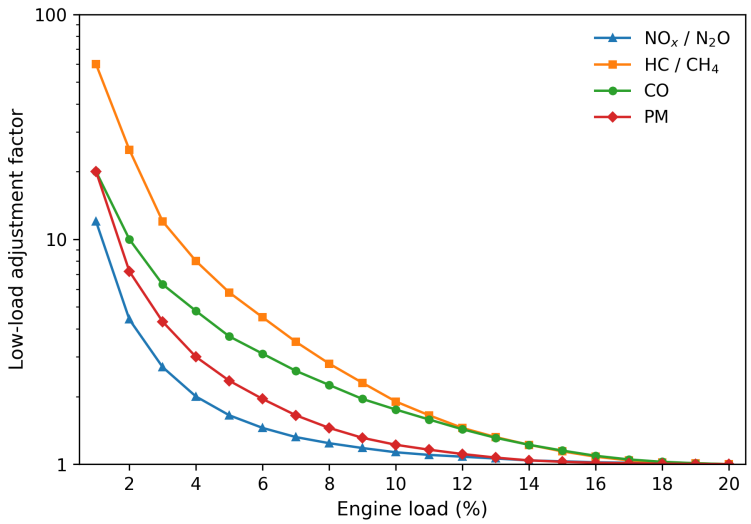


Figure S4. Low-load adjustment factors applied to main-engine emission factors in SEIMv2.3.

Notes. Low-load adjustment factors applied to main-engine emission factors as a function of estimated engine load. Baseline emission factors are retained at engine loads of 20% or higher. At engine loads below 20%, pollutant-specific low-load adjustment factors are applied to NOx/N_2O, HC/CH_4, CO, and PM according to (ICF International, 2009).

(2) In Supplementary Information, we reorganized Table S2 to summarize the baseline energy-based emission factors, whereas revised Table S3 to summarize the fuel properties, sulfur contents, and fuel-based emission factors used in the fuel-switching module. We also explained when and how LLAF was applied.

Table S2. Baseline energy-based emission factors applied in SEIMv2.3 (unit: g/(kWh))

Fuel	Engine	Engine type	NOx				$PM_{2.5}$	SOx		HC	CO	CH4	N2O	CO2	
			Tier 0	Tier I	Tier II	Tier III		Tier 0	Tier I–III					Tier 0	Tier I–III
HFO	ME	SSD	18.1	17.0	14.4	14.4	1.39	9.404	8.896	0.632	0.54	0.010	0.031	576.090	544.950
	ME	MSD	14.0	13.0	10.5	10.5	1.39	9.912	9.404	0.527	0.54	0.010	0.034	607.230	576.090
	ME	HSD	10.0	9.8	7.7	7.7	1.39	10.420	9.912	0.527	0.54	0.010	0.030	638.370	607.230
	AE	AE	11.2	11.2	11.2	11.2	1.40	10.420	9.912	0.400	0.54	0.010	0.040	638.370	607.230
	Boiler	BE	2.1	2.1	2.1	2.1	1.42	17.283	17.283	0.100	0.20	0.002	0.040	1058.760	1058.760
VLSFO (0.5% S)	ME	SSD	18.1	17.0	14.4	14.4	0.31	1.715	1.617	0.632	0.54	0.010	0.030	561.050	528.990
	ME	MSD	14.0	13.0	10.5	10.5	0.31	1.813	1.715	0.527	0.54	0.010	0.030	593.110	561.050
	ME	HSD	10.0	9.8	7.7	7.7	0.31	1.862	1.813	0.527	0.54	0.010	0.034	609.140	593.110
	AE	AE	11.2	11.2	11.2	11.2	0.31	1.862	1.813	0.400	0.54	0.010	0.036	609.140	593.110

	Boiler	BE	2.1	2.1	2.1	2.1	0.31	3.136	3.136	0.100	0.20	0.002	0.049	1025.920	1025.920
MGO (0.1% S)	ME	SSD	18.1	17.0	14.4	14.4	0.20	0.342	0.323	0.632	0.54	0.010	0.030	561.050	528.990
	ME	MSD	14.0	13.0	10.5	10.5	0.20	0.362	0.342	0.527	0.54	0.010	0.030	593.110	561.050
	ME	HSD	10.0	9.8	7.7	7.7	0.20	0.371	0.362	0.527	0.54	0.010	0.034	609.140	593.110
	AE	AE	11.2	11.2	11.2	11.2	0.20	0.371	0.362	0.400	0.54	0.010	0.036	609.140	593.110
	Boiler	BE	2.1	2.1	2.1	2.1	0.20	0.626	0.626	0.100	0.20	0.002	0.049	1025.920	1025.920
LNG	ME	SSD	1.3	1.3	1.3	1.3	0.02	0.005	0.004	0.500	1.30	2.500	0.020	475.750	335.500
	ME	MSD	1.3	1.3	1.3	1.3	0.02	0.005	0.004	0.500	1.30	5.500	0.020	475.750	335.500
	ME	HSD	1.3	1.3	1.3	1.3	0.02	0.005	0.004	0.500	1.30	5.500	0.020	475.750	335.500
	AE	AE	1.3	1.3	1.3	1.3	0.02	0.005	0.005	0.500	1.30	5.500	0.020	475.750	429.000
	Boiler	BE	1.3	1.3	1.3	1.3	0.03	0.009	0.009	0.105	0.20	0.040	0.020	783.750	783.750

Notes: The values listed in Table S2 are baseline energy-based emission factors. For each AIS record, the main-engine load was estimated from vessel speed using the propeller law. When the estimated load was below 20%, the baseline emission factor was multiplied by the pollutant-specific low-load adjustment factor shown in Fig. S6; at loads of 20% or higher, the adjustment factor was set to 1. Therefore, the effective emission factor used in the calculation varied with engine load.

Table S3. Fuel properties and fuel-based emission factors used for fuel switching

Fuel	Carbon content (m/m%)	$EFCO_2$,f (g/gfuel)	Sulfur content (m/m%)	$EFSO_2$,f (g/gfuel)
HFO (2.43% S)	0.8493	3.1140	2.6	0.0510
VLSFO (0.5% S)	0.8744	3.2060	0.5	0.0098
MGO (0.13% S)	0.8744	3.2060	0.13	0.0025
LNG	0.7500	2.7500	~0	0.0000317

MINOR COMMENTS

Q5. Graphical abstract

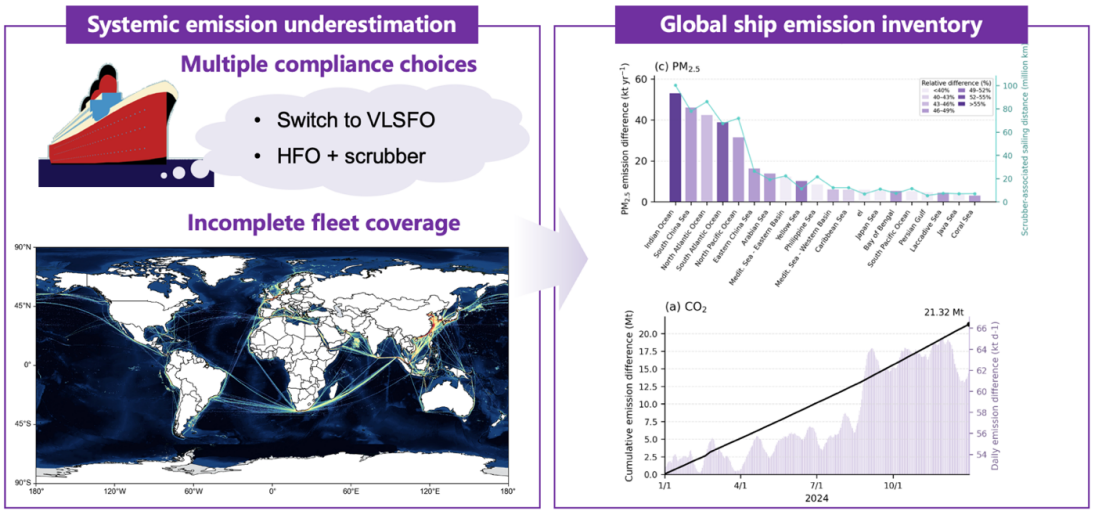
Objects in graphical abstracts need to be enlarged to allow for better illustration.

Response:

Thank you for this helpful suggestion. We have revised the graphical abstract by enlarging the main objects and simplifying the layout. Specifically, the ship-compliance illustration, fleet-description map, and inventory comparison panels were enlarged, while minor graphical elements and overly detailed labels were removed to improve readability.

Revision:

In page 2, we revised the graphical abstract:



Q6. Text revision

row 27: is the 1.7 rise of atmospheric emissions referring to CO2? specify.

Response:

Thank you for pointing this out. We confirm that the reported 1.7-fold increase refers specifically to global shipping CO_2 emissions, as reported in the cited study. To avoid ambiguity, we have revised the corresponding sentence in the manuscript to explicitly state "CO_2 emissions" instead of the more general term "atmospheric emissions".

Revision:

In INTRODUCTION *(page2, line28)*, we changed 'atmospheric' to 'CO_2':

As one of the hard-to-abate sectors (Sharmina et al., 2021), global shipping experienced a threefold increase in cargo turnover and a 1.7-fold increase in carbon dioxide (CO_2) emissions from 1970 to 2021 (Wang et al., 2025).

Q7. Text revision

IMO should always be written in capital letters.

Response:

Thank you for pointing this out. We have corrected the capitalization of IMO throughout the manuscript to ensure that it is consistently presented in capital letters, including those in the citation.

Q8. Implications for missing new vessels

Row 54-57: the implications of missing new vessels are discussed but it is not really clear to me why this is a challenge and why this is relevant, please specify.

Response:

Thank you for this valuable comment. We agree that the implications of missing newly active vessels were not sufficiently explained in the original manuscript.

Ship emission estimation relies on combining AIS-derived activity data with vessel technical specifications, such as deadweight tonnage, build year, engine power, ship type, and fuel information, which are used to assign engine loads and emission factors. However, when newly active vessels are absent from ship technical databases, their missing specifications have to be filled using default or assumed values. These default values may differ substantially from the actual characteristics of individual vessels, leading to large errors in engine power estimation, emission factor assignment, and consequently ship-level emissions. This makes missing new vessels a relevant challenge for near-real-time emission inventories.

To reduce such errors, our framework supplements missing vessel parameters using machine-learning models based on available physical information, such as vessel length and width, to predict deadweight tonnage, engine power, and ship type. These predicted parameters provide more vessel-specific inputs than direct default assignment and thus help improve the accuracy of emission estimation for newly active vessels. We have revised the manuscript to clarify this point.

Revision:

(1) In INTRODUCTION *(page3, para4)*, we explained why missing new vessels would become a challenge in generating accurate emission inventory:

Unlike ship activity data that can be processed at daily resolution, ship technical databases are generally updated with substantial delays (Yi et al., 2024a). Newly identified ships often lack key technical specifications and are represented by proxy values, which may affect engine power estimation, load-factor correction, and emission factor assignment.

(2) In METHOD *(page6, para2)*, we added the necessity of parameter completion:

(2) Static AIS lack main engine power and DWT information, which is also critical for matching emission factors. Without additional parameter completion, these

missing specifications would have to be represented by default or proxy values, which may introduce considerable uncertainties in engine power estimation and subsequent emission factor assignment. Through extensive literature review,…

Q9. Emission calculation formula

row 74: authors say formula is largely consistent, what differed? what does this mean?

Response:

Thank you for pointing this out. We agree that the phrase "largely consistent" was ambiguous. Our intention was to indicate that the fundamental emission calculation equations remain identical to those in our previous study, whereas the methodological advances of the present work lie in the daily activity preprocessing, dynamic ship technical database, and fleet compliance module.

Revision:

(1) In METHOD *(page4, para3),* we have revised the corresponding sentence to explicitly state that the core emission calculation equations are identical to those in our previous study:

The specific formula used in this study to calculate ship emissions are identical to those reported by (Yi et al., 2024a) and provided in Supplementary Information.

