

Interactive comment on “An Automated Baseline Correction Protocol for Infrared Spectra of Atmospheric Aerosols Collected on Polytetrafluoroethylene (Teflon) Filters” by A. Kuzmiakova et al.

Anonymous Referee #1

Received and published: 24 February 2016

Kuzmiakova et al present an automated smoothing splines (SSB) method to perform baseline corrections for PTFE (Teflon) filters analyzed by FT-IR. They apply this method to 794 filter samples collected by the US IMPROVE network. Results using the new SSB method are compared to analysis of the same filters using the manual polynomial baseline correction (PB). The two baseline correction methods yield similar results, and the SSB method has the advantage that it removes user intervention and potential biases associated with human intervention.

The paper is well written and topically relevant to AMT. It will be of most interest to

C1

other researchers who analyze PTFE filters using FT-IR. The method presented in this manuscript may potentially be of interest to all users of FT-IR or other spectroscopic methods. I suggest publication after addressing the comments below.

(1) In Section 2 (Methods), the authors go into great detail to explain the smoothing spline baseline correction. As part of their explanation, they introduce a significant amount of nomenclature. Many of the quantities they introduce seem to have multiple variants (e.g., EDF*, EDF_T, EDF_A, etc). The end result is an extremely detailed description of the method that is nearly impenetrable, as it requires readers to mentally juggle all of the different (and in my case, often unfamiliar) variable names and meanings. I strongly suggest that the authors include a glossary to define all of the various nomenclature, and that they take specific care to differentiate the various flavors of certain quantities (e.g., EDF as noted above).

(2) As a specific example of the previous comment, I am confused by the difference between EDF* and EDF_T. From Section 2.3.2 it seems that EDF* is determined from minimizing NAF (page 9, line 27). However the last line in that section (last line on page 9) introduces EDF_T, and I cannot find an explanation for EDF_T prior to that sentence. Thus, the use of EDF_T in section 3.1 is confusing. On Page 11, Line 21, EDF_A is used, again without prior introduction. While I understand in general terms the goal of using EDF in the fitting, I cannot tell the difference between EDF*, EDF_T, and EDF_A from the text.

(3) The authors are intimately familiar with the analysis of PTFE filters. Unfortunately, they assume that their readers have an equal level of expertise. For example, many readers may not be familiar with OC and EC determinations from PTFE filters, as the referenced papers were only recently published (most readers are likely much more familiar with OC/EC analysis of quartz filters). Thus some additional explanation of OC/EC analysis of PTFE filters is warranted. Likewise, some more background on clustering and its application for PTFE filters would be helpful.

C2

(4) The authors argue that the automated SSB method is preferential because it allows bulk, presumably fast analysis or large sets of filters. Thus it would be useful for the authors to note the time required by the existing PB method (per filter or for the full set of 794 filters) versus the automated SSB method.

(5) Figure 10 shows performance of the calibration set versus the test set, however all other figures seem to indicate the full set of 794 filters. Was a test set only used for OC/EC analysis? I think that a better explanation of OC/EC analysis of PTFE filters (comment 3) could clear up my confusion.

(6) For the data shown in Figures 7-10 - Is the same EDF used for every filter, or is it adjusted automatically for each? If it is the former, do the authors expect that the general EDF guidelines given here (e.g., EDF between 2 and 4) will apply universally? Or will each operator/sample collector need to determine EDF at the outset of analysis?

(7) Page 11, Line 22 - "extensive number of knots" for forming the spline - how exactly are the number and location of knots determined?

(8) Figure 7 may understate the clustering differences between PB and SSB. The total difference for the 5-cluster solution is only 10%, but for certain clusters (e.g., Type IV), the difference is much larger.

(9) Figure 9 compares integrated peak areas for different functional groups after baseline correction with SSB and PB. The figure implies that the same integration method or code is used in each case. The authors should specify whether or not this is true.

(10) Equation 1: what are x and y"?

(11) Page 10, Line 1 - particle size is labeled as micro-gram, not micro-meter

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2015-384, 2016.