



Supplement of

Measurement report: Six-year DOAS observations reveal post-2020 rebound of ship SO₂ emissions in a Shanghai port despite low-sulfur fuel policies

Jiaqi Liu et al.

Correspondence to: Shanshan Wang (shanshanwang@fudan.edu.cn) and Bin Zhou (binzhou@fudan.edu.cn)

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

S1. Overview of Vessel Activity in the WSW Channel.

To provide background information on local ship traffic conditions relevant to the observed SO₂ variations, this section summarizes key characteristics of vessel activity in the WSW channel based on AIS data from 2018 to 2023.

Figure S1 presents the temporal evolution of daily vessel numbers in the channel, including total ships, moving ships, and stationary ships. Seasonal reductions in traffic are evident around the time of the Chinese New Year each year, reflecting holiday-related slowdowns. Throughout the period, the overall number of ship traffic shows a gradual increasing trend. The vessel type composition is also illustrated, showing that cargo ships and passenger boats have remained the predominant categories.

Figure S2 shows daily statistics of the main engine (ME) and auxiliary engine (AE) power of vessels passing through the channel. The ME power is generally much higher than AE power, reflecting the dominant role of propulsion engines in energy consumption and emissions. The large standard deviations in both ME and AE power reflect the diversity of ship types in the WSW channel—ranging from large cargo ships and cruise vessels (with ME power up to 50,000–70,000 kW) to small fishing and harbor boats (tens of kW). In recent years, the upper percentiles of both ME and AE power have increased, suggesting a growing presence of larger or higher-powered vessels in the area.

Vessel speed is another relevant operational parameter. Although instantaneous speed can vary significantly within a single ship's trajectory, it is observed that the maximum speed of vessels operating in this region can reach up to 52.6 knots. At the same time, many ships remain stationary near the shore or move slowly within the channel, typically maintaining speeds around 5–6 knots.

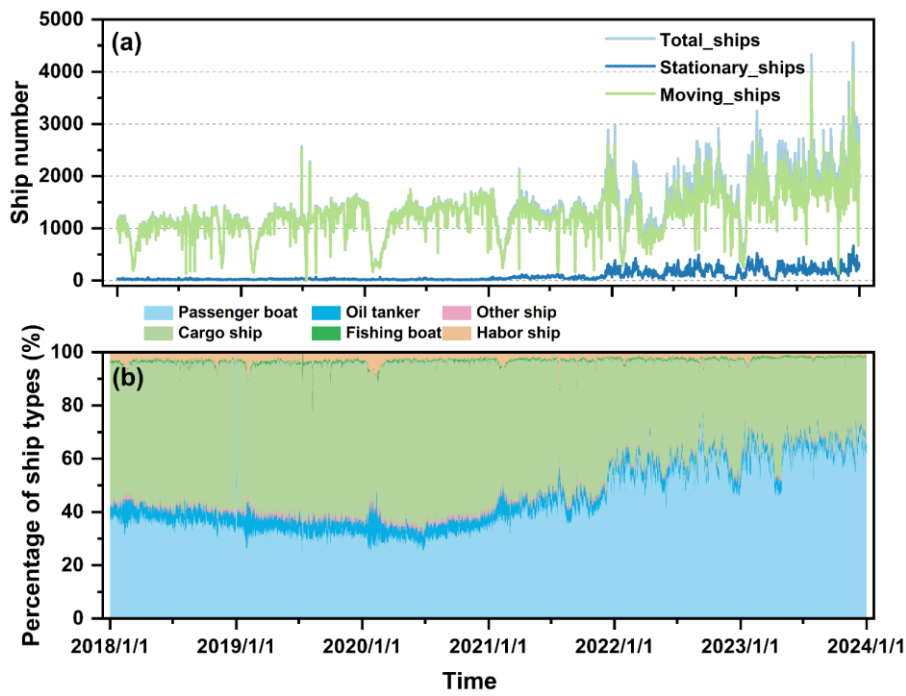


Figure S1. Temporal dynamics of daily ship traffic and ship type composition in the WSW channel (2018–2023). (a) Daily number of total ships, moving ships, and stationary ships detected from AIS records. (b) Percentage composition of different ship types over time, including passenger boats, cargo ships, oil tankers, shipping boats, harbor ships, and other vessels.

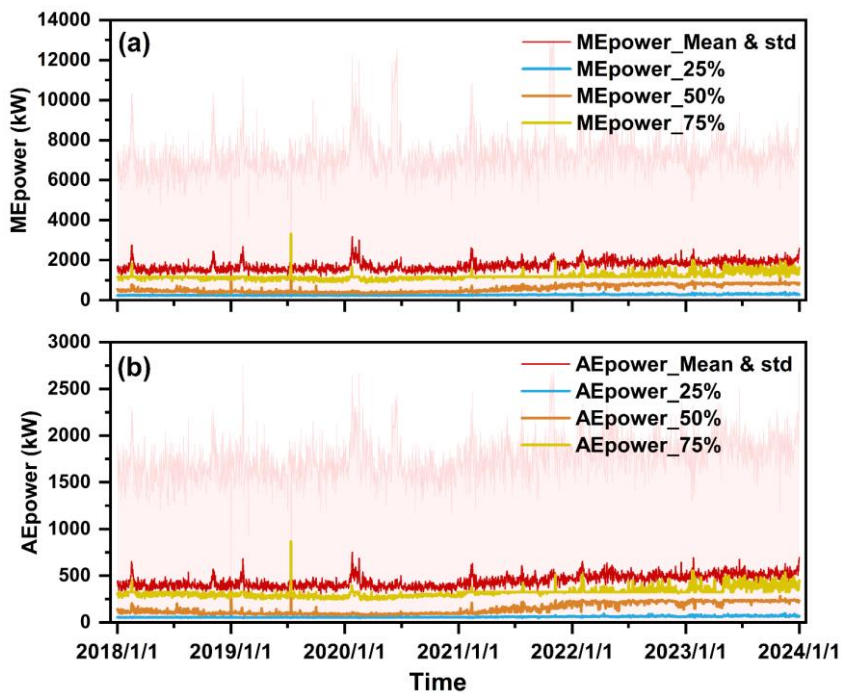


Figure S2. Temporal statistics of daily main engine and auxiliary engine power of vessels in the WSW channel (2018–2023). (a) Time series of main engine (ME) power, showing the mean $\pm$ standard deviation (shaded area) and the 25th, 50th, and 75th percentiles of power (kW). (b) Time series of auxiliary engine (AE) power, showing the mean $\pm$ standard deviation (shaded area) and the 25th, 50th, and 75th percentiles of power (kW).

S2. DOAS measurements and spectral data processing.

The active DOAS system for measuring SO₂ in Wusong Warf (WSW) is placed in the Wusong Maritime Safety Administration and the opposite of Huangpu River (the arrow of reflection mirror). While the system for Fudan University (FDU) is the Environmental Science Building and the southwest corner of campus (the arrow of reflection mirror). Both systems equipped with a 150 W Xenon lamp (Hamamatsu Photonics (China) Co., Ltd) and spectrometers (B&W Tek) with wavelength ranges of 185~440 nm. WSW's DOAS equipment can also observe the concentration changes of NO₂, HONO, O₃ and HCHO in the channel. To eliminate the effect of dark currents and atmospheric scattering light on the measurement, we subtracted the background spectrum collected by blocking the lamp's emitted light from the normally collected spectrum during the measurement spectrum sampling period. The average time resolution was 5~6 min for FDU's systems and 1~2 min for WSW's system. The observation site characteristics and the experimental setup for SO₂, NO₂, HONO, HCHO and O₃ measurements have been described in detail in previous studies (Liu et al., 2024; Guo et al., 2020).

The principle of DOAS is to utilize the narrow-band absorption characteristics of trace gases to determine the gas composition and to infer the concentration of trace gases. DOAS measurements of SO₂ have been mentioned in many studies and is a mature monitoring and analysis method (Wang et al., 2019; Jin et al., 2016; Cheng et al., 2019). In this study, the absorption cross sections of SO₂, HCHO, NO₂, O₃ and the solar spectrum were involved in SO₂ spectral fitting, using wavelength range of 299~308 nm. Based on the signal-to-noise ratio at a given optical path length and integration time, the detection limit in WSW is about 0.13 ppbv for SO₂, 0.51 ppbv for NO₂, 2.51 ppbv for O₃, 1.10 ppbv for HCHO, respectively. While the detection limit for SO₂ in FDU is 0.11 ppbv. The spectral fitting information of all the trace gas can be referred to **Table S1**.

