Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-404-RC2, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "iFit: An intensity based retrieval for volcanic SO_2 from scattered sunlight UV spectra" *by* Ben Esse et al.

Anonymous Referee #2

Received and published: 5 March 2019

Remote sensing of volcanic SO_2 emissions via Differential Optical Absorption Spectroscopy (DOAS) has become a major tool in volcanology. DOAS analyses differences between a measured light spectrum and a background spectrum. The background spectrum is usually recorded by the same instrument in the temporal proximity to the measured spectrum but in another viewing direction. Using such a background spectrum allows DOAS for an automatic correction of most instrumental effects, stratospheric effects, and shared tropospheric effects. As a drawback, this method is insensitive for a contamination of the background with e.g. volcanic SO_2 and DOAS then potentially underestimates the absolute volcanic SO_2 emissions. Salerno et al. (2009) and Lübcke et al. (2016) presented alternative approaches which use a synthetic background spectrum in order to detect/circumvent a possible background contamination.

Esse et al. present another approach *iFit* using a synthetic background spectrum. Advances beyond the state of the art are desirable and AMT is in principle a suitable journal for such a topic. It is however not clear to me whether their proposed approach poses a substantial advance. In particular, it is not possible for me to assess their approach (1) because the approach is not described in sufficient detail and (2) because its performance is not set in contrast to the state of the art. In addition, some wrong or ambiguous statements on DOAS raises concerns whether Esse et al. apply DOAS in the best available way.
In conclusion, I can not support the publication of the manuscript in the presented form. For publication, *iFit* has to be described - first of all and at the very least - in such a way that the reader is able to reproduce their results. Second, I highly recommend a quantitative and more comprehensive comparison of *iFit* with the approach from Lübcke et al. (2016) and with the standard DOAS approach in order to provide evidence that *iFit* in fact provides an advance beyond the state of the art. Thirdly, I recommend to neglect all redundant and subjective statements from the manuscript in order obtain maximum clearness. Please find below a detailed list of the major objections.
Finally, I expect that tackling these major objections would result in massive changes of most parts of the manuscript. In consequence, a second subsequent review appears to be mandatory which focusses on the then provided spectroscopic details of their approach.

1 No comprehensive description of their method is provided

I am aware that Esse et al. uploaded their used data and code written in python. The proposed approach has to be reproducible in principle with any programming language/platform/python version. Therefore, this review does not assess the python code but exclusively assesses the provided manuscript.
Esse et al. sketch the architecture of their approach in Figure 2, however, mathemati-

cal descriptions or measurement instructions are not provided for each step. Therefore, iFit can not be reproduced by the reader. In particular, the following steps have to be provided/clarified:

1. Page 6, lines 27-29: *"A model spectrum is then built on the high resolution model grid, which typically has a spacing of 0.01 nm..."* How is this model spectrum build? By a convolution with a Gaussian ILS with a FWHM of 0.01 nm?

2. How are the effects of the dark current and "bias" determined and corrected? Remark: the latter is typically called "offset" rather than "bias" (see e.g. Platt and Stutz, 2008).

3. Flat spectrum: Esse et al. retrieved the "flat spectrum" by a simple averaging. In contrast, Lübcke et al. (2016) retrieved the instrument effects (flat spectrum but also further instrument effects such as temperature effects) by a Principal Component Analysis (PCA). For me, the PCA approach appears to be more comprehensive. See also my comment 2.3. Please motivate the choice of a simple averaging instead.

4. Background polynomial: Is equation (10) correct? Physically, all light attenuation effects are on the same footing, i.e. both absorption and scattering effects are summands in the argument of the exponential function. In principle, the scattering effects can of course be written in the presented way e.g. as $P(\lambda) = \exp(P^*(\lambda))$ where $P^*(\lambda)$ is the "real" broad-band scattering polynomial. But this is strictly different from the polynomial as it is used in DOAS. In particular, its coefficients will be different to the polynomial $P(\lambda)$ denoted in equation (3). Please clarify.

5. Ring spectrum: First, I have analogous doubts concerning $R(\lambda)$ in equation (10) and equation (9). Second, please make more explicit how is the Ring spectrum retrieved.

6. Sky spectrum: How is it constructed? By adding the absorption effects of background O_3 and NO_2? If yes, how is the correct background amount determined?

7. Plume spectrum: Is also O_3 included in the fit step from sky spectrum to plume spectrum? If not, how are the diurnal variations of stratospheric O_3 contributions and assessed and corrected? Furthermore, why is particularly BrO (but no other gases) included in the SO_2 fit scenario? The BrO absorption is rather negligible in the typical SO_2 fit ranges.

