Reply to comments by Anders Svensson:

> ***Note that our responses are indented, bold and italicized***

The manuscript is concerned with estimating the uncertainties related to deriving the Aerosol Optical Depth (AOD) from the coloring of the sky in paintings made at the time of large volcanic eruptions. As such, the manuscript provides a critical comment to existing publications that are aiming at deriving such values. Using an atmospheric radiative transfer model, the manuscript presents a number of sensitivity studies by varying a number of parameters that may lead to uncertainties in the color-based AOD estimates. The manuscript is well written and easy to follow and the various uncertainties brought to the table appear relevant.

> ***Reply: We thank the reviewer for his positive and constructive review.***

I do not have comments on the radiative transfer model for which I am not an expert, but I have two suggestions for the authors to consider:

My first point concerns the 'true' error related to AOD estimates from paintings. For historical volcanic eruptions, there are ice-core based estimates of the sulfate deposition in both Greenland and Antarctica that provide an independent estimate of the stratospheric sulfate aerosol loading, which in turn can be translated into an stratospheric AOD (Gao et al., 2007;Gao et al., 2008;Sigl et al., 2015). Can those estimates be applied to give an independent estimate of the accuracy of the paintings derived AOD? Of course, the ice-core estimates can be questioned themselves, but a reasonable agreement between the two independent approaches would nevertheless suggest that both methods are providing AOD estimates that are in the right order of magnitude, at least. Likewise, a large disagreement between the two methods would suggest that at least one of them has very large uncertainties. Maybe this comparison has already been done in another study? It seems like a quite obvious comparison to make?

> ***Reply: This is a very good idea and a comparison with independent data sets (including AOD estimates based on ice cores) has actually already been made by Zerefos et al. (2014) (see their Figures 4 and 5). This comparison shows significant differences between the AOD values estimated from the different techniques (including analysis of ice cores and of historic transmission measurements). The agreement between the datasets is reasonable for the major eruptions with AOD values exceeding 0.1. But for the weaker eruptions the AODs estimated from the paintings are systematically larger than the ones from the other data sets. Figure 5 in Zerefos et al. (2014) shows comparisons of 50-year mean AODs from the different data sets and the AODs estimated from the paintings are systematically about 1 order of magnitude larger than the other data sets.***
>
> ***Zerefos et al. (2014) also list in Appendix D their AOD values together with values from different other studies. In order to investigate the differences, we determined averaged AOD values considering only the years for which data from the two datasets to be compared are available. We obtained the following results:***
>
> ***Zerefos vs. Robertson:***
> ***$AOD_{Zerefos}$ = 0.174 ± 0.107 (Mean value and standard deviation)***
> ***$AOD_{Robertson}$ = 0.017 ± 0.038***
>
> ***Zerefos vs. Crowley & Unterman:***
> ***$AOD_{Zerefos}$ = 0.198 ± 0.121***

$AOD_{Crowley}$ = 0.031 ± 0.058

Zerefos vs. Sato:
$AOD_{Zerefos}$ = 0.180 ± 0.093
AOD_{Sato} = 0.0175 ± 0.022

Zerefos vs. Stothers:
$AOD_{Zerefos}$ = 0.226 ± 0.124
$AOD_{Stothers}$ = 0.032 ± 0.044

Figure 1 below depicts all the datasets listed in Appendix D of Zerefos et al. (2014). Apparently, the mean AOD values do differ by about an order of magnitude. As the Figure below demonstrates, the differences reach two orders of magnitude in some cases. It is also obvious that the AOD values estimated from the historic paintings are almost all larger than 0.1. Significantly smaller values are not inferred.

A brief discussion on the differences between the different AOD data sets compared in Zerefos et al. (2014) was included in section 4 of the revised version of our manuscript.

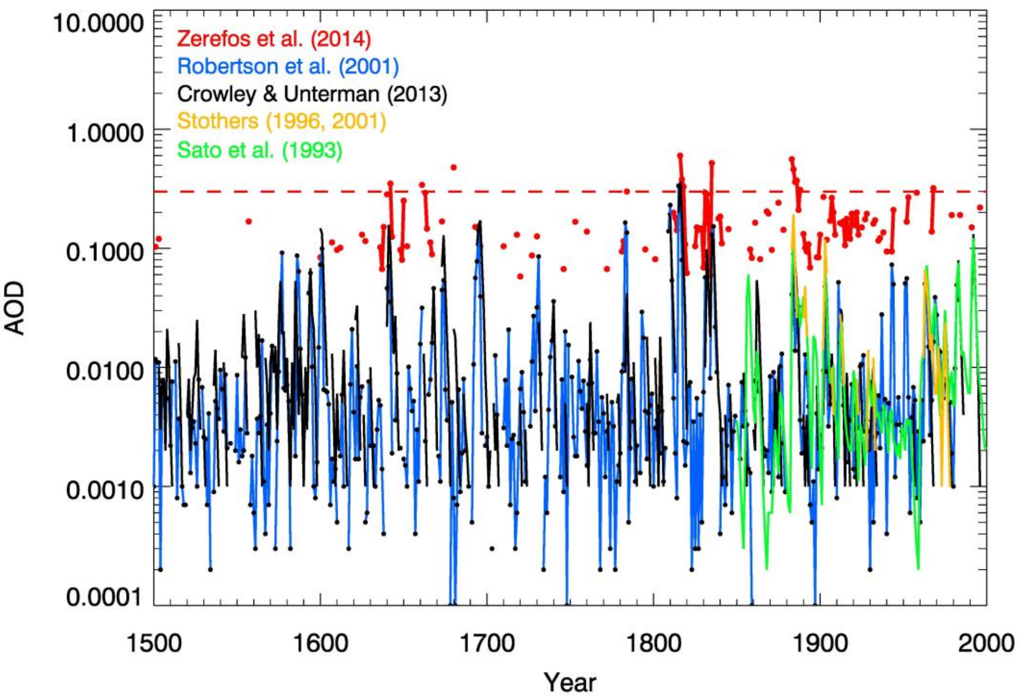

Figure 1: Comparison of AOD values retrieved from historic paintings (in red) with independent estimates, as described in our reponse. All data points were taken from the table in the appendix of Zerefos et al. (2014). The red dashed lines corresponds to AOD = 0.3.