Table S1. The detection limits of DOAS retrieval and the analytical residual.

Observed Station	Trace gas	Fitting window (nm)	absorption cross sections	Polynomial degree	Detection limits	Residuals
WSW	SO ₂	299~308	SO ₂ (Vandaele et al., 2009), NO ₂ (Voigt et al., 2002), HONO (Stutz et al., 2000), HCHO (Meller and Moortgat, 2000), and solar spectrum (Kurucz, 1984)	5	0.13 ppbv	0.00054
	NO ₂	365.3-380.4	NO ₂ (Voigt et al., 2002), HONO (Stutz et al., 2000), HCHO (Meller and Moortgat, 2000), and solar spectrum (Kurucz, 1984)	5	0.51 ppbv	0.00043
	O ₃	280.6-290.6	O ₃ (Voigt et al., 2001a; Voigt et al., 2001b), SO ₂ (Vandaele et al., 1998), HCHO (Meller and Moortgat, 2000), and NO ₂ (Voigt et al., 2002)	5	2.51 ppbv	0.00154
	HCHO	313~341	HCHO (Meller and Moortgat, 2000), NO ₂ (Voigt et al., 2002), SO ₂ (Vandaele et al., 1998), O ₃ (Voigt et al., 2001a), HONO (Stutz et al., 2000)	5	1.10 ppbv	0.00057
FDU	SO ₂	299~308	SO ₂ (Vandaele et al., 1998), NO ₂ (Voigt et al., 2002), HONO (Stutz et al., 2000), HCHO (Meller and Moortgat, 2000), and solar spectrum (Kurucz, 1984)	5	0.11 ppbv	0.00045

S3. Machine learning data input, model tuning, and performance evaluation.

We selected eight representative models as candidates, including one non-ensemble algorithms DecisionTreeRegressor (DTR), and seven ensemble models of ExtraTreeRegressor (ETR), RandomForestRegressor (RF), GradientBoostingRegressor (GBR), BaggingRegressor (BR), AdaBoostRegressor (ABR), XGBRegressor (XGB) and LGBMRegressor (LGBR). Those above algorithms are integrated in SCIKIT-LEARN (sklearn), an open-source machine learning library written in PYTHON (Hackeling, 2017). After performing grid search and 5-fold cross-validation, XGB and ETR were selected as models for filling missing data and de-meteorologizing according to their Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2).

XGB is an optimized distributed gradient enhancement library designed for efficiency, flexibility, and portability. It implements machine learning algorithms in the Gradient Boosting framework. Gradient Boosting, proposed by Friedman (2001), is a large class of algorithms in Boosting. Its idea is borrowed from gradient descent, and its basic principle is to train newly added weak learners based on the negative gradient information of the loss function of the current model and then combine the trained learners into the existing model in an accumulative form. The ETR model not only randomly selects data samples when constructing each decision tree, but also randomly selects eigenvalues to be partitioned when the nodes split. This extreme randomness reduces the variance of the model, thus obtaining a better ability to reduce overfitting and improve model stability than Random Forest models (Gall et al., 2011).

When training the model to fill the missing SO_2 values at FDU, seven meteorological data (from ERA5), time-series NO_2 concentration data observed by DOAS at FDU (representing the emission trends most closely associated with SO_2 variations within the FDU campus environment), and SO_2 data from two nearby stations, Hongkou Station and Ynagpu Station, during the same period (providing a macroscopic perspective of SO_2 variation across Shanghai's land areas) were utilized. With model R^2 is 0.86, RMSE is 0.42 ± 0.13 .

When training the model to fill the missing SO_2 values at WSW, three categories of input features were incorporated to comprehensively capture environmental influences from different sources: meteorological conditions, ship emissions, and urban land-based emissions. Specifically, these consisted of: seven meteorological variables from the ERA5 reanalysis dataset; co-measured pollutant data (including HCHO, HONO, O_3 , and NO_2) obtained via DOAS at the WSW site—which facilitated indirect capture of ship emission signals through cross-species learning; and meteorologically normalized SO_2 data from the FDU site (Deweathered_FDU), representing background variations associated with urban land-based emissions. The model achieved an R^2 of 0.76 and an RMSE of 0.65 ± 0.21 . The completed SO_2 concentration time series is presented in **Figure S3**.

The selection of predictor variables to represent ship emissions involved multiple rounds of testing and evaluation. Initial attempts to incorporate AIS-derived indicators, such as vessel counts and hourly bottom-up emission inventories within a 4 km radius around the WSW site, showed no significant correlation with observed SO_2 concentrations at the hourly scale—their inclusion resulted in negligible improvement in model performance. This outcome is attributed to the fact that AIS-based ship number do not capture distinctions in ship type, size, or operational status. For raw bottom-up emission inventories, it's spatially aggregated and cannot be readily matched to the high temporal resolution of hourly LP-DOAS measurements. Consequently, the approach shifted toward using co-measured pollutants (NO_2 , HCHO, HONO, O_3) obtained at the same WSW site, which are strongly influenced by ship activities.

Figure S4 presents the residual error plots and their frequency distribution between the predicted and observed SO_2 concentrations for both sites. **Figure S5** shows the scatter plots of the predicted versus observed SO_2 , along with the correlation coefficients (R^2). The results demonstrate that the mean residuals are negligible (0.0032 ppbv at

WSW and 1.16×10^{-5} ppbv at FDU). The majority of daily residuals (59.36% at WSW and 86.9% at FDU) fall within ± 0.2 ppbv, and the high R^2 values (above 0.9) confirm a strong model-observation agreement at both locations.

To evaluate the performance of the machine learning-based gap-filling algorithm, a point-to-point comparison was conducted between predicted and observed SO_2 concentrations. The evaluation used an independent validation dataset from 2024, consisting of 641 hourly measurements obtained during naturally continuous observation windows in January, February, and March. As shown in **Figure S6**, the gap-filled SO_2 concentrations (Predicted SO_2) demonstrate strong agreement with observed SO_2 . The model achieved an R^2 of 0.84, with an RMSE of 0.41 ppbv and MAE of 0.29 ppbv. The overall mean observed SO_2 concentration was 1.42 ppbv, compared to a predicted mean of 1.38 ppbv. The model accurately reproduced observed values across different concentration ranges: within the 1–3 ppbv interval, the predicted mean (1.74 ppbv) was nearly identical to the observed mean (1.75 ppbv), and for higher concentrations (3–5 ppbv), the predicted mean (3.88 ppbv) remained close to the observed value (3.66 ppbv). The model's ability to capture short-term SO_2 episodes—critical for characterizing ship plumes—was also evaluated. Among all data points, 1.25% exceeded 5 ppbv. For these high-concentration events, the predicted mean was 4.71 ppbv compared to an observed mean of 5.45 ppbv. The predicted maximum (5.94 ppbv) closely matched the observed maximum (6.08 ppbv). Although the reconstruction of peak concentrations shows a slight underestimation—likely due to the lower frequency of high-concentration events limiting training examples—the model overall captures the temporal variations in SO_2 concentrations well in the waterway environment.

Deweathered models used seven meteorological data (from ERA5) and time-related variables (Unix time, Julian day, and day of the week) to capture emission patterns. ERA5 meteorological data from the same hour within 14 days before and after each target time was used, providing 1,276 sets of meteorological data for training on the Observed_WSW and Observed_FDU series, as showed in **Figure S7**. The repeated input of over 1,000 sets of real historical meteorological data can be considered sufficient to effectively account for the impacts of various real-world meteorological conditions. For the Deweathered models constructed for both WSW and FDU, the RMSE, R^2 , are reported as 0.41 ± 0.12 and 0.79 in WSW, 0.32 ± 0.14 and 0.85 in FDU, respectively.

