

Supporting information for:

Addressing systemic underestimation in global ship emissions from fleet growth and fuel compliance

Authors

Weiwei Zhang ^{1,2}, Zhaofeng Lv ¹, Wen Yi ¹, Tingkun He ¹, Bensheng Xiao ¹, Qiang Zhang ³, Huan Liu ^{1,2*}

Affiliations

¹ State Key Laboratory of Regional Environment and Sustainability, School of Environment, Tsinghua University, Beijing 100084, China

² Institute for Carbon Neutrality, Tsinghua University, Beijing 100084, China

³ Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China

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

Emission calculation principles of SEIMv2.3

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 (1)-(3):

$$E_{ME,i,n,p} = P_{ME,i,n} \times EF_{ME,i,p} \times LLA_{F_{i,n,p}} \times \Delta T_{i,n} \times 10^{-6} \quad (1)$$

$$E_{AE,i,n,p} = P_{AE,i,n} \times EF_{AE,i,p} \times \Delta T_{i,n} \times 10^{-6} \quad (2)$$

$$E_{B,i,n,p} = P_{B,i,n} \times EF_{B,i,p} \times \Delta T_{i,n} \times 10^{-6} \quad (3)$$

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 LLA_F 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. 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 currently²⁸. 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 speed³. 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. (4):

$$P_{ME,i,n} = P_{ref,i} \times LF_{i,n} = \frac{P_{ref,i} \times \left(\frac{D_i}{D_{ref,i}}\right)^{0.66} \times \left(\frac{v_{i,n}}{v_{ref,i}}\right)^3}{\eta_w \times \eta_f} \quad (4)$$

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 n th 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 n th 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 Study⁸ as well as the National Standard for General Diesel Fuel of the People's Republic of China¹³. **Table. S2** exhibits the emissions factors based on fuel consumption for CO₂, SO₂, 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. (5):

$$EF_e = EF_f \times SFC \quad (5)$$

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 relationship with the main engine load, peaking at around an 80% load factor. Eq. (6) 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 .

$$SFC_{ME,i} = SFC_{base,ME,i} \times (0.455 \times LF_i^2 - 0.71 \times LF_i + 1.28) \quad (6)$$

Model performance of different algorithms for update STSD

For model selection, Huang et al.²⁹ 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](#).

Supplementary figures and tables

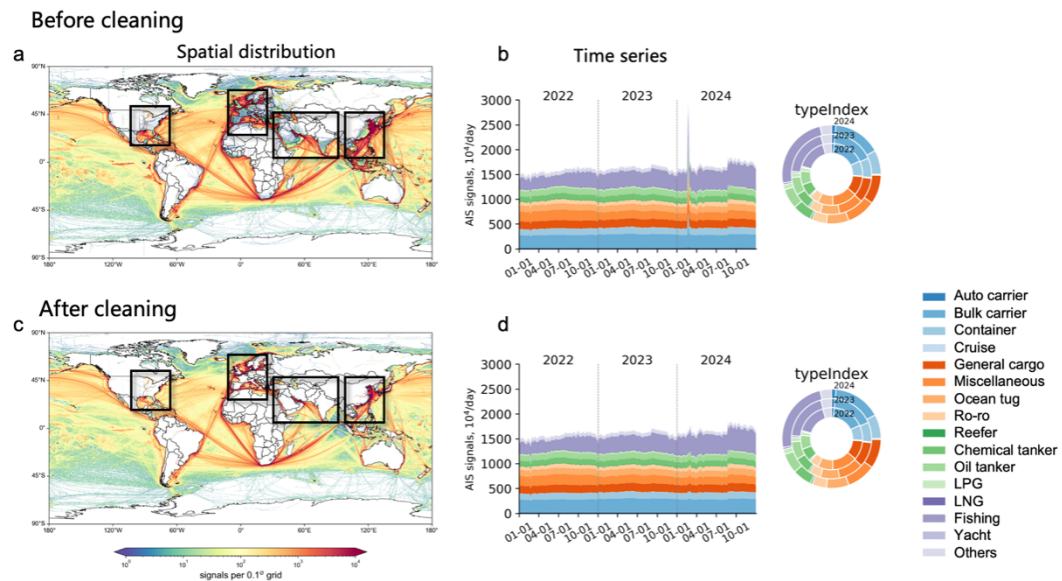


Figure S1. The spatial distribution and time series of AIS before and after preprocessing

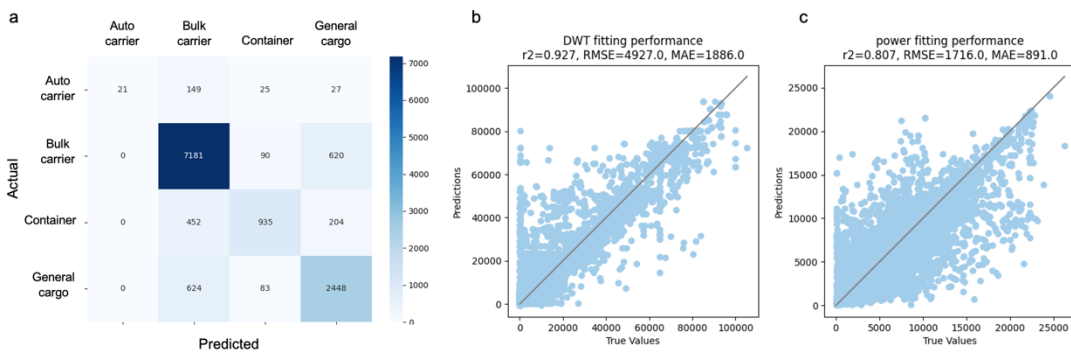


Figure S2. Fitting performance of cargo ship classification (a), DWT prediction (b), and main engine power prediction (c).

