Supplement of Atmos. Chem. Phys., 26, 16373–16387, 2026
https://doi.org/10.5194/acp-26-16373-2026-supplement
© Author(s) 2026. CC BY 4.0 License.



Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Addressing systematic underestimation in a global ship emission inventory through improved fleet representation and fuel compliance characterization

Weiwei Zhang et al.

Correspondence to: Huan Liu (liu_env@tsinghua.edu.cn)

The copyright of individual parts of the supplement might differ from the article licence.

S1. Emission calculation principles of SEIMv2.3

The mathematical formulation of the emission calculation is identical to that described by (Yi et al., 2024). Therefore, differences from (Yi et al., 2024) arise from updated input data 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. The emissions from the main engine, auxiliary engine, and boiler are simulated using Eqs (S1)-(S3):

$$E_{ME,i,n,p} = P_{ME,i,n} \times EF_{ME,i,p} \times LLAF_{i,n,p} \times \Delta T_{i,n} \times 10^{-6} \quad \text{(S1)}$$

$$E_{AE,i,n,p} = P_{AE,i,n} \times EF_{AE,i,p} \times \Delta T_{i,n} \times 10^{-6} \quad \text{(S2)}$$

$$E_{B,i,n,p} = P_{B,i,n} \times EF_{B,i,p} \times \Delta T_{i,n} \times 10^{-6} \quad \text{(S3)}$$

where the capital letters E (t) represents the emissions; EF ($gkWh^{-1}$) is the emission factor; P (kW) is the output power; T (h) is the time interval of two subsequent AIS signals; and LLAF is the low load adjust factor, which is applied only when the main engine load factor is lower than 20 %. The subscripts ME, AE, and B represent the main engine, auxiliary engine, and boiler, respectively; i represents an individual ship; n represents the nth AIS signals in the sequence; p represents species of GHGs or air pollutants.

For auxiliary engines and boilers' power demand, this model adopts the recommended values from the IMO Fourth and Third Greenhouse Gas Study reports (IMO, 2020). Due to the lack of information, this study did not consider the impact of other auxiliary devices on board, which are considered to be insignificant contributors to overall vessel energy consumption currently(Eirik Ovrum, 2024). The real-time output power of main engine depends highly on its load, which can be calculated in real time by changes in the ship's speed over ground obtained from AIS data. According to the propeller law, the main engine load factor is the cube of the ratio of the ship's actual speed to its design speed (Liu et al., 2016). The model also considers the influence of draft, weather, and hull fouling through parameterization schemes introduced in our last version. The formula for calculating the real-time power of the main engine can be found in Eq. (S4):

$$P_{ME,i,n} = P_{ref,i} \times LF_{i,n} = \frac{P_{ref,i} \times \left(\frac{D_i}{D_{ref,i}}\right)^{0.66} \times \left(\frac{v_{i,n}}{v_{ref,i}}\right)^3}{\eta_w \times \eta_f} \quad \text{(S4)}$$

where $P_{ref,i}$ (kW) is the maximum engine output power of the main engine of the ship i, and $LF_{i,n}$ represents the main engine load factor of the ship i at the nth AIS signals in the sequence. D_i represents the average draft; $D_{ref,i}$ represents the designed draft; $v_{i,n}$ (knot) represents the speed over ground of the ship i at the nth AIS signals in the sequence; $v_{ref,i}$ (knot) represents the design speed of ship i, obtained from the static technical profiles; and η_w represents the weather correction factor and η_f the fouling correction factor, both of which are unitless.

