General comments

The manuscript presents a timely and policy-relevant study that quantifies shipping emission factors and apportions organic aerosol sources in Toulon, France, a coastal port city. The authors report detailed EFs for a wide range of pollutants, including SO_2, NO_x, CO, CH_4, black carbon, organics, PAHs, and particle number concentrations. By distinguishing EFs across vessel types and operational modes (e.g., arrivals vs. departures) and comparing results to pre-regulation studies, the work offers clear evidence of the evolving impact of marine fuel regulations on air quality. The ability to resolve vessel-specific EFs and link them to auxiliary engine power, fuel type, and plume characteristics is a notable strength of the study and represents a valuable dataset for emission inventory improvement and regulatory evaluation. The integration of detailed HR-ToF-AMS measurements with Positive Matrix Factorization factors that distinguish between road- and shipping-related emissions of hydrocarbons represents a valuable and novel contribution to the field, particularly in light of the IMO2020 regulation and the forthcoming SECA Med implementation.

That said, the manuscript would benefit substantially from a comprehensive round of language editing, as many sentences could be rewritten for clarity, conciseness, and smoother flow. The readability of the manuscript would benefit from the consistent use of the Oxford comma, which helps to clarify lists and prevent ambiguity. Additionally, contractions (e.g., "don't" in Line 152) should be avoided in formal scientific writing to maintain a professional tone throughout the text. Section 3.2 would benefit from more direct references to the figures to help guide the reader through the results. Currently, the discussion lacks clear connections to the visual data, which makes it difficult to follow the interpretation and significance of the findings. Overall, the study is technically sound and presents novel results, but significant editorial revision is needed to ensure the presentation matches the quality of the science.

Response:

We sincerely thank the reviewer for this very positive and constructive overall evaluation of our work.

In response to these general comments, we conducted a thorough editorial revision of the entire manuscript to enhance clarity, conciseness, and stylistic consistency. Contractions have been removed, the Oxford comma has been applied systematically, and sentence structure has been refined to improve readability and scientific flow.

We also reinforced the connection between the discussion and the visual material, especially in Section 3.2, by explicitly referencing the relevant figures and clarifying their interpretation.

All editorial and stylistic revisions suggested by the reviewer have been implemented consistently throughout the manuscript. These changes are highlighted in blue in the revised version for transparency.

Reviewer comment:

Lines 186–188. "The pie chart representing the median PM1 chemical composition shows the *following proportions: 53 % organics, 16 % sulfate, 7 % ammonium, 2 % nitrate and 21 % BC." Could you provide the standard deviations to provide insight of the variability?*

Response:

The standard deviations of each PM_1 chemical species have now been calculated and are reported both in the text and in the corresponding figure caption. The values are as follows: organics (±13 %), sulfate (±9 %), ammonium (±4 %), nitrate (±1 %), and BC (±14 %). This addition highlights that organics and BC are the dominant and most variable components of PM_1. The pie chart continues to represent the median composition, which better reflects the central tendency of the campaign dataset while minimizing the influence of high-concentration events.

Reviewer comment:

Lines 188–191. "The most intense PM_1 peaks are associated with ship arrivals and departures ..." Could you quantify the difference between the median and these most intense peaks? How much is this increase in PM_1, PN, Org, BC, NO_x relative to sulfate? Relative and absolute differences would be good to list here.

Response:

We quantified the difference between (i) the median concentration computed across all identified plumes and (ii) the maximum value reached during the most intense plume for PM_1, PN, Org, BC, NO_x and SO_4^{2-}. For each species, we report both the absolute increase (max – median) and the relative increase (max / median).

Using the plume medians and the corresponding maximum values, we obtain:

- PM_1: increase of +21 µg.m^{-3}, factor 3.1, max 152 µg.m^{-3}
- PN: increase of +2.4 × 10^{4} cm^{-3}, factor 3.4, max 6.0 × 10^{4} cm^{-3}
- Org: increase of +15 µg.m^{-3}, factor 5.1, max 112 µg.m^{-3}
- BC: increase of +2.9 µg.m^{-3}, factor 3.6, max 14 µg.m^{-3}
- NO_x: increase of +123 µg.m^{-3} (≈ 65 ppb), factor 3.3, max 470 µg.m^{-3}
- SO_4^{2-}: increase of +0.27 µg.m^{-3}, factor 1.20, max 6.8 µg.m^{-3}

