SAGE III on ISS Version 5.21 Release Notes

This statement applies to SAGE III/ISS Version 5.21 (Solar Level 1B, Solar Level 2, and Lunar Level 2) data products.

<u>Note on version number increment</u>:
Starting with the events on May 31, 2022, the version number reported in the data products is incremented from 5.20 to 5.21. This change reflects an upgrade of the operating system and software compilers used to generate the products. While the Level 1 and Level 2 algorithms remain the same, the underlying upgrades introduce perturbations to the processing stream which manifest as small product differences between the two versions. These differences have been determined to occur without bias or trend, and at magnitudes smaller than reported uncertainty for the vast majority of data points.

<u>Reader Software:</u>
Both IDL and Python reader software are available to access the products in the native format. New readers are required for this release. The product formats have changed slightly, so new readers are being provided. Readers for previous versions of SAGE III data are not compatible with this version. However, the new version of the IDL and Python data readers are backwards compatible with the previous version of the data product. Note that Version 5.21 product files in HDF format are not supported by the readers, though v5.1 HDF files are still supported.

<u>General Comments</u>:
This release contains the second official release of SAGE III/ISS products that are suitable for use in both validation and research studies for some data products (see the Data Product User's Guide for details). Version 5.21 replaces Version 5.1 and the public are invited to continue to make comments on the new release.

Vertical profiles of ozone, nitrogen dioxide (NO_2), and water vapor (H_2O) concentrations as well as multi-wavelength aerosol extinction coefficient are included in the solar Level 2 data product files. Three ozone profiles are available in this release of the solar products: a UVbased mesospheric product (i.e., "Ozone_Mes" in the product files) and two Chappuis-based products. One Chappius-based product uses a spectrally-focused spectral fitting retrieval (i.e., "Ozone_MLR") while the other uses a broad-spectrum retrieval scheme that is similar to that of SAGE II (i.e., "Ozone_AO3"). Composite ozone and retrieved temperature/pressure products are not included in the V5.21 data set. Vertical profiles of ozone, nitrogen dioxide (NO_2), and nitrogen trioxide (NO_3) concentrations are included in the lunar Level 2 product files. Chlorine Dioxide (OClO) from lunar occultation are not included in this release. The channel wavelengths used in solar and lunar retrievals are available in the Data Product User's Guide (DPUG).

Loss of Events
The mid-inclination orbit of the ISS periodically results in high solar beta angles ($|\beta|>60°$) that make solar occultation measurements impractical. Additionally, events are occasionally not acquired due to obstructions of the Sun or Moon by the ISS or its components. There are brief periods during which SAGE III measurements are not taken due to unfavorable ISS configuration or activities (e.g., abnormal orientation, Extravehicular Activity (EVA) or 'new'

space vehicle arrival). A space vehicle of one kind or another is always docked at the ISS and generally presents no harm to SAGE III.

A number of other special cases occur with varying frequency and may affect data quality. These include situations where 1) the pointing system (hexapod) was unable to move the instrument into the requested position for the event, 2) the contamination window was closed for the event, 3) the ISS time correction parameter was invalid, but could be corrected by interpolation, 4) large mechanical vibrations occurred, 5) a line-of-sight, structural blockage was detected and mitigated during the exoatmospheric portion of the event, 6) spectral calibration could not be performed (likely due to a blockage), and/or 7) a solar eclipse occurred during a portion of the event. Such cases are indicated in the product files by the issuance of a bit flag.

Broad Changes from v5.1 to v5.21:

As is common with any new version of data, there are a long list of minor changes and bug fixes, but there are also a few significant changes that broadly affect the data products. Those changes are listed here. Changes specific to an individual product are detailed in the Solar or Lunar product notes later.

