Intercomparison of Pandora Stratospheric NO_2 Slant Column Product with the ~~NIWA~~ NDACC-Certified M07 ~~NDACC Standard~~Spectrometer in Lauder, New Zealand

T. N. Knepp[1,2], R. Querel[3], P. Johnston[3], L. Thomason[2], D. Flittner[2], and J. M. Zawodny[2]

[1]Science Systems and Applications Inc., Hampton, VA, 23666 United States
[2]NASA Langley Research Center, Hampton, VA, 23681 United States
[3]National Institute of Water and Atmospheric Research, Lauder, Central Otago, New Zealand

Correspondence to: T. N. Knepp
(travis.n.knepp@nasa.gov)

Abstract. In September 2014, a Pandora multi-spectral photometer operated by the SAGE-III project was sent to Lauder, New Zealand to operate side-by-side with the National Institute of Water and Atmospheric Research's (NIWA) Network for Detection of Atmospheric Composition Change (NDACC) ~~standard~~ certified zenith slant column NO_2 instrument to allow intercomparison between the two instruments, and for evaluation of the Pandora unit as a potential SAGE-III validation tool for stratospheric NO_2. This intercomparison spanned a full year, from September 2014 – September 2015. Both datasets were produced using their respective native algorithms using a common reference spectrum (i.e. 12:00 on 26-February 2015). Throughout the entire deployment period both instruments operated in a zenith-only observation configuration. Though conversion from slant column density (SCD) to vertical-column density (VCD) is routine (by application of an air mass factor), we limit the current analysis to SCD only. This omission is beneficial in that it provides ~~a strict intercomparison of~~ an intercomparison based on similar modes of operation for the two instruments and the retrieval algorithms as opposed to introducing an ~~AMF-dependence~~ air mass factor dependence in the intercomparison as well. It was observed that the current hardware configurations and retrieval algorithms are in good agreement (R > 0.95). The detailed results of this investigation are presented herein.

1 Introduction

The Stratospheric Aerosol and Gas Experiment (SAGE) missions have provided a legacy of high-quality solar occultation measurements ~~of stratospheric ozone and aerosol for over 3~~ for vertically profiling stratospheric O_3 and UV/VIS/NIR aerosol extinction coefficients from the upper troposphere into the mesosphere for more than three decades (**????**). ~~SAGE-III/Meteor, operated aboard the Russian Meteor-3M platform between February 2002 and March 2006.~~ These observations have formed a crucial component for understanding ozone trends, and how stratospheric chemistry and aerosol influence ozone mixing ratios and climate. An updated version of ~~this instrument~~ the SAGE instrument (hereafter referred to as SAGE-III) was integrated into the International Space Station (ISS) in March 2017 with routine observations ~~expected to start~~ starting in April. The SAGE-III project will focus on reassessing the state of stratospheric O_3 recovery and provide requisite aerosol observations for climate and ozone models. To this end, the standard data products for this mission are aerosol extinction coefficients, aerosol optical depth, O_3, H_2O, and NO_2 mixing ratios. For an overview of the instrument and products see **?**.

As with any new instrument, a significant post-launch activity is planned to validate the accuracy and precision of the data products, and provide validated datasets to end users. While the key SAGE-III species measurements are validated using well-known and characterized instruments, one important product remains difficult to measurevalidate: NO_2. NO_2 is important due to its role in partitioning stratospheric odd hydrogen, providing a chemical pathway for conversion of ozone-destroying species to their reservoir forms (e.g. halogen species as discussed by **?**), and may be responsible for up to 70 % of the stratospheric ozone loss (**?????**). The quality of the NO_2 retrievals also impacts the quality of short wavelength aerosol extinction coefficient measurements as well as, to a lesser extent, ozone.

Observations are made over a large range of latitudes depending on season and the details of the orbit but only at two latitudes on a given day (where the spacecraft crosses the terminator or (given the question) each sunrise and sunset encountered by the spacecraft (one of each per orbit). Due to the unique viewing geometry of SAGE-II and the rapid variability of NO_2 across the solar terminator, NO_2 measurements from previous SAGE missions (SAGE-II and SAGE-III/Meteor) proved to be challenging. For SAGE-III/Meteor, NO_2 is often validated using measurements from other space-based instruments that generally do not fully match the SAGE viewing geometry, location and/or time. While a chemistry model can correct for some of these differences, generally these comparisons leave significant questions regarding the NO_2 data quality. Given the variability and relative sparsity of observations, Pandora provides a unique capacity to be carried to a measurement location rather than only providing data when an observation occurs near a fixed site. This enables observations from places that are challenging for the SAGE instrument particularly where strong gradients across the tropopause may occur (like the tropics) or other observations of opportunity (i.e. various field campaigns).

An alternative method that provides some corroboration to the SAGE-III measurement quality is comparison with ground-based Differential Optical Absorption Spectroscopy (DOAS, e.g. **?**) or Fourier Transform Spectroscopy (FTS, e.g. **?**) measurements of the column NO_2 using zenith-looking instruments that measure scattered light across the ultra-violet and visible wavelengths. These observations can be used to infer, among other species, column NO_2 as a function of solar zenith angle (SZA). Zenith-viewing observations when SZA $\approx 90°$ are analogous to solar occultation measurements of NO_2. However, observation of stratospheric NO_2 is challenging at many locations due to the preponderance high levels of tropospheric NO_2 from anthropogenic human-derived sources. Therefore, measurement sites that are in locations that are considered "background-level" are advantageous.

The National Institute of Water and Atmospheric Research (NIWA) Lauder, New Zealand site provides the desired required tropospheric background-level conditions for observation of stratospheric NO_2. The Lauder group has a long history of providing high-quality observations (**??**), and is generally considered a standard for stratospheric NO_2 observationsand O_3 (**??**). Data collected at the Lauder site have been used to infer data quality for SAGE II NO_2 and were used to identify and help correct a time-dependent error in those observations (**?**). For the new SAGE-III mission, observations by the NIWA instrument will be useful for understanding NO_2 data quality. However, since the challenges of making space-based measurements is often latitude dependent, a single site will not provide all the corroborative data needed to make a robust assessment of data quality.