The emission factors applied in the study is mainly based on the IMO Fourth Greenhouse Gas

Study (IMO, 2020), U.S. Environmental Protection Agency reports (Starcrest Consulting Group, 2022), and the National Standard for General Diesel Fuel of the People's Republic of China (GB 252-2015). Table. S2 exhibits the emissions factors based on fuel consumption for CO_2, SO_2, and $PM_{2.5}$. Energy-based emission factors are calculated based on fuel-based emission factors as well as specific fuel consumption (SFC, kWh/kg fuel), using Eq. (S5):

$$EF_e = EF_f \times SFC \quad \text{(S5)}$$

SFC is the fuel consumption per unit of work performed by a ship, mainly related to the fuel calorific value and engine efficiency (%). During ship operations, energy efficiency typically follows a quadratic rationship with the main engine load, peaking at around an 80% load factor. Eq. (S6) is used to calculate the SFC for main engines based on the SFC under the optimal operating condition (SFC_{base}) and main engine load of the ship i.

$$SFC_{ME,i} = SFC_{base,ME,i} \times (0.455 \times LF_i^2 - 0.71 \times LF_i + 1.28) \quad \text{(S6)}$$

S2. Model performance of different algorithms for update STSD

For model selection, (Huang et al., 2024) conducted preliminary experiments comparing several candidate algorithms, including Logistic Regression, Support Vector Machines, Random Forest, and XGBoost. The tree-based ensemble models, particularly XGBoost, consistently outperformed others in the validation. Consequently, XGBoost was chosen for the ship type classification task due to its highest accuracy, and both XGBoost and Random Forest were selected for the regression tasks based on their respective strengths in prediction accuracy and robustness. The final models, features, and hyperparameters used for complementing the three categories of parameters are listed in Table S1.

To assess whether an alternative ship database could provide the technical information required by SEIMv2.3, we compared the STSD with Sea-web in terms of fleet coverage and the availability of key variables (Table S5 and Table S6). The comparison shows that Sea-web provides extensive ship-level information, but its fleet definition and the fields available to this study do not fully match the requirements of the STSD/SEIMv2.3 inventory. In particular, main engine power, a key input to ship-level emission estimation, was not available among the Sea-web fields examined in this study. Thus, Sea-web provides complementary ship information but cannot directly replace the STSD for our emission calculations.

S3. Characteristics of observed scrubber-equipped vessels

To improve the representation of scrubber-based sulfur compliance pathways, a vessel-level scrubber database was incorporated as an observational constraint. The database was cleaned and

matched with the ship fleet using MMSI. After removing invalid records, 7,546 vessels with valid MMSI information were retained, among which 6,138 vessels were identified as fitted with scrubbers. The recorded installation years ranged from 2007 to 2024 (Fig. S7), covering both the early adoption period of scrubber technology and the post-IMO 2020 sulfur regulation period. The observed scrubber-equipped fleet included 6,138 vessels, representing 723.4 million tonnes of DWT and 483.8 million GT. Bulk carriers (1,945 vessels, 31.69%), container ships (1,672 vessels, 27.24%), and tankers (1,261 vessels, 20.54%) were the dominant categories, together accounting for 79.5% of identified scrubber installations (Fig. S8). Based on this observed composition, FLSCCM gap filling was restricted to four major scrubber-relevant ship categories: bulk carriers, container ships, oil tankers, and chemical tankers.

The observed scrubber fleet also showed a strong size preference (Fig. S9). Vessels with DWT above 50,000 tonnes accounted for approximately 69.5% of identified installations, including 28.67% between 100,000 and 200,000 tonnes and 18.67% above 200,000 tonnes. Therefore, a DWT threshold of 50,000 tonnes was introduced for database-unmatched vessels in the FLSCCM gap-filling procedure to retain the majority of observed scrubber-equipped vessels while avoiding excessive allocation among vessel groups with limited representation in the observed fleet. Smaller vessels with confirmed scrubber installations were retained through the observed pathway and were not excluded by this threshold. Although scrubber adoption was concentrated among large vessels, smaller vessels were also present in the observed database, with approximately 30.5% of identified scrubber-equipped vessels having DWT below 50,000 tonnes.