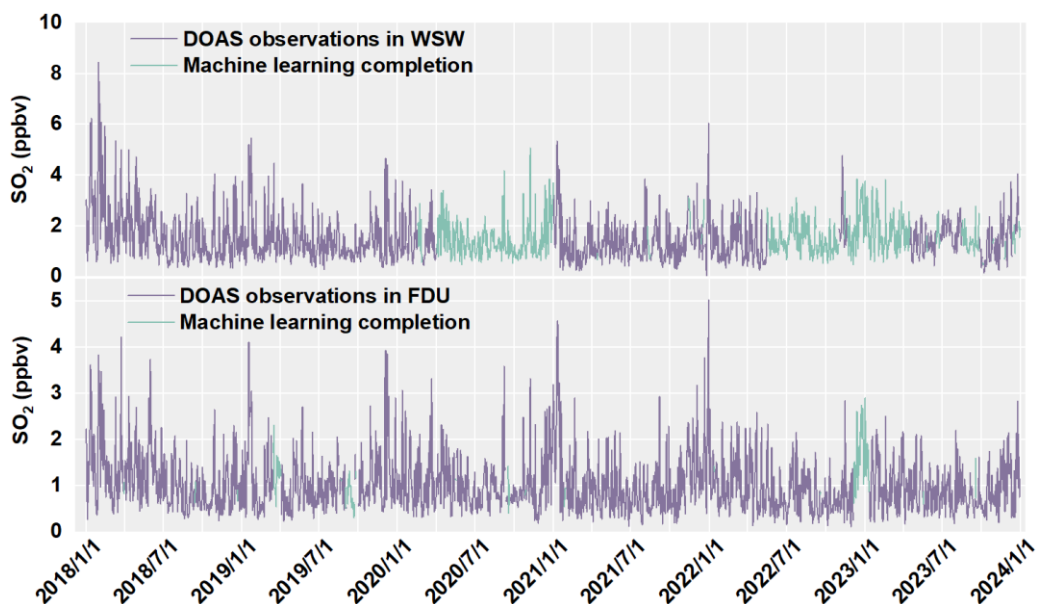


Figure S3. Time series of SO_2 for WSW and FDU complemented by DOAS observations and machine learning models.

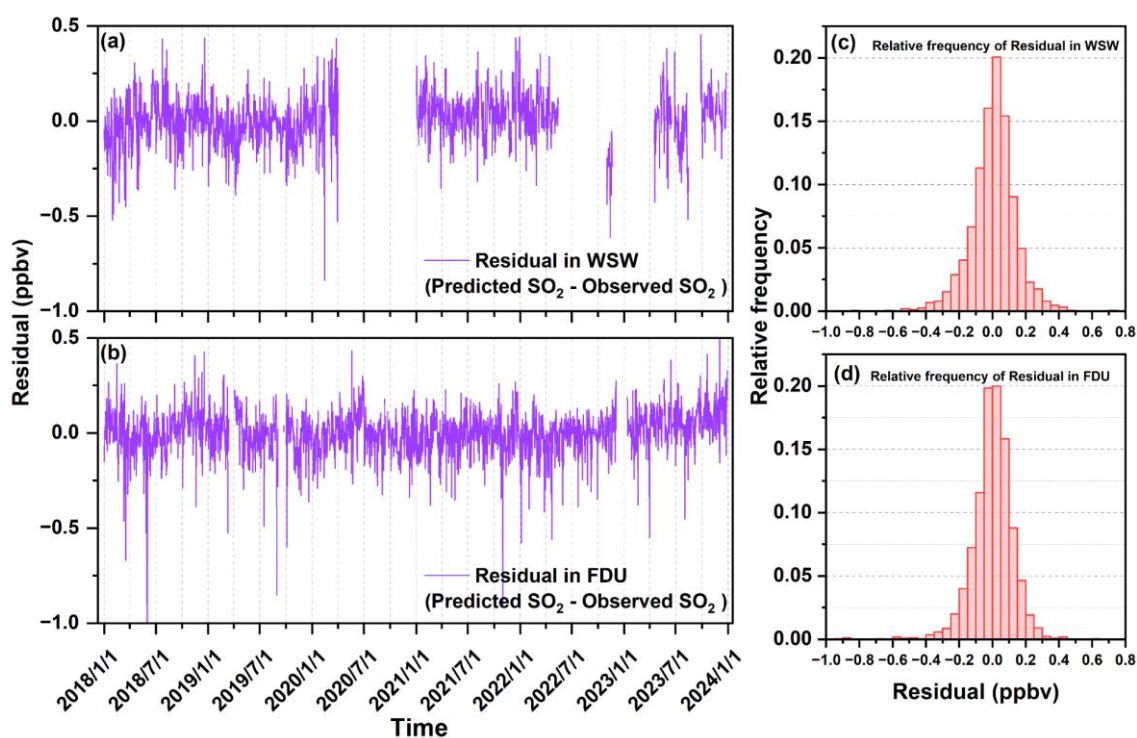


Figure S4. Time series and frequency distribution of residuals (Predicted SO_2 – Observed SO_2) at the daily mean scale for (a, c) WSW and (b, d) FDU during 2018–2023.

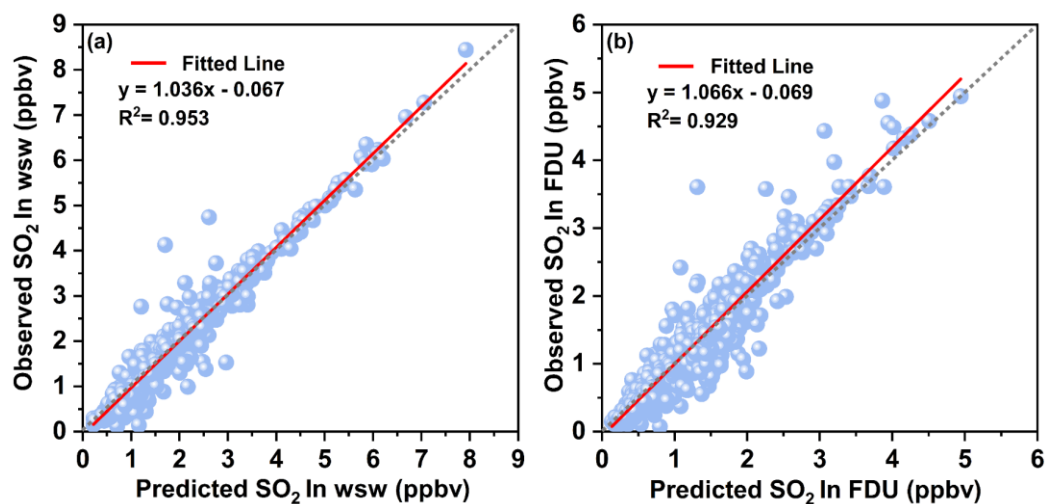


Figure S5. Scatter plots between predicted and observed SO_2 concentrations at the daily mean scale for (a) WSW and (b) FDU.

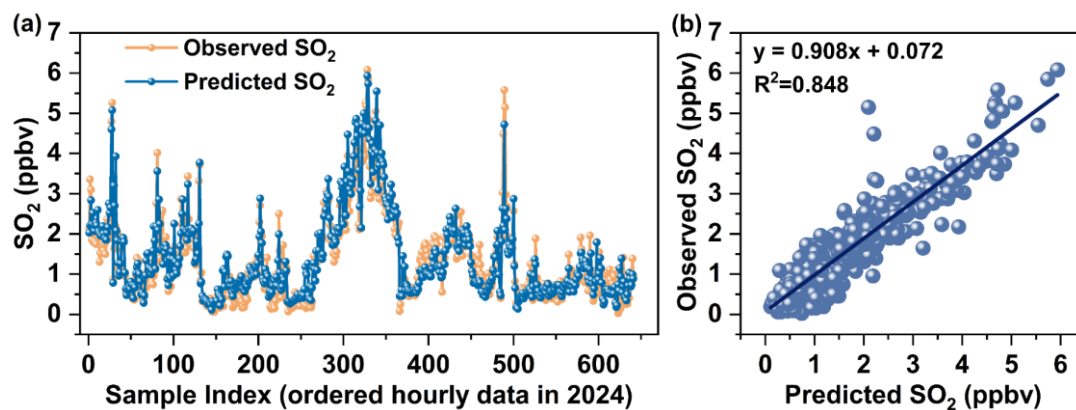


Figure S6. Comparison between observed and machine learning-predicted hourly SO_2 concentrations at WSW in 2024. (a) Temporal variation using ordered sample index. (b) Regression plot showing strong agreement ($R^2 = 0.848$) between predicted and observed values.