As a result, the SAGE-III group has acquired a Pandora unit (**??**) with the hope of using it as a portable system for providing corroborative data that can be deployed at sites of opportunity, for instance low latitudes, throughout the SAGE-III/ISS mission. To date, Pandoras have not established a record for measuring NO_2 where the column is dominated by the stratosphere rather than a polluted troposphere so an evaluation of the capabilities of this instrument in this regard is necessary. Herein, we report the results of a comparison of observations by ~~the NIWA M07~~ a NIWA owned/operated instrument and the SAGE-III Pandora unit when operated side-by-side between September 2014 and September 2015 at the NIWA facility in Lauder, New Zealand.

2 Instrumentation

2.1 LaRC Pandora

Pandora is a sun-viewing spectrometer that was initially developed for validation of the Ozone Monitoring Instrument (OMI) aboard the Aura satellite (**???**), and has proven to be sensitive to fluctuations in boundary layer NO_2 over short time periods (**?**). Due to Pandora's potential for retrieving stratospheric gas column densities (i.e. operating in zenith orientation during twilight hours) it has been evaluated as a potential validation instrument for the SAGE-III mission.

A detailed description of the instrument has been provided by **?**. Briefly, the Pandora model used in the current study consisted of: 1. an optical head (mounted on a two-axis tracker capable of moving through 360° azimuth and ~~180~~90° zenith) containing filter wheels for controlling polarization and radiant flux; 2. a single-strand, multi-mode fiber-optic cable with 400 μm core diameter and numerical aperture of 0.22 to transmit photons to the spectrometer; 3. a temperature stabilised Avantes spectrometer (model number ULS2048x64, 280 – 525 nm) with a 50 μm slit, focal length of 75 mm, and resolution on the order of 0.6 nm; 4. laptop computer for instrument control and data logging. The improved optics and spectrometer of this model enabled the instrument to record solar spectra from lunar reflectance and scattered radiation, ~~thereby enabling acquisition of twilight spectra for estimating~~ which has spurred investigation regarding its ability to accurately estimate the slant-column density (SCD) ~~stratospheric component and potential for vertical profiling of select species~~ of stratospheric species from twilight spectra.

The Pandora retrieval algorithm was previously described in **?**~~and **?**~~, **?** and **?**, with relevant cross-section details presented in Table ~~??~~1. Briefly, spectral fitting is performed using laboratory-measured absorption cross sections and implement shift-squeeze functions to fit the observed spectra with the solar reference spectrum's Fraunhofer line structure (for zenith observations an instrument-observed solar-reference spectrum was used from the spectrum recorded at 12:00 (local time) on 26-February 2015), with a fourth-order polynomial applied for removal of aerosol and Rayleigh scattering effects.

Though Pandora was developed to operate in a Sun-tracking mode and has undergone numerous revisions to allow data collection in sky (i.e. scattered irradiance for elevation scans) and moon observation modes, ~~in this study the Pandora~~ the instrument's capability of making accurate twilight observations remained unknown. Part of the motivation of the current study was to evaluate Pandora's ability to make reliable twilight observations, thereby demonstrating its applicability to SAGE-III validation. To this end, Pandora only operated in the zenith-observation mode to allow direct intercomparison with the zenith-oriented NIWA instrument.

~~Gas~~	Instrument	~~Cross Section T (K) λ Range (nm)~~ Setting
O_3 Cross Section	Pandora	**??** (225 K, 300 – 330 nm)
	M07	**?** (218 ~~300~~ K, 428 – ~~330~~ 469 nm)
NO_2 Cross Section	Pandora	**?** (220 K, 400 – 485 nm)
	M07	**?** (220 ~~400~~ K, 428 – ~~485~~ 469 nm)
O_4 Cross Section	Pandora	**?** (262 K, 400 – ~~454.42~~ 454 nm)
	~~Pandora~~ M07	~~?~~ **?** (262 ~~454.43~~ K, 428 – ~~485~~ 469 nm)
Ring	~~M07~~ Pandora	~~?~~ **?**
	~~262~~ M07	~~400 – 454.42~~ NDACC recommended pseudo cross section **?**
Polynomial Order	Pandora	4^{th}
	M07	3^{rd}

Table 1. Relevant ~~cross-section~~ retrieval details for the two instruments under study.~~Species that required multiple cross-section sources to cover the required wavelength range (e.g. O_4) have multiple cross sections listed and their respective wavelength domains.~~

2.2 NIWA spectrometer

The NIWA instrument (M07) is a zenith-oriented instrument used for measuring stratospheric slant column NO_2. M07 is the current instrument contributing to the continuous time series of stratospheric NO~~2~~ $_2$ from Lauder that started in 1980, and is part of the Network for Detection of Atmospheric Composition Change (**??**). The NDACC-certified M07 instrument has been described previously (**???**) as has the STRATO retrieval software that was build in-house (**?**). Briefly, M07 is a Czerny-Turner monochromator (320 mm focal length, ≈0.8 nm resolution, F/5 entrance field of view, 1 mm wide slit) with a bi-alkali photocathode photomultiplier detector. The scanning mechanism ~~has been~~ was modified to provide fast scanning with a long lifetime and smooth wavelength motion. The instrument is mounted in a temperature controlled cabinet on a rotating table following the line of the sun-zenith plane and a Glan-Thompson polariser is used in front of the entrance ~~silt~~ slit to provide polarised zenith measurements. Similar to the Pandora, the cross sections used for retrievals are listed in Table ~~??~~1.

2.3 Uncertainties

Within the scope of the current work, the dominant source of uncertainty for each instrument, during twilight conditions, was statistical uncertainty due to limited light throughput. In this regard, Pandora is inferior to M07 (*vide supra*) as it was initially designed for direct-sun observations. Other sources of uncertainty that have less impact within the current analysis are slant-column amount in the chosen reference spectrum, fitting settings such as NO_2 temperature, and retrieval technique.