The construction-year distribution further indicated a strong age dependence of scrubber adoption (Fig. S10). Among the 6,138 observed scrubber-equipped vessels, 5,507 vessels (89.7%) were built in or after 2006, accounting for 94.1% of the total DWT, whereas only 631 vessels (10.3%) were built before 2006. Therefore, vessels built before 2006 represented a relatively small older-ship tail of the observed scrubber fleet. For database-unmatched vessels, FLSCCM considered ships built in or after 2006 as eligible candidates when reliable construction-year information was available. This age constraint was not applied to observed scrubber-equipped vessels because database records provide direct vessel-level evidence of installed scrubber capability. Alternative treatments for missing construction-year information were evaluated in the sensitivity analysis.

Based on these observed fleet characteristics, FLSCCM was formulated as an observed-constrained gap-filling module rather than an independent scrubber prediction model. Vessels identified in the observed scrubber database were retained as ship-level evidence and excluded from further inference, while only database-unmatched vessels were screened using ship type, DWT, and construction-year constraints before compliance pathway assignment. This approach reduces

potential overestimation from unrestricted statistical allocation while maintaining the ability to reconstruct missing compliance pathways in the global fleet.

To further evaluate the FLSCCM reconstruction against an independent source, we examined the DWT and age distributions of ships recorded in IMO GISIS and assigned to the HFO-with-scrubber pathway by FLSCCM in the 2024 SEIM fleet (Fig. S10). The two datasets show broadly similar DWT distributions and are both dominated by ships aged 5–19 years. However, differences remain in ship-age composition, and GISIS itself does not provide a complete global scrubber registry. The comparison therefore serves as a fleet-level structural check of the reconstructed compliance pathway rather than validation of the absolute number or scrubber status of every individual ship.

S4. Sensitivity analysis of FLSCCM structural assumptions

To evaluate how uncertainty in the structural assumptions used for FLSCCM gap filling propagates into the reconstructed fleet size and emission estimates, 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.

Overall, although these structural assumptions substantially affect the number of vessels assigned to the HFO-with-scrubber pathway, their influence on global $PM_{2.5}$ and BC emission estimates remains comparatively limited across the tested scenarios.

Supplementary figures and tables

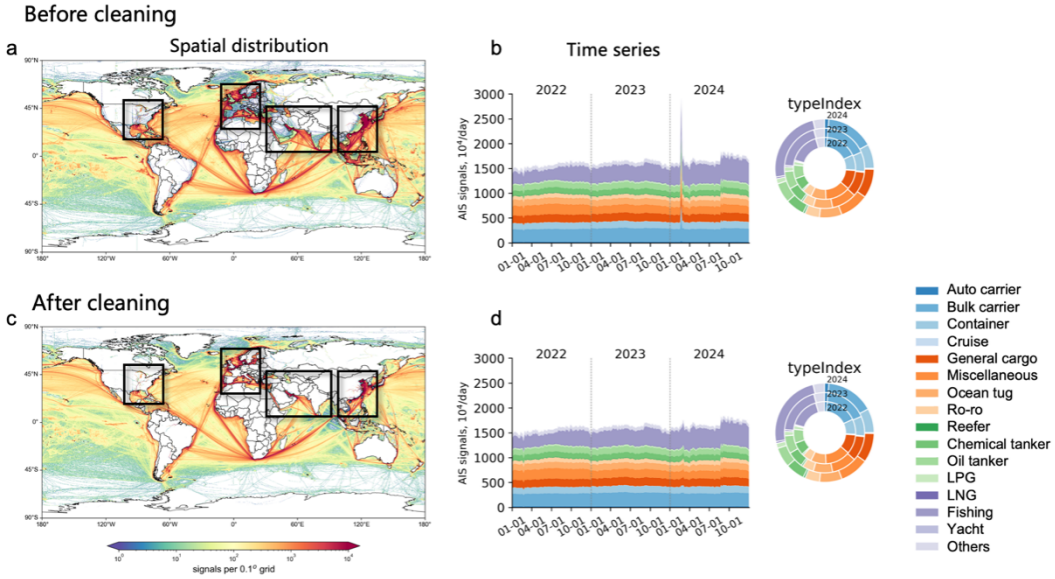


Figure S1. The spatial distribution and time series of AIS before and after preprocessing