(2) In Supplementary Information, we further highlighted that the core emission equations were identical to Yi et al. (2024a) and that the differences occurred in updated input data preprocessing and module-level assignments:

Emission calculation principles of SEIMv2.3

The mathematical formulation of the emission calculation is identical to that described by Yi et al. (2024a). Therefore, differences from Yi et al. (2024a) arise from updated input data preprocessing and module-level assignments rather than changes in the core emission equations. SEIMv2.3 calculates atmospheric emissions for every ship by every two subsequent AIS signals…

Q10. Annual data preprocess

row 103: Do the authors know how annual data is treated? is that unprocessed? how is the overlap handled?

Response:

Thank you for this valuable comment. We admitted that the preprocessing procedures for the annual AIS dataset were not clearly described in the original manuscript.

Annual AIS datasets are not used without preprocessing. Similar to the daily AIS datasets, they undergo ship identification and quality-control procedures before being used for emission estimation. The main difference lies in the characteristics of the two datasets and the corresponding preprocessing strategies. For the annual AIS dataset, the interval threshold used to identify temporally abnormal AIS signals is set to a much larger value because the data are collected less frequently. In addition, signal transmission delay is not considered during annual preprocessing, since annual datasets do not aim to preserve the exact temporal sequence of ship activities. In contrast, the daily AIS preprocessing framework is specifically designed for near-real-time applications and therefore includes additional procedures to identify delayed AIS transmissions and reconstruct the chronological sequence of ship movements.

We would also like to clarify that the annual and daily AIS datasets are processed independently and are used in different model applications. Only one dataset is used in each emission calculation; therefore, there is no overlap or double counting between the two approaches.

Based on your suggestion, we have clarified the differences between the preprocessing procedures for annual and daily AIS datasets in the revised manuscript.

Revision:

(1) In METHODS *(page4, line106)*, we have expanded the description of the preprocessing procedures for the annual and daily AIS datasets. We also further explained the difference between daily and annual AIS preprocessing:

Therefore, rigorous pre-processing was necessary especially for daily near-real-time emission calculation.
Annual AIS data used in previous SEIM were also preprocessed before emission calculation, while the present daily framework differs from the annual preprocessing mainly in its treatment of temporal anomalies and delayed signal

transmission. For the annual AIS dataset, temporal anomaly filtering was conducted with a larger interval threshold because the signals were collected at lower frequency.
For the daily AIS dataset, firstly, to ensure the stability of daily input data, …

(2) In METHOD *(page4, line111)*, We also further clarified the difference between annual and daily AIS pre-processing, in handling the transmission delay:

Unlike annual AIS preprocessing, the daily framework explicitly accounts for transmission delays because daily AIS data received on a given day may contain signals with timestamps from both the current and previous days.

Q11. Implications for scrubber installation

Scrubber installations are also underreported, please discuss implications for this.

Response:

When scrubber-equipped vessels are not identified in technical databases, their use of high-sulfur fuel oil combined with exhaust gas cleaning may not be adequately represented, leading to biases in the assumed fuel mix and associated emission factors. This issue affects not only SO_2 estimates, but potentially also PM2.5, black carbon, and particle number.

Moreover, complete installation records alone cannot fully resolve this issue, because the actual operation of scrubbers may vary across voyages and operating conditions. Globally consistent and continuously available information on whether scrubbers are operating, particularly during ocean-going voyages, remains limited. Consequently, assuming that an installed scrubber is operated throughout all relevant periods may also introduce errors into emission estimates.

We have therefore added a discussion clarifying that incomplete scrubber installation and operational data may affect both atmospheric emission estimates and the broader environmental assessment of shipping emissions.

Revision:

In DISCUSSION *(page16, line361)*, we added a discussion clarifying that incomplete information on both scrubber installation and actual operation can

affect the emission calculation:

Finally, incomplete scrubber installation records may bias $PM_{2.5}$, BC, and PN emission calculation because unidentified scrubber-equipped vessels may be treated as using low-sulfur fuels. Besides, installation records alone do not fully indicate whether scrubbers are actually operated during the whole voyages, which introduces further uncertainty.

Q12. Reference

row 150-152: add reference for the emission factors.

Response:

Thank you for this helpful suggestion. The sources of the emission factors were already provided in the Supplementary Information. However, we agree that the corresponding references were not explicitly indicated in the main text, which may have caused unnecessary ambiguity. We have therefore added the relevant references in the revised manuscript and clarified that the emission factors are primarily based on the IMO Fourth Greenhouse Gas Study(IMO, 2020) and the National Standard for General Diesel Fuel of the People's Republic of China.

Revision:

In METHOD *(page4)*, we added the source and reference for emission factors:

… and provided in Supplementary Information. Emission factors for different ship types and fuels are listed in Table S2 and Table S3, primarily based on the IMO Fourth Greenhouse Gas Study (IMO, 2020) and U.S. Environmental Protection Agency reports (Starcrest Consulting Group, 2022).

Q13. Scrubber installation determination

row 167-170: Does the DWT threshold ensure that cruise vessels are included. There is a DNV report on ship types with scrubbers showing that a large fraction of container vessels and cruise ships have installed scrubbers. Speed will also determine fuel consumption and thus motivate installation of scrubbers.

Response:

We agree that the original description of the DWT threshold was insufficient and could give the impression that scrubber adoption was determined solely by vessel size. In the revised manuscript, we improved our previous FLSCCM with existing scrubber installation data. The DWT threshold is no longer used as an independent scrubber prediction criterion. Instead, FLSCCM adopts an observed-first

framework, where vessel-level scrubber records obtained from Clarksons Research are first incorporated as observational constraints, and the DWT threshold is applied only to database-unmatched vessels during residual gap filling.

We analyzed the characteristics of the observed scrubber-equipped fleet using the vessel-level scrubber database. The results showed that scrubber installations were concentrated among large vessels, with approximately 69.5% of observed scrubber-equipped vessels having DWT above 50,000 tonnes. Therefore, a 50,000 DWT threshold was introduced as a structural constraint for gap filling. Importantly, this threshold does not exclude smaller vessels with confirmed scrubber installations; observed scrubber-equipped vessels were directly retained regardless of their DWT.

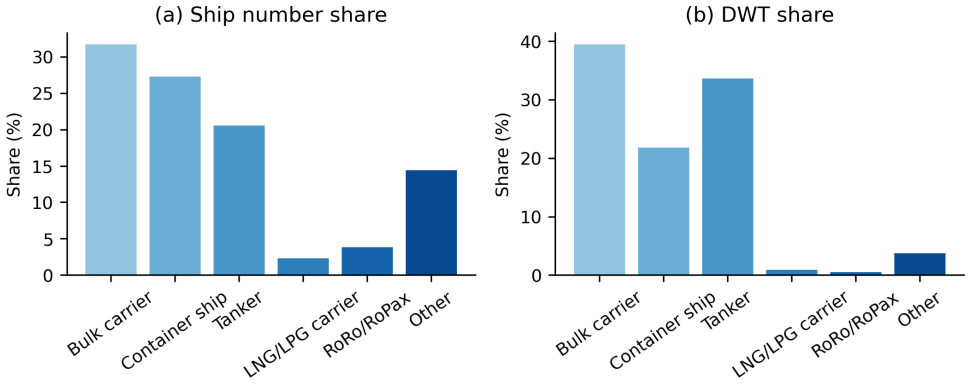


Figure S7. Ship type distribution of scrubber-equipped fleet in 2024

We further evaluated the uncertainty associated with this threshold through sensitivity analyses using alternative DWT thresholds (40,000 and 60,000 DWT). The resulting changes in $PM_{2.5}$ and BC emissions were limited, demonstrating the robustness of the revised FLSCCM to reasonable variations in the screening criterion.

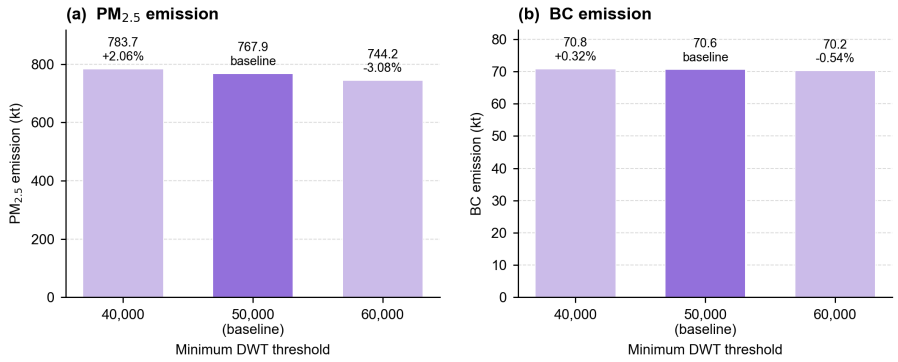


Figure S12. Sensitivity of emission estimates to minimum DWT screening threshold

Revisions:

In METHOD (*page8, para3*), we explained FLSCCM's methodology for ships unmatched with scrubber installation records:

Analysis of the scrubber installation database in Supplementary Information showed that scrubber adoption was concentrated among major cargo ship categories and larger vessels. Bulk carriers, containers, and tankers together accounted for approximately 79.5% of identified installations (Fig. S8). Therefore, gap filling was restricted to bulk carriers, containers, oil tankers, and chemical tankers. In addition, ship size and construction year were introduced as screening criteria because scrubber installation is economically favoured for ships with higher fuel consumption and sufficient remaining service life. Database-unmatched ships were considered eligible candidates only when they had a DWT of at least 50,000 tonnes and a construction year later than 2006 when reliable information was available. These criteria were applied only to ships without scrubber installation records during the residual gap-filling procedure and did not affect ships with confirmed scrubber installations. The threshold settings were determined based on the characteristics of the observed scrubber-equipped fleet. For vessels identified or assigned as following the HFO-with-scrubber pathway, SEIMv2.3 applies a baseline $PM_{2.5}$ removal efficiency of 37% and a BC removal efficiency of 30% in the emission calculation. Uncertainties associated with scrubber assignment criteria and removal efficiencies are further evaluated through sensitivity analyses in the Supplementary Information and the Discussion section.

In SUPPLEMENTARY INFORMATION, we removed the original Fig. S4. Meanwhile, we added Fig. S6-S9 to illustrate the fleet composition of scrubber-equipped ships and Fig. S11-S12 to present the sensitivity analysis of different DWT and build year threshold.

Q14. Figure 4

Figure 4 "fleet number by weight bin", not possible to see the higher weight bins. Also, can the higher representation of new-found vessels in Bulk and container say something about reporting/control of these segments in general? Oil tankers are perhaps rigorously documented with strict vetting procedures etc while bulk/container are less controlled?

Response:

Thank you for this helpful suggestion. We have revised Fig. 4 into a more readable donut-chart format. The inner ring represents the proportion of newly identified vessels relative to the original STSD fleet, while the outer ring illustrates the distribution of newly identified vessels by ship type, build period, and weight bin. This design allows both fleet expansion and structural characteristics of newly identified vessels to be visualized simultaneously.

We agree that the higher representation of newly identified vessels in bulk carriers and container ships may provide information on reporting and operational visibility across ship segments.

However, we would not interpret this pattern as evidence that bulk carriers and container ships are less controlled. Bulk carriers and container ships are subject to general SOLAS requirements, including mandatory AIS carriage for cargo ships above relevant gross-tonnage thresholds. In addition, bulk carriers are covered by SOLAS Chapter XII on additional safety measures, while containerized shipping is subject to SOLAS requirements such as verified gross mass before loading (IMO, 2024b, a). Therefore, the observed pattern is more likely related to fleet growth and AIS-based detectability rather than weaker regulatory control.

In our framework, newly identified vessels are primarily captured from AIS signals and then supplemented with technical information, rather than being identified from reporting-based records alone. Therefore, the larger number of newly identified bulk carriers and container ships may reflect both their fleet growth and their relatively high detectability in AIS-based updating. Bulk carriers have a large baseline fleet size, so a similar relative increase can lead to a large absolute number of additional vessels. For container ships, the recent expansion of containerized shipping (UNCTAD, 2021) may also contribute to the larger number of newly identified vessels.