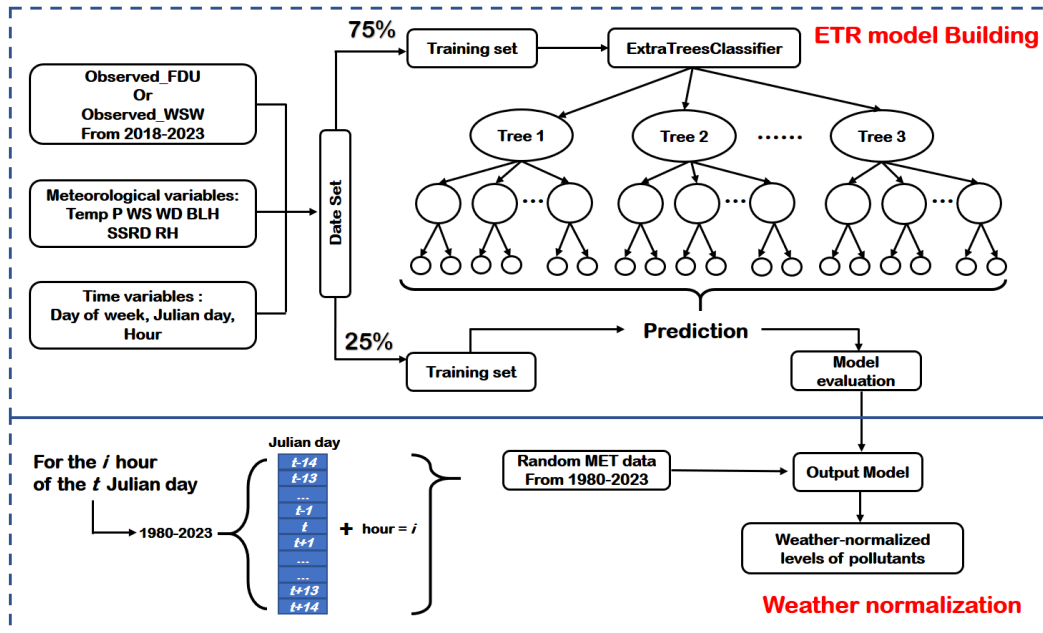


Figure S7. Deweathered model training. First train the models of Observed_WSW and Observed_FDU with temporal and meteorological feature parameters respectively for 6 years; then use the models to train the meteorological parameters for specified time periods from 1980 to 2023 and normalize the results.

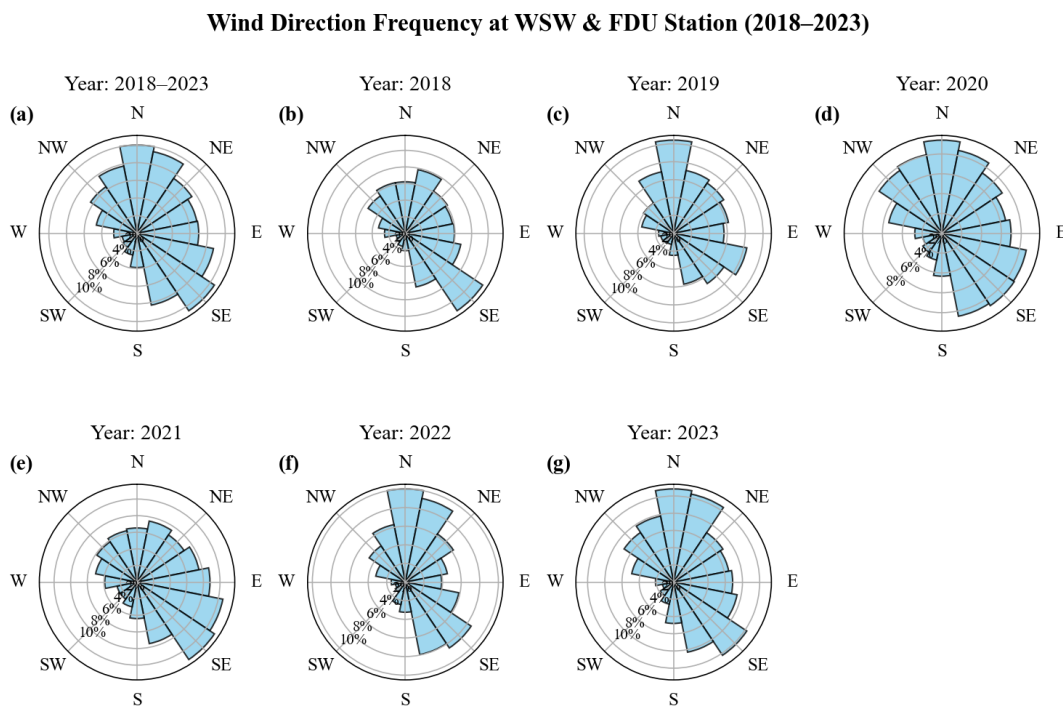


Figure S8. Wind direction frequency distribution at WSW and FDU station from 2018 to 2023. (a) The aggregated wind distribution for all years. (b)–(g) The show annual wind patterns from 2018 to 2023. Wind direction is plotted in polar coordinates with percentage frequency indicated by concentric circles.

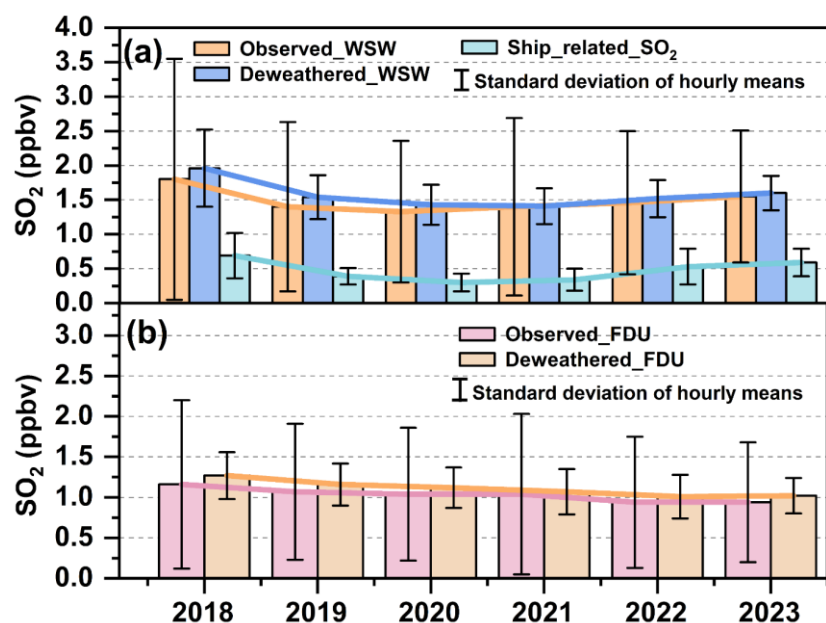


Figure S9. Yearly average SO₂ concentrations at two sites from 2018 to 2023. (a) Observed and deweathered SO₂ concentrations at the WSW site, with the contribution of ship-related SO₂. The orange bars represent observed SO₂ concentrations (Observed_WSW), the blue bars represent deweathered SO₂ (Deweathered_WSW), and the green line with stars shows ship-related SO₂. (b) Observed and deweathered SO₂ concentrations at the FDU site. The pink bars represent observed SO₂ (Observed_FDU), while the orange bars represent deweathered SO₂ (Deweathered_FDU). Error bars represent the standard deviation across hourly mean values.