3 Mode of operation and location

The Pandora unit was deployed to the NIWA station in Lauder, Central Otago, New Zealand (45.038 S, 169.68 E, 370 m ASL) to run side-by-side with the NIWA-operated M07 spectrometer. Both instruments performed retrievals using a common reference spectrum (collected on 26-February, 2015 12:00 local time) as observed by the respective instruments. It is worth noting that, other than the Pandora's fixed zenith observation state~~)~~, both instruments were operated in their normal states, not in a customized operation mode, and both used their standard retrieval algorithms.

New Zealand is generally an atmospherically clean environment, with pollution levels that can be considered as background level (e.g. approximately an order of magnitude below urban centers in the continental United States). As a point of reference, NO_2 retrievals (VCD) over New Zealand and the continental United States from the OMI (Level 3, version 3 algorithm) are presented on the same scale in Figs. 1 and 2. It is observed that aside from some western states (e.g. Nevada and Oregon), the U.S. rarely experiences ~~similar~~ similarly clean conditions as New Zealand. Furthermore, Fig. 2 displays a downward trend in overall column NO_2 over the Chemistry and Physics of the Atmospheric Boundary Layer Experiment (CAPABLE) station located at NASA's Langley Research Center in Hampton, VA (37.103 N, -76.387 E, 5 m ASL) that is being driven by a decreasing tropospheric column, while NO_2 over Lauder has remained consistent since 2005. There is no corresponding change in stratospheric NO_2 for either site.

Statistics describing the variability in NO_2 over both sites were broken into three categories (total-column, stratospheric and tropospheric contributions) and are presented in Table 2. The statistics presented in Table 2 are similar for both sites when scaled according to corresponding column density (i.e. despite the total-column standard deviation being significantly different for both sites the relative error ($\sigma/\overline{x}$) remains similar). Despite these similarities, the tropospheric variability remains different for the two sites indicating a higher degree of variability over the continental US. The tropospheric contribution and variability remains significantly higher over CAPABLE as compared to Lauder with approximately 55.7 % of the NO_2 column residing in the troposphere over the CAPABLE site, and only 13.5 % over Lauder. These differences are driven by the ubiquity of local sources in the eastern United States as compared to central New Zealand.

Figures 1 and 2 and Table 2 demonstrate that not only is the NO_2 column significantly higher over the continental US, but the Lauder column is dominated by the stratospheric contribution. One effect of differing source strengths is seen in Fig. 2. Being in different hemispheres, the two sites should be approximately six months offset in their seasonal cycle, though the total-column time series shows the two sites are in phase. However, the stratospheric contribution for the two sites (panel ~~b~~(b)) remained approximately six months out of phase as expected. This can be explained by the difference in tropospheric and stratospheric chemistry. Under normal, moderately-polluted conditions, tropospheric chemistry is sufficiently perturbed to force the tropospheric NO_2 column density out of phase with the stratosphere. This shows up in the data as an approximately six month offset. Since the northern hemisphere site is dominated by tropospheric NO_2 (as shown in Fig. 2 and Table 2), and the southern hemisphere site has significantly less tropospheric contribution, the two sites ~~were~~ are in phase with one another in the total-column panel. Nitrogen dioxide SCDs from the ground instruments (Fig. 2, panel (b)) show the surface instruments are accurately detecting the stratospheric seasonal cycle (i.e. are in phase with the OMI stratospheric column over Lauder).

Since Lauder provides a clean, background-level, environment with few local or regional anthropogenic emission sources, it provides ideal conditions for observation of stratospheric species and evaluation of the Pandora system for detecting stratospheric NO_2 and as a possible validation tool for current and upcoming satellite missions that focus on stratospheric chemistry.

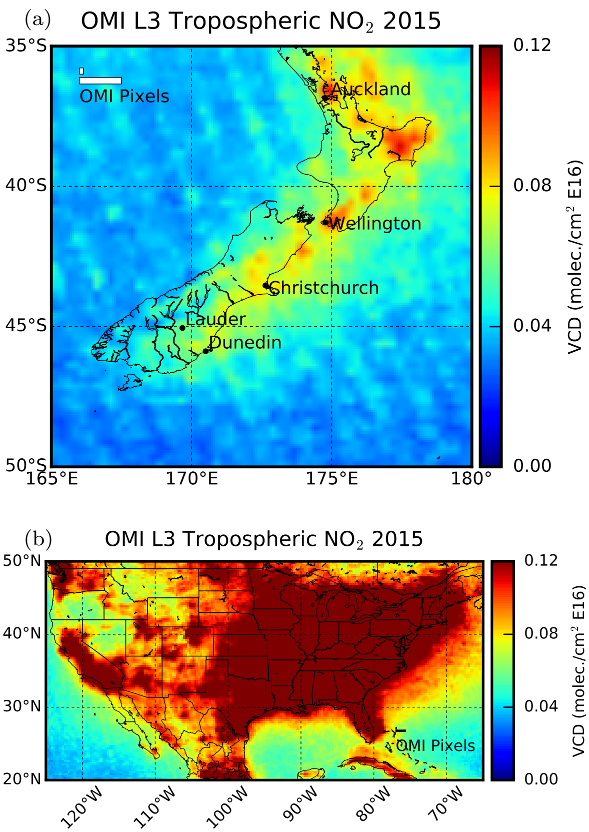

Figure 1. Annual average for OMI NO_2 (L3, v3.0) maps over New Zealand and North America. OMI pixel sizes (nadir and swath edge) are represented by the white boxes within each panel. For comparison purposes, both plots were put on the same color scale.