To compare these enhancements with sulfate, we normalized each relative increase to the sulfate factor. The resulting ratios are:

- PM_1: 2.6 × sulfate
- PN: 2.9 × sulfate
- Org: 4.2 × sulfate
- BC: 3.0 × sulfate
- NO_x: 2.8 × sulfate

Change in the manuscript: (lines 214 - 218)

"When comparing the plume median to the maximum concentration reached during the most intense events, PM_1 *increased by a factor of about 3.1 (≈ +21 μg.m* $^{-3}$*), PN by 3.4 (≈ +2.4 ×* 10^4 *cm* $^{-3}$*), Org by 5.1 (≈ +97 μg.m* $^{-3}$*), BC by 3.6 (≈ +11 μg m* $^{-3}$*), and* NO_x *by 3.3 (≈ +123 μg m* $^{-3}$*). In contrast, sulfate increased only by a factor of about 1.2 (≈ +0.27 μg m* $^{-3}$*), confirming its much weaker variability during peak plumes."*

Reviewer comment:

Line 237. "The NO_x *EF from ferries show significant variability, ranging from 5.2 g/kgfuel to 72.5 g/kgfuel." While it is great to provide the range of values, to best capture the variability, the standard deviation should be provided.*

Response:

We have now calculated and included the standard deviations of the emission factors in both the main text and Table 2. Reporting the mean, median, range, and standard deviation for each pollutant provides a more comprehensive representation of the variability observed among different vessels and operational modes.
This information has been explicitly mentioned in all EFs' paragraphs.

Reviewer comment:

Line 254. "though engine design (Tier I for D, Tier II for A, B, C, E)." Can you explain what the "though engine design" is and what each tier means? This is not explained in the main text nor supplement.

Response:

The phrase "though engine design" referred to differences in engine generation and it is related to IMO NO_x emission standards (Tiers I–III) defined under MARPOL Annex VI. These tiers establish maximum allowable NO_x emissions per unit of engine power as a function of rated speed and year of construction. Tier I, applicable to engines built after 2000, allows 17.0 g kWh^{-1} at 130 rpm, decreasing to 9.8 g kWh^{-1} at 2000 rpm. Tier II (after 2011) tightens these limits to 14.4 g kWh^{-1} at 130 rpm and 7.7 g kWh^{-1} at 2000 rpm. Tier III, applicable to engines installed within NO_x Emission Control Areas (NECAs)—a subset of Emission Control Areas (ECAs)—after 2016, further reduces these limits to 3.4 g kWh^{-1} at 130 rpm and 2.0 g kWh^{-1} at 2000 rpm.
In our dataset, ship D is equipped with a Tier I engine, while ships A, B, C, and E are fitted with Tier II engines.

Change in the manuscript: lines (51 - 53)

"In addition to the limits on fuel sulfur content, MARPOL Annex VI also establishes progressive NOx emission standards, known as Tier I, Tier II, and Tier III. These tiers define the maximum allowable NOx emissions per unit of engine power as a function of the engine's rated speed."

Reviewer comment:

Lines 275–276. "This CO EF, ranging from 0.03 g/kgfuel to 1.90 g/kgfuel (median of 0.28 g/kgfuel), show small variation across ship types or vessel size." Remove "This". "show" should be *"shows", but is the implication of small variation across ship types or vessel size based on the listed range? The upper limit of the range is over 6 times the median and the median is over 9 times the lower limit of the range. This tells me that it is highly variable (though, standard deviation should also always be included when discussing variability), so what would support the argument that there is small variation across ship types or vessel size?*

Response:
We thank the reviewer for catching the error and the inconsistency in interpretation. The statement indeed referred to BC rather than CO. We agree that the expression "small variation" was misleading.

The sentence has been corrected as follows (lines 307 - 311):

"The BC EFs exhibit a broad range of values (0.03–1.90 $g{\cdot}kg^{-1}$ fuel), primarily due to a few isolated plumes characterized by unusually high values. These outliers are likely linked to transient conditions such as low engine load or incomplete combustion, rather than systematic differences among ship types or vessel sizes. When excluding these extreme points, the interquartile range (0.09–0.42 $g{\cdot}kg^{-1}$ fuel) suggests that most of the plumes fall within a relatively consistent range across vessels under typical operating conditions."