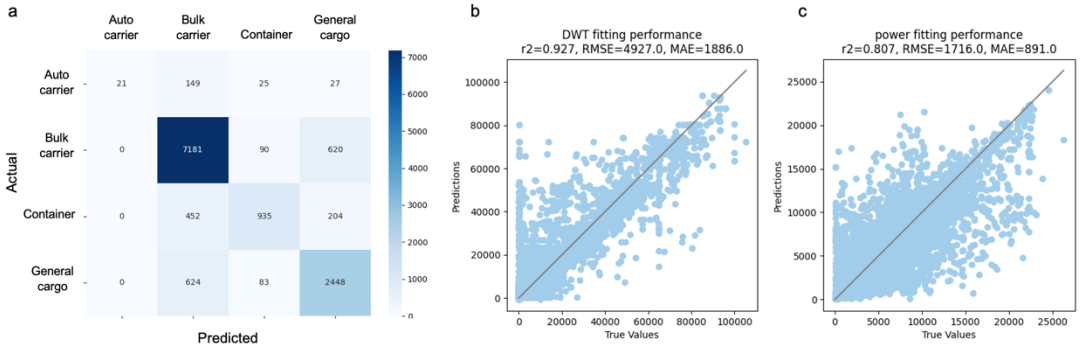


Figure S2. Fitting performance of cargo ship classification (a), DWT prediction (b), and main engine power prediction (c).

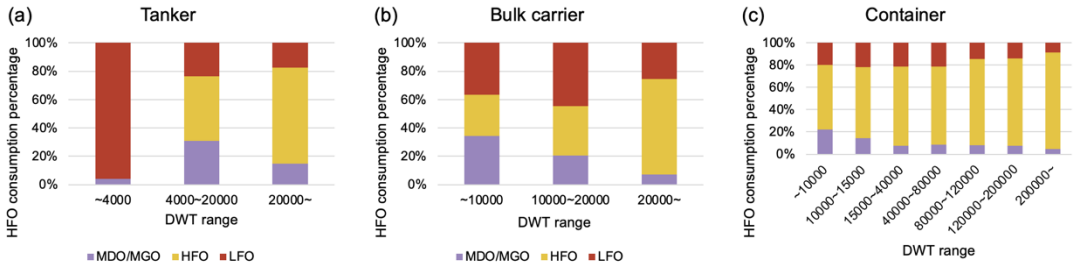


Figure S3. Distribution of fuel consumption by DWT range in 2022 for tanker (a), bulk carrier (b), and container ship (c), based on IMO-reported fuel consumption statistics (IMO, 2024).

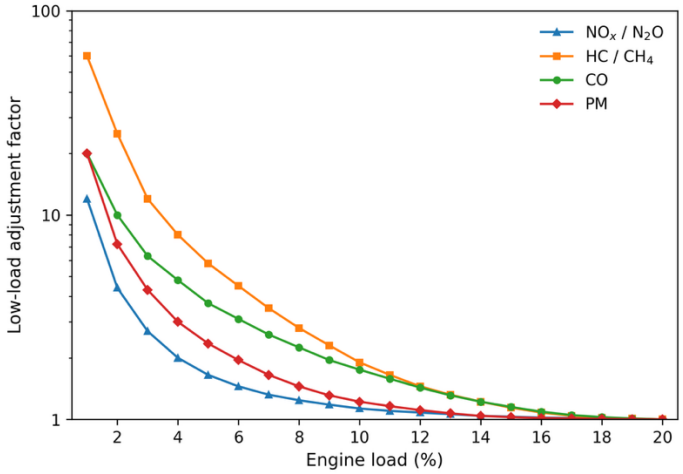


Figure S4. Low-load adjustment factors applied to main-engine emission factors in SEIMv2.3.

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).