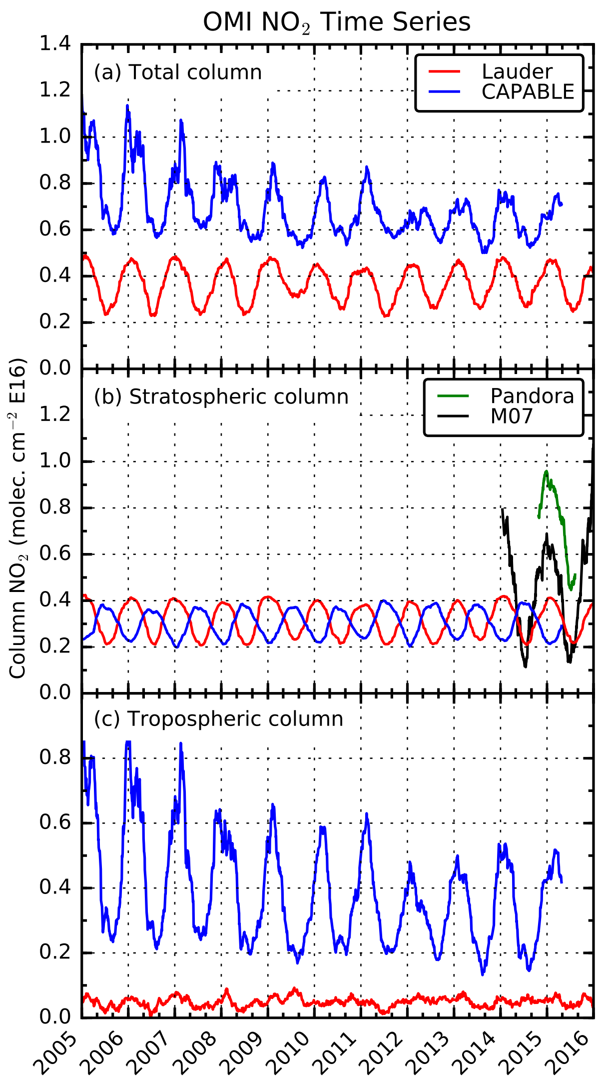

Figure 2. ~~Time-series~~ Time series plots for total- (a), stratospheric- (b), and tropospheric-column (c) NO_2 data products from OMI (L3, v3.0, VCD) over Lauder (red) and CAPABLE (blue). OMI data were filtered to remove cloud fractions greater than 20 % and overpasses greater than 50 km from the site. A seven-day normally-weighted rolling mean was applied to smooth the plots and remove higher-frequency fluctuations. Pandora (green) and M07 (black) data presented in panel (b) are slant-column densities.

4 Intercomparison

Pandora and M07 data were filtered to remove points where the retrieval uncertainty was greater than 10 % of the retrieved value followed by resampling to five-minute means to allow direct, temporally-aligned, intercomparisons. ~~For plotting purposes,~~

Parameter	Lauder	CAPABLE
Total-column ($\overline{x}$)	0.37	0.70
Total-column (σ)	0.08	0.14
Total-column ($\sigma/\overline{x}$)	0.21	0.19
Stratosphere ($\overline{x}$)	0.32	0.31
Stratosphere (σ)	0.07	0.06
Stratosphere ($\sigma/\overline{x}$)	0.21	0.19
Stratospheric Fraction (%)	86.5	44.3
Troposphere ($\overline{x}$)	0.05	0.39
Troposphere (σ)	0.01	0.17
Troposphere ($\sigma/\overline{x}$)	0.29	0.45
Tropospheric Fraction (%)	13.5	55.7

Table 2. Statistics regarding stratospheric and tropospheric contribution and variability of NO_2 (VCD) as observed by OMI. All units are molec cm^{-2} E16.

some datasets were smoothed via a rolling mean (e.g. to remove higher-frequency fluctuations). Unless otherwise noted, all intercomparisons and analyses were carried out using 5 min averaged data.

4.1 Aggregate analysis

An aggregate analysis was performed on the resampled data by binning the SCD according to SZA for a visual evaluation of correlation the correlations as seen in Fig. 3. It is observed that the correlation is generally poor during pre-sunrise/twilight hours (i.e. when SZA>92.5°), but improves with decreasing SZA where it peaked peaks at 80-85° . Within each sub-panel (Table 2). At 95 % confidence, all R^2 values in Table 2 are significantly different except when comparing the 87.5-90.0° bin with either the 85.0-87.5° or 80.0-85.0° bins; however, the 80.0-85.0° and 87.5-90.0° bins remain statistically different. Within each panel of Fig. 3 the data were are color coded to correspond to the SZA range within each sub-panel and provide insight into how the short-term change in SZA influenced correlationagreement. As an example, in panel a (a) it is observed that data collected at higher SZA (red-shaded points) were further from the one-to-one line than data collected at lower SZA (blue-shaded points). Analogously, it is observed in panels (e-g) that as SZA decreased, so too did the degree of correlation. Therefore, we can conclude that the Sun's zenith angle played a role in the degree of agreement between the two instruments, though this dependence cannot be separated from day-to-day chemical variability.

It remains clear that the two instruments had have strikingly good agreement for zenith angles greater than 70°, as supported by Fig. 3 and Table 3. Below 70° the correlation dropped rapidly (by almost fifteen percentage points between bins). However, when considering data collected within the SZA range most relevant to stratospheric retrievals (i.e. (8085,92.5], see Table 4) the mean percent difference remained below 10 %, with R^2 >0.95. From a satellite-validation perspective, this bodes well for