Reviewer comment:
Lines 285–286. "The highest values are observed for ferries A, B and D characterized by powerful auxiliary engines (more than 6,000 kW cumulated for each ferry)." Is it possible to correlate auxiliary engine power with Org EF?

Response:
We calculated the correlations between organic aerosol (Org) emission factors and both auxiliary and main engine power for all maneuvers as well as for arrivals and departures separately (Table S3).

No significant correlation was found between Org EFs and either auxiliary or main engine power when across all maneuvers ($R^2 < 0.15$). During arrivals, a weak negative trend was observed with auxiliary power was observed (slope = $-1.86 \times 10^{-4} \pm 2.16 \times 10^{-4}$ g kg^{-1} fuel kW^{-1}, $R^2 = 0.20$), whereas during departures, no clear relation emerged ($R^2 \approx 0$). The correlation with main engine power remained low in all cases ($R^2 \leq 0.33$).

These results suggest that organic aerosol emissions are not governed by installed engine power, but rather by operational and ship-specific parameters such as fuel combustion efficiency.

All regression results are summarized in Table S4 and the corresponding plots are presented in Figure S6, which includes SO_2 and Org EFs as a function of auxiliary, main, and total engine power for arrivals, departures, and combined maneuvers.

Change in the manuscript: (lines 319 - 322)

"However, no significant correlation was found between Org EFs and either auxiliary, main, or total engine power ($R^2 < 0.20$), indicating that fuel composition, combustion efficiency, and transient operating conditions exert a stronger influence than engine size. Regression statistics

for arrivals, departures, and combined maneuvers are summarized in Table S4, and the corresponding relationships are illustrated in Figure S6."

Reviewer comment:

Lines 516–517. "The overall fraction of OA related to transport (road and maritime) is quite high accounting for almost one third of the total OA (29.2 %)." "Quite high" relative to what? Any reference?

Response:

The expression "quite high" referred to the fact that the fraction of OA attributed to transport sources (road + maritime) in our PMF analysis (29.2 % of total OA) is near the upper end of values reported for coastal and port environments.

Although few studies directly apportion OA from shipping, comparable receptor-model analyses report lower or similar values. Broader receptor-model studies including both inorganic and organic species generally report total shipping contributions between 5 % and 20 % of $PM_{2.5}$ or PM_{10} (Wu2019, Bove2016, Pandolfi2011, Minguillon2008). Therefore, the 29 % of OA attributed to transport in Toulon falls at the upper end of values reported literature for port or near port environments. These comparison values and references have now been added directly to the manuscript to clarify the context and substantiate the interpretation.

Addition to the manuscript: (lines 564 - 567)

"The overall fraction of OA related to transport (road and maritime) account for 29.2 % of total OA. Broader receptor-model studies including both organic and inorganic species generally estimate 5–20 % of PM2.5 or PM10 from maritime sources (Wu et al., 2019; Bove et al., 2016; Pandolfi et al., 2011; Minguillón et al., 2008)."

Reviewer comment:

Section 2. Can you define the criteria for a plume? How long are each of the plumes? How well does linear interpolation of EFs represent the background? Can you provide any metrics that capture the uncertainty? Perhaps a figure of a plume event which captures the duration of the plume and the variability of EFs would suffice.

Response:

In this study, a plume is defined as a transient enhancement in particle number concentration (CPC) exceeding at least twice the local background, with a mean wind direction between 130° and 290° (from southeast to northwest) and a wind direction standard deviation below 30°. The start and end times are determined from concurrent increases in CPC and CO_2, adjusted to the instrumental time steps. Typical plume durations range from a few minutes up to about 20 minutes, depending on ship distance and meteorological conditions.

The background concentration is determined by linear interpolation between pre- and post-plume periods, which provides results consistent with other background estimation methods (rolling or median; Volent et al., 2025). Each emission factor (EF) is automatically calculated and then manually validated to ensure that only genuine ship plumes are retained, with clearly defined start and end times for each event. The plume boundaries are individually adjusted for each pollutant to account for slight desynchronization between instruments, ensuring accurate integration of excess concentrations.