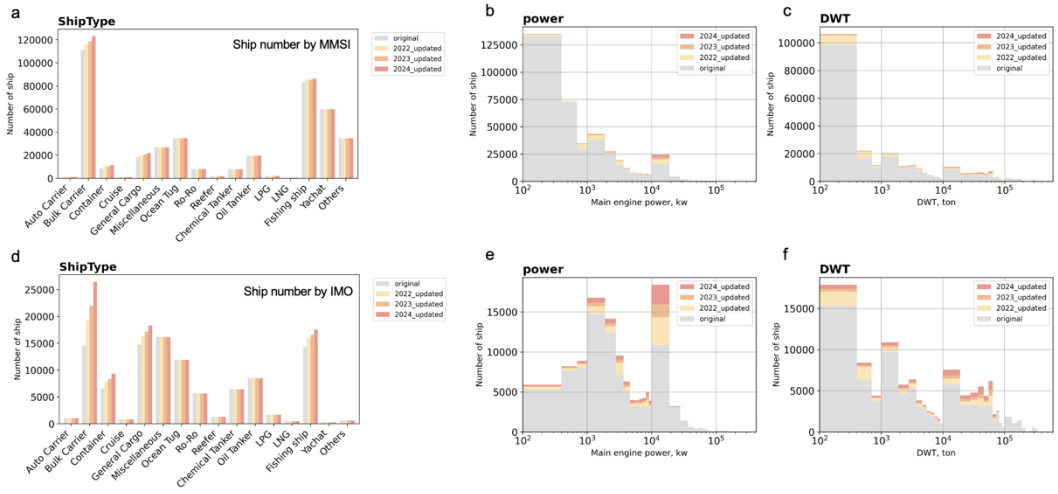


Figure S5. The composition of ship types (a), main engine power (b), and dead tonnage weight (c) for annual evolution of STSD with ship number by MMSI. The composition of ship types (d), main engine power (e), and dead tonnage weight (f) for annual evolution of STSD with ship number by IMO.

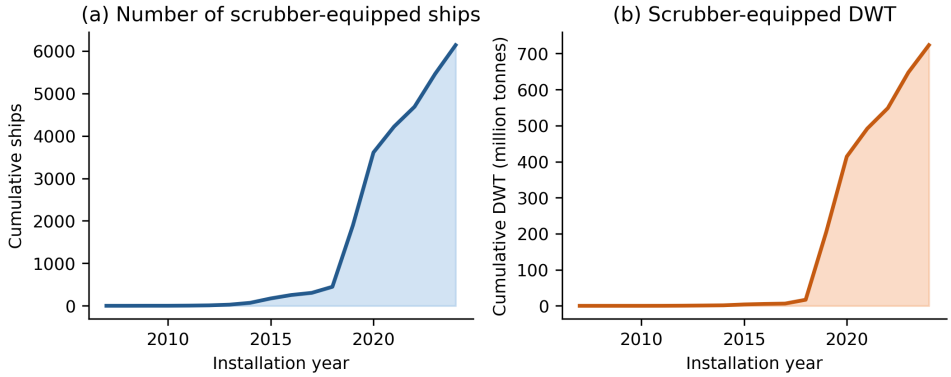


Figure S6. Global scrubber fleet growth in cumulative number (a) and DWT (b) obtained from Clarksons Research World Fleet Register.