8. Instrument line shape function ILS: *"The ILS was measured using a mercury lamp to be* $0.50\,(\pm 0.01)\,nm$*"* (page 8, line 27). The uploaded data does provide ILS (only) at 302 nm (instrument H15972) with a FWHM of 0.58 nm and at 301 nm (instrument FLMS02101) with a FWHM of 0.60 nm. Please provide the full mercury spectra for a presented instruments. (Remark: both uploaded ILS have indeed an about Gaussian shape. A super-Gaussian model proposes exponents of (only) 2.3 and 2.1 and the asymmetry is rather small as well.)
   The ILS is a unique property of the instrument, although it varies in general with temperature and wavelength. Accordingly, for a given instrument (and similar temperature) all convolution operations have to use one identical ILS. Applying different ILS can cause a significant decrease in accuracy. Furthermore, all compared spectroscopic should apply an ILS retrieved at the same wavelength in order to be consistent. Ideally, this wavelength is chosen in the wavelength range, e.g. at 315 nm. For practical reasons, the mercury line either at 302 nm or at 334 nm should be chosen. Please clarify which measurement results for the ILS are used. I propose to add a table to the manuscript which lists all instruments used in this study and their spectroscopic properties.
   However, later a modelled ILS with *"an ILS width of 0.56 nm"* (page 10, line 9) has been used instead of the measured ILS. Please clarify why instead of the exact measurement results an apparently wrong ILS is used at this step.

9. How are the wavelength shift and stretch determined and corrected?

10. Furthermore, physical-logical the wavelength shift and stretch should be actually applied on the measured spectrum in order to correct for temperature-driven variations of the instrument during the measurement. Analogously, the "flat spectrum" should be applied on the simulated spectrum in order to correct for the instrument effects of the real instrument. I can imagine that these inconsistencies are mathematically identical and may thus lead to the same results. Please clarify why/whether these inconsistencies are required.

11. What actually does "perform fit"? Is it an ordinary DOAS fit? Or is it iteratively minimising $\tau = \log(\frac{I_{right\,hand\,side}}{I_{left\,hand\,side}})$ by means of varying the SO_2 column density?

12. Is a stray light correction applied?

2 Missing comparison with the state of the art

A comparison with the state of the art is required to provided evidence that *iFit* adds substantial value to the literature. Thereby, it should made clear under which scenarios *iFit* may improve the state of the art and under which scenarios it does not.

1. The manuscript is motivated by the possible underestimation of the SO_2 slant column density in a volcanic gas plume. However, no *iFit* results for such a scenario have been provided. Please explain this inconsistency.

2. According to the list of literature provided in the manuscript (and to my knowledge), the approach from Lübcke et al. (2016) is the current state of the art to face such background contamination issue. I highly recommend a direct quantitative comparison of these two methods when applied on the same data (ideally contaminated data) in order to reveal the major differences.

3. Esse et al. propose to use *iFit* for evaluating data recorded by permanent monitoring stations. Monitoring stations are typically not temperature controlled in order to improve their robustness and to lower their power consumption (see e.g. Galle et al., 2010). *iFit* has been tested exclusively for temperature stabilised instruments and Esse et al. concluded that *"care must be taken if the spectrometers are not temperature stabilised"* (page 11, line 9). Accordingly, Esse et al. have not provided evidence that *iFit* is suitable for monitoring stations. This is in particular in sharp contrast to the approach from Lübcke et al. (2016) which presented their results for contaminated data from monitoring stations.

4. I agree with Esse et al. that also a comparison with standard DOAS (recorded background spectrum) appears to be mandatory. For the arguments stated in the very first paragraph, I expect that a standard DOAS approach performs in general (i.e. for non-contaminated scenarios) better than *iFit*. In contrast, the narrative in the current manuscript is rather one-sided, highlighting possible problems in the DOAS approach only. I highly recommend that any subjective valuations are neglected from the technical manuscript parts (e.g. "methods", and "results"). Differences between *iFit* and DOAS should be discussed later in the "discussion" part where all evaluating statements should be supported by (quantitative) evidence.

5. Esse et al. conclude *"the lack of a requirement for a reference spectrum means that iFit would be especially well suited to deployment in permanent scanning stations"* (page 11, line 14). Permanent scanning stations scan typically from horizon to horizon and thus automatically recorded the reference spectrum. Applying *iFit* thus does not provide any gain in measurement time *in particular* at permanent scanning stations. Furthermore, probably only few permanent measurement stations are at all affected by background SO_2 contamination. Accordingly, possible benefits from *iFit* are limited to those stations. Please limit your conclusions with respect to those scenarios where you can provide evidence that

iFit at least does not perform more poorly than the alternative approaches.

3 Curious statements on DOAS

Several strictly wrong or curious/ambiguous statements create some doubts whether Esse et al. apply DOAS on the state of the art level. Namely:

1. Page 4, lines 11-19: Is there any need to discuss the option of a high-pass filter? Is a high pass filter used in *iFit* or for the DOAS retrieval in this manuscript? If not, this paragraph appears to be redundant. The figures 4, 5, 10 show results of the *total* absorption cross section.

2. Page 4, line 26: *"The ILS can either be a mathematical function (such as a Gaussian)..."* The ILS is a property of the instrument and has in general an arbitrary shape. Although it can be indeed often *approximate in good agreement* by a Gaussian line shape function, the real ILS itself is not a mathematical function!

3. Page 4, lines 28-31: Equation 6 is not true! The convolution operation and the scalar multiplication operation are commutative!