This combined automatic and manual validation ensures that each EF corresponds to a well-defined emission event. A representative plume example showing the background interpolation, plume boundaries, and variability in measured pollutants has been added in Figure S5 in the Supplementary Information.

Change in the manuscript: (lines 156 - 167)

"In this study, a plume was defined as a transient enhancement in particle number concentration (CPC) exceeding at least twice the local background, observed under a mean wind direction between 130° and 290° (from southeast to northwest) and a wind direction standard deviation below 30°. The start and end times were determined based on concurrent increases in PN and CO_2 concentration, adjusted to the instrumental time resolution. Plume durations ranged from a few minutes to approximately 20 minutes, depending on the ship's distance from site and the prevailing meteorological conditions.

Each emission factor (EF) was automatically calculated using an emission factor calculation tool (Le Berre et al., 2024), based on the carbon mass balance method, and subsequently manually validated to ensure that only genuine ship plumes were retained, with clearly defined start and end times for each event. The plume boundaries were individually adjusted for each pollutant to account for slight desynchronization between instruments, thereby ensuring accurate integration of excess concentrations. This combination of automatic and manual validation ensures that each EF corresponds to a well-defined transient emission event. An illustrative example of a plume, including the background interpolation and pollutant variability, is provided in Figure S5."

Reviewer comment:
Section 3.2. A literature table in the supplement with all the references cited in the main text for each EF would be helpful.

Response:
A new table (Table S5) has been added to the Supplementary Information, compiling all literature emission factor (EF) values cited in the main text. The table lists, for each pollutant, the reported EF ranges, measurement type (on-board, in-plume, or test-bench), fuel type, engine tier, and corresponding references. This addition provides a comprehensive overview of the datasets used for comparison and later discussed in Section 3.2.

Reviewer comment:

Section 3.2.1. If more powerful auxiliary engines lead to more SO_2 emissions, can you provide a correlation of engine power to SO_2 EF? How much of an increase in engine power (kW) results in an increase in EF of SO_2?

Response:
We recalculated the correlations between SO_2 emission factors (EFs) and both auxiliary and main engine power, considering all maneuvers as well as arrivals and departures separately (Table S3).

Over all maneuvers, a moderate correlation was found between SO_2 EF and auxiliary engine power (slope = $1.48 \times 10^{-4} \pm 7.14 \times 10^{-5}$ g kg^{-1} fuel kW^{-1}, $R^2 = 0.35$), while the correlation with total installed engine power was negligible ($R^2 = 0.00$).
When separating the operational modes, correlation with auxiliary engine power strengthened during departures (slope = $2.25 \times 10^{-4} \pm 1.05 \times 10^{-4}$, $R^2 = 0.60$) and weakened during arrivals (slope = $7.03 \times 10^{-5} \pm 9.77 \times 10^{-5}$, $R^2 = 0.15$).
Correlations with main engine power were generally weaker ($R^2 < 0.4$).

On average, an increase of 1,000 kW in auxiliary power corresponds to an increase of approximately 0.15 g kg^{-1} fuel in SO_2 EF during departures.
These results confirm that auxiliary engines are the main contributors to sulfur emissions during maneuvering phases, consistent with their dominant operation while vessels are within the port area.
All regression results are summarized in Table S4, and the corresponding plots are presented in Figure S6, which includes SO_2 and Org EFs as a function of auxiliary, main, and total engine power for arrivals, departures, and combined maneuvers.

Change in the manuscript: (lines 252 - 258)

"The median SO_2 EFs exhibited a moderate correlation with auxiliary engine power across maneuvering ($R^2 = 0.35$). This relationship was stronger during departures ($R^2 = 0.60$) and weaker during arrivals ($R^2 = 0.15$). During departure, an increase of 1,000 kW in auxiliary engine power was associated with an average rise of approximately 0.15 g kg_fuel^{-1} in SO_2 EF. Correlations with main and total engine power were somewhat weaker ($R^2 < 0.4$), suggesting that auxiliary engines are the primary contributors to SO_2 emissions during maneuvering phases. Regression statistics for arrivals, departures, and combined maneuvers are summarized in Table S4, and the corresponding relationships are illustrated in Figure S6."

