

This is a novel approach to deriving stratospheric sulfate aerosol information from the SAGE series of instruments. Overall, the approach seems logical and appropriate. One of the major benefits of employing a statistical approach is that error estimates automatically ‘fall out’ of the analysis. One of the issues I perceive in the results is a possible bias to larger particle size, presumably from larger particles ‘leaking’ into the solution set, as I will discuss in more detail in specific comments provided below.

Another (probably unavoidable) issue with the approach is that it treats all aerosols as sulfates, even if they are smoke, PSCs, volcanic ash, ice... It might require a sophisticated user to interpret (or know when to ignore) parameters derived from events other than volcanoes that inject aerosols into the stratosphere, or when there are clouds (e.g., PSCs or cirrus). However, it is also true that the majority of aerosols in the stratosphere (outside of PSCs in the polar vortex) are likely sulfate, particularly when you consider background aerosols. The authors provide a warning regarding smoke, but it might be worth extending that warning to include any aerosols other than sulfates.

I will make the comment that sulfate aerosol particle size distribution information derived from SAGE could represent a valuable data set. Instruments that measure in limb-scattering mode (such as OSIRIS and OMPS) need to assume size distribution parameters in their analysis. The field seems to have settled on using parameters appropriate for “fine mode” sulfate aerosols (as determined by particle counter measurements) for analysis, based on the argument that fine mode aerosols typically significantly outnumber “coarse mode” aerosols in the atmosphere. However, larger particles provide dramatically higher contribution to the scattering signal, as this paper shows, and it is not appropriate to simply ignore larger particles when analyzing scattering-based measurements. The field (whether it realizes it or not) is in desperate need of more representative sulfate aerosol size distribution parameters if we want to determine the impact of sulfate aerosols on climate more accurately. The SAGE series of instruments represents a likely source for such information.

Specific comments:

line 126: it was observed that if $\sigma = 1.5$ (Fig. 4, panel b), $r_m = 75$ nm (i.e., background conditions)

These numbers are typical values ascribed to background sulfate aerosols, but as I said above, they are not representative of the PSD you would need to reproduce measurements involving scattering using a monomodal size distribution. They are representative of the fine mode in a bimodal size distribution (they are basically the same values you use for the fine mode in Figure 13b when considering the impact of having a bimodal distribution). You are free to use any numbers you want in your calculations, of course, but explicitly referring to these values as “background conditions” serves to perpetuate the notion that they are representative of the values you would determine when analyzing your background measurements with a monomodal distribution. They are not.

Looking at Figure 15 (Northern Hemisphere for SAGE II), one would estimate a typical value for r_m under background conditions to be somewhere in the vicinity of 200 nm. Looking at Figure 18 (Northern Hemisphere for SAGE III/ISS), one might estimate a value closer to 150 nm.

This paper makes the case that enhanced scattering from larger particles means one cannot ignore the coarse mode despite the lower number density relative to the fine mode. Suggesting that $r_m = 75$

nm, $\sigma = 1.5$ is representative of background aerosols undermines that argument because those numbers were chosen to be consistent with fine mode parameters from a bimodal distribution observed by particle counters. Using those numbers argues that the coarse mode can be ignored in measurements involving scattering.

In the end, the goal is not to accurately model the physical nature of atmospheric aerosols, but rather to model the optical characteristics of those aerosols, which is the important consideration for their impact on climate. If the real distribution is bimodal (or some higher order multimodal) and a monomodal analysis is employed, the derived parameters represent effective values for r_m and σ . The values of these effective parameters will be skewed mightily by the extreme sensitivity of scattering effectiveness as a function of particle size. Larger particles will contribute to the values of the effective parameters in a manner that is wildly out of proportion to their relative number density. That is a limitation of the measurement system that must be acknowledged if we hope to properly interpret the implications of the measurements.

Dealing with effective parameters will muddle the interpretation of the chemical impact of the aerosols (the weighting of the effective r_m to larger values will likely imply a larger aerosol volume than exists in reality). It will also impact the inferred greenhouse effect from absorption in the infrared (absorption and scattering have different sensitivities to particle size). However, those issues are worries for another day. A positive first step would be moving away from PSD parameters that are a good representation of the bulk physical characteristics for the majority of sulfate aerosols in the atmosphere (i.e., $r_m = 75$ nm and $\sigma = 1.5$, corresponding to only the fine mode aerosols) to effective PSD parameters that provide a more accurate representation of the optical response of all (e.g., both fine mode and coarse mode if the distribution is bimodal) sulfate aerosols in the atmosphere (i.e., r_m larger and σ TBD but probably a bit larger to encompass both modes in the bimodal distribution).

Line 215: For example, if the reported errors of the extinction coefficients were all within 5% and the inferred mode radius was 100 nm then, per Fig. 7 (c), we know that on median the inferred mode radius is $\approx 5\%$ too high, and that 90% of the time the inferred value is within $\pm 15\%$ of the target value.