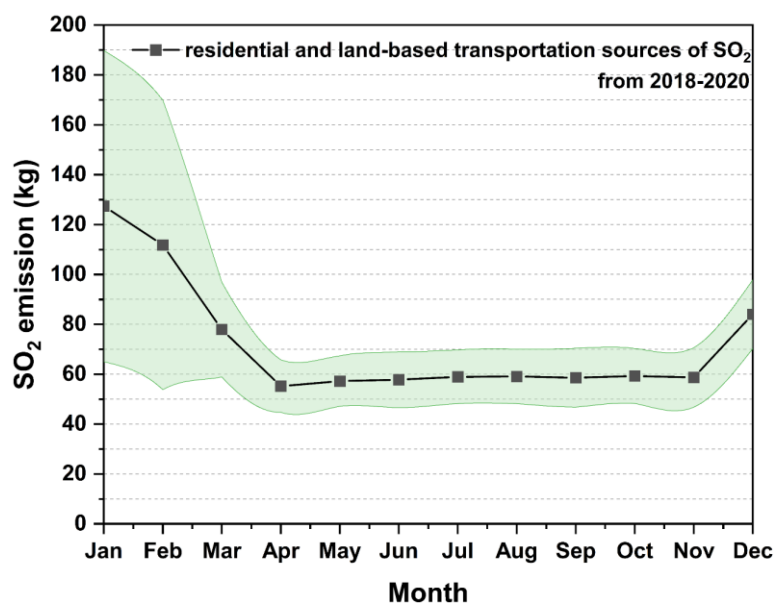


Figure S10. Monthly average values of SO₂ from residential and land-based transportation sources in Shanghai from 2018-2020 in China's Multi-Resolution Emission Inventory (MEIC).

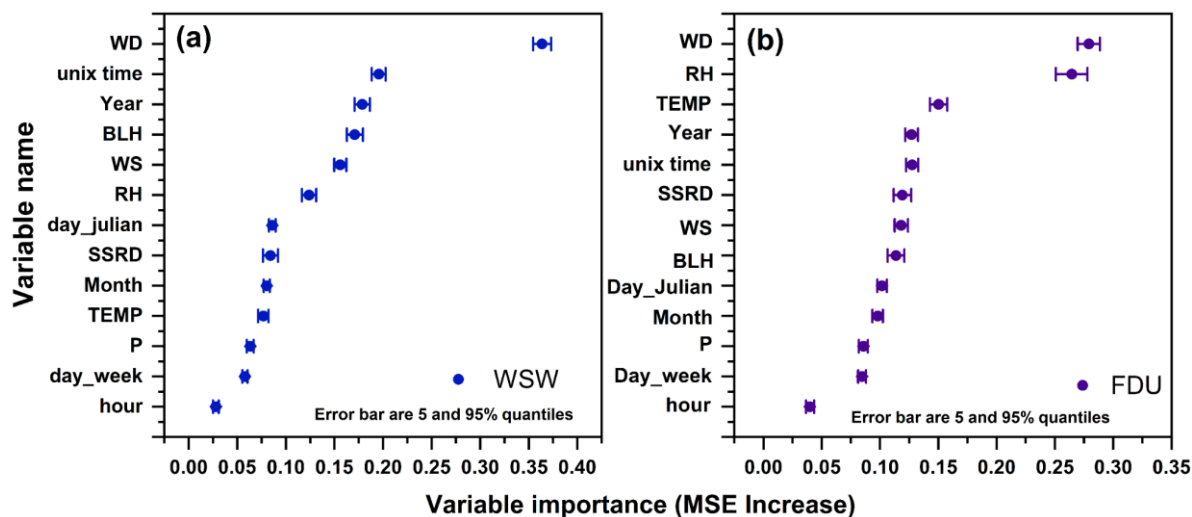


Figure S11. Variable importance plot for SO₂ at (a) WSW and (b) FDU between 2018 and 2023 calculated by 50 ETR models. The Mean Squared Error (MSE) increase quantifies how much predictive accuracy depends on each variable; a higher value denotes greater importance.

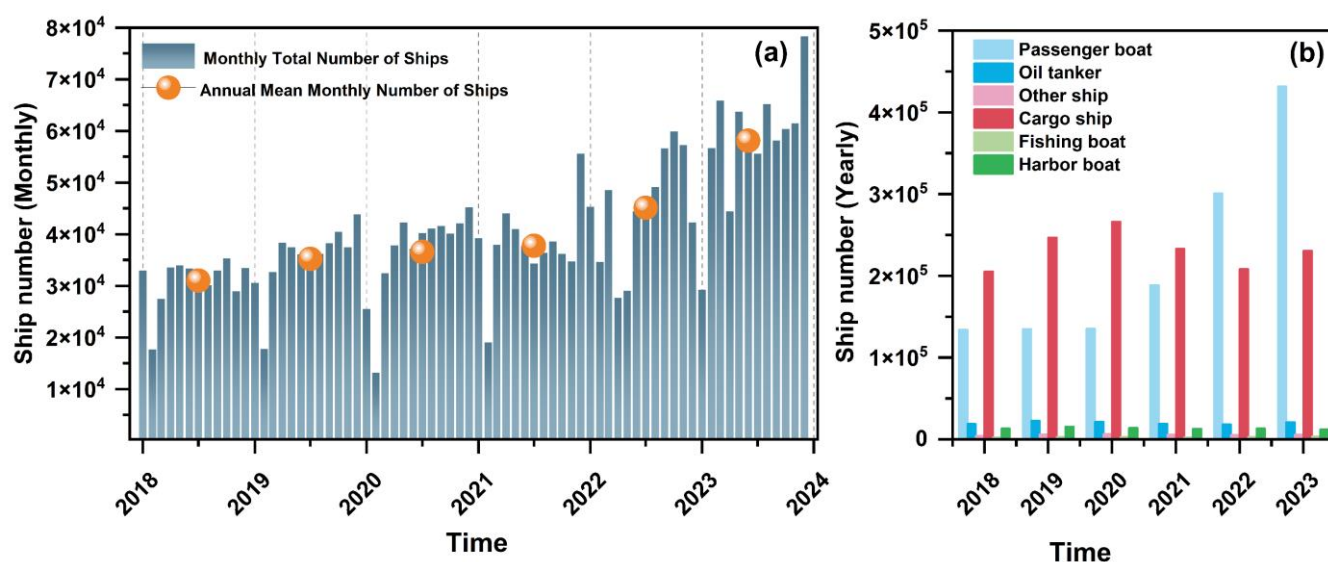


Figure S12. Annual variation of shipping activity in the channel from 2018 to 2023. (a) Monthly total number of ships and annual mean values. (b) Yearly ship number by ship type (cargo, oil tanker, passenger boat, fishing boat, and harbor boat). (For a more robust parameter of activity, a ship emission inventory (Section S5 in the Supplement) was created, incorporating ship number, type, ME & AEpowers, and speed for comparison with Ship_{related}SO₂).

S4. Matching SO₂ Peaks with Ships.

The process of matching SO₂ peaks with ships involves extracting the occurrence time of SO₂ plume signals identified by the BEAD_s algorithm and searching for vessel trajectories in the cleaned AIS dataset that coincide with the plume occurrence time and intersect the light path. AIS ship track matching logic is as follows:

A straight-line equation is derived using the latitude and longitude coordinates of the DOAS emitter and the reflector as waypoints. The plume time t_{plume} is input into the system to determine, for any vessel position, whether the condition $(t_1 < t_{\text{plume}} < t_2) \cap (t_2 - t_1 < 10\text{mins}) \cap (\text{speed}_1 \neq 0) \cap (\text{speed}_2 \neq 0) \cap (\text{trajectory crossing the light path})$ is satisfied. For any given peak, multiple vessels may potentially be matched. Assuming uniform velocity change for the vessel between t_1 and t_2 . As showed in **Figure S13**.