Reviewer comment:
Section 3.4.2. Could you provide an explanation for the diurnal cycle of the cooking organic aerosol lacking meal time peaks in mass? Could this factor rather just have a continental origin rather than cooking?

Response:
The absence of a lunchtime peak in the COA diurnal cycle is explained by the wind configuration during midday hours: between 11:00 and 14:00, winds predominantly came from the south to southwest (port and sea sectors), whereas the main restaurant district is located to the north–northwest of the monitoring site. As a result, lunchtime cooking emissions were rarely transported to the receptor.

In contrast, a modest increase is observed in the evening, which is consistent with more frequent northerly and northwesterly winds during 18:00–21:00 that occasionally place the site downwind of the restaurant area. This is confirmed by wind-sector analyses restricted to meal-time windows (Figure S15), showing that only 0.78 % of winds during lunch originated from the restaurant sector, compared to 2.05 % during dinner.

Regarding the factor's identity, its mass-spectral features (enhanced m/z 55 relative to 57, low f44) can support a primary cooking signature and differ from both hydrocarbon-like traffic emissions and secondary OOA. The factor is also temporally decoupled from the afternoon OOA maximum, which argues against a continental or regional origin.

Change in the manuscript: (lines 515 - 522)

"The muted lunchtime COA signal results from the site being systematically upwind of the restaurant district, located to the north–northwest of the station. During 11:00–14:00, winds predominantly originated from the south to southwest (port and sea sectors), preventing the transport of cooking emissions to the site. A slight evening enhancement is nonetheless observed, consistent with the higher occurrence of northerly and northwesterly winds during 18:00–21:00, which intermittently place the site downwind of the restaurant area. Wind roses restricted to meal-time periods (Figure S15) confirm this pattern, with only 0.78% of lunchtime winds and 2.05% of evening winds originating from the restaurant sector. These conditions explain both the absence of a midday peak and the weak but detectable evening COA contribution."

Reviewer comment:

Section 3.5. How does the PNSD when selecting the 10 most intense peaks of each factor compare to the PNSD when each factor makes up a certain threshold of the mass? Can you validate this method in finding a representative PNSD in other ways?

Response:

Using a fixed mass-fraction threshold (e.g., requiring a factor to exceed 30–50 % of OA) is not suitable in our case because several PMF factors never reach such high contributions. Applying a uniform threshold would either exclude most occurrences or select time windows influenced by multiple overlapping sources.

We therefore used the ten most intense events (top-10 peaks) for each factor, corresponding to periods when the factor contribution is locally maximal and least affected by other sources. This approach ensures that the selected particle number size distributions (PNSDs) reflect the dominant conditions of each source while minimizing cross-influence.

This choice is also consistent with the SMPS time resolution (≈ 2 min per scan), which cannot fully capture the very short plumes (1–3 min) typically resolved by the AMS. Focusing on the strongest, temporally separated events (≥ 10 min apart) guarantees that the averaged PNSDs are representative of each factor's typical behavior.
With faster instruments such as the EEPS (1 s resolution), it would be possible to integrate number size distributions directly into a joint PMF framework, providing more robust PNSDs directly linked to the factors. We mention this as a future development for improving temporal and size-resolved source apportionment.

Change in the manuscript: (lines 616 - 623)

"Since several PMF factors contributed only to 20–30% of the total OA, applying a fixed mass-fraction threshold would have introduced bias in the PNSD selection, favoring overlapping or mixed-source events. To avoid this, the ten most intense events (top-10 peaks) were selected for each factor. These events represent locally dominant periods that are minimally influenced byother sources and separated by at least ten minutes. This selection criterion is consistent with the temporal resolution of the SMPS (2 min per scan), which constrains its ability to resolve short-lived plumes often captured by the AMS. The use of higher time-resolution instruments, such as the EEPS (1 second resolution), would allow direct incorporation of particle number size distributions into the PMF framework. This is a direction we intend to pursue in future work to enhance the characterization of source-specific temporal and size-dependent variability."

Reviewer comment:

Figure 1. The caption lists chloride as a species that is plotted, but according to the legend, it is missing from the figure. Please either include it in the figure or remove from caption and provide reasoning for its absence.

Response:

Chloride represents a negligible fraction of PM_1 (0.2 ± 1.0 %) throughout the campaign, well below the visual resolution of the pie chart. For this reason, it was omitted from the plotted species to improve figure readability. We have now removed the reference to chloride from the figure caption and specified in the text that its contribution is minor (<1 %).