Looking at Figure 7c, I would suggest we can infer more than this. There is a strong tendency for the analysis to generate a larger value for r_m . For r_m below ~ 200 nm, this is evident even in the average values, with a clear high bias introduced into the inferred value. For r_m above ~ 250 nm, the effect manifests as a tall upper 'whisker' stretching the fringe of the distribution in the solution set to include larger particles than were present in the pseudo-SAGE data used as the target.

Conversely, the values for inferred distribution width σ in Figure 7f are generally biased low (other than for $\sigma \leq 1.2$). This is likely not a coincidence. There is an inverse relationship between r_m and σ . If you increase one of the parameters, you can compensate for that in the calculation by decreasing the value of the other parameter. It looks like there is some aspect of the sensitivity study that trades a smaller distribution width for a larger mode radius. For smaller r_m (< 250 nm), it doesn't appear to be a straight swap, though, because in Figure 7o the inferred values for effective radius r_e are also biased high. I would call this problematic. Even though the bars in Figure 7o cross 1.0 and the offsets could therefore be deemed to be statistically insignificant, there is a clear bias introduced. These are synthetic

data where you know the truth. I assume that the indicated errors are uncertainty levels assumed in the Σ matrix rather than random errors added to the set of extracted data.

My interpretation would be that the analysis approach tends to allow a number of solutions at larger r_m to leak into the solution set. If it happens in the synthetic data, I expect it would affect the analysis of real measurements as well (i.e., create a bias toward larger r_m). Is there any way to adjust the analysis to reduce that bias?

Line 268: Since the H₂SO₄ content of atmospheric aerosols is, ultimately, unknown we note that this situations adds an unknown element to the analysis

The H₂SO₄ content in the droplet will be governed by thermodynamic equilibrium. The H₂O vapor pressure for the droplet should equal the ambient partial pressure of H₂O in the atmosphere (Steele and Hamill 1981, [https://doi.org/10.1016/0021-8502\(81\)90054-9](https://doi.org/10.1016/0021-8502(81)90054-9)). If one knows the ambient temperature and H₂O concentration, an estimate of the H₂SO₄ content can be calculated (Bernath et al 2023, <https://doi.org/10.1016/j.jqsrt.2023.108520>). If there is higher H₂O pressure in the ambient air, the droplet should take up H₂O to find equilibrium (thereby decreasing wt% H₂SO₄). Conversely, if the H₂O pressure in ambient air is lower, there would be evaporation from the droplet to achieve equilibrium.

Line 280: the particles were composed of 75% H₂SO₄ with $r_m=75$ nm, and $\sigma=1.5$ (i.e., background conditions)

Another reinforcement of fine mode parameters supposedly being representative of background aerosols. A caveat of some sort would be appropriate, unless you truly believe the numbers are representative (despite what your later results suggest).

Section 4.3: observations of stratospheric smoke

I don't disagree with what was done in this section, but it may underestimate the impact of smoke on the PSD parameters. Unlike the liquid droplet sulfate aerosols, smoke particles are not constrained to be spherical, which could significantly affect its scattering characteristics.

Also, aged smoke particles apparently turn glassy, which would presumably change its optical constants, possibly outside the range of combined BC and BrC employed in the study, although for all I know turning glassy might push the optical constants closer to sulfate aerosol values, which would reduce the systematic errors.

Figure 12: calculating values for N

When you are working with extinction ratios, it is not clear how number density (N) is being derived. Information on number density cancels when you calculate the ratios. Even if you were working with straight extinction values (rather than ratios), you would only be able to determine the column density of

the particles along the line of sight (the average number density times the path length through the sulfate aerosols), not the number density itself.

A discussion of adding number density to LUTs occurs in Section 5. I assume that is involved. The ratio of number density for the two modes in a bimodal distribution will affect the shape of the spectrum, so I can see how that ratio could be determined in the bimodal analysis. Nothing in the paper indicates how it is being done in a monomodal analysis where you are working with extinction ratios.

Line 339: However, the shorter wavelengths show an interesting behavior in that their extinction increased rapidly followed by a flattening and slight decrease in extinction, which is subsequently followed by another rapid increase.

This behavior is actually associated with resonances in the Mie scattering for certain particle sizes. Note that the curve for 384 nm in Figure 13 has two peaks, one near where r_2 is roughly equal to the wavelength (around 384 nm), and one where r_2 is roughly twice the wavelength (around 768 nm). Similarly, the curve for 520 nm has a peak where r_2 is near the wavelength (520 nm). If you were to extend the plots to higher wavelength, you would presumably see a second peak somewhere around 1040 nm.

Line 441: what if the second mode has a different composition than the first?

Droplets with different sizes will have slightly different H_2O vapor pressures because it will be impacted by the degree of curvature of the surface, but I would expect the composition of two droplets under the same conditions (temperature and ambient H_2O concentration) to have relatively comparable compositions.

Line 479: under enhanced aerosol load (e.g., $k_{1020} > 1E-4 \text{ km}^{-1}$) the majority of the SAD and r_e estimates were within the $\pm 30\%$ uncertainties stated in Thomason et al. (2008)

We can also note from Figure 17 that the values of r_e from this study are systematically higher than the v7.0 results, which is perhaps another indication that this approach yields a high bias, particularly at smaller mode radius, as was (to me) suggested by the sensitivity study results in Figure 7.