Figure S13. Schematic Diagram of the Matching Between Ship AIS Data and SO₂-rich plume signals. With ship's positions at two locations (Location₁ and Location₂) are shown along with their respective timestamps (t_1 and t_2) and speed (Speed₁ and Speed₂). The bidirectional arrows indicate the light path of DOAS. t_{plume} indicates when the SO₂-rich plume appeared. Base map: © OpenStreetMap contributors 2025. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

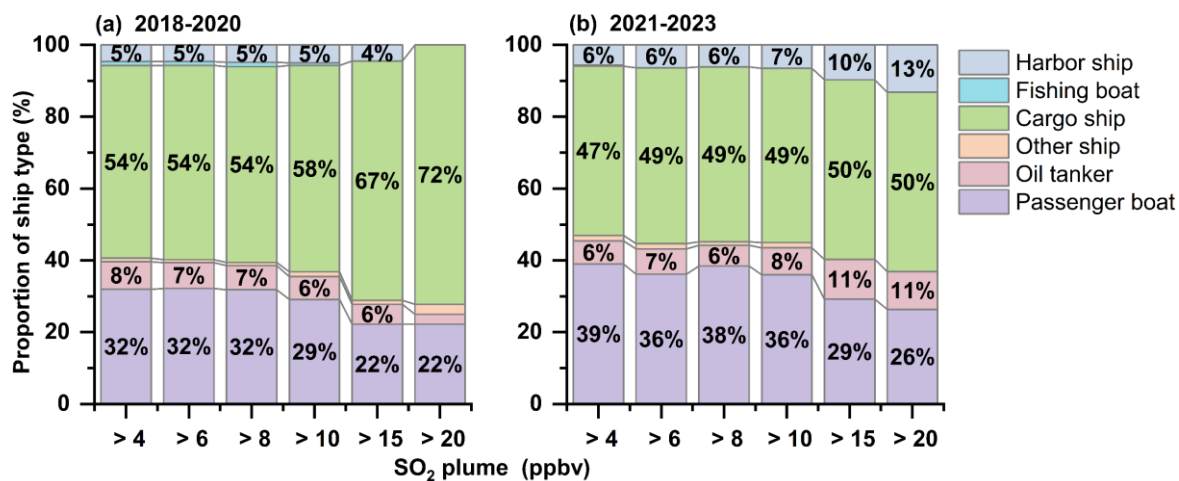


Figure S14. Proportional distribution of Fastest-Matching ship types corresponding to different peak SO₂ values during (a) the policy adjustment period (2018–2020) and (b) the policy stabilization period (2021–2023). The ship types include harbor ships, fishing boats, cargo ships, passenger boats, oil tankers, and other ships.

S5. Comparison Between Observational Data and AIS-Based Ship Emission Inventory.

In the paragraph of this appendix, we compared Ship_related_SO₂ derived from DOAS observations with those estimated by traditional bottom-up ship emission inventories, discussed the similarities and differences in outcome trends between the two approaches, and identified the underlying causes. AIS data provides detailed information on ship activities and is commonly used for calculating ship emission inventories on large spatiotemporal scales (Mao et al., 2020; Zou et al., 2020).

The reason for employing a comprehensive ship emission inventory from AIS, rather than relying on any single ship parameter (e.g., ship number, engine power, or speed), is as follows: While parameters like ship number, main engine power, and speed are valuable indicators, they are independently insufficient to accurately represent actual SO₂ emissions. This is because emissions are the product of a complex interplay of these factors. For instance: A high-powered ship moving slowly may emit similarly to a lower-powered ship at high speed; A stationary ship using its auxiliary engine for onboard services may emit more than a ship maneuvering at low speed with its main engine at idle; Simply counting all vessels equally ignores the vast differences in emission potential between a large container ship and a small fishing boat.

Therefore, a bottom-up emission inventory methodology was adopted (**Section S6 in the Supplement**). This approach synthesizes the key parameters derived from AIS data—including ship type, instantaneous position and speed, and installed main and auxiliary engine power—into a holistic framework. By applying standardized emission algorithms and fuel sulfur content assumptions, this inventory translates dynamic ship activity into estimated hourly SO₂ emissions.

The scatter plots in **Figure S15** illustrate the correlation (R^2) between ship emission inventory-based SO₂ emissions and the 14-day mean SO₂ concentrations based on observation at the WSW site. In the process of removing meteorological influences and land-based emissions, the correlation between the ship emission inventory and SO₂ concentrations progressively improves step by step. For the period from 2018 to 2020, the R^2 increases from 0.064 (Observed_SO₂) to 0.154 (Deweathered_SO₂), and further to 0.32 (Ship_related_SO₂). Similarly, for the period from 2021 to 2023, the R^2 rises from 0.043 (Observed_SO₂) to 0.163 (Deweathered_SO₂), and ultimately reaches 0.54 (Ship_related_SO₂). This trend underscores the effectiveness of the combined meteorological normalization and land-based emissions subtraction processes in refining our understanding of Ship_related_SO₂ contributions. Compared with directly observed_SO₂, the emissions inventory explains the trend of Ship_related_SO₂ changes better.

Figure S16 illustrates the 14-day mean variations of Ship_related_SO₂ concentrations and ship emission inventory in the WSW from 2018 to 2023. During the policy adjustment period (2018–2020), both the Ship_related_SO₂ and the corresponding SO₂ emissions in the inventory showed a gradual decline. If all ships had complied with the low-sulfur fuel policy, SO₂ emissions from ships would have shown a sharp decrease at the early stage of policy implementation, as illustrated in **Figure S16c**. However, due to the presence of non-compliant ships (as discussed in Sections 3.2 and 3.3), the reduction in SO₂ emissions from ships has been a gradual process, as shown in **Figure S16a**. While the consistency between Ship_related_SO₂ and the inventory improved during the policy stabilization period (2021–2023) in **Figure S15f**, which means that the fuel use of ships is closer to the policy requirements.

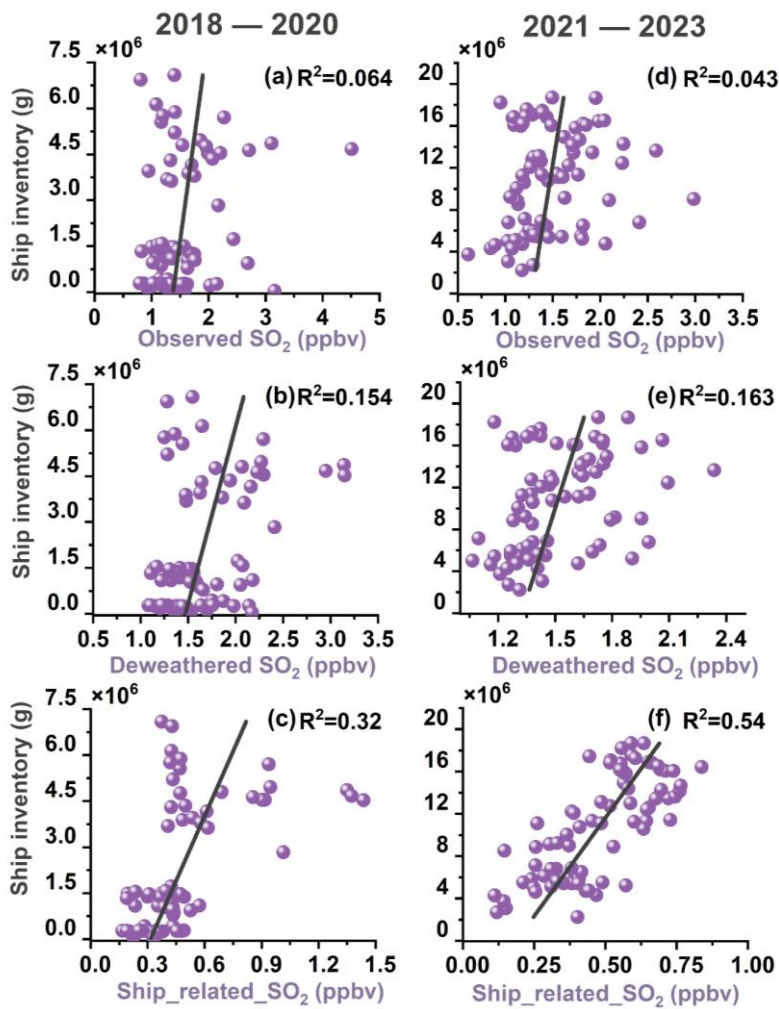


Figure S15. Correlations between 14-day mean SO₂ concentrations (x-axis) at WSW site and ship SO₂ inventory (y-axis), divided into three categories: (a, d) Observed_SO₂ concentrations, (b, e) Deweathered_SO₂ concentrations, and (c, f) Ship_related_SO₂ concentrations. (a–c) correspond to the policy adjustment period from 2018 to 2020, while panels (d–f) represent the policy stabilization period from 2021 to 2023.

