Response to Reviewer #1

We thank the reviewer for his/her comments, which have helped us to improve the manuscript. Our responses to the specific queries follow below. Our responses are given in blue text.

1. General comments

This manuscript is not easy to understand (read), and is little bit long. Authors should clearly state the strengths of this study in the Introduction and state the important findings of this study at elsewhere. They should rewrite the whole manuscript to become more clear explanations.

We now state more explicitly some of the important findings of our manuscript in the abstract, and have aimed to clarify the manuscript throughout.

There are too many abbreviations such as SSD, LSF, and rBC. Please limit the abbreviations to the general terms or necessary terms. May be about half or less of them are acceptable.

We have used abbreviations that are common in the field associated with ships and ship emissions and in the black carbon literature. SSD, LSF and rBC are all common abbreviations, and we thus retain them here.

Abstract: It is not easy to understand. The description in the Conclusion is better to be understood. So please reconstruct the Abstract so as to clearly state the main objective and key findings.

It is not clear to us what specifically the reviewer finds unclear. We have nonetheless attempted to clarify and have added two sentences, in particular, to make clearer what has been done. Specifically:

"These measurements have been used to assess the level of agreement between these different techniques for the determination of BC emission factors from ship plumes. Also, these measurements greatly expand upon the number of individual ships for which BC emission factors have been determined during real-world operation."

This manuscript shows the emission factor (EF) only in g/kg-fuel. However, the EF in g/km (per travel distance) may be also (or more) useful for the readers. Also EF in g/km could be more variable with engine type and fuel type and engine load. So I wonder how the results would become when these data is analyzed for the EF in g/km. Please state regarding comments in the manuscript. Of course it is good to show some results of EF in g/km. Authors should clearly describe that EF can be estimated only in g/kg-fuel due to the measurement and estimation protocol at the Introduction or Experimental methods.

We report our measured EF's only in g/kg-fuel because this is the parameter that is most directly accessible to us via measurement. To convert the EFs to g/km may be useful, but would add considerable uncertainty as we do not know the fuel economy of the vessels in question, which is likely

a function of vessel speed and operating conditions, in any case. To avoid having to make unnecessary assumptions in the reporting of our measured values, we continue to use only g/kg-fuel.

Experimental Method: Please clearly state the following two points at the beginning of this section: (1) Estimation concept of the EF. How the EF is estimated form ambient measurement? (2) Feature and limitation of the estimation method. Regarding (2), for example, please describe the followings: (a) How the authors can distinguish between the plume from own vessel (R/V Atlantis) and the plume from other vessel? (b) The definition of "plume" (e.g. darkness, concentration, distance from the target vessel). (c) How they can catch or detect plume? (d) What times or what seconds the measurements were conducted in a single plume? How the authors selected or extracted of the effective data from lots of the measured data?

The "plume intercept" method has been widely used in previous studies. The EF is estimated from the increases in the concentration above background levels for the pollutant of interest relative to the corresponding increase in the concentration of CO2 above background. Plumes are identified based on the relative location of the Atlantis to the target vessel, the relative wind direction, and the observation of an increase in the pollutant concentration and CO2 above background levels. This was already stated: "Plumes were identified based on the relative ship positions and winds, and by a noticeable increase in [CO2], babs or [BC] above background." Plumes from other ships are easily distinguishable from self-plumes (i.e. those from the Atlantis) based on knowledge of the relative wind direction and the fact that we sample from in front of the main stack of the Atlantis. We have provided additional references to other plume intercept studies to make the overall methodology clearer.

We discuss the EF determination method at the appropriate location in the experimental methods, namely in the section titled "Emission Factor Determination." We do not agree that it is necessary to state this at the beginning of the overall experimental section.

The time-scale associate with each plume is variable, ranging from about 1 minute to many minutes. This is now stated.

3.1 BC measurement technique comparison, P24687, L14-19: Were the EFs obtained from the four different measurements, though the SP2 values are systematically lower than the other three instruments. It seems reasonable that only three data is used to obtain average EF.

The reviewer asks an important question. The reported averages exclude the SP2. We have now made this clear by stating "In all cases, the single ship averages were determined from only the PAS, PSAP and SP-AMS measurements because of the likely low bias in the SP2 measurements, discussed above."

Table S1: Are the data, which the ship speed is zero, for the ships at anchor at harbor? If so, these data may have to be removed from the EF estimation because the engine condition is different from that in running condition and the plume can be strongly affected by the other sources (ship and onshore sources).

The EFs determined for most ships with zero speed are for those anchored at sea, not at harbor. Regardless, as we are confident that all plumes are from a single vessel (based on wind direction and wind speed data) separate from "other" sources, we do not agree that it is necessary to exclude such ships. We have added a clarifying sentence to the section "Influence of Engine Load": "Ships that have zero speed were typically anchored offshore, not at harbour, which is important since at many California ports ships at harbour operate on shore power, which would lead to a reduction in emissions."

Figs.3&4: Please clearly state the reason of the large variation of EF. Is it mainly from measurement and estimation error or from vessel itself? If the cause is vessel itself, please describe the main cause of the variation.

The variability is because we are considering emissions from many individual vessels that are operating under a variety of conditions. This is now stated explicitly in the main text: "The observed variability is a result of sampling over many individual vessels."

Abstract: Please write "R/V Atlantis" as "research vessel (R/V) Atlantis" because some readers may not understand it is a vessel.

Done.

P24676, L14, Abstract: EF of BC is the main result of this study. So it is better to show more detailed data such as the range (minimum to maximum) or regarding values.

We have modified the abstract to read "The geometric average EFBC (excluding outliers) determined from over 71 vessels and 135 plumes encountered, was 0.31 ± 0.31 g-BC (kg-fuel)-1, where the standard deviation represents the variability between individual vessels."

P24676, L6, Abstract: Please show the concentration of sulphur in fuel where the "Relatively Low sulphur" is stated.

We now state "...utilizing relatively low sulphur fuels (average fuel sulphur content of 0.4%, 0.09% and 0.03% for vessels operating slow speed, medium speed and high speed diesel engines, respectively)." And also "...operating on high sulphur fuels (average fuel sulphur content of 1.6%)."

P24676, L7-11, Abstract: Two phrases start with "which" are confusing. So please omit those phrases or divide that sentence in two. Or the description in the Conclusion is better to be understood.

"Which" here is used in a grammatically correct manner, and so we have decided to retain the statement.

P24676, L19-21, Abstract: it is not easy to understand what difference is stated by the words "significant difference". This sentence may be able to be omitted.

"Significant difference" means "statistically significant difference", as stated (and discussed explicitly in the manuscript). Also, as this is a key finding of this work, we feel that it must be retained in the abstract.

Fig.1 and Table1: Please show in the captions which instrument is shown in y-axis.

We have updated the captions to indicate that the first listed instrument is always on the ordinate (y-axis). However, as we are using an ODR fit there is much less dependence in the derived fit parameters than if we had used a standard linear least squares method.

Fig. 3: Please explain the meanings of the value in parenthesis in the caption.

We have added the sentence "The number of ship encounters is listed in parentheses for each engine or vessel type."