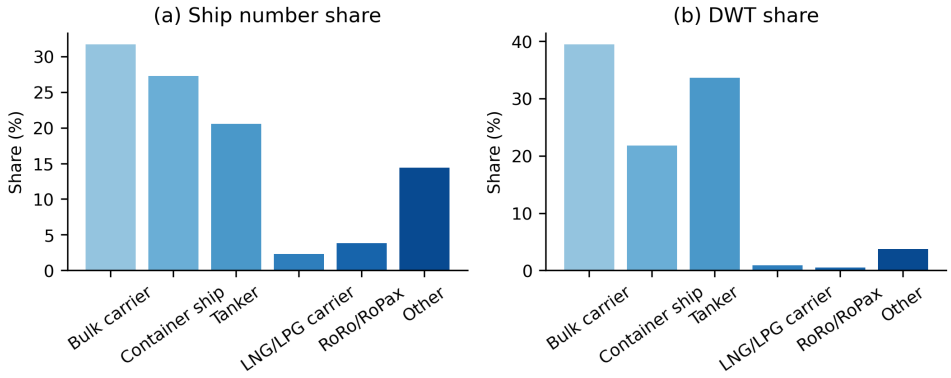


Figure S7. Ship type distribution of scrubber-equipped fleet in 2024

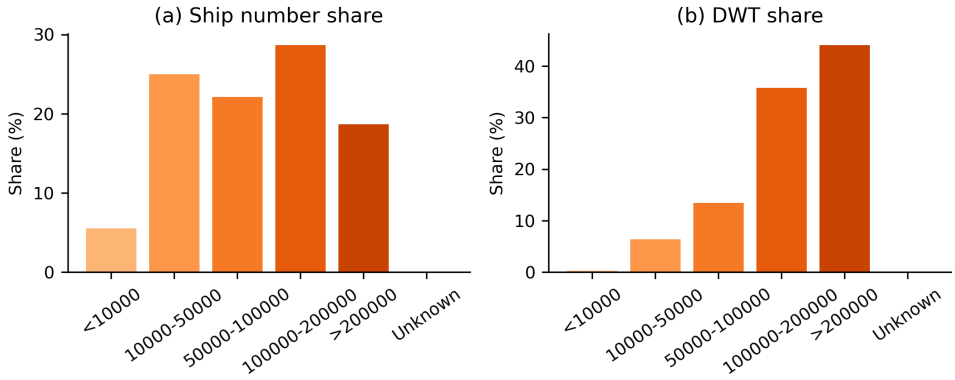


Figure S8. Ship DWT distribution of scrubber-equipped fleet in 2024

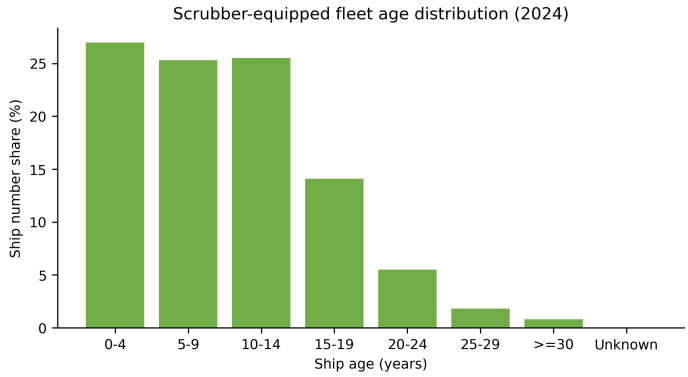


Figure S9. Scrubber-equipped fleet age distribution in 2024

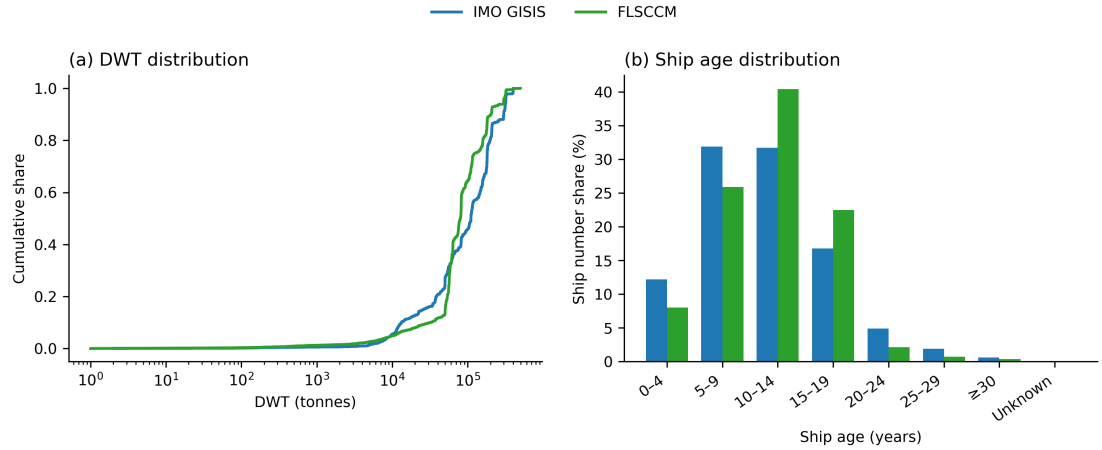


Figure S10. Comparison of the DWT and age distributions of scrubber-equipped vessels recorded in IMO GISIS and identified by FLSCCM in the 2024 SEIM fleet. (a) Cumulative distributions of vessel DWT. (b) Vessel-number shares by age category.