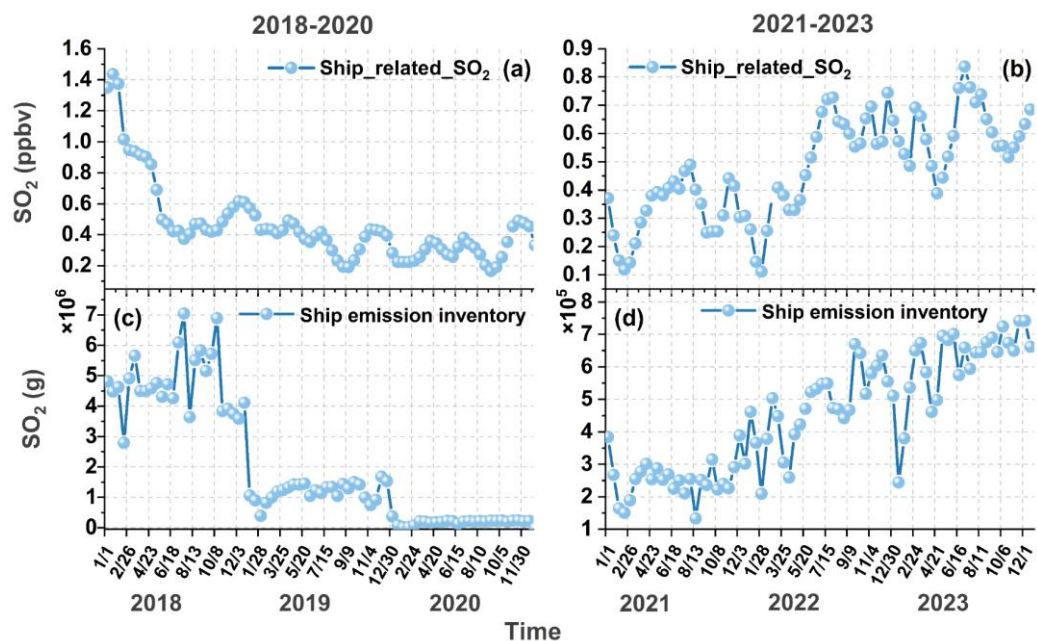


Figure S16. 14-day mean variations of Ship_related_SO₂ concentrations and emission inventory in the WSW channel from 2018 to 2023. (a) and (b) represent the 14-day mean Ship_related_SO₂ derived from observations for 2018–2020 and 2021–2023, respectively. (c) and (d) show the corresponding 14-day mean SO₂ emissions from the ship emission inventory during the same periods.

S6. Calculation of ship emission inventory based on AIS.

The ship atmospheric pollutant emission inventories used in this study were calculated using a bottom-up ship emission model based on AIS data. The model was improved upon the studies by Fan et al. (2016) and Feng et al. (2019). The basic framework and specific parameters of the model are introduced as follows.

The model is primarily divided into three parts: data preprocessing, inventory calculation, and gridded output. The model has specific format requirements for the AIS data used. Since the raw AIS data contains extensive information, it is necessary to selectively extract the required information and preprocess it before inputting it into the model for further calculations. The raw AIS data includes dynamic and static information about ships. The dynamic information required by the model mainly includes the following: AIS message type (msgtype), Maritime Mobile Service Identity (MMSI), navigation status (status), current speed (speed), ship's turn direction (turn), current longitude (lon), current latitude (lat), timestamp (second), actual heading (heading), and date (day). The static information primarily includes AIS message type (msgtype), Maritime Mobile Service Identity (MMSI), International Maritime Organization (IMO) number or ship name (shipname), ship type (shiptype), distance from the positioning antenna to the bow (to_bow), distance to the stern (to_stern), distance to the port side (to_port), distance to the starboard side (to_starboard), and the current maximum static draught (draught). Before calculations, the obtained data must be classified and processed. Additionally, the ship's static information needs to be matched with corresponding basic ship information in the database.

After preprocessing the AIS data, the data is input into the calculation model one by one to compute the ship pollutant emissions. In this study, the calculation of atmospheric pollutant emissions from ships is divided into two parts: the main engine and the auxiliary engine. The calculation formula for the pollutant emissions from the main engine, E_m , is shown in Equation (S1), while the calculation formula for the pollutant emissions from the auxiliary engine, E_a , is shown in Equation (S2):

$$E_m = mP \times mLF \times T \times mEF \times mCF \times mLLAM \times mFCF \quad (S1)$$

$$E_a = aP \times aLF \times T \times aEF \times aFCF \quad (S2)$$

$$mLF = (AS/MS)^3 \quad (S3)$$

Those parameters are:

mP/aP : Ship main engine/ auxiliary engine power, kW;

mLF/aLF : Main engine /auxiliary engine load factor;

AS : Current speed of ship, knots;

MS : Max speed of ship, knots;

mEF/aEF : Main engine/auxiliary engine emission factor, g/kWh;

mCF : Main engine control factor;

T : Time step, h;

$mLLAM$: Main engine low load adjustment factor;

$mFCF/aFCF$: Main engine/ auxiliary engine fuel control factor.

After separately calculating the pollutant emissions from the main and auxiliary engines of the ships, a gridding program is applied to aggregate the total pollutant emissions from ships within each grid area. The results are then compiled into the required ship emission inventory for subsequent analysis.

A more detailed description of the emission inventories used in this study can be found in previous articles (Fan et al., 2016; Yuan et al., 2023; Feng et al., 2019).

S7. Limitations and Uncertainties.

Although this study provides valuable insights into the contribution of maritime shipping to ambient SO₂ in Shanghai, several limitations and uncertainties should be acknowledged.

From a data perspective, an additional source of uncertainty lies in the background subtraction method, which assumes that the FDU site accurately represents the urban land-based SO₂ level. In China, stringent emission control policies have led to a substantial reduction in land-based SO₂, and our long-term meteorology-adjusted analysis at FDU confirms that its background concentrations have already declined to relatively low levels with only minor interannual variability. Nevertheless, some degree of spatial heterogeneity in urban SO₂ emissions is unavoidable. As a result, the land-based contributions at FDU and WSW may still differ slightly, introducing potential bias in the background subtraction. However, such uncertainties are unlikely to affect the robustness of our analysis at broader temporal scales (e.g., monthly averages).

From a measurement coverage perspective, another source of uncertainty arises from the limited measurements at the WSW site in 2020 and between July 2022 and July 2023, during which predicted values were used to fill missing periods. Our validation analysis shows that the gap-filling model reproduces long-term SO₂ variations reliably, with a mean residual of -0.0032 ppbv over 2018–2023 (**Figures S4, S5, Section S3 in the Supplement**), although short validation samples (e.g., in 2024) suggest that biases of up to -0.04 ppbv may occasionally occur. Even if the concentrations during 2020–2022 were uniformly adjusted by this margin, the main interannual trends—a decrease from 2018 to 2020 followed by an increase from 2021 to 2023—would remain unchanged. We note, however, that the absence of measurements may reduce the number of high-SO₂ plumes captured during these years. Because our plume-related analyses in Sections 3.2 and 3.3 are based on relative contributions rather than absolute plume counts, this influence is expected to be limited, but some degree of bias cannot be fully excluded.

From a model perspective, the Deweathered approach relies on the choice of input variables and on the assumption that meteorological impacts can be fully captured by the ERA5 parameters and time-related covariates. Other relevant factors, such as local-scale turbulence or unmeasured meteorological drivers, may not be fully represented.

From the experimental design perspective, an important source of uncertainty in this study arises from the vertical sampling geometry of the DOAS system. The light path was located approximately 10 m above ground level, with the observation site itself about 6 m above mean sea level. Tidal variation (1–4 m) and ship stack heights mean that the intercepted section of the SO₂ plume could vary between individual events—capturing different segments of the vertical plume profile depending on stack height and tidal level.

However, the DOAS setup and tidal conditions remained broadly consistent during the entire 2018–2023 period, and vessel types and traffic patterns did not experience abrupt structural changes. Therefore, this geometric uncertainty is systematic and comparable across years, and is unlikely to bias the interannual patterns observed in the plume concentration distributions. Our analysis focuses on the relative frequency of plumes within specific concentration ranges and their temporal trends, rather than on deriving absolute emission rates for individual vessels.