Reviewer comment:

Figures 2 and 4. These figures are not referenced anywhere within the main text, which suggests that either (1) they are not significant enough for the main text and should be in the supplement or (2) more text is needed to incorporate these figures. Since Figure 3 is referenced in Section 3.2.5, I suggest that Figures 2 and 4 should be referenced in the other sections in Section 3.

Response:

References to Figures 2 and 4 have now been added in the corresponding subsections of Section 3, where gaseous (Figure 2) and particulate (Figure 4) emission factors are discussed. These figures are central to the interpretation of the results and thus remain in the main text.

Reviewer comment:

Table S2, S6, S7, S11, S12, S13. These tables are not referenced anywhere within the main text.

Response:

References to Tables S2, S6, S7, S11, S12, and S13 have now been added in the relevant sections of the main text. Each table is now explicitly cited where its data are discussed (e.g., S2 in Section 3.2.1 for ship and engine characteristics, S6–S7 in Section 3.4 for PMF factor comparisons, and S11–S13 in Section 3.5 for PNSD and OA factor analyses).

Reviewer comment:

Has work been done to see if the increase in shipping emissions is affected by the number of ships? This is especially relevant as your results section discusses the need to address fleet evolution.

Response:

In this study, we report emission factors (EFs), which represent pollutant mass or number emitted per unit of fuel consumed by an individual vessel. EFs are therefore ship-specific, normalised quantities, and they do not depend on the number of ships operating in the port. By contrast, shipping emissions correspond to the total amount emitted over a given period, which indeed scales with the number of ship calls, vessel size, operating time, and traffic composition.

Because our analysis is based on a single summer campaign, the study does not investigate the relationship between total shipping emissions and ship numbers. According to the Toulon Harbour Authority, ferry activity has increased slightly since 2019 (≈ +9 %), but this trend mainly reflects the introduction of larger vessels rather than an increase in ship calls.

The variability observed in our dataset therefore reflects differences in vessel characteristics—including fuel type, engine technology, and the presence or absence of EGCS—rather than changes in traffic frequency. Assessing how total port-area emissions evolve with ship numbers would require a multi-year analysis combining EFs, detailed AIS traffic data, and vessel-specific activity profiles, which is beyond the scope of the present work but represents a valuable direction for future studies.

Reviewer comment:
Are there other meteorological parameters such as relative humidity and temperature that are available? Is there any consideration as to how these might affect measurements?

Response:

Meteorological parameters including wind speed, wind direction, and temperature were continuously monitored at the site using a Tridi USA-1 (Metek) ultrasonic anemometer equipped with temperature sensors. This information has been added to Table 1 in the manuscript.

Within the observed range, no significant correlation was found between emission factors and temperature, as the short duration of individual plumes (typically < 20 min) limits the impact of meteorological variability on the derived EFs. The potential influence of temperature on the gas–particle partitioning of semi-volatile organic compounds such as PAHs has been further discussed in Section 3.3, following Reviewer #2's related comment.

Reviewer comment:
Technical Comments.

Response:
We thank the reviewer for the detailed editorial and technical suggestions. All recommended corrections have been implemented throughout the manuscript to improve clarity, consistency, and grammatical accuracy. Specifically:

- Acronyms and particle size definitions (UFP, PM_1, $PM_{0.1}$) have been explicitly defined at first mention.

- Minor grammatical and typographical corrections have been applied (e.g., verb agreement, punctuation, spacing, and article usage).
- All figure and table captions have been checked for consistency with the corresponding legends and symbols.
- Redundant or repetitive sentences (e.g., the background estimation explanation in Lines 123–124) have been removed to streamline the text.
- Unit notations (e.g., “nm” instead of “nanometers”) and reference formatting (e.g., placement of parentheses in citations) have been standardized.
- Sentences have been restructured where suggested to improve readability and adherence to scientific writing norms.
- All specific line edits mentioned by the reviewer (Lines 27–588) have been implemented as indicated, including revised phrasing, citation style, and typographical corrections.

We appreciate these thorough and constructive comments, which have helped ensure stylistic uniformity and improve the overall clarity and precision of the manuscript.