4. Page 6, lines 29-30: *"A wavelength-shift is a common correction in DOAS..."*. This is correct, however, when the spectra are wavelength-calibrated prior to the DOAS fit this shift is typically in the order of ± 0.001 nm. Do Esse et al. refer to the additional wavelength shift parameter which is usually allowed between the measurement spectrum and the absorption cross-sections in order to partially compensate for the convolution with a (slightly) wrong ILS? Anyway, this wavelength shift is typically limited to ± 0.2 nm rather than ± 2 nm. A wavelength shift of 2 nm appears to be absurdly large. Please clarify.

5. Page 7, line 27: *"The wavelength region used for these results (304 - 320 nm) is common to most scattered sunlight retrievals of SO_2."* This is not true. In DOAS - the predominant remote sensing technique for volcanic SO_2 - the used wavelength range starts almost exclusively at 310 nm (e.g. Lübcke et al., 2016), 312 nm (e.g. Theys et al., 2017; Kern and Lyson, 2018), 314 nm (e.g. Lübcke et al., 2014; Dinger et al., 2018), or 326 nm (e.g. Hörmann et al., 2013). The reason for these lower limits is, that the applied approximation in DOAS are only justified as long as the "absorbance" is not much above 0.1. For the data presented in Figure 10, this means the DOAS retrieval should start not lower than at 314 nm. This limitation does not have to hold for an intensity based fit, however, Esse et al. have to make sure that all presented DOAS data are retrieved for an absorbance below 0.1. Otherwise their DOAS results would be too low and a quantitative comparison between *iFit* and DOAS therefore flawed.

6. Page 8, line 31: *"In particular, use of higher wavelengths leads to an overestimation in the retrieved SO_2, possibly due to the reduced strength of the SO_2 absorption spectrum at higher wavelengths"*. This interpretation of the findings is not supported by further evidence. Furthermore, the *"reduced strength of SO_2 absorption"* can be expected to result in a larger fit error but there is no obvious reason why this should cause a less accurate result. In fact, I would interpret the findings other way round: the lower the wavelength the larger is the underestimation in SO_2 due to saturation effects and a decreasing solar background radiation (see Platt and Stutz, 2008). At least for DOAS the "absorbance" should be below 0.1 in order to keep the applied approximations justified. With these fundamental limitations in mind, I consider the results for the wavelength range from 310-320 nm the most accurate. Furthermore, I expect that starting at 314 nm or 326 nm would give larger and even more accurate results in particular for the cells above 500 ppmm. In consequence, *iFit* would overestimate the SO_2 slant column density. Please provide evidence for your interpretation. In particular, I

highly recommend to present (additionally or exclusively) DOAS results when the wavelength range starts at least at 310 nm.

4 Some minor formal objections

- Inconsistent use of brackets (e.g. compare equation 7b and equation 8)
- The (slant) column densities are sometimes denoted by a_i and sometimes by α. They have to be denoted by the same consistent letter throughout the manuscript. Furthermore, they should always hold the index.
- I_0 and I_0^* appears in several forms throughout the manuscript. They should be consistently denoted by a strictly constant sign.

Additional references

Dinger, F., Bobrowski, N., Warnach, S., Bredemeyer, S., Hidalgo, S., Arellano, S., Galle, B., Platt, U., and Wagner, T.: Periodicity in the BrO/SO2 molar ratios in the volcanic gas plume of Cotopaxi and its correlation with the Earth tides during the eruption in 2015, Solid Earth, 9, 247–266, doi:10.5194/se-9-247-2018, 2018.

Galle, B., Johansson, M., Rivera, C., Zhang, Y., Kihlman, M., Kern, C., Lehmann, T., Platt, U., Arellano, S., and Hidalgo, S.: Network for Observation of Volcanic and Atmospheric Change (NOVAC)âĂŤA global network for volcanic gas monitoring: Network layout and instrument description, Journal of Geophysical Research: Atmospheres, 115, doi:10.1029/2009JD011823, 2010.

Hörmann, C., Sihler, H., Bobrowski, N., Beirle, S., Penning de Vries, M., Platt, U., and Wagner, T.: Systematic investigation of bromine monoxide in volcanic plumes from space by using the GOME-2 instrument, Atmospheric Chemistry and Physics, 13, 4749–4781, doi:10.5194/acp-13-4749-2013, 2013.

Kern, C. and Lyons, J.: Spatial distribution of halogen oxides in the plume of Mount Pagan volcano, Mariana Islands, Geophysical Research Letters, 45, 9588–9596, doi.org/10.1029/2018GL079245, 2018

Lübcke, P., Bobrowski, N., Arellano, S., Galle, B., Garzón, G., Vogel, L., and Platt, U.: BrO/SO2 molar ratios from scanning DOAS measurements in the NOVAC network, Solid Earth, 5, 409–424, doi:10.5194/se-5-409-2014, 2014.

Theys, N., De Smedt, I., Yu, H., Danckaert, T., van Gent, J., Hörmann, C., Wagner, T., Hedelt, P., Bauer, H., Romahn, F., Pedergnana, M., Loyola, D., and Van Roozendael, M.: Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis, Atmos. Meas. Tech., 10, 119-153, https://doi.org/10.5194/amt-10-119-2017, 2017.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-404, 2018.