N	Pandora ($\overline{x}$)	M07 ($\overline{x}$)	Diff. ($\overline{x}$)	Diff. (σ)	% Diff.	Ratio ($\overline{x}$)	Slope	Intercept	R^2	SZA Range
1135	6.488	11.648	-5.160	4.00	44.3	0.602	0.226	3.860	0.128	(92.5,95.0]
848	7.072	7.933	-0.861	0.915	10.9	0.907	0.820	0.569	0.933	(90.0,92.5]
1208	4.470	4.767	-0.297	0.463	6.2	0.955	0.866	0.340	0.955	(87.5,90.0]
752	2.914	3.025	-0.111	0.289	3.7	0.991	0.850	0.344	0.958	(85.0,~~90.0~~87.5]
1660	1.895	1.855	0.040	0.199	2.2	1.076	0.841	0.336	0.960	(80.0,85.0]
1260	1.223	1.088	0.135	0.149	12.4	1.214	0.787	0.367	0.920	(75.0,80.0]
235	1.003	0.848	0.155	0.129	18.3	1.289	0.737	0.378	0.899	(70.0,75.0]
151	0.806	0.649	0.157	0.108	24.3	1.323	0.702	0.351	0.752	(65.0,70.0]

Table 3. Summary of aggregate statistics for the Pandora/NIWA intercomparison using the standard algorithms and parameters for each instrument. Differences and ratios are relative to the NIWA instrument (i.e. Pandora - M07, Pandora/M07). Statistics were generated using data that were resampled to 5 min means; no further smoothing or binning was applied.

~~Season~~ Statistic
~~Summer~~ N
~~Winter~~ SZA Mean
~~Summer~~ SZA Stdev
~~Winter~~ SZA Min
~~Summer~~ SZA Max
~~Winter 284 3.139 3.103 0.036 0.239 1.2 1.025 0.932 0.248 0.960 Summer 469 3.138 3.340 -0.202 0.271 -6.1 0.952 0.870 0.233 0.952 Winter 167 1.994~~

Table 4. ~~Summary of seasonal~~ Solar zenith angle statistics ~~for~~ from all (4072) SAGE-III overpasses between 16-March 2017 and 12-August 2017. Solar zenith angles were calculated with respect to a surface instrument's viewing geometry based on the ~~Pandora~~SAGE-III observation time and surface latitude/~~NIWA intercomparison~~longitude.~~Similar to Table 3.~~

future validation efforts of stratospheric NO_2 as >95 % of the inter-instrument variability is accounted for without further correction at zenith angles most relevant to stratospheric observation.

The ~~decreased~~ decrease in correlation at lower SZA's (i.e., as the Sun approaches solar noon) was driven by an apparent offset in the Pandora retrieval at lower SCD's (≈0.55x10^{16} molec cm^{-2}) where Pandora seems to lose sensitivity and accuracy, as seen in panels (e-h) of Fig. 3. A similar "tailing" effect due to decreased sensitivity was observed by **?** when comparing Pandora NO_2 VCD's to in-situ observations and is likely due to the instrument's accuracy limit ~~(≈0.3x10^{16} molec cm^{-2} as stated in (?))~~and light throughput. Therefore, 0.55x10^{16} molec cm^{-2} is considered to be the lower-limit of detection for the current instrument. However, due to the M07's larger slit width and longer focal length, it has more throughput and greater sensitivity than the Pandora~~. This allows~~, thereby allowing the M07 to continue making reliable measurements up to SZA of 95°.

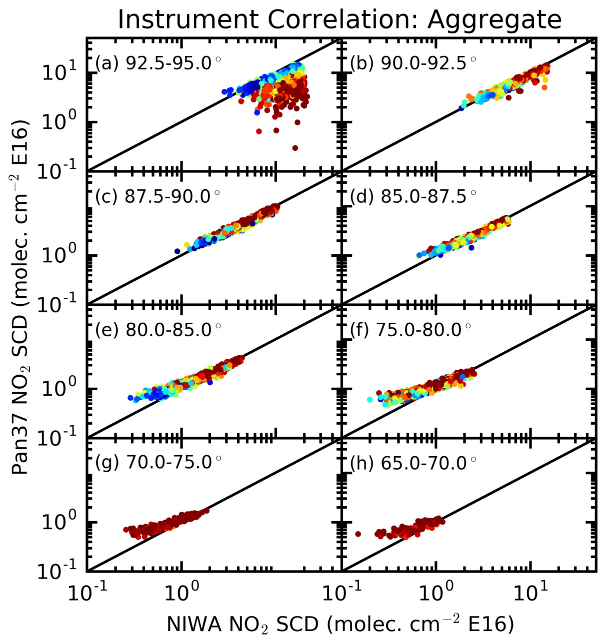

Figure 3. Correlation plots for data collected by the Pandora and M07 instruments. Data were resampled to five-minute averages, and color coded according to SZA within the specified bin range (i.e. red colors represent the ~~upper-SZA~~ upper SZA limit, blue represent the ~~lower-bounds~~ lower bounds for each sub-panel). Values within the figure legends indicate SZA ranges.

4.2 Seasonal dependence

To better understand seasonal variability seen within the datasets, the data were broken into two major seasons: austral summer (DJFM) and austral winter (JJAS). Seasonal-correlation plots were generated (Fig. 4)~~, and showed~~ ; they show nearly identical behavior ~~as~~ to the aggregate (Fig. 3) ~~, though~~ with most of the tailing behavior ~~was isolated~~ being attributed to winter conditions ~~.~~ in agreement with the seasonal cycle depicted in Fig. 2.

SCD and statistical time series plots (Fig. 5) were generated to evaluate the seasonal dependence of both instruments and the inter-instrumental statistics over the year-long operation period. The SCD time series was generated by first binning the data by SZA followed by calculating daily means, which were then smoothed via a 7-day rolling mean. Statistical time series presented in panels (e-af) were generated by resampling each dataset to 5 min averages (i.e., forcing both datasets onto a common date/time index) followed by calculation of the specified statistic on a day-by-day basis, which was then smoothed via a 7-day rolling mean.

Both instruments displayed the expected diurnal (elevated at large SZA, reduced at smaller SZA) and seasonal (elevated NO_2 in austral summer (DJFM), followed by reduced levels in winter (JJAS)) trends in NO_2 SCD (see also Table 5). This is in agreement with the expected diurnal behavior (**?**) and the observed satellite seasonality (**?** and Fig. 2).