If a quantitative estimation of individual vessel emissions were to be conducted, obtaining the actual stack height of ships would be crucial. Unfortunately, such information is not contained in the AIS system. A feasible solution would be to integrate camera-based observations to capture photographs of vessels passing through the light path at moments of elevated SO₂ signals, allowing stack height and plume geometry to be determined more accurately. This is a direction our group intends to pursue in future work to further reduce the uncertainties associated with vertical sampling geometry.

References

- Cheng, Y., Wang, S., Zhu, J., Guo, Y., Zhang, R., Liu, Y., Zhang, Y., Yu, Q., Ma, W., and Zhou, B.: Surveillance of SO₂ and NO₂ from ship emissions by MAX-DOAS measurements and the implications regarding fuel sulfur content compliance, *Atmos. Chem. Phys.*, 19, 13611-13626, <http://10.5194/acp-19-13611-2019>, 2019.
- Fan, Q., Zhang, Y., Ma, W., Ma, H., Feng, J., Yu, Q., Yang, X., Ng, S. K. W., Fu, Q., and Chen, L.: Spatial and Seasonal Dynamics of Ship Emissions over the Yangtze River Delta and East China Sea and Their Potential Environmental Influence, *Environmental Science & Technology*, 50, 1322-1329, <http://10.1021/acs.est.5b03965>, 2016.
- Feng, J., Zhang, Y., Li, S., Mao, J., Patton, A. P., Zhou, Y., Ma, W., Liu, C., Kan, H., Huang, C., An, J., Li, L., Shen, Y., Fu, Q., Wang, X., Liu, J., Wang, S., Ding, D., Cheng, J., Ge, W., Zhu, H., and Walker, K.: The influence of spatiality on shipping emissions, air quality and potential human exposure in the Yangtze River Delta/Shanghai, China, *Atmos. Chem. Phys.*, 19, 6167-6183, <http://10.5194/acp-19-6167-2019>, 2019.
- Friedman, J. H.: Greedy function approximation: a gradient boosting machine, *Ann. Stat.*, 1189-1232, 2001.
- Gall, J., Yao, A., Razavi, N., Gool, L. V., and Lempitsky, V.: Hough Forests for Object Detection, Tracking, and Action Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33, 2188-2202, <http://doi.org/10.1109/TPAMI.2011.70>, 2011.
- Guo, Y., Wang, S., Gao, S., Zhang, R., Zhu, J., and Zhou, B.: Influence of ship direct emission on HONO sources in channel environment, *Atmospheric Environment*, 242, 117819, <https://doi.org/10.1016/j.atmosenv.2020.117819>, 2020.
- Hackeling, G.: Mastering Machine Learning with scikit-learn, Packt Publishing Ltd 2017.
- Jin, J., Ma, J., Lin, W., Zhao, H., Shaiganfar, R., Beirle, S., and Wagner, T.: MAX-DOAS measurements and satellite validation of tropospheric NO₂ and SO₂ vertical column densities at a rural site of North China, *Atmospheric Environment*, 133, 12-25, <https://doi.org/10.1016/j.atmosenv.2016.03.031>, 2016.
- Kurucz, R. L.: Solar Flux Atlas from 296 to 1300 nm, National Solar Observatory Atlas, 1, <https://doi.org/10.1017/S0074180900035427>, 1984.
- Liu, J., Wang, S., Zhang, Y., Yan, Y., Zhu, J., Zhang, S., Wang, T., Tan, Y., and Zhou, B.: Investigation of formaldehyde sources and its relative emission intensity in shipping channel environment, *Journal of Environmental Sciences*, 142, 142-154, <https://doi.org/10.1016/j.jes.2023.06.020>, 2024.
- Mao, J., Zhang, Y., Yu, F., Chen, J., Sun, J., Wang, S., Zou, Z., Zhou, J., Yu, Q., Ma, W., and Chen, L.: Simulating the impacts of ship emissions on coastal air quality: Importance of a high-resolution emission inventory relative to cruise- and land-based observations, *Science of The Total Environment*, 728, 138454, <https://doi.org/10.1016/j.scitotenv.2020.138454>, 2020.
- Meller, R. and Moortgat, G. K.: Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225-375 nm, *J. Geophys. Res.: Atmos.*, 105, 7089-7101, <https://doi.org/10.1029/1999JD901074>, 2000.
- Stutz, J., Kim, E., Platt, U., Bruno, P., Perrino, C., and Febo, A.: UV-visible absorption cross sections of nitrous acid, *J. Geophys. Res.: Atmos.*, 105, 14585-14592, <https://doi.org/10.1029/2000JD900003>, 2000.
- Vandaele, A. C., Hermans, C., and Fally, S.: Fourier transform measurements of SO₂ absorption cross sections: II.: Temperature dependence in the 29000–44000cm⁻¹ (227–345nm) region, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110, 2115-2126, <https://doi.org/10.1016/j.jqsrt.2009.05.006>, 2009.

- Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A., and Coquart, B.: Measurements of the NO₂ absorption cross-section from 42 000 cm⁻¹ to 10 000 cm⁻¹ (238–1000 nm) at 220 K and 294 K, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 59, 171-184, [https://doi.org/10.1016/S0022-4073\(97\)00168-4](https://doi.org/10.1016/S0022-4073(97)00168-4), 1998.
- Voigt, S., Orphal, J., and Burrows, J. P.: The temperature and pressure dependence of the absorption cross-sections of NO₂ in the 250-800 nm region measured by Fourier-transform spectroscopy, *J. Photochem. Photobiol., A*, 149, 1-7, [https://doi.org/10.1016/s1010-6030\(01\)00650-5](https://doi.org/10.1016/s1010-6030(01)00650-5), 2002.
- Voigt, S., Orphal, J., Bogumil, K., and Burrows, J.: The temperature dependence (203-293 K) of the absorption cross sections of O₃ in the 230-850 nm region measured by Fourier-transform spectroscopy, *J. Photochem. Photobiol., A*, 143, 1-9, [https://doi.org/10.1016/S1010-6030\(01\)00480-4](https://doi.org/10.1016/S1010-6030(01)00480-4), 2001a.
- Voigt, S., Orphal, J., Bogumil, K., and Burrows, J. P.: The temperature dependence (203–293 K) of the absorption cross sections of O₃ in the 230–850 nm region measured by Fourier-transform spectroscopy, *Journal of Photochemistry and Photobiology A: Chemistry*, 143, 1-9, [https://doi.org/10.1016/S1010-6030\(01\)00480-4](https://doi.org/10.1016/S1010-6030(01)00480-4), 2001b.
- Wang, Y., Dörner, S., Donner, S., Böhnke, S., De Smedt, I., Dickerson, R. R., Dong, Z., He, H., Li, Z., Li, Z., Li, D., Liu, D., Ren, X., Theys, N., Wang, Y., Wang, Y., Wang, Z., Xu, H., Xu, J., and Wagner, T.: Vertical profiles of NO₂, SO₂, HONO, HCHO, CHOCHO and aerosols derived from MAX-DOAS measurements at a rural site in the central western North China Plain and their relation to emission sources and effects of regional transport, *Atmos. Chem. Phys.*, 19, 5417-5449, <http://doi.org/10.5194/acp-19-5417-2019>, 2019.
- Yuan, Y., Zhang, Y., Mao, J., Yu, G., Xu, K., Zhao, J., Qian, H., Wu, L., Yang, X., Chen, Y., and Ma, W.: Diverse changes in shipping emissions around the Western Pacific ports under the coeffect of the epidemic and fuel oil policy, *Science of The Total Environment*, 879, 162892, <https://doi.org/10.1016/j.scitotenv.2023.162892>, 2023.
- Zou, Z., Zhao, J., Zhang, C., Zhang, Y., Yang, X., Chen, J., Xu, J., Xue, R., and Zhou, B.: Effects of cleaner ship fuels on air quality and implications for future policy: A case study of Chongming Ecological Island in China, *Journal of Cleaner Production*, 267, 122088, <https://doi.org/10.1016/j.jclepro.2020.122088>, 2020.