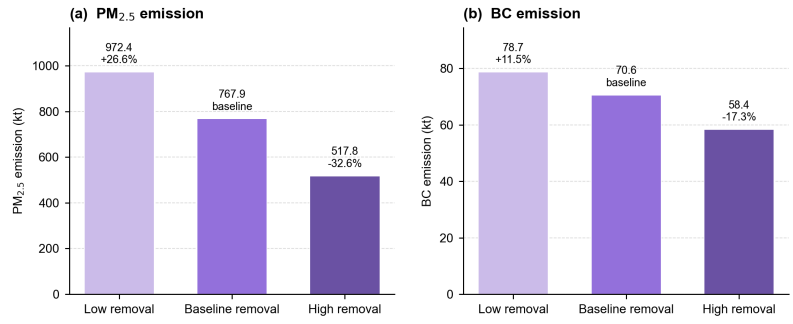


Figure S11. Sensitivity of $PM_{2.5}$ and BC emissions to scrubber removal efficiency assumptions

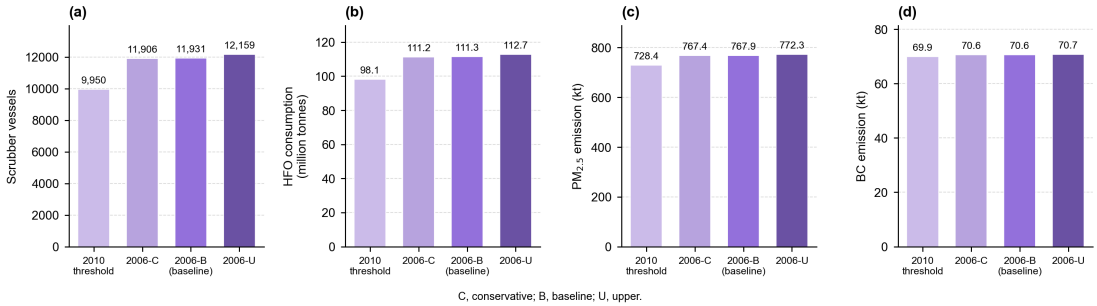


Figure S12. Sensitivity of reconstructed scrubber pathways to vessel eligibility criteria

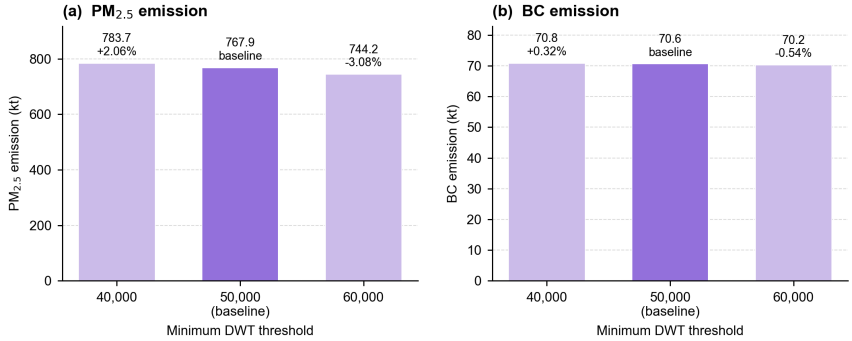


Figure S13. Sensitivity of emission estimates to minimum DWT screening threshold

Table S1. Algorithms, features, and hyperparameters chosen for predicting cargo ship type, DWT, and main engine power

