Reviewer 1:

First of all, we would like to acknowledge the comments provided by Reviewer#1, which have help us improve our manuscript. Clarifications of the issues below have been included in the MS.

Line 18.- Is there is an official recommendation? – Please provide a reference

Thanks for the advice. The official recommendation is EN 14662 Standard. Ambient air - Standard method for the measurement of benzene concentrations - Part 3: Automated pumped sampling with in situ gas chromatography, 2015. However, we have decided to eliminate the complete sentence from the Abstract as it is not relevant.

Line 33.- Please be more pragmatic in the resolution of this problem. Apart from requesting the manufacture intervention, it would be useful to present a list of measurements to carry out by the user in order to minimize or avoid this problem.

We appreciate the suggestion from the reviewer. This matter has been clarified in the MS: we have added information in the Result and Discussion section about measurements to carry out by the user in order to minimize this problem. The final wording is "[…] *This approach would require continuous measurements of TCM in air and a knowledge of how TCM deviates measurements from its real value, which in turn, requires carrying out tests similar to those presented in this paper with dynamic dilution systems in controlled test atmospheres. This measure could not be easy to apply for economic and technical reasons so the whole responsibility must not only fall on the network managers. It seems reasonable that the manufacturers of the equipment tackle actions for solving this problem –or, at least, for reducing the extent of the interference in their measurements-, since they have the required technology and equipment. In any case, users of this type of equipment should be aware of the problem to try to minimise it. The discussion of this issue in the appropriate forum (e.g. the European Committee for Standardisation) seems also pivotal to reduce the uncertainty in benzene measurements by GC-PID in presence of TCM concentrations*".

Line 49/50.- change degrees to ºC:

It has been done. Thank you for your suggestion.

Line 59.- Remove double endpoint

Double endpoint has been deleted. Thank you for your recommendation.

Line 70.- Why is it compared with hydrocarbons of similar molecular weight? This is not an indicator of stability in the atmosphere

Thank you for this interesting comment. We agree with the reviewer and accordingly, the mention to hydrocarbons of similar molecular weight has been removed. It was a comment recommended by a previous journal reviewer.

Line 76 .- Please make reference to the corresponding legislation

We have added the reference: Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe, Off. J. Eur. Communities, 152, 1–43.

http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:152:0001:0044:EN:PDF

Line 127.- Please provide appropriated reference.

Thank you for your suggestion. The following reference has been added: Senum, G. I.: Quenching or enhancement of the response of the photoionization detector, J. Chromatogr. A, 205(2), 413–418, doi:10.1016/S0021-9673(00)82668-0, 1981.

Line 223.- Who was the certifying body?

We appreciate the suggestion from the reviewer; we have included in the MS the certifying bodies, which were the respective manufacturers of the gas mixtures. The mixtures were certified according to Standard ISO 6141:2000.

What are the uncertainties of the final generated concentrations?

This is an interesting point. Flow rates were continuously controlled and the expanded uncertainty of the generated concentrations was calculated from the standard uncertainty of the gas mixtures in the gas cylinders and the standard uncertainties of the flow rates. We have added a new paragraph in the Experimental set-up section: "*The expanded uncertainties of all generated concentrations of pollutants were estimated from the standard uncertainties of the high concentration gas cylinders and the standard uncertainties of the gas flow rates. In all cases, the final expanded*

uncertainty was less than 5%, according to the limit established in Standard EN 14662-3".

Line 356.- This seems a relevant item to be reported in the conclusions to be considered in the EN standard.

We agree with the reviewer. We have modified the Conclusion section accordingly: "The research described in this article has determined that TCM causes a significant interference in the measurement of benzene by GC-PID. This interference is negative, that is, readings of benzene are below their real ambient values, which may originates a mismanagement of the air quality of a location with presence of TCM in its air in relation to benzene".

Line 474.- Although these biases seem very high, it would be of interest to demonstrate that they are significant compared to the measurement uncertainties by considering the whole experimental setup.