Wavelength Map

We have revisited the CCD wavelength registration and spectral point spread functions (PSFs) using a combination of on-orbit spectral calibration data combined with the solar spectrum and retrieved product data compared with spectral model output (mostly for water vapor). The default wavelength map (i.e., central wavelength of each pixel column on the CCD) has changed slightly from the previous version. Perhaps more importantly, the PSFs have been reevaluated and are generally a little narrower through the visible and broader in the UV and near-IR than before. This has a small impact on the overall amounts of retrieved products with changes of less than a percent for ozone up to a few percent for water vapor.

O4 Spectroscopy

The spectroscopic data base for the oxygen dimer (O_4) was incorrectly utilized due to a preprocessing error in the SAGE retrieval code, resulting in anomalously high ozone values and anomalously low 520, 602, and 676 nm aerosol values below ~20 km in v5.1. This has been corrected in v5.21.

Smoothing

We have removed all vertical smoothing that was previously applied to the solar data products (most often a 1-2-1 smoothing scheme), making the retrieved vertical resolution of all data products ~0.75 km (reported on a 0.5 km grid). Naturally, this will degrade the precision of the as-reported profiles between versions 5.1 and 5.21. Users are encouraged to explore the tradeoff between precision and vertical resolution by smoothing the as-reported products depending upon their intended use for the data.

Altitude Registration of Meteorological Profiles

The reported profiles of temperature and pressure (not the retrieved profiles that are not part of this version) come directly from the MERRA2 Reanalysis data that is used for the SAGE retrieval algorithm. However, the MERRA2 data files report temperature, pressure, and geopotential height. The SAGE retrieval algorithm converts the geopotential height to geometric altitude, which is ultimately reported in the product data file. It was recently discovered (Wang et al., 2020) that this conversion used an older and less robust approximation for Earth's surface gravity that has been updated to the WGS84 model in v5.21.

NO2 Clearing

There was a bug in the retrieval algorithm prior to v5.21 where the contribution of NO_2 to the line-of-sight optical depths was not removed below the bottommost altitude where NO_2 was retrieved. This has been corrected for v5.21, with small impacts to ozone and wavelength-dependent impacts to aerosol (more noticeable at shorter wavelengths). Since the correction only affects data below the retrieved NO_2 profile, the impacts are typically seen only in the troposphere.

Event IDs

All SAGE data sets have provided some form of event identifier. For previous versions of SAGE III data, this took the form of an eight-digit integer where the first six digits were the orbit number and the last two digits represented the event type (e.g., sunrise/sunset = 10/20). However, the orbit number is actually a computed value, making it possible for the Event ID to change from one version number to the next if the pre-processing software is updated. This is the case for version 5.21. To ensure continuity going forward, we have created a new Event ID based on the date of the event, the sequence in which the event was collected on that day (e.g., the fifth event collected), and the event type in the form: YYYYMMDDEETT. The EE value increments for any event type (i.e., solar occultation, lunar occultation, limb scatter measurement, or calibration events). The TT indicator is no longer a number (making the new Event ID a string) and identifies the type of measurement: SR/SS for sunrise/sunset, MR/MS for moonrise/moonset, and LS for limb scatter. An important note is that the date in the new Event ID is based on the UTC time in the raw telemetry data file when the instrument began its data acquisition process such that the date in the Event ID may not match the date field in the product data file that is tied to the 20 km sub-tangent point. This helps to ensure that the new Event ID will never change during future reprocessing. The old orbit number-based event identifier is still listed in the product files and, to ensure continuity between version 5.1 and version 5.21, any old Event ID that has changed is forced to take on its v5.1 value. However, all future versions beyond v5.21 will not enforce this backwards continuity (we may remove the old Event ID entirely). This means the old Event IDs will be continuous between versions 5.1 and 5.21 and the new Event IDs will be continuous between version 5.21 and all future versions.