Function	Algorithm	Features	n estimator	Learning rate	Max depth
Cargo ship classification	XGBoost	Width, length, perimeter, naive area	/	0.1	3
DWT prediction	XGBoost	Width, length	100	0.1	3
Main engine power prediction	Random Forest	Width, length	100	0.1	3

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

Fuel	Engine	Engine type	NOx				$PM_{2.5}$	SO_2		HC	CO	CH4	N2O	CO_2	
			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. Energy-based emission factors for BC (g/kWh)

Main engine load (%)	HFO SSD	HFO MSD/HSD	HFO AE/BE	MDO/MGO SSD	MDO/MGO MSD/HSD	MDO/MGO AE/BE	LNG
(0,5]	0.077	0.746	0.080	0.002	0.610	0.060	0.003
(5,10]	0.060	0.381	0.080	0.001	0.280	0.060	0.003
(10,20]	0.047	0.195	0.080	0.001	0.130	0.060	0.003
(20,30]	0.040	0.132	0.080	0.001	0.080	0.060	0.003
(30,40]	0.037	0.099	0.080	0.001	0.060	0.060	0.003
(40,50]	0.033	0.081	0.080	0.001	0.050	0.060	0.003
(50,60]	0.032	0.068	0.080	0.001	0.040	0.060	0.003
(60,70]	0.030	0.058	0.080	0.001	0.030	0.060	0.003
(70,80]	0.028	0.051	0.080	0.000	0.030	0.060	0.003
(80,90]	0.028	0.046	0.080	0.000	0.020	0.060	0.003
(90,100]	0.026	0.041	0.080	0.000	0.020	0.060	0.003

Table S4. 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

Table S5. Availability of ship technical variables required by SEIMv2.3 in the STSD before and after machine-learning completion and in Sea-web

Variable required by our model	STSD before ML	STSD after ML	Sea-web
Ship type	Partial	✓	✓
GT	Partial	✓	✓
DWT	Partial	✓	✓
Year built	Partial	✓	✓
Main engine power	Partial	✓	X

Table S6. Fleet coverage, temporal scope, size, and accessibility of the STSD, SEIMv2.3 fleet subset, Sea-web, and UNCTAD data

Database	Vessel coverage	Temporal scope	Approx. number	Accessibility
STSD	Historical and current vessel records used to construct SEIM	Current + historical	440,000	Based on accessible/public sources
Ships extracted from STSD for SEIM calculation	Active vessels in the target year	2024	132,000	Research inventory
Sea-web	Propelled seagoing merchant ships ≥100 GT, including active, newbuildings, casualties and lost/broken-up vessels	Current + historical	>200,000	Commercial subscription
UNCTAD	Merchant fleet ≥100 GT	2024	108,789	Public

Reference

Eirik Ovrum, T. L., Marius, Leisner, Eirill Bachmann Mehammer, Ola Gundersen Skåre, Henrik Helgesen, Øyvind Endresen: Maritime forecast to 2050: Energy transition outlook 2024, DNV, 2024.

Huang, I.-L., Lee, M.-C., Nieh, C.-Y., and Huang, J.-C.: Ship Classification Based on AIS Data and Machine Learning Methods, Electronics, 13, 98, 2024.

ICF International: Current Methodologies in Preparing Mobile Source Port-Related Emission Inventories, 2009.

IMO: Fourth IMO GHG Study – Final Report, 2020.

Liu, H., Fu, M., Jin, X., Shang, Y., Shindell, D., Faluvegi, G., Shindell, C., and He, K.: Health and climate impacts of ocean-going vessels in East Asia, Nat. Clim. Chang., 6, 1037-1041, 10.1038/nclimate3083, 2016.

Yi, W., Wang, X., He, T., Liu, H., Luo, Z., Lv, Z., and He, K.: High-resolution global shipping emission inventory by Shipping Emission Inventory Model (SEIM), Earth Syst. Sci. Data Discuss., 2024, 1-31, 10.5194/essd-2024-258, 2024.