Conversely, recent studies suggest that shadow-fleet activities are predominantly associated with oil tankers, which may reduce their detectability in AIS-based fleet updating (Caprile, 2024; U.S. Department of the Treasury, 2025). Therefore, the smaller apparent increase in oil tankers should not necessarily be interpreted as evidence of more complete fleet records alone. Although this remains a plausible explanation, we do not have sufficient quantitative evidence to assess its

contribution in the present study.

Revision:

In RESULTS *(page 10)*, we revised Fig. 4:

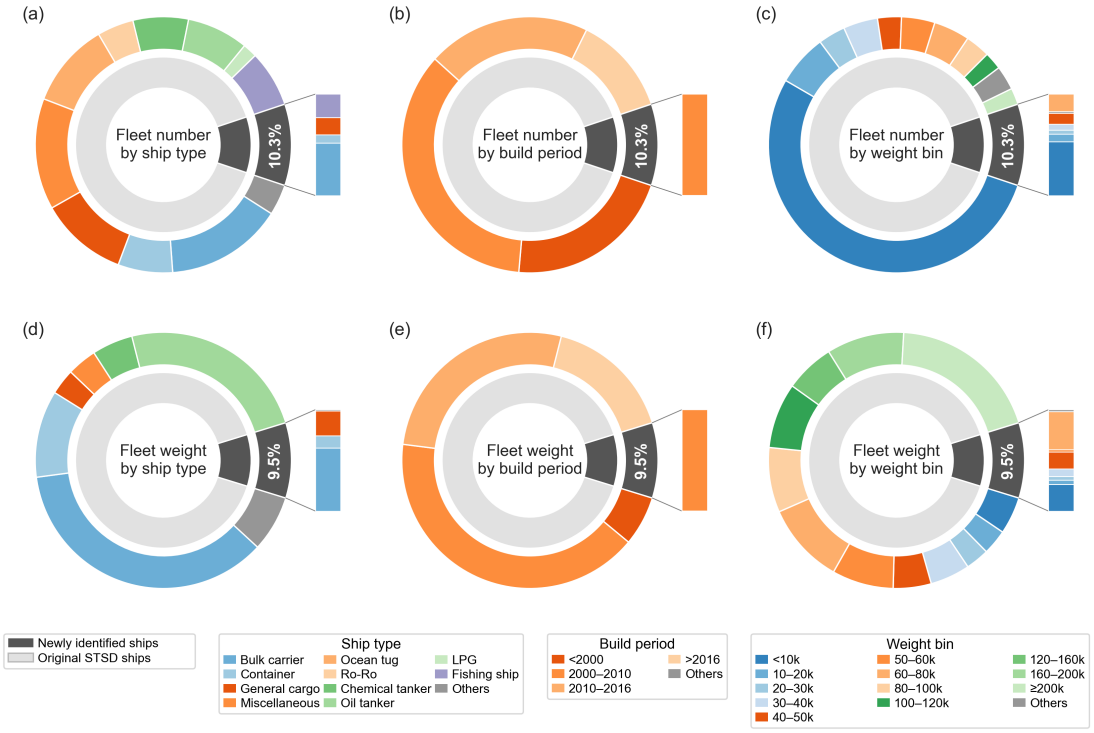


Figure 4: The composition of ship number and total weight of the active global fleet by SEIMv2.3 and SEIMv2.2. **Panels (a–c) show fleet number by ship type, build period, and weight bin, respectively, while panels (d–f) show the corresponding distributions of total fleet weight. The inner ring distinguishes vessels contained in the original Ship Technical Specification Database (STSD; light grey) from vessels newly identified through the dynamic STSD updating procedure (dark grey). The outer coloured ring shows the category composition of the original STSD fleet, and the linked stacked bar shows the composition of newly identified vessels using the same colour scheme. "Newly identified" refers to vessels previously missing from the STSD and does not necessarily indicate newly built ships. Fleet number is counted using unique valid IMO numbers.**

Q15. Validation

EDGAR and CEDS should be described in text.

Response:

Thank you for this comment. We agree that EDGAR and CEDS were insufficiently introduced in the original manuscript. We have now defined both datasets at their first occurrence and briefly described their scope and relevance to the comparison. EDGAR, the Emissions Database for Global Atmospheric Research, is a global

anthropogenic emission inventory that provides sector-resolved estimates of greenhouse gases and air pollutants. CEDS, the Community Emissions Data System, provides consistent historical anthropogenic emissions of reactive gases, aerosols, and greenhouse gases by sector. Their international shipping estimates were used here as independent global datasets for comparison with the temporal trends derived from SEIMv2.3.

Revision:

In RESULTS *(page11, line242)*, we added definitions and brief descriptions of EDGAR and CEDS at their first occurrence in the Results section. We also clarified that their international shipping estimates were used as independent global datasets for evaluating the temporal trends derived from SEIMv2.3.

… using different models from 2019 to 2024. For comparison, we used international shipping emissions from the Emissions Database for Global Atmospheric Research (EDGAR), a global sector-resolved inventory of anthropogenic greenhouse gases and air pollutants, and the Community Emissions Data System (CEDS), a consistent historical inventory of sectoral anthropogenic emissions. These datasets provide independent global estimates against which the temporal trends in SEIMv2.3 can be evaluated.

Q16. Text revision

The only GHG included is CO2 right? If yes, be specific and change to CO2 instead of GHG in text.

Response:

Thank you for pointing this out. In this study, greenhouse gas (GHG) emissions include not only CO_2 but also CH_4 and N_2O, which are estimated following the same emission calculation framework. Therefore, the use of the term GHG in the manuscript is intentional rather than referring exclusively to CO_2. To avoid potential ambiguity, we have revised the relevant text to explicitly clarify that GHG emissions in this study comprise CO_2, CH_4, and N_2O.

Revision:

In RESULTS *(page11, para1)*, we clarified that GHG emissions in this study comprise CO_2, CH_4, and N_2O at GHG's first appearance in the RESULTS Section:

Figure 6 shows the global ship emissions of nine atmospheric pollutants and greenhouse gases (GHG), including CO_2, CH_4, and N_2O, calculated using different models from 2019 to 2024.

Q17. Share of global shipping emissions

How does the increase of CO2 emission change the global shipping contribution of CO2-emissions? What is this equivalent to?

Response:

Thank you for this insightful comment. International shipping contributes approximately 2.9% of global anthropogenic CO_2 emissions. Therefore, a 5.4% increase in estimated shipping CO_2 emissions corresponds to an increase of approximately 0.16% in the contribution of shipping to global anthropogenic CO_2 emissions.

Although this change is relatively small from the perspective of the global carbon budget, this does not diminish the importance of accurately quantifying shipping emissions. A relatively small percentage correction corresponds to a substantial amount of CO_2 emissions at the global scale and is important for policy evaluation and tracking progress toward decarbonization targets. In addition, unlike CO_2, shipping emissions contribute disproportionately to air pollution in coastal regions; therefore, improving the completeness of ship emission inventories is particularly important for regional air-quality assessment.

Based on this suggestion, we have clarified these implications in the revised manuscript.

Revision:

In RESULTS *(page12, para1)*, we clarified the implication of the underestimated CO_2 emissions and emphasized that, although their share in total shipping emissions is relatively small, their inclusion improves the completeness of the inventory and supports policy-relevant emission tracking:

Although the proportion of underestimated emissions relative to total shipping emissions was relatively low, the inclusion of newly identified ships enabled us to capture emission changes caused by fluctuations in maritime trade and provided a

more complete basis for monitoring shipping decarbonization and evaluating mitigation policies.

Q18. Text revision

row 252: should it really say "only 29%"? 29% sounds like a lot

Response:

Thank you for pointing this out. We agree that the word “only” was inappropriate in this context because an underestimation of 29% is still substantial. We have therefore revised the sentence to present the comparison in a neutral manner without implying that the underestimation in the South China Sea is negligible.

Revision:

In RESULTS *(page14)*, we revised the sentence as follow:

For BC emissions, the underestimated portion reached 39% and 38% in the Indian Ocean and Red Sea, respectively, and was also as high as 29% in the South China Sea.

Q19. Particular number

row 301: elaborate more on the particle number discussion, many more references on that topic and the impact of scrubbers on particle number.

Response:

Thank you for this valuable suggestion. We agree that the original discussion of particle number (PN) emissions was too brief and did not sufficiently explain the implications of the IMO 2020 sulfur regulation or the role of scrubbers.

In the revised manuscript, we substantially expanded this discussion from three aspects. First, we clarified that changes in PM cannot be directly translated into changes in PN or particle size distributions. Second, we reviewed recent studies on the response of PN emissions to sulfur regulation and fuel switching. These studies show that although sulfur reduction generally decreases PN emissions, the magnitude of the reduction depends strongly on fuel type and combustion conditions, and therefore remains less certain than the response of PM. Third, we expanded the discussion of scrubbers. Recent onboard and laboratory

measurements indicate that scrubbers modify PN emissions in a particle-size-dependent manner, with different responses for total PN, nonvolatile PN, and larger soot-containing particles. Consequently, PM reductions cannot be directly used to infer PN reductions.

Based on these findings, we further clarified that the present study improves the representation of PM emissions through the fleet low sulfur compliance choice module, but does not explicitly simulate PN emissions or particle size distributions. We therefore identified explicit representation of PN emissions under different fuel-compliance pathways as an important direction for future development.

Revision:

In DISCUSSION *(page17, para3)*, we expanded the discussion of PN emissions and scrubbers' impact on them:

… His findings indicated that the IMO low sulfur regulation did not affect PM2.5 emissions. In addition to $PM_{2.5}$ mass emissions, PN emissions deserve further attention because ship exhaust contains large numbers of ultrafine particles that are not well represented by particle mass alone. Kuittinen et al. (2021) argued that ship emissions remain a significant source of anthropogenic PN emissions after 2020. Measurements in the Baltic Sea SECA showed that reducing fuel sulfur content from 1.5% to 0.1% decreased plume PN concentrations by 27% and ambient PN concentrations by 32% (Seppälä et al., 2021). Meanwhile, ship plumes can still contribute up to 19% of PN concentrations on days affected by shipping lanes (Kivekäs et al., 2014). The effect of scrubbers on PN is also size- and configuration-dependent. Onboard measurements on a modern cruise ship showed that a hybrid scrubber reduced total PN and nonvolatile PN, but no significant reduction was observed for particles larger than 50 nm, which commonly include BC and metal-containing particles (Kuittinen et al., 2024).

Q20. Uncertainties

row 302-307: this should be discussed earlier in discussion and elaborated on: why does uncertainties appear? How can they be reduced? Is there a way of validating at single-ship level?

Response:

Thank you for this valuable suggestion. We agree that the original uncertainty

discussion was too brief. In the revised manuscript, we expanded the discussion about uncertainty on three aspects: (1) the main sources of uncertainty, (2) possible approaches to reduce them, and (3) the feasibility of single-ship validation.

Specifically, we explained that uncertainties mainly arise from machine-learning-based supplementation of missing ship specifications and from the statistical allocation of fuel-compliance pathways in the Fleet Low Sulfur Compliance Choice Module. We also discussed that these uncertainties could be reduced through improved ship specification databases, larger training datasets, more complete scrubber and fuel-use information, and better access to ship-level operational data. Finally, we clarified that direct single-ship validation remains challenging because onboard fuel-consumption records and continuous emission measurements are rarely publicly available, but future validation could use onboard measurements, noon reports, EU MRV, or IMO DCS data where available.

Revisions:

In DISCUSSION *(page16, line356)*, we expanded the discussion about uncertainties, taking into consideration about what you mentioned in Q2, Q3, and Q11:

Despite the improvements introduced in SEIMv2.3, several uncertainties remain. First, missing vessel specifications supplemented through machine learning may introduce errors at the individual-ship level, although these uncertainties are expected to decrease after aggregation at regional and global scales. Second, FLSCCM reconstructs sulfur compliance pathways using vessel-level scrubber records combined with constrained statistical allocation because global ship-specific fuel-use observations are unavailable. Third, FLSCCM does not distinguish ULSFO and MGO, and fuels of 0.1% sulfur are represented using MGO emission factors. This simplification is expected to have limited influence on SO_2, but it may lead to some underestimation of $PM_{2.5}$ and BC emissions where residual-type ULSFO is used. Therefore, uncertainties may arise from incomplete scrubber records, assumed scrubber operation, and simplified representation of ULSFO using MGO emission factors. These uncertainties mainly affect $PM_{2.5}$ and BC estimates, while their influence on SO_2 remains limited.