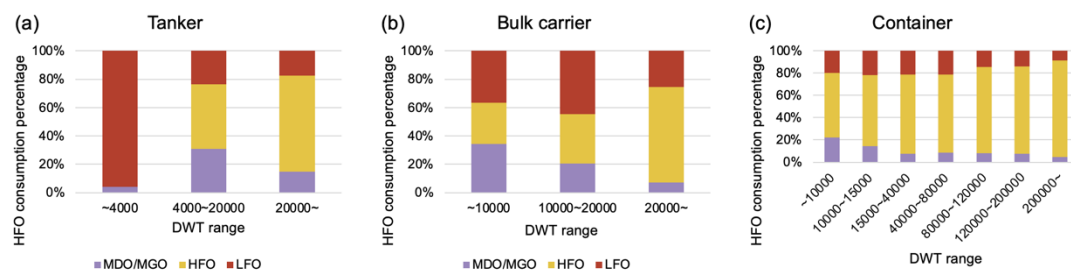


Figure S3. Fuel consumption for different DWT range in 2022 for tanker (a), bulk

carrier (b), and container (c).

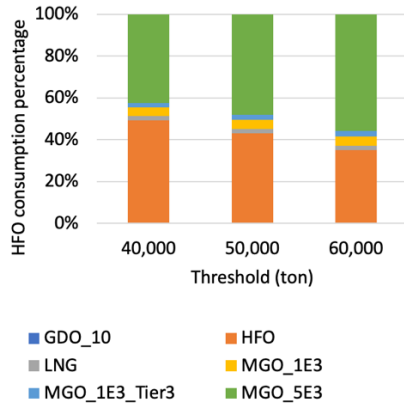


Figure S4. HFO consumption percentage under different thresholds

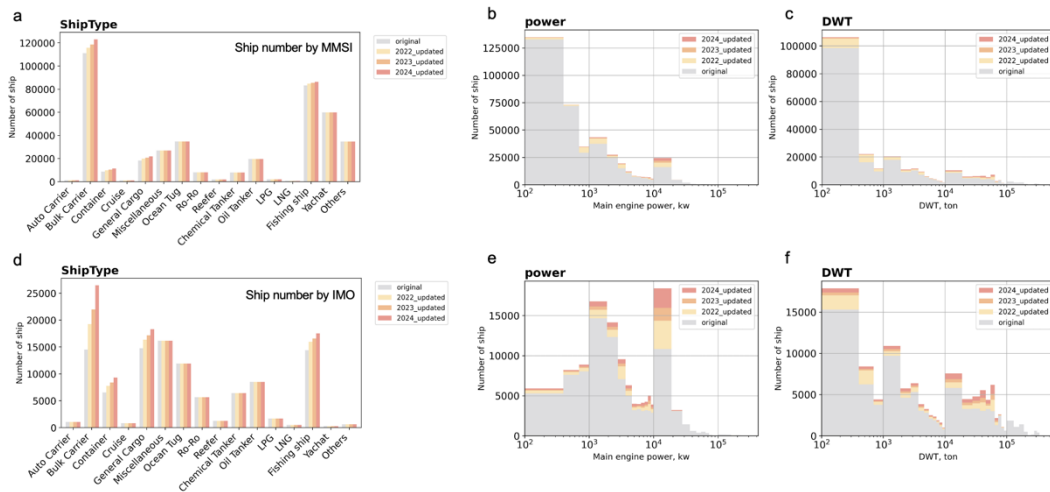


Figure S5. The composition of ship types (a), main engine power (b), and dead tonnage weight (c) for annual evolution of STSD with ship number by MMSI. The composition of ship types (d), main engine power (e), and dead tonnage weight (f) for annual evolution of STSD with ship number by IMO.

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. Emission factors for specific air pollutants and GHGs applied in this study

(unit: g/kwh)

	Main Engine			LNG	Auxiliary Engine		Boiler	
	SSD	HFO/MGO			HFO/MGO /GDO	LNG	HFO/MGO /GDO	LNG
		MSD	HSD					
NOx (Tier0)	18.1	14	10					
NOx (Tier1)	17	13	9.8	1.3	11.2	1.3	2.1	1.3
NOx (Tier2)	14.4	10.5	7.7					
NOx (Tier3)	3.4	2.6	2					
CO	0.54	0.54	0.54	1.3	0.54	1.3	0.2	0.2
HC	0.632	0.527	0.527	0.5	0.4	0.5	0.1	0.105
CH4	0.01	0.01	0.01	2.5/5.5a	0.01	5.5	0.002	0.04
N2O	0.03	0.03	0.03	0.02	0.04	0.02	0.045	0.02

Table S3. Chemical component of ship-used fuel and corresponding fuel-based emission factors for CO₂, SO₂, and PM_{2.5}.

Fuel	Carbon content (m/m%)	EFCO _{2,f} (g/gfuel)	Sulfur content (m/m%)	EFSO _{2,f} (g/gfuel)	EFPM _{2.5,e} (g/kwh)		
					Main Engine	Auxiliary Engine	Boiler
HFO (2.43% S)	0.8493	3.114	2.6	0.051	1.39	1.4	1.42
VLSFO (0.5% S)	0.8744	3.206	0.5	0.0098	0.31	0.31	0.31
MGO (0.5% S)	0.8744	3.206	0.5	0.0098	0.31	0.31	0.31
MGO (0.13% S)	0.8744	3.206	0.13	0.0025	0.2	0.2	0.2
LNG	0.75	2.750	~0	0.0000317	0.02	0.02	0.03