Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-247-AC2, 2018

Atmospheric Measurement Techniques Discussions

Open Access

Interactive comment on "Identification of spikes associated with local sources in continuous time series of atmospheric CO, CO_2 and CH_4" *by* Abdelhadi El Yazidi et al.

Abdelhadi El Yazidi et al.

abdelhadi.el-yazidi@lsce.ipsl.fr

Received and published: 6 January 2018

We would like to thank referee#1 for the valuable comments and her/his time to review this paper. Author's comments are presented below.

1- Referee's comment: Language and structure The article is well written, though too lengthy and it contains quite some minor but still sloppy errors that should be corrected. For this I included a list of minor corrections at the end of this review. An additional check of the text by a native speaker would be beneficial to the paper. 1- Author's response: We thank the referee for pointing out the list of minor errors. We have corrected all the suggested errors.

2- Referee's comment: Here and there the text is too long. It is a useful exercise, but not rocket science, so could be dealt with also by a shorter text. I propose to shorten section 3.3 and 3.4 with 30-50%. 2- Author's response: As suggested by the referee we have shortened the sections 3.3 and 3.4.

3- Referee's comment: An important issue is that the paper only handles two methods as the COV method is discarded right away. The text should be revised to better reflect this. 3- Author's response: We agree with the referee that the COV method was not developed as much as the other two (SD and REBS) in the paper. The COV method showed its robustness to filter out spikes, yet it requires setting the spike percentage a priori. In the abstract (line 32) we explained that the COV method will be excluded from the case study analysis"Analysis of the results of each method leads us to exclude the COV method due to the requirement to arbitrarily specify an a-priori percentage of rejected data in the time series, which may over- or under-estimate the actual number of spikes.". In line 189 we said that this method is not automatic, and in figure S2 in the supplement, we presented that we could select the same spike as the SD and the REBS method if we set the right a-priori percentage used as an input for the COV method. In line 302 "Because of this limitation for automatic spike detection we have discarded the COV method from further tests for the selection of the most reliable method for spike detection." we explained why we discarded the method for the Pic du Midi and Finokalia case study comparisons. 3-Author's changes in manuscript: We have added this sentence to the manuscript. Line 97: The study will focus more on the SD and the REBS method since they are fully-automatic and they do not require any a-priori information for the implementation.

4- Referee's comment: I would suggest to move the first paragraph of the conclusions in section 4 to replace parts of the introduction and summary, as this is the best introduction text to the paper. 4- Author's response: We agree that the first paragraph of conclusion could be used in the introduction. 4-Author's changes in manuscript: We have moved the first paragraph of the conclusion in introduction line 75. We have added few lines in the conclusion: line 429 “The recent increase in the number of studies that have been applied to study the spatial representativeness of GHG observations demonstrate the need to define efficient and reliable methods for the identification spikes related to local contamination sources. Three methods based on the standard deviation calculation were compared in order to provide an objective algorithm for the GHG data spike detection.”

5- Referee’s comment: General comments The topic is very relevant for improving the quality of ambient greenhouse gas observations by a regional network like the ICOS atmosphere network in Europe by an automated procedure, additionally to human manual quality control. The methodology used is sound but not spectacular. The two spike detection methods tested are very basic and relatively straightforward techniques that have proven their usefulness in air quality applications. It would have been useful to also look into more sophisticated methods that apply Fourier transform Savitkzy-Golay (1964) filters or wavelet transforms (e.g. Wee et al, 2008) to achieve this end. I would like to see some good arguments whether and why this has not been considered. 5- Author’s response: We thank referee#1 for reflecting this issue. We looked into different methodology before selecting these methods. Considering the wavelet transforms (Wee et al, 2008) and the Fourier transform methods, both methods showed their robustness for filtering out spikes in earlier studies. However, the two methods require the same conditions: the signal should be continuous, and smooth. In our case none of these conditions are full filled. Considering that the measurements are regularly interrupted due to different reasons (e.g. calibration, flushing time after switching from sampling level to another, power or internet outage), we had to select a method that handles time-series with data gaps. Moreover, if we apply a Fourier Transform method on continuous measurements, we will get a signal composed with different frequencies only. All information that varies with the time will be lost. In other words, we can analyze what happen (spikes to be filtered out), without knowing when this happens. In fact, we consider that having also the time information related to the spike would allow us to better understand the origin of the contamination (as presented in the manuscript for the Pic du Midi and Finokalia cases).

6- Referee's comment: I agree with referee #1 that it would be good to refer to the percentages of hours detected than the absolute number. 6- Author's response: We have changed in table 6 the percentage of the impacted hours to the percentage of the whole time series. We agree that it is interesting to show the percentage of hours instead of the number. Figure S7 is now completed by results for all sites. 6-Author's changes in manuscript: Table 6 is updated. Figure S6 is updated to figure S7.

7- Referee's comment: It would be good to state in the text more clear that avoiding spikes is more important than filtering them out and detection of spikes should always be followed by looking to the cause of the spikes in order to try to minimize them further. 7- Author's response: We agree that it is a major issue, and we have added in the conclusion (lines from 456 to 461) a sentence to emphasis this point. This problem was also reflected in the Pic du Midi case, where the use of the spike detection method confirms the need to move the inlet by 200m from its original location, in order to avoid the frequent contaminations. The aim of this study is to evaluate the methods in order to select the one that will be implemented in the ICOS Atmospheric Thematic Center Quality Control (ATC-QC) software to perform daily spike detection. In a future study, the chosen method will be coupled with meteorological data, such as wind direction, in order to try to assess the origin of the contaminations. 7-Author's changes in manuscript: (lines from 456 to 461): However, even if the implementation of an automatic algorithm can successfully identify short-term spikes due to local contaminations, it is important to note that the priority in the selection of a background site should be to avoid as much as possible the occurrence of such spikes. In the case where the spikes can not be totally avoided, it is then important to try to understand their cause and look for possible actions to minimize them. The modification of the air intake at the Pic du Midi, described in this study, is a very good example of what can be done once the origin of spikes is understood.

8- Referee's comment: It is good to see from this paper that the contribution of the spikes, in general, is low on the average signal observed, except for the PDM site with the obvious problem of the nearby pollution source. The 4 sites chosen for the paper are said to be representative for the ICOS atmosphere station network, but neither of them is a continental tall tower within 100 km or an urban region. It would also be interesting to see how the spike detection results vary for the vertical gradient along a tall tower where the footprint of the measurements varies from local for low sampling heights to more regional for the higher elevations. 8- Author's response: The OPE station presented in this study is a 120m tall tower site located in a rural area, which is representative of most of the future ICOS tall towers. The measurements are carried out at three sampling heights 10, 50, and 120m. Every hour the ambient air is sampled for 20 min alternatively at the three levels. We have run the SD method separately on the three levels. The method detects for CH_4 a percentage of 1.6%, 1.9% and 1.8% of contaminated data, and 1%, 1% and 1.1% for CO_2 respectively for 10m, 50m, and 120m. The comparison between the three sampling height did not show a significant gradient along the tall tower. Moreover, the selected spikes did not occur during the same hours for the three levels. This is mainly related to the difference in the sampling time along the tall tower.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-247, 2017.