To quantify the influence of scrubber-related uncertainty on particulate emission estimates, we conducted sensitivity analyses using different scrubber removal efficiency assumptions. Under the low-, baseline-, and high-removal scenarios,

corresponding to $PM_{2.5}$ removal efficiencies of 10%, 37%, and 70% and BC removal efficiencies of 10%, 30%, and 60%, global $PM_{2.5}$ emissions varied by +26.6% and −32.6% relative to the baseline case, while BC emissions varied by +11.5% and −17.3%. Despite this uncertainty, all scenarios consistently show higher particulate emissions than inventories that neglect or incompletely represent HFO-with-scrubber pathways, indicating that the identified underestimation is robust to plausible variations in scrubber removal performance.

In SUPPLEMENTARY INFORMATION, we added:

S4. Sensitivity analysis of FLSCCM structural assumptions

To evaluate the uncertainty associated with structural assumptions in FLSCCM gap filling, sensitivity analyses were conducted by varying the DWT threshold and construction-year treatment. For DWT sensitivity, the baseline age treatment (construction year threshold ⩾2006) and scrubber particle removal efficiencies were maintained, while three DWT thresholds (40,000, 50,000, and 60,000 DWT) were tested.

The observed scrubber fleet remained unchanged across scenarios (4,285 vessels with active AIS records included in the 2024 emission calculation), as DWT thresholds were applied only to database-unmatched vessels. In contrast, FLSCCM-inferred scrubber vessels decreased from 8,638 to 5,639 as the threshold increased from 40,000 to 60,000 DWT, resulting in total HFO consumption changes from 116.7 to 103.3 Mt. Relative to the baseline threshold (50,000 DWT), $PM_{2.5}$ emissions changed by +2.1% and −3.1%, while BC emissions changed by +0.3% and −0.5%, under the 40,000 and 60,000 DWT scenarios, respectively. These results indicate that the estimated particulate emissions were relatively insensitive to the tested DWT threshold range.

Additional sensitivity tests evaluated the influence of incomplete construction-year information. Extending the construction-year threshold from ⩾2010 to ⩾2006 increased the identified scrubber fleet from 9,950 to 11,931 vessels and increased HFO consumption from 98.1 to 111.3 Mt. However, $PM_{2.5}$ and BC emissions increased by only 5.4% and 1.0%, respectively. Within the ⩾2006 scenarios, alternative treatments of missing construction-year information resulted in small variations, with $PM_{2.5}$ emissions ranging from 767.4 to 772.3 kt and BC emissions ranging from 70.6 to 70.7 kt. These results suggest that the tested treatments for incomplete vessel-age information have limited influence on the global emission estimates.

Response to Reviewer #2:

OVERALL

The paper describes model updates to the Shipping Emission Inventory Model and their consequences for global ship emission data. The authors give the impression that their findings are quite general and not specific to their own model development only. I think this paper would be much better suited for a model development journal like Geoscientific Model Development than for Atmospheric Chemistry and Physics. In any case, it is important that the authors make clear from the beginning that they mainly speak about correcting deficiencies in their own model system rather than general findings about global shipping emission calculations. Also, the so-called "single compliance choice assumption" for selecting a technology to comply with global sulfur limits is a simplification that is potentially highly inaccurate and should not be used if better information is available from a ship technical database.

The paper also needs major improvements in the English language and a more precise description of the findings. A real discussion of the uncertainties would also help to make the paper more relevant for a wider audience.

Response:

We sincerely thank the reviewer for this constructive and insightful comment. We agree that the original manuscript did not sufficiently clarify the scope of our conclusions and the limitations of the Fleet Low-Sulfur Compliance Choice Module (FLSCCM). Our intention was not to imply that all global shipping emission inventories exhibit the same magnitude of underestimation. Instead, this study identifies and quantifies two specific sources of uncertainty within the SEIM framework, including incomplete fleet representation and simplified sulfur compliance pathway assumptions.

In the revised manuscript, we have therefore clarified that the reported emission differences represent the improvements achieved by SEIMv2.3 relative to SEIMv2.2 rather than universal biases across all shipping emission models. Furthermore, we have substantially revised the FLSCCM framework by incorporating vessel-level scrubber installation records as observational constraints. The revised observed-first framework retains database-identified scrubber-

equipped vessels and applies structurally constrained gap filling only to vessels without observed information, thereby avoiding unnecessary reliance on statistical assumptions when direct observations are available.

We have also added sensitivity analyses to evaluate uncertainties associated with scrubber pathway reconstruction and particle removal assumptions. The results demonstrate that the identified particulate emission underestimation remains robust under plausible variations in scrubber removal efficiencies and improve the uncertainty characterization and applicability of the proposed framework.

Major Specific Comments

Q21. Line 27

Line 27: You need to specify which species you refer to. The number is different for CO2, NOx, SO2 … .

Response:

Thank you for pointing this out. Specifically, the reported 1.7 times rise refers to CO_2 emissions, while other pollutants show different changes. In the revised manuscript, we have specified the pollutant when presenting the emission increase and avoided using a generalized description.

Revisions:

In INTRODUCTION *(page2, para1)*, we changed 'atmospheric' to 'CO_2':

As one of the hard-to-abate sectors (Sharmina et al., 2021), global shipping experienced a threefold increase in cargo turnover and a 1.7-fold increase in carbon dioxide (CO_2) emissions from 1970 to 2021 (Wang et al., 2025).

Q22. Line 31

Line 31: The number of 17% is highly dependent on the global development of CO2-emissions in all non-shipping sectors. You need to mention this

Response:

Thank you for pointing this out. We have revised the sentence to clarify that this value represents the projected contribution of shipping to global anthropogenic CO_2 emissions under a business-as-usual scenario, rather than an absolute increase in shipping emissions alone.

Revisions:

In INTRODUCTION *(page2, para1)*:

Under a business-as-usual scenario without additional mitigation measures, international shipping was projected to account for approximately 17% of global CO_2 emissions by 2050 (Cames et al., 2015), which reflects both the future development of international shipping emissions and the emission trajectories of other sectors.

Q23. Line 34-37

Lines 34-37: You need to say between which years these reductions took place, i.e. when the policies were implemented.

Response:

Thank you for this helpful suggestion. We have revised the text by adding the implementation years of the Turkish Straits/Marmara Sea sulfur control measures and China's DECA policy.

Revisions:

In INTRODUCTION *(page2, para1)*:

Studies evaluating regional sulfur regulations have demonstrated their effectiveness in reducing ship-related SO_2 emissions. The sulfur emission control measures implemented in the Marmara Sea and Turkish Straits since 2006 reduced SO_2 concentrations by 90% in Istanbul (Viana et al., 2015) while China's Domestic Emission Control Area (DECA) policy, implemented progressively during 2016–2019, achieved an emission reduction by 33% within 200 nautical miles away from the Chinese mainland's territorial sea baseline (Wang et al., 2021).

Q24. Specify ship emission models

Line 42/43 and line 45: Which ship emission model(s) are meant here? It seems you speak about ship emission models in general, but you do not mention any.

Response:

Thank you for this helpful comment. In the revised manuscript, we have clarified this point by explicitly referring to representative AIS-based ship emission models, including STEAM, MariTEAM, and MoSES, and by avoiding generalized statements implying that all inventories use identical assumptions.

Revisions:

In INTRODUCTION *(page3, para1)*, we added:

Therefore, existing models such as STEAM, MariTEAM, and MoSES generally

rely on vessel characteristics, fleet statistics, regulatory assumptions, or predefined scenarios to estimate fuel use and emission-control practices (Schwarzkopf et al., 2021; Jalkanen et al., 2009; Kramel et al., 2021). These assumptions introduce two major uncertainties in global ship emission inventories: incomplete representation of the evolving fleet and insufficient characterization of fuel compliance pathways.

Q25. Number of scrubber-equipped ships

Line 50: It is unclear how all of these numbers go together. When 65% of the fuel sales are (high sulfur) HFO, there must be many ships being equipped with scrubbers. What is the number of scrubber ships in your model system? It must be much higher than 800. And how does the number of 65 % HFO agree with Fig. S4, which shows lower values for HFO for three ship size classes?

Response:

Thank you for raising this important point. We agree that the original manuscript did not clearly distinguish between fleet-level fuel statistics, regional scrubber observations, and the vessel-level scrubber representation used in SEIMv2.3, which may have caused confusion.

The reported 65% HFO share from IMO fuel statistics represents a broad fuel category at the global fleet level and cannot be directly converted into the number of scrubber-equipped ships, because HFO consumption may include different residual fuel categories and does not provide ship-level compliance information. Similarly, the 800 scrubber-equipped vessels reported by Lunde Hermansson et al. (2024) represent a regional observation in the Baltic Sea rather than the global scrubber-equipped fleet.

In the revised SEIMv2.3 framework, ship-level scrubber installation records from Clarksons Research were incorporated as observational constraints. For the 2024 emission calculation, 4,285 vessels with effective scrubber installation records and active AIS activities were identified. For vessels without observed scrubber records, the FLSCCM applied structurally constrained gap filling based on vessel type, DWT, and construction year. This procedure additionally assigned 7,647 vessels to the HFO-with-scrubber pathway. Therefore, the total number of scrubber-associated vessels represented in SEIMv2.3 was 11,932 in 2024.

We have revised the manuscript to clarify the distinction between IMO-reported fuel statistics, regional scrubber observations, vessel-level scrubber installation

records, and modeled compliance pathways.

Revisions:

These revisions have been incorporated in Section 2.3 (*page8*) and the Discussion section *(page16)*.

Q26. Clarification of newly identified vessels and distinction from newly built ships

Line 56: What are "newly build ships" in your terminology? Are these ships younger than 5 years old? Why should their number increase over time?

Line 196: "SEIMv2.3 primarily supplemented activities of ships constructed between 2000 and 2010": So these are not "new" ships (i.e. recently built) as you stated before.

Line 204: Again, you talk about "new ship deployment", while the age of most ships was between 14 and 24 years in 2024 as you state in line 196.

Figure 5: You should make clear that these are signals from ships that were not in your database before. It does not mean that these ships are new and it also does not mean that other ship emissions models did not include these ships.

Line 275/276: Again, you should make clear that this increase of ships in the STSD does not mean that all of these are newly built ships.

Response:

Thank you for these valuable comments. We agree that the terminology used in the original manuscript was ambiguous and could lead to misunderstanding.

In this study, the term "newly identified ships" refers to vessels detected from daily AIS data but absent from the previous ship technical specification database (STSD). These ships do not necessarily represent newly built ships. Instead, they mainly represent previously missing vessels whose technical information was not included or updated in the existing database.

The number of newly identified vessels may increase over time because global ship fleets continuously evolve, whereas static technical databases may not be updated at the same temporal resolution as operational AIS records. Newly constructed ships and delayed database updates may all contribute to discrepancies between active AIS-based fleets and static technical databases.

The age distribution analysis further supports this interpretation. SEIMv2.3

primarily supplemented vessels constructed between 2000 and 2010, indicating that the dynamic STSD update mainly improved the representation of existing vessels missing from the previous database rather than only capturing recently constructed vessels.

We have revised the terminology throughout the manuscript and replaced ambiguous expressions such as “new ships”, “newly built ships”, and “new ship deployment” with “newly identified ships”, “previously missing ships”, or “dynamic STSD updates” where appropriate.

Revisions:

In INTRODUCTION *(page3, para4)*, we clarified the terminology of ‘newly identified ships’ :

Another challenge facing ship emission inventories is the timely incorporation of newly identified ships into ship technical databases. In this study, newly identified ships include both newly delivered vessels and existing vessels absent from the original static database.