Event Processing

The SAGE retrieval algorithm requires external meteorological inputs (i.e., vertical profiles of temperature and pressure) to process the data and uses the MERRA-2 Reanalysis to do so. MERRA data are released in monthly intervals and so SAGE data are processed in monthly blocks by interpolating MERRA data to the SAGE event times and locations. MERRA-2 data files have a 3 hour temporal resolution, meaning that any SAGE events in the final 3 hours of the month cannot use interpolated meteorological data. In v5.1, the meteorological data was not properly utilized in this temporal window when a new month of data was processed. For v5.21, processing of the events in the final day of a new month will be delayed until the following month to maintain consistent use of the MERRA-2 monthly data files.

Solar Product Notes:

The v5.21 data products have been screened by the SAGE III/ISS team in order to remove failed events and/or specific product profiles. Most of these are due to platform-related issues such as blockages and severe platform disturbances but a few fail due to unknown instrument and/or retrieval algorithm anomalies. As a matter of practice only the most severe failed events are removed and some thought into filtering data for less egregious data anomalies may be necessary for the user. In v5.21, we are implementing an automated "QA" process for excluding data from release rather than the "by eye" method used for previous versions. We welcome any feedback from the public and request that potentially anomalous profiles be reported so that we can improve this process for future versions.

Ozone: The "MLR" and "AO3" v5.1 ozone products were thoroughly validated in Wang et al. (2020). While the MLR and AO3 profiles are generally in good agreement, the AO3 product has less noise in the upper stratosphere (above 40 km) and is thus the recommended ozone product. The SAGE III/ISS AO3 ozone data are of sufficient quality to be suitable for scientific research including trend studies. Differences between v5.1 and v5.21 resulted in only minor changes (<0.5% in the stratosphere from a small upward shift in altitude registration) and thus should not change that usage assessment.

The mesospheric (i.e., "Ozone_Mes") product has not changed appreciably from v5.1 and is still affected by out-of-band stray light, resulting in biases, and a potential correction has not yet been completed.

In the past, SAGE III has provided a composite ozone product. In version 5.21, this product is left blank as the SAGE Team feels that more thought into the construction of this product is required. It will not be released until the mesospheric ozone product is corrected.

Aerosol Extinction Coefficient: The v5.1 aerosol extinction coefficient products have already been used for a number of studies related to validation of other instruments (e.g., Kar et al., 2019; Rieger et al., 2019; Chen et al., 2020), documenting the impacts of recent volcanic eruptions and wildfires (e.g., Bourassa et al., 2019; Khaykin et al., 2020; Kloss et al., 2021) , and contributing to long-term studies of aerosol (e.g., Chouza et al., 2020; Kovilakam et al., 2020). Only one anomaly was noted in the literature, namely a negative bias in the aerosol channels coinciding/scaling with the ozone cross-sections (illustrated in Wang et al., 2020). This bias is most noticeable in the 520, 602, and 676 nm channels, while the remaining channels appear unaffected. A simple correction involves spectrally fitting the aerosol extinction coefficient spectrum at non-affected channels to interpolate to affected channels.

The v5.21 release introduces one significant change to the aerosol products. Since there are two stratospheric ozone products, there are two separate ways to compute aerosol. The first is as a residual from the MLR ozone calculation (i.e., removing all trace gas species' contributions from the line-of-sight optical depths) and the second is explicitly as part of the AO3 ozone calculation. Previous versions of the SAGE III data used the first method. However, the validated AO3 ozone product is now the recommended data to use for scientific studies so we thought it more consistent to use the second method. Additionally, preliminary analyses show the v5.21 AO3-based aerosol product to be very similar to the v5.1 MLR-based aerosol product, with a small mitigation of the negative biases in the lower stratosphere in the mid-visible wavelength channels.

Nitrogen Dioxide: The v5.1 NO_2 products have undergone some preliminary validation in Dube et al. (2020). The largest change between v5.1 and v5.21 comes from the new wavelength map, resulting in a ~5% decrease in overall NO_2 in the stratosphere.