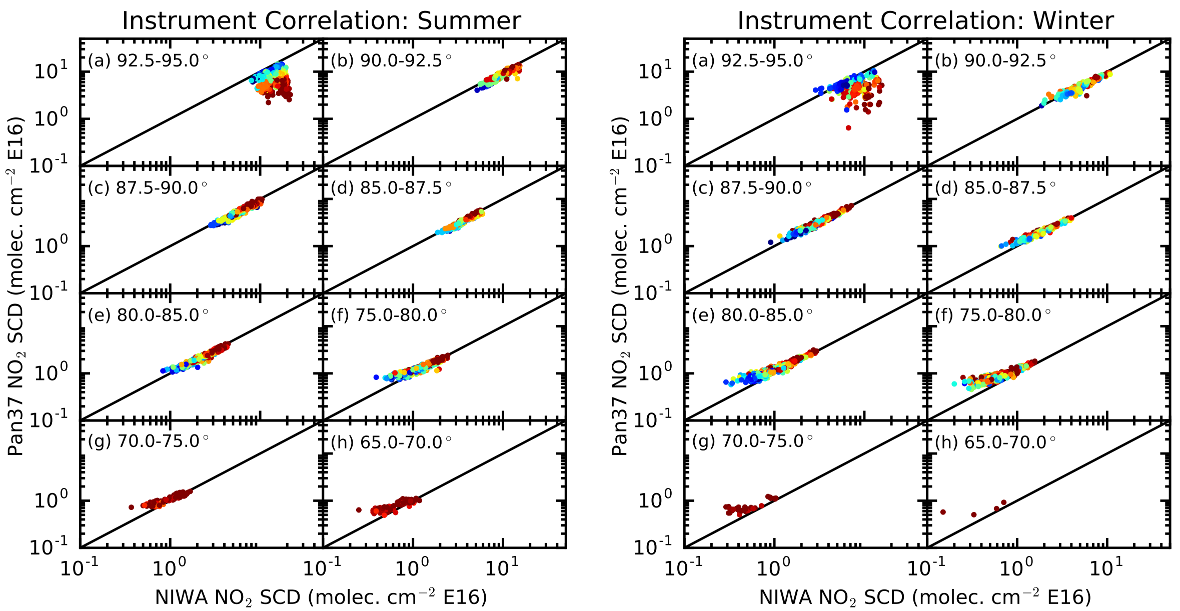

Figure 4. Correlation plots for data collected by the Pandora and M07 instruments broken into austral summer/winter. Similar to Fig. 3, data were resampled to five-minute averages, and color coded according to SZA within the specified bin range

Inter-instrumental statistics and seasonal dependence were further evaluated. It was observed that the two products tended to have good agreement throughout the year (generally with ±10 %, see Table 5 and Fig. 5 panels (m-p)), with maximal differences at ~~extreme~~ high SZAs (i.e. >2.5° below the horizon, panel ~~a) or at extreme low~~ (a)) or at very low NO_2 (i.e. below the Pandora's sensitivity cutoff, as demonstrated in the tailing behavior of Fig. 3). Further, no seasonal dependence on R^2 was observed as R^2 remained high throughout the year (>95 %, Table 5 and Fig. 5 panels (m-p)).

Other statistics presented in Table 5 and Figure 5 show a slight seasonal dependence in the measured values. An interesting seasonal and SZA dependence was observed in the ratio and slope data in Table 5 in that the wintertime ratios and slopes were always larger than their summertime counterparts (excluding the pre-sunrise data), and can be most clearly seen in the ratio and difference data in Fig. 5. Ideally, the ratio and inter-instrument offset would remain constant regardless of season, though this was not observed. What is observed is a disproportionate increase in the Pandora-measured SCD (i.e. increasing difference and ratio) from summer to winter compared to M07. Even after removing data where SCD<$1x10^{16}$ the wintertime ratios remain disproportionately high (not shown), therefore this cannot be attributed to Pandora's low-SCD trailing affect. While the source of this seasonal dependence remains unknown, the observed seasonality changed slow enough that the correlations and regressions were not significantly influenced.