Line 503: The southern hemisphere was influence by more events including transport from the 2017 Canadian wildfire

It is surprising that you would choose not to mention Australian Black Summer (or Australian New Year, as it is also called) wildfires from 2019/2020, labeled Aw in Table 7, the most dramatic Southern Hemispheric event in recent years other than the Tonga eruption. It would certainly rate mention before long-range transport from the 2017 Canadian wildfires (more typically termed the Pacific Northwest pyroCb event). The impact of the 'Aw' event is clearly evident in r_m and σ in Figure 18. See, for example, the light green swath in the plot for σ that extends from the beginning of 2020 into early 2022, at which

point σ in the stratosphere turns dark green following the Tonga eruption. It looks like enhanced smoke aerosols from Aw might have persisted for a couple of years. This would be an opportunity to warn readers that because these are smoke aerosols, not sulfate, the values of the parameters will have systematic errors.

SAGE II versus SAGE III/ISS

No comparison is made between the SAGE II results (Figures 15 and 16) and the SAGE III/ISS results (Figures 18 and 19). Later (in the Conclusion), the statement is made that the SAGE II recorded was dominated by large particles, but I would suggest it is a stretch to claim that background aerosols were unusually large for the entire duration of the SAGE II data record. There is presumably an aspect of differences in analysis approach for the two instruments contributing here.

In general, the SAGE III/ISS results show more structure, which is likely expected from the additional spectral information available. As alluded to in the Conclusion, SAGE II appears to yield larger values of r_m for background aerosol conditions. It is worth asking whether that is a byproduct of using only one extinction ratio in the analysis (condition #0). It might be instructive to analyze the SAGE III/ISS data with the same approach as was used for SAGE II (using condition #0) and look at the differences from the nominal results (using the hybrid conditions 5/6/15). That would give a sense of potential systematic errors in the SAGE II results arising from the limited spectral data employed in the analysis.

Line 523: Here, the plume of largest particles were centered at ≈ 20 km near the equator and descended to lower altitudes toward the higher latitudes.

The difference between the peak altitude near the equator versus that at higher latitudes is not a direct indication of aerosol sedimentation. As was mentioned earlier in the paragraph, Brewer Dobson circulation will naturally transport air to lower altitude as it moves poleward, even in the absence of sedimentation. To get a sense of the descent, you should observe the location of the > 300 nm plume at a particular latitude (e.g., 30°S or 45°S) and note how it changes over time.

Minor comments and typos:

line 45: As an occultation measurement, the SAGE instrument peers through hundreds, sometimes thousands, of kilometers of atmosphere

It depends on how you define the extent of the atmosphere, I suppose. For measurement purposes, I would define the overall height of the atmosphere according to the range over which atmospheric extinction contributes significantly to the signal. With that consideration, I would say you are looking through hundreds of kilometers, not thousands.

line 95: “distribution width”

This term was never explicitly defined. Looking at Equation 2, some people might interpret distribution width as σ , while others might interpret it as $\ln(\sigma)$. The latter choice might be the more logical interpretation. σ would be literally interpreted as the width of the distribution in $\ln(r)$ space. σ is defined in the paper as the geometric standard deviation, and the text shifts into referring to it as distribution width without comment.

Table 1: Range of H₂SO₄ composition

Ideally, when determining a statistical projection of the solution set, your ‘basis set’ would span the full set of expected conditions. At low temperatures and/or high H₂O concentrations, wt% H₂SO₄ can be 50% or lower (Bernath et al 2023, <https://doi.org/10.1016/j.jqsrt.2023.108520>). These would fall outside the chosen basis set.

line 110: Logically, 1 can infer

Traditionally, it would be written as “one can infer”

line 205: to determine how well to 2 matched

Typo. “...how well the two matched”?

Figure 7: Plots are being provided for VD, a quantity that has yet to be mentioned in the text and doesn’t get mentioned for some time.

Line 254: LUTS

LUTs

Line 268: Since the H₂SO₄ content of atmospheric aerosols is, ultimately, unknown we note that this situations adds an unknown element to the analysis

situations -> situation

Line 335: $r_1=75$ nm, $\sigma_1=1.45$, $r_2=310$ nm, $\sigma_1=1.05$

$\sigma_2=1.05$

Caption of Figure 15: $k_{520}/k_{1020} \leq 1.4$

$$k_{520}/k_{1020} \leq 1.4$$

Caption of Table 7: Table includes labels used to identify events in Figs. 18 and 19.

And Figures 15 and 16

Line 503: The southern hemisphere was influence by...

influence -> influenced.

Line 521: due to differing deposition rates

deposition -> sedimentation?

Line 523: Here, the plume of largest particles were centered at ≈ 20 km

were -> was (plume is singular).

Line 567: Overall the 2 records were in good agreement.

2 -> two