Water Vapor: The v5.1 vertical profiles of H_2O concentration have been thoroughly validated in Davis et al. (2021) and Park et al. (2021), showing good agreement with other instruments as well as capturing patterns of stratospheric variability. A few anomalies noted in those studies have been mitigated and/or improved through retrieval algorithm changes. These changes include: 1) improved stability resulting in a significant reduction (though not elimination) of failed water vapor retrievals with large negative values as well as "keel-over" profiles; 2) switching from retrieving on a 1 km grid that is later interpolated to a 0.5 km grid to retrieving on a 0.5 km grid; and 3) fixing a bug in the choice of CCD pixel groups used for the retrieval that caused a noticeable sensitivity of the water vapor retrieval to enhanced aerosol loading. In addition to these improvements, the changes in the PSFs resulted in a roughly 0.3 ppm increase in stratospheric water vapor. With the combination of previous validation studies and improvements to the data, v5.21 water vapor should be considered suitable for scientific studies.

Lunar Product Notes:

Ozone: The vertical profiles of ozone concentration have undergone some preliminary validation and the show generally good agreement with other instruments. The biggest change between v5.1 and v5.21 is the impact on the O_2-based altitude registration offset. Ozone derived from lunar measurements appeared to have a positive altitude bias in v5.1 that has now been improved in v5.21.

Nitrogen Dioxide: The vertical profiles of NO_2 concentration are a research product with limited evaluation. While individual profiles have significant noise, averaging multiple profiles (e.g., over ~24 hours) to improve the precision shows lunar NO_2 behaves similar to the solar NO_2 product.

Nitrogen Trioxide: The vertical profiles of NO_3 concentration are a research product with the user cautioned on the use of this data. This product tends to be rather noisy and, under most conditions, requires substantial averaging to produce a meaningful profile.

References:

Bourassa et al.: Satellite limb observations of unprecedented forest fire aerosol in the stratosphere, JGR, doi:10.1029/2019JD030607, 2019.

Chen et al.: Evaluation of the OMPS/LP stratospheric aerosol extinction product using SAGE III/ISS observations, AMT, doi:10.5194/amt-13-3471-2020, 2020.

Chouza et al.: Long-term (1999–2019) variability of stratospheric aerosol over Mauna Loa, Hawaii, as seen by two co-located lidars and satellite measurements, ACP, doi:10.5194/acp-20-6821-2020, 2020.

Davis et al.: Validation of SAGE III/ISS solar water vapor data with correlative satellite and balloon-borne measurements, JGR, doi:10.1029/2020JD033803, 2021.

Dubé et al.: Accounting for the photochemical variation in stratospheric NO_2 in the SAGE III/ISS solar occultation retrieval, AMT, doi:10.5194/amt-14-557-2021, 2021.

Kar et al.: CALIPSO level 3 stratospheric aerosol profile product: version 1.00 algorithm description and initial assessment, AMT, doi:10.5194/amt-12-6173-2019, 2019.

Khaykin et al.: The 2019/20 Australian wildfires generated a persistent smoke-charged vortex rising up to 35 km altitude, Commun Earth Environ 1, doi:10.1038/s43247-020-00022-5, 2020.

Kloss et al.: Stratospheric aerosol layer perturbation caused by the 2019 Raikoke and Ulawun eruptions and their radiative forcing, ACP, doi:10.5194/acp-21-535-2021, 2021.
Kovilakam et al.: The Global Space-based Stratospheric Aerosol Climatology (version 2.0): 1979–2018, ESSD, doi:10.5194/essd-12-2607-2020, 2020.
Park et al.: Near-global Variability of Stratospheric Water Vapor observed by SAGE III/ISS, JGR, doi:10.1029/2020JD034274, 2021.
Rieger et al.: A multiwavelength retrieval approach for improved OSIRIS aerosol extinction retrievals, JGR, doi:10.1029/2018JD029897, 2019.
Wang et al.: Validation of SAGE III/ISS solar occultation ozone products with correlative satellite and ground based measurements, JGR, doi:10.1029/2020JD032430, 2020.