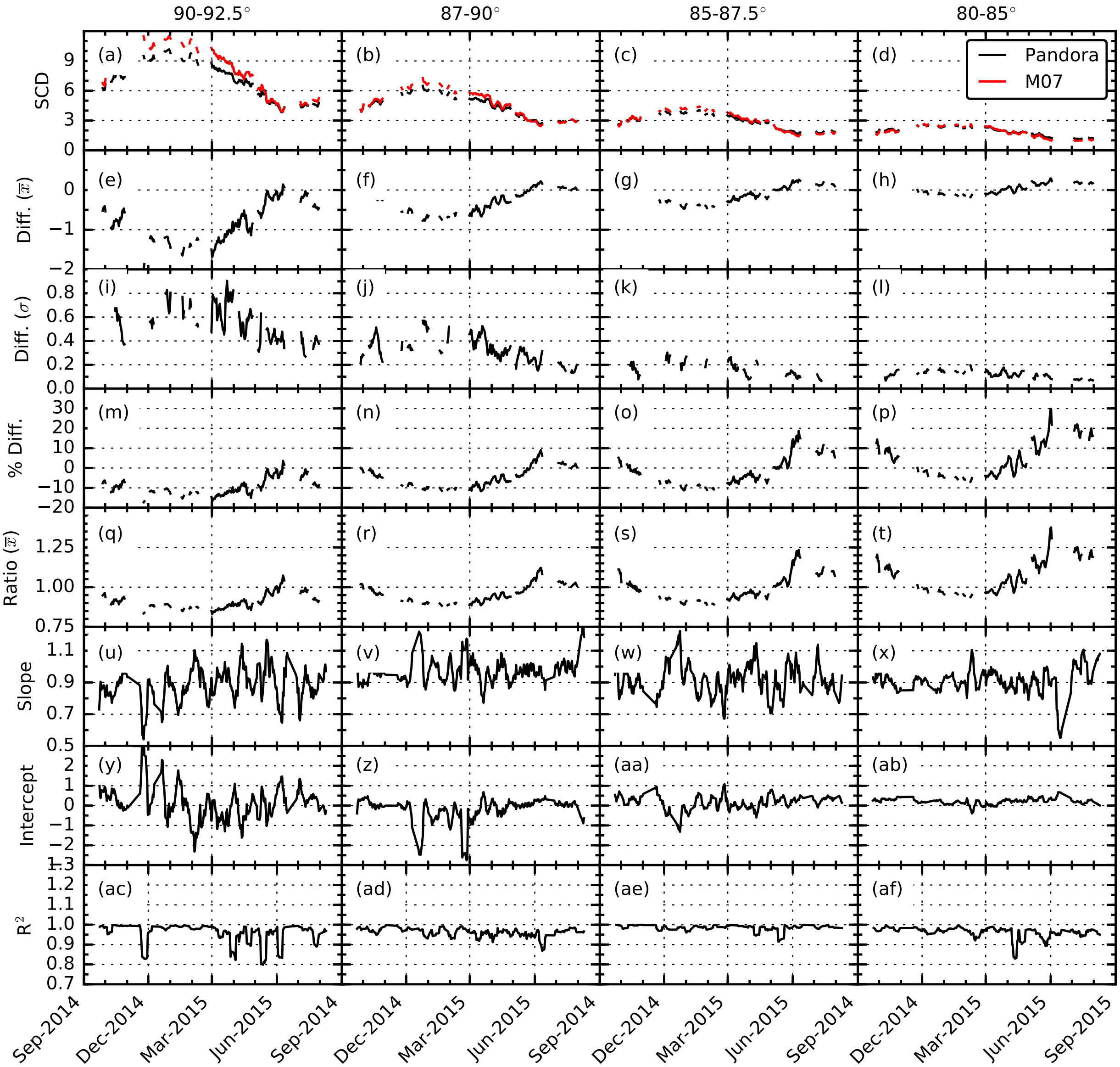

Figure 5. Time series for NO_2 SCD and daily statistics binned by solar-zenith angle. Data were smoothed by a seven-day rolling mean. Panel descriptions: a-d: SCD for both instruments broken up by SZA; e-h: mean SCD difference (Pandora - M07); i-l: standard deviation of differences; m-p: percent SCD difference; q-t: SCD ratio (Pandora/M07); u-x: line of best fit slope (Pandora vs. M07); y-ab: line of best fit intercept; ac-af: R^2 coefficient of correlation. A seven-day normally-weighted rolling mean was applied to smooth the plots and remove higher-frequency fluctuations.

Season	N	Pandora ($\bar{x}$)	M07 ($\bar{x}$)	Diff. ($\bar{x}$)	Diff. (σ)	% Diff.	Ratio ($\bar{x}$)	Slope	Intercept	R^2	SZA Range
Summer	694	6.957	12.735	-5.778	4.056	-45.4	0.582	0.181	4.651	0.071	(92.5,95.0]
Winter	258	4.972	8.129	-3.157	2.716	-38.8	0.660	0.140	3.835	0.045	
Summer	513	7.744	8.795	-1.051	0.880	-11.9	0.888	0.831	0.440	0.926	(90.0,92.5]
Winter	202	4.839	5.124	-0.285	0.577	-5.6	0.961	0.829	0.593	0.917	
Summer	729	4.837	5.269	-0.432	0.477	-8.2	0.928	0.870	0.252	0.947	(87.5,90.0]
Winter	284	3.139	3.103	0.036	0.239	1.2	1.025	0.932	0.248	0.960	
Summer	469	3.138	3.340	-0.202	0.271	-6.1	0.952	0.870	0.233	0.952	(85.0,87.5]
Winter	167	1.994	1.849	0.145	0.162	7.8	1.106	0.900	0.330	0.953	
Summer	1044	2.031	2.050	-0.020	0.198	-1.0	1.025	0.839	0.311	0.954	(80.0,85.0]
Winter	366	1.322	1.135	0.188	0.124	16.5	1.224	0.902	0.298	0.943	
Summer	872	1.254	1.146	0.108	0.146	9.5	1.170	0.774	0.368	0.925	(75.0,80.0]
Winter	227	0.906	0.672	0.234	0.106	34.8	1.439	0.891	0.307	0.868	
Summer	173	1.019	0.878	0.141	0.132	16.0	1.259	0.702	0.403	0.905	(70.0,75.0]
Winter	36	0.736	0.521	0.215	0.104	41.3	1.516	0.782	0.328	0.775	
Summer	132	0.798	0.639	0.158	0.105	24.8	1.318	0.701	0.349	0.752	(65.0,70.0]
Winter	4	0.670	0.445	0.225	0.143	50.4	1.951	0.601	0.402	0.693	

Table 5. Summary of seasonal statistics for the Pandora/NIWA intercomparison. Similar to Table 3.

5 Conclusions

The Pandora instrument was collocated with an ~~NDACC-standard~~ NDACC certified instrument (M07 spectrometer) at the NIWA station in Lauder, New Zealand over the period of one year. Spectra from each instrument were processed ~~on~~ using separate algorithms to calculate the NO_2 SCD throughout the day, but with a focus on twilight periods. We showed that the two instruments and algorithms were well correlated (R^2 >0.95) throughout the entire intercomparison period, and that time of year had minimal impact on the ~~results~~correlation. Further, it was shown that, within a specified SZA bin, a change in SZA influenced the correlation (e.g. Figs. 3 and 4).