In other sections, we replaced ambiguous expressions with “newly identified ships”.

Q27. Update STSD with daily AIS

Line 66/67: Why wasn’t this done before? And: Is the technical information included in the AIS data detailed enough to improve the emission model? Typically, the AIS data is not as detailed as that in a ship characteristics database.

Response:

Thank you for this important comment. Previously, the static ship technical database was used as the primary fleet description because it provided relatively comprehensive vessel specifications. However, as ship fleets evolve, static databases may gradually miss newly entering or previously unrecorded vessels, leading to incomplete fleet representation. In the revised analysis, we found that neglecting these newly identified ships resulted in an unrealistic declining trend in global ship emissions, inconsistent with previous studies and observed changes in maritime activity. Therefore, we introduced an AIS-assisted updating framework to improve fleet completeness.

We also agree that AIS data alone do not provide all technical parameters required for emission calculations. In our framework, AIS data are not used to replace

dedicated ship databases. Instead, AIS-based records are first used to identify missing vessels and provide available information. Missing or coarse technical parameters, including ship type, DWT, and main engine power, are subsequently reconstructed using machine-learning models trained on available vessel information such as length and width. The prediction performance of these models is described in Section 2.2 and Table S1, including ship-type classification accuracy (0.82), DWT prediction performance ($R^2 = 0.92$), and main engine power prediction performance ($R^2 = 0.80$). We have revised the manuscript to clarify this workflow.

Q28. Daily AIS-based activity processing

Line 77: It is unclear what you mean by AIS data that is updated daily instead of yearly. Isn't the AIS data always historic, i.e. collected over a longer time and then used as input for SEIM?

Response:

Thank you for this important clarification. Our intention was to describe the temporal resolution and operational capability of the SEIMv2.3 framework. Specifically, AIS data are collected and processed at a daily resolution, allowing the framework to update ship activity and calculate emissions with only a short delay (approximately two days) when applied operationally. The emission results presented in this manuscript are reported up to 2024 because this study focuses on demonstrating and validating the framework using a complete historical period, rather than providing a real-time monitoring product. We have revised the manuscript to clarify that SEIMv2.3 provides a near-real-time emission estimation capability based on daily AIS processing and dynamic fleet updating.

Revisions:

In METHOD *(page4, para4)*,

… three key improvements: (1) Daily AIS-based activity processing. AIS records are processed at a daily temporal resolution rather than annual aggregation, enabling ship activity reconstruction and emission estimation with a short delay and providing the technical basis for near-real-time emission updates. Detailed methods are provided in Sect. 2.1.

Q29. Ship identification

Line 81: You should explain how ships were identified in SEIM v2.2? Is there a ship characteristics database that represents the fleet composition in a certain year?

Response:

Thank you for this comment. We agree that the overview of methodological improvements was not sufficiently clear and may have caused confusion regarding the role of AIS data in SEIMv2.2 and SEIMv2.3. We have revised the overview section to clarify that daily refers to the temporal resolution of AIS processing and fleet updating rather than real-time AIS availability. Besides, the detailed technical parameter reconstruction procedure, including machine-learning-based supplementation of missing specifications, is described in Sect. 2.2.

We have the Ship Technical Specification Database (STSD) to represent the total fleet composition. Then we use ships' MMSI or IMO numbers to match AIS signals with our STSD to gain the active fleet composition of a certain year.

Revisions:

In METHOD *(page4, para4)*,

(2) Daily identification of previously missing ships. In SEIMv2.2, ship activity records were matched with our Ship Technical Specifications Database (STSD), which may become incomplete as new ships enter operation or previously unrecorded ships become active. SEIMv2.3 uses daily AIS activity records to identify previously missing ships and update the STSD, with detailed procedures described in Sect. 2.2.

Q30. Data source

Line 98: Please explain where or from whom you get the AIS data.

Response:

The AIS data used in this study were obtained from a commercial AIS data provider and include global dynamic and static AIS records. Dynamic AIS data provide vessel trajectories, including timestamps, positions, and navigation information, while static AIS data provide basic vessel information, such as IMO number, MMSI, ship type, length, width, and draft. We have clarified the AIS data source and description in Section 2.1 of the revised manuscript.

Q31. Three-day AIS processing

Line 112: When you need to perform a three-day AIS processing to get reliable data, how does this agree with your "near-real-time" approach?

Response:

Thank you for this important comment. We agree that the original description did not sufficiently clarify the relationship between three-day AIS processing and the near-real-time capability of SEIMv2.3. Here, the three-day AIS processing window does not represent a three-day delay in emission reporting, but rather a trajectory reconstruction procedure that uses temporally adjacent AIS signals to compensate for delayed or incomplete position records.

In the context of global ship emission inventories, near-real-time refers to the capability of updating emission estimates within a short time lag based on continuously collected ship activity data, rather than real-time vessel monitoring at the operational level. After accounting for the required AIS preprocessing window, SEIMv2.3 can theoretically provide ship emission estimates with only a short delay (approximately several days), which represents a substantial improvement over conventional inventories based on monthly or annual activity aggregation.

We have revised the manuscript to clarify this definition and avoid confusion between near-real-time emission inventory updating and real-time AIS monitoring.

Revisions:

In METHOD *(page5, para3)*, we clarified our definition of 'near-real-time' emission calculation:

… actual sailing trajectories. This preprocessing window is required for trajectory reconstruction from delayed AIS transmissions, while SEIMv2.3 enables near-real-time emission updates with a short delay after AIS data acquisition. The overall effect…

Q32. Information gap between STSD and AIS

Line 128: When I understand this correctly, STSD contains more information, e.g. on engine power, compared to AIS data. How do you solve this problem?

Response:

Thank you for this important comment. We agree that AIS and STSD contain different types of information and that AIS cannot directly replace the technical database. In SEIMv2.3, daily static AIS data serves as a source for identifying vessels that are active but missing from the existing STSD.

For these previously missing vessels, available information from AIS (e.g., vessel

identifiers, dimensions, and ship-related attributes) is combined with information from existing STSD records to reconstruct missing technical parameters using machine-learning models (explained in *Sect2.2 page6* and *Supplementary Information S2. Model performance of different algorithms for update STSD*). The reconstructed parameters are then incorporated into the updated STSD for emission calculation.

Q33. Figure2

Fig. 2: Why are there only 5 ship types? In line 131 you say that SEIM has 14 ship types. And why are there more than 400,000 ships in the database when there are only approx. 100,000 being active (see line 82)? Why is there a sharp increase in the number of identified new ships between July and October 2024?

Response:

Thank you for this helpful comment.

The five ship categories shown in Fig. 2 represent the dominant categories identified from daily static AIS information for newly detected vessels. Other ship categories newly identified in the AIS data occurred in relatively small numbers and were therefore not displayed separately. These categories represent the initial classification provided by static AIS rather than the final 14 ship types used in SEIMv2.3.

The cumulative number of more than 400,000 ships in Fig. 2 represents the historical accumulation of vessels identified during the daily updating process, rather than the number of active ships in a specific year. Some vessels may exist in the technical database but have no AIS activity during a given year due to reasons such as maintenance, repair, lay-up, or inactivity. Therefore, the cumulative database size is expected to exceed the annual active fleet size.

The increase in newly identified ships during July–October 2024 was caused by a higher number of newly received AIS static records during this period, which improved the detection of previously missing vessels. It does not represent a sudden increase in global ship fleet size.

Revisions:

In METHOD *(page6)*, we revised the caption of Fig.2:

Figure 2: The composition of ship types (a), deadweight tonnage (b), and main

engine power (c) for daily newly detected ships. The ship categories shown here represent the dominant categories identified from processed AIS static information for newly identified ships and do not correspond directly to the 14 emission categories used in SEIMv2.3.

Q34. AIS cleaning

Line 90: What is meant by "well cleaned"? How much data is omitted during the cleaning process?

Response:

Thank you for this helpful comment. We agree that the term "well-cleaned" was not sufficiently specific. In the revised manuscript, we have clarified this term by describing the quality-control objectives and quantitative effects of the AIS preprocessing procedure.

Specifically, a well-cleaned AIS dataset refers to a dataset after removing unreliable signals while preserving valid vessel activities. The preprocessing procedure was designed to eliminate abnormal records, including invalid geographic coordinates, physically inconsistent vessel movements, and temporal anomalies, while retaining realistic signals such as vessel activities along rivers. As illustrated in Fig. 1 and Fig. S1, the preprocessing effectively removed abnormal signals drifting over land, reduced unrealistic inland signal accumulation, and prevented abnormal temporal increases in AIS signal numbers.

The preprocessing retained 91.6–92.2% of the raw AIS signals, corresponding to an average removal rate of approximately 8.2%.

Q35. Line142

Line 142: What you call here the "information of [a] newly detected ship" is a guess (or interpolation) based on existing information. Similar methods were already applied before in models like MoSES (Schwarzkopf et al., 2021) and STEAM (Jalkanen et al. 2009,2012).

Response:

Thank you for this important clarification. We agree that the term "supplemented information of newly detected ships" was imprecise and may imply that the technical parameters were directly observed from AIS data. In fact, AIS provides partial vessel information, while missing technical specifications were reconstructed using machine-learning models trained on ships with known

characteristics. Similar parameter estimation approaches have been applied in previous ship emission models, including MoSES (Schwarzkopf et al., 2021) and STEAM (Jalkanen et al., 2009, 2012).

The novelty of our approach is not the parameter estimation itself, but the integration of daily AIS-based vessel identification with technical parameter reconstruction to dynamically update the ship technical database and improve fleet completeness. We have revised the manuscript to clarify this point.

Revisions:

In METHOD *(page7, para1)*, we added:

Ultimately, the reconstructed technical information of newly detected ships was merged with the original STSD, forming a dynamically updated STSD as shown in Fig. 2.

Q36. Emission factor for BC

Line 150-153: BC is not included in Table S3, so what are the reductions based on? Is BC taken as a constant fraction of PM2.5? Is this a correct assumption? BC emissions should depend strongly on combustion conditions rather than on S-content or ash-content of the fuel.

Response:

Thank you for this important comment. We agree that the original manuscript did not clearly present the basis of BC emission estimation, which may have caused confusion. BC emissions in SEIMv2.3 are not derived as a constant fraction of $PM_{2.5}$ emissions. Instead, BC emissions are calculated using independent emission factors that vary with fuel type, engine category, and engine load conditions.

In the revised manuscript, we reorganized the supplementary emission factor tables to improve clarity. Specifically, the emission factor information for BC has been added as Table S3, while the original Table S3 has been moved to Table S4. The revised tables now separately provide emission factors for different pollutants and fuel types, including BC, $PM_{2.5}$, and SO_2-related parameters.

We have also revised the description in Sect. 2.3 to clarify that reductions in BC emissions after switching from HFO to low-sulfur fuels result from differences in fuel- and combustion-condition-dependent emission factors rather than directly from sulfur content changes.

Revisions:

In SUPPLEMENTARY INFORMATION, we added Table S3 Energy-based emission factors for BC (g/kWh)

Q37. LNG option

Line 158-160: Isn't there LNG as shipping fuel included in the technical ship descriptions (STSD)?

Response:

Thank you for this helpful comment. We agree that LNG-fueled vessels are included in the ship technical database (STSD) and are considered in the emission calculation framework. However, the purpose of FLSCCM is to reconstruct sulfur compliance pathways after the IMO 2020 global sulfur limit, rather than to predict all possible future fuel transitions.

Although LNG is an alternative compliance pathway, its adoption has remained limited compared with low-sulfur fuel switching and HFO use with scrubbers, particularly for the existing global fleet. Therefore, FLSCCM only allocates the dominant post-2020 sulfur compliance pathways between low-sulfur fuels (VLSFO/MGO) and HFO with scrubbers. LNG-fueled vessels identified in the STSD continue to be calculated using LNG-specific fuel consumption and emission factors in the emission inventory.

Revisions:

In METHOD (page8, para1):

(3) Using alternative fuels such as LNG, which can substantially reduce emissions but currently represent a limited share of global marine fuel consumption. Therefore, FLSCCM focuses on the allocation of the two dominant sulfur compliance pathways: low-sulfur fuel switching and HFO with scrubbers.