References:

Crowley, T. J. and Unterman, M. B.: Technical details concerning development of a 1200 yr proxy index for global volcanism, Earth Syst. Sci. Data, 5, 187–197, doi:10.5194/essd-5-187-2013, 2013.

Robertson, A., Overpeck, J., Rind, D., Mosley-Thompson, E., Zielinski, G., Lean, J., Koch, D., Penner, J., Tegen, I., and Healy, R.: Hypothesized climate forcing time series for the last 500 years, J. Geophys. Res., 106, 14783–14803, 2001.

Sato, M., Hansen, J. E., McCormick, M. P., and Pollack, J. B.: Stratospheric aerosol optical depths 1850–1990, J. Geophys. Res., 98, 22987–22994, 1993.

Stothers, R. B.: Major optical depth perturbations to the stratosphere from volcanic eruptions: pyrheliometric period 1881–1960, J. Geophys. Res., 101, 3901–3920, 1996.

Stothers, R. B.: Major optical depth perturbations to the stratosphere from volcanic eruptions: Stellar extinction period 1961–1978, J. Geophys. Res., 106, 2993–3003, 2001.

My second point concerns an overall error estimate for the historic color painting method for estimating the stratospheric aerosol optical thickness based on the uncertainties introduced in the present manuscript. In the manuscript, we are provided with numerous figures showing the AOD sensitivity to factors such as particle size distribution, wavelength, solar zenith angle, albedo, azimuth angle, etc. All of those dependencies certainly leave the impression 'that the uncertainties of the estimated aerosol optical depths are so large that the values have to be considered highly questionable', as mentioned in the abstract. However, how large are the uncertainties 'typically' in a real-case scenario? If we add up all of the uncertainties using a realistic range of values for the parameters discussed in the study, do we then end up with 5% or 50% uncertainty on the final result? If the total uncertainty is in the range below say 50%, the method may still be applicable, eg if there are several paintings of the (sky of the) same eruption that may provide independent evidence. If the final uncertainty estimate is large however, say above 50%, the entire approach of using paintings for estimating the AOD becomes questionable. Therefore, some kind of summary providing a combined uncertainty from all of the discussed parameters would be quite helpful. Also, an estimate of the relative uncertainty contribution from each investigated parameter would be helpful again using a realistic range of parameters. If possible, some uncertainty estimates/ranges could be provided in a table? This may provide some useful guidance for future studies of what knowledge is needed to make constrained AOD estimates from paintings. Maybe, in some cases, there is independent evidence of say the position of the Sun or the time of the day when the picture was painted? Likewise, we may become wiser in the future about what to expect from the particle size distribution related to large volcanic eruptions. Thus - wearing an optimistic hat - it could be that some of the uncertainty ranges discussed in the manuscript could be significantly reduced or even eliminated for specific paintings/eruptions?

Reply: Another good idea, which was also suggested by reviewer #3. We thought about a total error estimate already when writing the manuscript and decided to omit it, because we believe it is not possible to reduce the complexity of the problem and the impacts of the different effects and parameter uncertainties in a single number. Also, the large differences between the AODs in Zerefos et al. (2014) and the independent data sets discussed above question the possibility of a total error estimate. We do think, however, that it is possible to carry out a total error estimate for individual paintings, if some of the critical parameters are well known or can be sufficiently constrained. Also following the comments by reviewer #3 we adjusted the basic message of the paper and we now state that it may be possible to infer quantitative information on the AOD from individual paintings, whose history is well known and for which all relevant parameters are known sufficiently accurately.

We understand that it would be desirable to have a total error estimate or at least a paragraph describing the issues related to provide one. We added a discussion to the results section, providing a justification for our assumption that a total error estimate is not possible. The following aspects are important here.

A major problem – for paintings showing scenes with the sun below the horizon – is the strong SZA-dependence of the red-green colour ratio (see Fig. 3 of the revised version of the paper). As already pointed out in the paper, without exact knowledge of the SZA, an estimation of the AOD is not possible. Even if the painting is finished in, e.g. 30 minutes during the sunset, the SZA (and with it the colours) will have changed significantly.

Regarding the problems discussed in section 4 (How realistic were the colours on the day the painting was finished & how did the colours change over time), we strongly believe that general estimates of these effects are not possible and should not be attempted, in order to avoid implying a false sense of reliability of the results. Again, for an individual painting robust error estimates might be possible, but not for the entirety of all paintings.

The effects of uncertainties in the assumptions of the amount (and profile) of ozone or the surface albedo are rather small and not the main issues.

References:

Gao, C. C., Robock, A., and Ammann, C.: Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models, Journal of Geophysical Research-Atmospheres, 113, D23111, 10.1029/2008jd010239, 2008.

Gao, C. H., Oman, L., Robock, A., and Stenchikov, G. L.: Atmospheric volcanic loading derived from bipolar ice cores: Accounting for the spatial distribution of volcanic deposition, Journal of Geophysical Research-Atmospheres, 112, 10.1029/2006jd007461, 2007.

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., Salzer, M., Schüpbach, S., Steffensen, J. P., Vinther, B. M., and Woodruff, T. E.: Timing and climate forcing of volcanic eruptions for the past 2,500 years, Nature, 523, 543-549, 10.1038/nature14565, 2015.