As stated before, the uncertainty of the generated gas mixtures was below 5%. In addition, a lack-of-fit test was performed in order to test the accuracy of the readings. For this, after calibration of the analysers, several gas mixtures of benzene in air with different concentrations ranging from 0 to 50 μg/m^3 were measured. Relative differences between the readings and the reference concentrations were calculated and, in all cases, were below 10%. This value is much lower than the biases that occur in the readings when there is TCM in air ambient (34, 44 and 70 % when there is a TCM concentration of 0.7, 1.4 and 4.5 μg/m^3). Moreover, other potential influencing parameters such as temperature or pressure were kept constant during the experiments, allowing to conclude that the high biases obtained in readings when TCM was added to the mixture are due to the presence of this substance. The following paragraph has been included in the MS: "*In order to ensure that the biases obtained in these and subsequent experiments were only due to the interfering compounds tested, sample and surrounding temperatures, sample pressure and voltage were kept constant during all experiments. A lack-of-fit test was performed in order to test the accuracy of the readings. For this, after calibration of the analysers, several gas mixtures of benzene in air with different concentrations ranging from 0 to 50 μg/m^3 were measured. Relative differences between the readings and the reference concentrations were calculated and, in all cases, were below 10%, much lower than the values reported in the Result and Discussion section*".

Line 486 .- Please consider my comments on L33:

We have added a new paragraph in the “Results and discussion section” (please see comment on line 33): Also, in “Conclusions” we have added a similar conclusion. The final wording is the following: “*Interestingly enough, it is established in part 3 of the standard EN 14662:2015 that the managers of the air quality monitoring network are responsible for determining the presence of TCM in the area where benzene is measured. If detected, they must act to eliminate the effect of the interferent. However, this approach would require continuous measurements of TCM in air and a knowledge of how TCM deviates measurements from its real value, which in turn, requires carrying out tests similar to those presented in this paper with dynamic dilution systems in controlled test atmospheres. This may entail economic and technical issues so manufacturers of the chromatographs should try to solve this problem as they have greater technical and scientific capacity than network managers. In any case, all these issues should be discussed in the appropriate forum (e.g. the European Committee for Standardisation) in order to improve the uncertainty of benzene measurements and, thus, the management of the air quality”*.

Line 614.- Why are Ucorg and Vtest (%) reported only in this Table?

Parameters U_{Corg} and V_{test} (%) were only reported in Table 3 because this table contained the data obtained according to concrete indications of the EN 14662 Standard, which requires the calculation of V_{test} and U_{Corg} in order to compare and verify its acceptability with the performance criterion (<5%) established in such standard.

We consider that using the relative error is more logical and useful to evaluate the deviations than the V_{test} and U_{Corg} parameters. For this reason, in the rest of the Tables, which contained results of tests proposed by ourselves and not contemplated in the Standard, such parameters were not included (U_{Corg} and V_{test}). However, we have decided to merge former Tables 3 and 4 into a single new table (Table 2) in order to save space, and we have decided to calculate parameters V_{test} and U_{Corg} for the results presented in former Table 4 to maintain the symmetry of new Table 2.

Line 649.- Why are the results of analyzer II not reported?

The Analyzer II belongs to the official surveillance network from the Government of Región de Murcia and it is operating continuously in a monitoring station. Therefore, we only had such equipment for a limited time in our lab. Given that both analysers exhibited a similar performance in the first set of experiments carried out, we

considered that the results obtained with Analyser I would be representative of both of them. In addition, a reproducibility test was carried out in the lab. Both analysers were subject to measure simultaneously a gas mixture containing 5 µg/m^3 nominal benzene in zero air. The reproducibility (in µg/m^3) was calculated as:

$$Reproducibility = \sqrt{\frac{\sum d_i^2}{2n}}$$

where d_i is the i difference in readings between Analyser I and II and n is the total number of measurements (6 in our case). The value obtained was 0.067 µg/m^3 when the average concentration of benzene in the reference gas mixture gas 4.6 µg/m^3, which means 1.5% bias. This value was considered low enough to perform the subsequent tests just with one analyser. All these results have been added to the MS.

Line 470 and 649.- What is the reproducibility of the Eq (15) between different analyzers?

Eq. (15) was only obtained for Analyser I, as only this analyser was subject to the tests in Section 2.2.2. Reproducibility has been calculated as detailed in the previous comment.