Q38. Figure S3

Line 164: Is Figure S3 based on IMO data? Please give the correct reference in the supplements.

Response:

Thank you for this comment. We confirm that Figure S3 was derived from IMO

fuel consumption statistics. The figure was generated by aggregating and visualizing the IMO-reported fuel consumption data for 2022 by ship type and DWT category. We have added the corresponding reference to the Supplementary Information and clarified the data source in the figure caption.

Revisions:

Figure S3. Distribution of fuel consumption by DWT range in 2022 for tanker (a), bulk carrier (b), and container ship (c), based on IMO-reported fuel consumption statistics (IMO, 2024).

Q39. FLSCCM methodology and sensitivity analysis

Line 167-170: This threshold is very artificial and most likely not realistic. It seems unrealistic that only big ships use scrubbers while smaller ones don't. Please compare to existing scrubber installation data. Also, Figure S4 shows that HFO consumption is reduced with the size of the ships. Please also explain what you mean with MGO_5E3 and MGO_1E3.

Line 254/255: Indeed, there are multiple choices for being compliant with the global low sulfur rules. Your model, that assumes that only the biggest ships use scrubbers while smaller ones use LSFO or VLSFO may not be very accurate in terms of spatial distribution of PM2.5 and BC emissions.

Lines 305-307: You could test the implications of your assumptions for the spatial and temporal emission distribution. For example, you could check the size distribution of all known scrubber ships and assume that smaller ships may also use scrubbers.

Response:

Thank you for this important comment. We agree that the original description of the DWT threshold was insufficient and could give the impression that scrubber adoption was determined solely by vessel size. In the revised manuscript, we improved our previous FLSCCM with existing scrubber installation data. The DWT threshold is no longer used as an independent scrubber prediction criterion. Instead, FLSCCM adopts an observed-first framework, where vessel-level scrubber records obtained from Clarksons Research are first incorporated as observational constraints, and the DWT threshold is applied only to database-unmatched vessels during residual gap filling.

We analyzed the characteristics of the observed scrubber-equipped fleet using the vessel-level scrubber database. The results showed that scrubber installations were

concentrated among large vessels, with approximately 69.5% of observed scrubber-equipped vessels having DWT above 50,000 tonnes. Therefore, a 50,000 DWT threshold was introduced as a structural constraint for gap filling. Importantly, this threshold does not exclude smaller vessels with confirmed scrubber installations; observed scrubber-equipped vessels were directly retained regardless of their DWT.

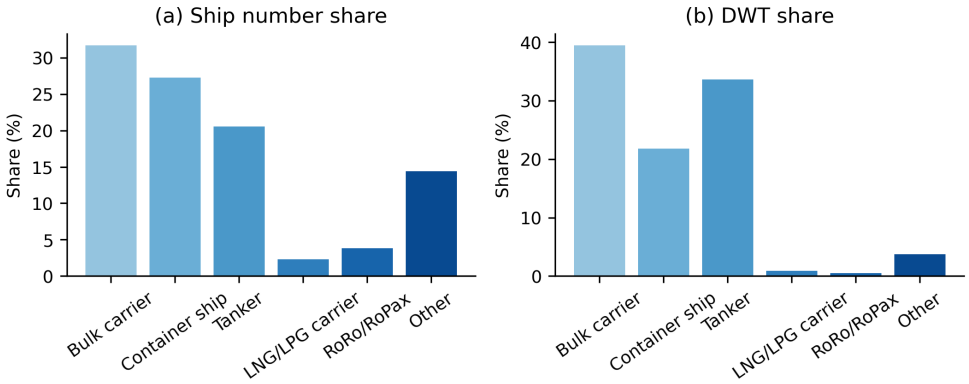


Figure S7. Ship type distribution of scrubber-equipped fleet in 2024

We further evaluated the uncertainty associated with this threshold through sensitivity analyses using alternative DWT thresholds (40,000 and 60,000 DWT). The resulting changes in $PM_{2.5}$ and BC emissions were limited, demonstrating the robustness of the revised FLSCCM to reasonable variations in the screening criterion.

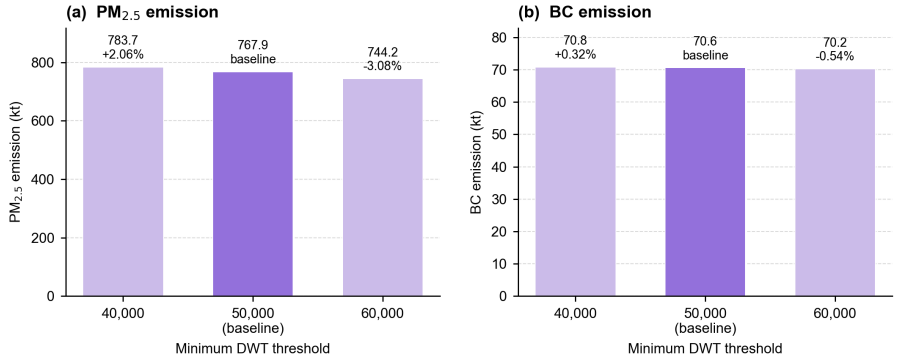


Figure S12. Sensitivity of emission estimates to minimum DWT screening threshold

Revisions:

In METHOD (page8, para3), we explained FLSCCM's methodology for ships unmatched with scrubber installation records:

Analysis of the scrubber installation database in Supplementary Information showed that scrubber adoption was concentrated among major cargo ship categories and larger vessels. Bulk carriers, containers, and tankers together accounted for approximately 79.5% of identified installations (Fig. S8). Therefore, gap filling was restricted to bulk carriers, containers, oil tankers, and chemical tankers. In addition, ship size and construction year were introduced as screening criteria because scrubber installation is economically favoured for ships with higher fuel consumption and sufficient remaining service life. Database-unmatched ships were considered eligible candidates only when they had a DWT of at least 50,000 tonnes and a construction year later than 2006 when reliable information was available. These criteria were applied only to ships without scrubber installation records during the residual gap-filling procedure and did not affect ships with confirmed scrubber installations. The threshold settings were determined based on the characteristics of the observed scrubber-equipped fleet. For vessels identified or assigned as following the HFO-with-scrubber pathway, SEIMv2.3 applies a baseline $PM_{2.5}$ removal efficiency of 37% and a BC removal efficiency of 30% in the emission calculation. Uncertainties associated with scrubber assignment criteria and removal efficiencies are further evaluated through sensitivity analyses in the Supplementary Information and the Discussion section.

In SUPPLEMENTARY INFORMATION, we removed the original Fig. S4. Meanwhile, we added Fig. S6-S9 to illustrate the fleet composition of scrubber-equipped ships and Fig. S11-S12 to present the sensitivity analysis of different DWT and build year threshold.

Q40. Line174

Line 171-174: While the relative values match well in 2020, this gets worse in the following years. What is the reason for the increase in HFO in the IMO data after 2020? Is this caused by an increase in scrubber installations?

Response:

Thank you for this insightful comment. We agree that the increase in IMO-reported HFO consumption after 2020 reflects the continued use of HFO-based compliance pathways, including vessels equipped with exhaust gas cleaning systems. However, IMO fuel statistics do not directly distinguish HFO consumption associated with scrubber operation from other residual fuel categories.

In the revised manuscript, we incorporated vessel-level scrubber installation records into FLSCCM to better represent the HFO-with-scrubber pathway. The

reconstructed fuel consumption was further compared with IMO statistics. The results show that SEIMv2.3 reproduces the overall fuel consumption magnitude and the major transition from HFO-dominated fuel use before IMO 2020 toward low-sulfur fuel pathways after 2020.

Differences in the relative HFO share after 2020 may arise from differences in fuel classification and methodological approaches. IMO statistics are based on aggregated fuel reporting, whereas SEIMv2.3 explicitly separates HFO-with-scrubber, VLSFO/MGO/MDO, and LNG pathways using vessel-level activity and technology information. Therefore, IMO statistics were used as an external fleet-level diagnostic rather than a direct calibration target.

Q41. Line 179 and Fig. S5

Line 179-189 and Fig. S5: It seems that for some ship types, no new ships are detected, e.g. RoRo, chemical tankers, … (Fig. S5). This may lead to uncertainties.

Response:

Thank you for this valuable comment. We agree that the number of newly identified vessels varies among ship categories and that limited identification in certain categories may introduce uncertainty. We further analyzed the relationship between newly identified vessels and the composition of the global active fleet.

In this study, newly identified vessels refer to vessels detected from daily AIS data but absent from the existing STSD, rather than newly built vessels. Therefore, the number of newly identified vessels depends on both the size of each ship category and the completeness of the original technical database.

For ship categories with limited newly identified vessels, such as Ro-Ro and chemical tankers, their contribution to the global active fleet is relatively small. In 2024, Ro-Ro and chemical tankers accounted for approximately 1.7% and 4.1% of active vessels, respectively. Therefore, the limited number of newly identified vessels in these categories primarily reflects their smaller fleet size rather than a systematic limitation of the identification framework.

In contrast, major cargo categories with larger fleet populations, such as bulk carriers, container ships, and general cargo ships, accounted for much larger fractions of the global active fleet (14.2%, 5.3%, and 12.5%, respectively) and constituted the dominant groups among newly identified vessels. Therefore, the dynamic STSD update primarily

improved the representation of ship categories with the greatest contribution to global shipping activity.

We have clarified this point in the revised manuscript and added the comparison between newly identified vessels and global active fleet composition in the Supplementary Information.

Q42. Figure 4

Figure 4: Why don't you use simple bar plots? They would be easier to read. And why is there no increase in "Fleet number_IMO" by weight bin for ships other than < 10kt?

Response:

Thank you for this helpful suggestion. We agree that simple bar plots provide a straightforward comparison of absolute numbers. However, the original purpose of Fig. 4 was not only to compare the number of vessels but also to illustrate the composition of newly identified vessels relative to the existing fleet. Therefore, instead of using bar plots, we have revised Fig. 4 into a more readable donut-chart format. The inner ring represents the proportion of newly identified vessels relative to the original STSD fleet, while the outer ring illustrates the distribution of newly identified vessels by ship type, build period, and weight bin. This design allows both fleet expansion and structural characteristics of newly identified vessels to be visualized simultaneously.

Revisions:

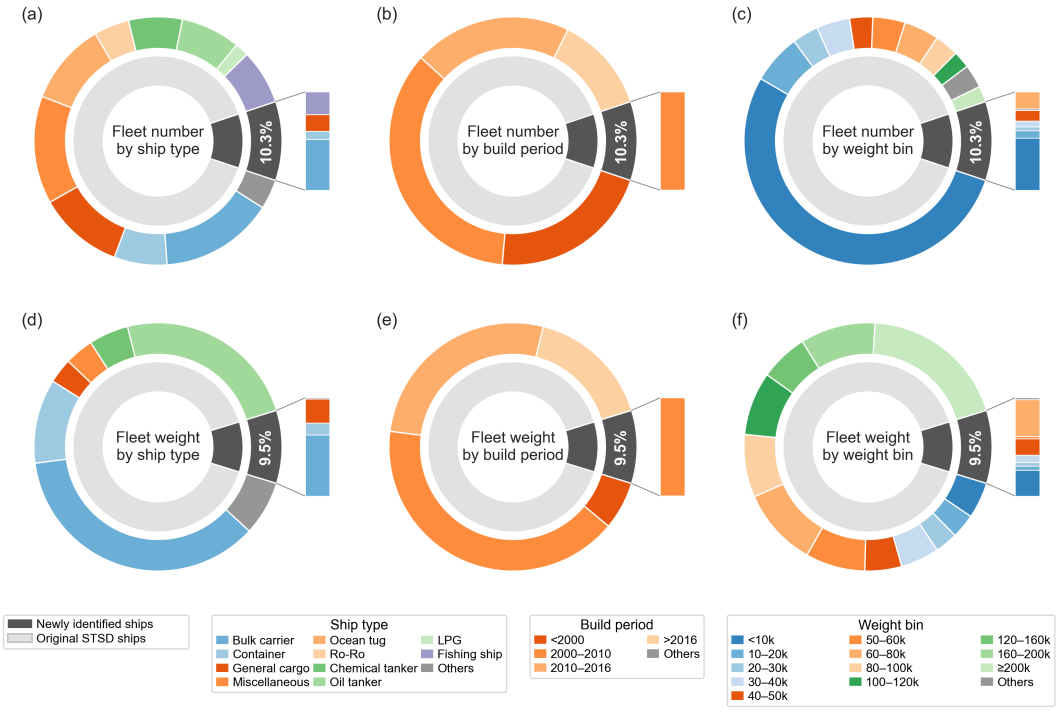


Figure 4: The composition of ship number and total weight of the active global fleet by SEIMv2.3 and SEIMv2.2. Panels (a–c) show fleet number by ship type, build period, and weight bin; panels (d–f) show the corresponding distributions of total fleet weight. The inner ring distinguishes ships contained in the original STSD (light grey) from ships newly identified through the dynamic STSD updating procedure (dark grey). The outer coloured ring shows the category composition of the original STSD fleet, and the linked stacked bar shows the composition of newly identified vessels using the same colour scheme.

Q43. Line215

Line 215: Which are these different models? Please introduce them here. EDGAR and CEDS are global emission models, not necessarily ship emission models. You should find out, which ship emission model is behind the data (e.g. STEAM).

Response:

Thank you for this important comment. We agree that the original wording was ambiguous because EDGAR and CEDS are global emission inventories rather than dedicated ship emission models. We have revised the description to clarify the methodological differences among the datasets used for comparison.

Specifically, the shipping emissions in EDGAR are based on the Ship Traffic Emission Assessment Model (STEAM3) framework for international shipping activity and spatial allocation, whereas CEDS is a global emission inventory based on a top-down fuel/activity approach rather than an AIS-based ship emission model. In addition, the comparison includes AIS-based ship emission inventories, including SEIMv2.2 and SEIMv2.3, to evaluate the impact of the proposed

framework improvements.

Revisions:

In RESULTS (*page11, para1*):

For comparison, available international shipping emission estimates from the Emissions Database for Global Atmospheric Research (EDGAR) and the Community Emissions Data System (CEDS) were used as independent global references to evaluate the temporal consistency of SEIMv2.3. For EDGAR, international shipping emissions were derived using the Ship Traffic Emission Assessment Model (STEAM3) framework (Crippa et al., 2024).

Q44. Figure 6

Figure 6: Please explain STSD and FLSCCM in the caption.

Revisions:

Figure 6: Comparison of annual global ship emissions results from SEIMv2.3, SEIMv2.2, and other emission inventories from 2019 to 2024. SEIMv2.2 results were sourced from (Yi et al., 2024a). SEIMv2.3 represents the method presented in this paper. EDGAR represents the emission results from Emissions Database for Global Atmospheric Research, sourced from (Crippa M., 2025), whose shipping emission results are based on the STEAM3 ship emission model. CEDS represents the emission results from Community Earth atmospheric Data Systems, sourced from (Hoesly, 2025). STSD-2024 and FLSCCM-2024 denote configurations incorporating updated fleet information and reconstructed sulfur compliance pathways, respectively.

Q45. CEDS

Lines 222/223: Why is there no PM2.5 data from CEDS?

Response:

Thank you for this comment. We agree that the original manuscript did not clearly explain the difference in pollutant coverage among comparison inventories. The available CEDS dataset used in this study does not provide a directly comparable $PM_{2.5}$ estimate for international shipping emissions. Therefore, CEDS $PM_{2.5}$ data were not included in Figure 6.

We have clarified in the revised manuscript that the comparison inventories have different pollutant coverage and that they were used only where comparable emission estimates were available.

Revisions:

In RESULTS (*page11, para1*):

For comparison, available international shipping emission estimates from the Emissions Database for Global Atmospheric Research (EDGAR) and the Community Emissions Data System (CEDS) were used as independent global references to evaluate the temporal consistency of SEIMv2.3.

Q46. Daily update of STSD, Fig. 7, and Fig. 8

Line 232: In order to improve SEIM, it is not necessary that ships are identified on a daily basis. It is only necessary that they are not neglected and that the technical information about the is reliable.

Line 256-265: See my comment regarding line 232: I cannot follow the argument that a daily update is necessary when you run the model in a hindcast mode as described here (i.e. calculating the 2024 emissions some time in 2025 or later). Therefore, I do not see that Fig 7 and Fig. 8 contain useful information.

Response:

Thank you for this important comment. We agree that the primary contribution of the updated framework is not the daily update frequency itself, but the improved completeness and timeliness of global ship fleet representation. For historical emission reconstruction, a periodically updated ship technical database may also reduce missing vessel information. However, the daily AIS-based updating framework provides the technical foundation for future near-real-time emission estimation by enabling continuous fleet information updates with short delays after AIS data acquisition.

Following this suggestion, we have revised the manuscript to clarify that the emission improvements in SEIMv2.3 mainly result from two aspects: (1) improved representation of previously missing ships through dynamic fleet database updating, and (2) improved characterization of sulfur compliance pathways through the FLSCCM, rather than from the daily temporal resolution alone.

We have also revised the interpretation of Fig. 7 and Fig. 8. Figure 7 now focuses on the spatial heterogeneity and potential sources of emission underestimation. Specifically, regional emission discrepancies are compared with sailing distances associated with newly identified ships and scrubber-equipped ships. The results show that CO_2 and NOx discrepancies are mainly associated with incomplete fleet

representation, whereas $PM_{2.5}$ and BC discrepancies are more strongly linked to incomplete representation of HFO-with-scrubber pathways. For example, the South China Sea contributed the largest shares of global CO_2 and NOx underestimation (14.2% and 14.4%, respectively), corresponding to high STSD-added vessel activity. In contrast, the Indian Ocean showed the largest contributions to $PM_{2.5}$ and BC underestimation (15.1% and 14.7%, respectively), associated with substantial scrubber-equipped vessel activity.

Regarding Fig. 8, we agree that the daily emission differences should not be interpreted as evidence that daily updating is required for retrospective emission inventories. Therefore, we have revised the text to clarify that Fig. 8 illustrates the temporal accumulation of emission discrepancies caused by incomplete fleet representation and simplified compliance pathway assumptions throughout 2024, rather than the direct benefit of daily updating frequency. The revised analysis shows that the emission discrepancies between SEIMv2.3 and SEIMv2.2 remained consistently positive and accumulated throughout 2024, reaching 21.3 Mt CO_2, 615.8 kt NOx, 350.3 kt $PM_{2.5}$, and 6.5 kt BC by the end of the year.

These revisions clarify that the major scientific implication of SEIMv2.3 is addressing systematic emission underestimation caused by incomplete fleet information and compliance pathway representation, while the daily updating capability provides a technical basis for future near-real-time emission estimation.

Revisions:

In RESULTS (page15), we revise Fig. 7 and Fig. 8:

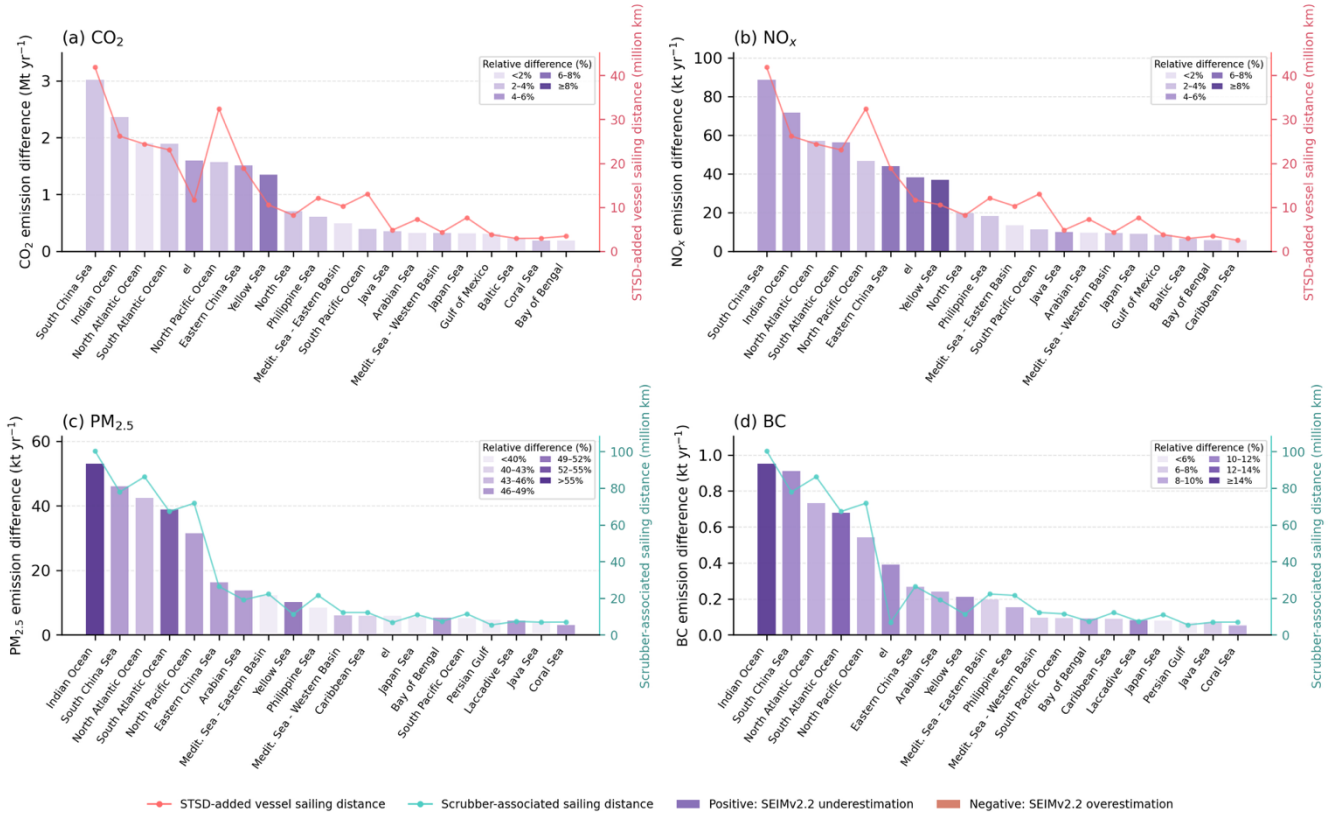


Figure 7. Ship $PM_{2.5}$, BC, CO_2, and NOx emission underestimation mitigated by SEIMv2.3 in different ocean regions in 2024.

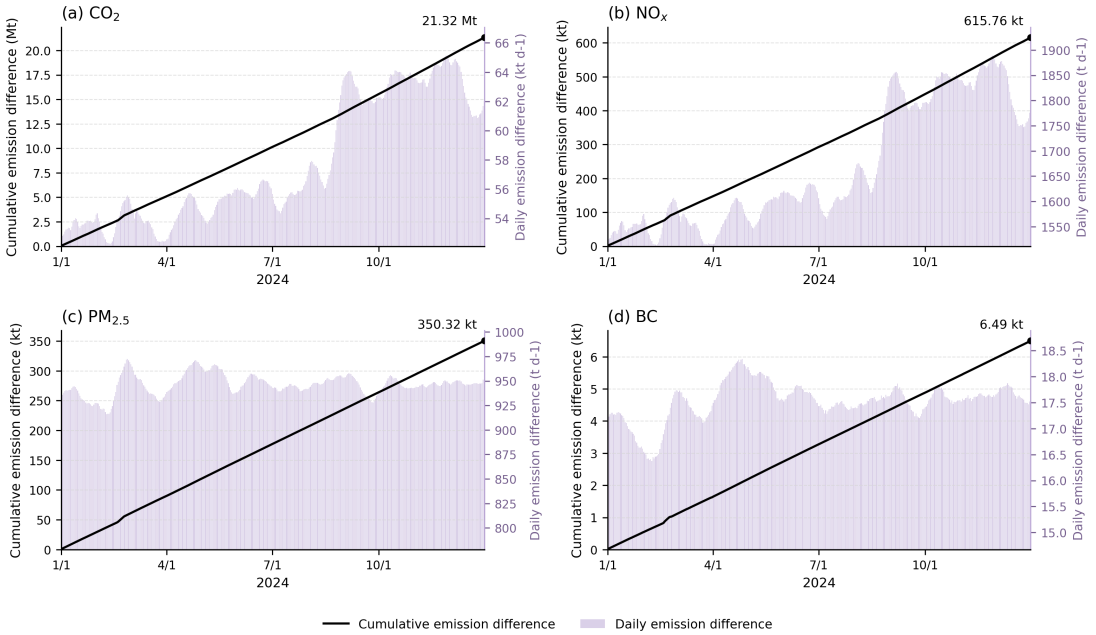


Figure 8. Temporal accumulation of emission discrepancies between SEIMv2.3 and SEIMv2.2 during 2024.

Q47. Line 281

Line 281/282: "by previous bottom-up method": What do you mean here? Do you refer to SEIMv2.2?

Response:

Thank you for this comment. Yes, "previous bottom-up method" refers specifically to the previous version of our model, SEIMv2.2. We agree that the original wording was ambiguous and could be interpreted as referring to other bottom-up ship emission inventories.

We have revised the text to explicitly state "SEIMv2.2" instead of "previous bottom-up method" in the revised manuscript.

Revisions:

Our results suggest a decrease by 48.2% in PM2.5 emission from 2020 to 2019, compared with the 77.2% reduction estimated by SEIMv2.2.

Q48. Line 290

Line 290: This "systemic emission underestimation" was only demonstrated for SEIMv2.2, not for any other ship emission model. You need to point this out.

Response:

We agree that the original wording was too broad and could imply that systematic emission underestimation has been demonstrated for all existing ship emission inventories. In this study, the quantified underestimation was directly demonstrated by comparing SEIMv2.3 with its previous version, SEIMv2.2.

However, the underlying causes identified in SEIMv2.2, including incomplete fleet representation and simplified fuel compliance assumptions, represent methodological challenges that may also affect other AIS-based bottom-up inventories when similar assumptions are applied. We have therefore revised the statement to clarify that our results demonstrate the impact of these limitations within the SEIM framework, while highlighting their broader relevance for ship emission inventory development.

Revisions:

In Discussion (page16, para2):

Overall, our study establishes a dynamically updated framework for global ship emission estimation and reveals substantial emission underestimation associated with incomplete fleet representation and simplified sulfur compliance assumptions in previous SEIM configurations.

Q49. Line 294

Line 294/295: Why do you explicitly mention satellite studies for verification?

There may be other possibilities.

Response:

Thank you for this valuable comment. We agree that satellite-based observations represent only one possible approach for evaluating shipping emission inventories and should not be presented as the only verification pathway. In the revised manuscript, we have removed the previous statement emphasizing satellite-based concentration inversion studies and instead discussed multiple independent sources of evidence, including IMO fuel consumption statistics, scrubber installation records, and $PM_{2.5}$ observational studies.

Minor Comments

Q50. Line 10

Line 10: what is a "hard-to-abate sector"? Please be more specific.

Response:

Thank you for this comment. We agree that the term "hard-to-abate sector" was not sufficiently explained in the abstract. We have revised the first sentence to clarify the characteristics that make shipping difficult to decarbonize, including high energy demand, long-lived assets, and limited near-term low-carbon alternatives.

Revisions:

In ABSTRACT:

As a sector with high energy demand, long-lived assets, and limited near-term low-carbon alternatives, global shipping is considered difficult to decarbonize and is subject to international and regional emission control regulations.

Q51. Line 14

Line 14/15: What is "near-real-time" here? One hour, one day, ...?

Response:

Thank you for this comment. We agree that the term "near-real-time" was not sufficiently defined in the original manuscript. In this study, near-real-time refers to the capability of updating ship emission estimates within a short delay after AIS data acquisition, rather than instantaneous real-time monitoring. The preprocessing framework reconstructs reliable vessel trajectories using recently available AIS signals, enabling emission estimation with an operational delay of several days.

We have clarified this definition throughout the revised manuscript and avoided implying that the inventory provides instantaneous emission monitoring.

Revisions:

In ABSTRACT:

Here, we established a dynamically updated Ship Emission Inventory Model that allows near-real-time emission calculation within several days after AIS data acquisition.

Q52. Line 20

Line 20/21: explain the "single compliance choice assumption". This is a very model-specific term.

Revisions:

In ABSTRACT:

Meanwhile, assuming that all ships switch to low-sulfur fuels after 2020 would lead to underestimation of …

Q53. Line 83

Line 83: explain STSD

Revisions:

In METHOD (page4, para4):

In SEIMv2.2, ship activity records were matched with our Ship Technical Specifications Database (STSD),

Q54. Line 93

Line 93: So you mean 113 million dynamic and 14 million static signals per day? Please specify this. Does this refer to a global coverage?

Response:

Thank you for this comment. We apologize for the ambiguity in the original description. Yes, the reported 113 million dynamic AIS signals and 14 million static AIS signals represent the daily global AIS data volume collected in this study. After processing daily AIS records over the study period, the total number of AIS signals reached approximately 45 billion records annually. We have revised the manuscript to explicitly clarify the temporal scale and global coverage of the AIS dataset.

Revisions:

In METHOD (page5, para2), we clarified:

In this study, we collected global daily AIS data from 2022 to 2024, comprising approximately 113 million dynamic signals and 14 million static signals per day. The daily AIS processing resulted in approximately 45 billion processed signals

annually.

Q55. Line 183

Line 183: Does the percentage relate to the number or the DWT?

Line 183: Similar to before: Does the percentage relate to the number or the installed power?

Response:

Thank you for this comment. We agree that the original description did not clearly specify the basis of these percentages. All percentages reported in this paragraph refer to the number of newly identified vessels, rather than their DWT or installed engine power. We have revised the text to explicitly state "by number" for ship type, DWT range, and main engine power distributions. The corresponding figure caption has also been clarified.

Revisions:

Regarding DWT range, newly identified vessels were predominantly distributed within the 10,000–100,000 DWT range, accounting for about 59% of newly identified vessels by number. Meanwhile, main engine power was largely concentrated in the 1,000–10,000 kW range, which included approximately 58% of newly identified vessels by number.

Q56. Text revision

Line 16: of the global ship fleet

Line 174: in the Summary and Discussion section

Line 282: please correct "Most underestimation"

Check the References for errors, e.g. line 416

Response:

Thank you for pointing this out. We have revised the relevant sentences to correct grammatical issues, added definite articles where necessary, and carefully checked the reference list for formatting and citation errors.

Reference

1. Caprile, A. L., Gabija: Russia's 'shadow fleet': Bringing the threat to light, European Parliamentary Research Service, Brussels, 2024.
2. ICF International: Current Methodologies in Preparing Mobile Source Port-Related Emission
3. Inventories, 2009.
4. IMO: Fourth IMO GHG Study – Final Report, 2020.
5. Verification of the Gross Mass of Packed Containers: https://www.imo.org/en/OurWork/Safety/Pages/Verification-of-the-gross-mass.aspx, last access: 5 July 2026.
6. Bulk Carrier Safety: https://www.imo.org/en/OurWork/Safety/Pages/BulkCarriers.aspx, last access: 5 July 2026.
7. Jalkanen, J. P., Brink, A., Kalli, J., Pettersson, H., Kukkonen, J., and Stipa, T.: A modelling system for the exhaust emissions of marine traffic and its application in the Baltic Sea area, Atmos. Chem. Phys., 9, 9209-9223, DOI 10.5194/acp-9-9209-2009, 2009.
8. Kivekäs, N., Massling, A., Grythe, H., Lange, R., Rusnak, V., Carreno, S., Skov, H., Swietlicki, E., Nguyen, Q. T., Glasius, M., and Kristensson, A.: Contribution of ship traffic to aerosol particle concentrations downwind of a major shipping lane, Atmos. Chem. Phys., 14, 8255-8267, 10.5194/acp-14-8255-2014, 2014.
9. Kramel, D., Muri, H., Kim, Y., Lonka, R., Nielsen, J. B., Ringvold, A. L., Bouman, E. A., Steen, S., and Strømman, A. H.: Global Shipping Emissions from a Well-to-Wake Perspective: The MariTEAM Model, Environmental Science & Technology, 55, 15040-15050, 10.1021/acs.est.1c03937, 2021.
10. Kuittinen, N., Timonen, H., Karjalainen, P., Murtonen, T., Vesala, H., Bloss, M., Honkanen, M., Lehtoranta, K., Aakko-Saksa, P., and Rönkkö, T.: In-depth characterization of exhaust particles performed on-board a modern cruise ship applying a scrubber, Science of The Total Environment, 946, 174052, https://doi.org/10.1016/j.scitotenv.2024.174052, 2024.
11. Schwarzkopf, D. A., Petrik, R., Matthias, V., Quante, M., Majamäki, E., and Jalkanen, J.-P.: A ship emission modeling system with scenario capabilities, Atmospheric Environment: X, 12, 100132, https://doi.org/10.1016/j.aeaoa.2021.100132, 2021.
12. Seppälä, S. D., Kuula, J., Hyvärinen, A. P., Saarikoski, S., Rönkkö, T., Keskinen, J., Jalkanen, J. P., and Timonen, H.: Effects of marine fuel sulfur restrictions on particle number concentrations and size distributions in ship plumes in the Baltic Sea, Atmos. Chem. Phys., 21, 3215-3234, 10.5194/acp-21-3215-2021, 2021.
13. Sharmina, M., Y., E. O., C., W., R., F., J., G. D. E. H., P., G., A., L., W., L. E., M., T., P., v. V. D., E., V. N., R., W. F., and and Le Quéré, C.: Decarbonising the critical sectors of aviation, shipping, road freight and industry to limit warming to 1.5–2°C, Climate Policy, 21, 455-474, 10.1080/14693062.2020.1831430, 2021.
14. Shi, Z., Endres, S., Rutgersson, A., Al-Hajjaji, S., Brynolf, S., Booge, D., Hassellöv, I.-M., Kontovas, C., Kumar, R., Liu, H., Marandino, C., Matthias, V., Moldanová, J., Salo, K., Sebe, M., Yi, W., Yang, M., and Zhang, C.: Perspectives on shipping emissions and their impacts on the surface ocean and lower atmosphere: An environmental-social-economic dimension, Elementa-Sci. Anthrop., 11, 10.1525/elementa.2023.00052, 2023.

15. Sofiev, M., Winebrake, J. J., Johansson, L., Carr, E. W., Prank, M., Soares, J., Vira, J., Kouznetsov, R., Jalkanen, J.-P., and Corbett, J. J.: Cleaner fuels for ships provide public health benefits with climate tradeoffs, Nature Communications, 9, 406, 10.1038/s41467-017-02774-9, 2018.
16. Treasury Intensifies Sanctions Against Russia by Targeting Russia's Oil Production and Exports: https://home.treasury.gov/news/press-releases/jy2777?utm_source=chatgpt.com, last access: 5 July 2026.
17. UNCTAD: Review of Maritime Transport 2021, https://doi.org/10.18356/9789210000970, 2021.
18. Wang, X., Liu, H., Zhang, J., Fu, X., Chen, D., Zhang, W., Yi, W., Lv, Z., Zhang, Q., and He, K.: Global shipping emissions from 1970 to 2021: Structural and spatial change driven by trade dynamics, One Earth, 10.1016/j.oneear.2025.101243, 2025.
19. Yi, W., Liu, H., Peng, L., Liu, Z., Luo, Z., Rovenskaya, E., Ng, S. H., Strelkovskii, N., Yan, R., Wang, X., Yang, Z., He, T., Zhang, W., Cai, F., Zhang, Q., and He, K.: Drivers and environmental impacts of Arctic shipping, Nature Reviews Earth & Environment, 7, 343-358, 10.1038/s43017-026-00790-2, 2026.