The Pandora instrument was shown to have a fundamental limitation due to so-called tailing effects ~~, which are a~~ where the instrument seems to lose sensitivity to changes in NO_2 slant-column density below 0.55x10^{16} molec cm^{-2}. The tailing effect is the product of the spectrometer's light throughput, signal-to-noise, and the overall system's precision and accuracy limits. Therefore, Pandora systems may experience sensitivity ~~problems under either~~ limitations under extreme-clean conditions~~(though the authors cannot identify a region that may be cleaner than southern New Zealand), or winter time near solar noon; neither of which influence twilight retrievals~~. However, the Pandora instrument may prove useful for SAGE-III validations ~~for SZA observations near~~ (SZA at time of overpass ≈90° ~~and possibly down to 85°. Lower SZA's may not be accessible~~

~~due to inconsistent curvature from site to site.~~ , Table 4). The SAGE-III project plans on deploying Pandora to sites of interest (ideally low-latitude, ~~tropospherically-clean~~ tropospherically clean environments) with balloon-launching capabilities for ongoing validation work. Due to Pandora's portability the instrument can also be quickly deployed in response to events of interest (e.g. volcanic eruptions).

6 Data availability

All data used within the current study and all code are available from the authors upon request. OMI data are available from the OMI team via http://avdc.gsfc.nasa.gov/index.php?site=2045907950.

Competing interests. The authors declare they have no conflict of interest.

Acknowledgements. P. Johnston and R. Querel were supported by NIWA as part of its Government-funded, core research. L. Thomason, D. Flittner, and J. M. Zawodny were supported by NASA's SAGE-III project. T. N. Knepp was supported by NASA's SAGE-III project through the STARS-III contract.

We appreciate the thorough review from referee #1. The manuscript has been updated to implement the recommendations as described below.

1. Section 1: More details on the SAGE missions necessary. For instance, please add some more detail about how the SAGE-III/Meteor instrumentation works (including a short description of its viewing geometry, overpass times, etc.), the key SAGE species measurements (besides NO2), and any other data for which SAGE is used. This reviewer is not familiar with this missions, suspects that not all readers will be familiar. Added detail will greatly help to provide context on why validation against Pandora is both necessary and desirable.

    (a) More detail was added to introduction.

1. What is the citation(s) for the NIWA M07 instrument being considered a standard for stratospheric NO2 measurements? This is unclear.

    (a) The NIWA group and their instruments have a long heritage of providing data of the highest quality. However, to label this an a "standard" is incorrect. The title and text have been updated to remove confusion in this regard.

1. The NIWA M07 instrument is specifically mentioned only within the last sentence of the introduction; is this the particular instrument that is considered a standard for NO2? Or was it chosen for this intercomparison for another reason, and if so, why? This instrument needs to be introduced along with NIWA rather than at the end of the introduction, to prevent confusion over why the M07 instrument was used.

    (a) Again, the title was updated to remove reference to the NIWA instrument as a community standard. Also, the text was changed to allow introduction of the NIWA instrument under the appropriate section.

1. Section 2.1: "Briefly, the Pandora model used in the current study consisted of. . ." is unclear; is this different from the "normal" working setup of the instrument, or the same? A note on this would be helpful. The statement at the end of the section (". . .the Pandora only operated in the zenith-observation mode. . ." also contributes to the lack of clarity.

    (a) The Pandora instruments have been evolving over time, so it is not accurate, at this point, to say there is a standard hardware configuration (e.g. it is not accurate to say they all have the same spectrometer specs etc.). However, there are general characteristics that remain consistent from version to version. Each instrument undergoes the same calibration procedure and the Pandora group have performed multiple intercomparisons between other Pandora units and other DOAS instrumentation (e.g. at the CINDI and CINDI-2 campaigns). Therefore, it is reasonable to trust in the instruments performance. We specify this specific instrument's hardware for clarity.

        What is different here is the *mode* of operation. Normally, Pandora instruments track the Sun or look away from the Sun to do elevation scans after the Sun is well above the horizon. In the current study we evaluate the instruments performance in a zenith-only observation mode, specifically during twilight hours. The text has been updated to elucidate this difference in the operation mode.

1. Section 3: ". . .both instruments were operated in their normal states, not in a customized operation mode. . ." - this gets back to the comment about Section 2.1 about whether Pandora was used the same it has been in previous studies (or not). This statement should be a reiteration of the mode of operation for Pandora (and M07) from Section 2, to make sure it is clear how these instruments were used (and how this does or does not differ from previous studies).

    (a) As noted above and in the text, the only difference between the mode of operation in the current study and past studies is in the orientation of the entrance optics (i.e. zenith only). The intent of the current study and previous studies is different. Under "normal" Pandora operation the instrument will either do elevation scans, Sun tracking, or some combination of the two. The intention of these normal operation modes is to either collect total VCDs (Sun tracking) or attempt some vertical profiling (elevation scans). While the elevation scans typically involve a zenith observation, these scans are carried out when the Sun is well above the horizon, not during twilight conditions.

Therefore, we cannot provide a comparison with previous studies as the current study is fundamentally different. The current study was carried out to determine whether the Pandora instrument is capable of making twilight observations with the entrance optics oriented in the zenith direction, as stated in the manuscript.

1. Section 3: the statistics thing (troposphere beings so different)

   (a) We are unable to determine what the reviewer means here. No changes made.

1. Section 3: Last sentence ("Since Lauder provides a clean, background-level,. . .") provides a clear statement of the motivation for this work that is not dependent on the specific SAGE mission. This should be perhaps mentioned earlier in the paper (maybe even the introduction after introducing Pandora and Lauder, NZ).

   (a) Given the context, we see manuscript as already meeting the reviewer's request. However, minor wording changes have been implemented within the introduction to bring this out.

1. – Section 4: What are the major retrieval uncertainties for Pandora and M07? These should be briefly described, in Section 2 where the two instruments are initially described. Also should make note of any other known limitations/issues related to the instruments or their retrievals.

   (a) Uncertainties now listed in added section.

1. What does it mean that some datasets were smoothed? Were both Pandora and M07 datasets smoothed, or portions of one or the other instrument's datasets? This statement is unclear. Also, why was five minutes chosen for the averaging time-why not 1 minute, for example?

   (a) We understand the lack of clarity of this statement and appreciate the reviewer bringing this to our attention. When dealing with long time-series data such as the OMI data presented in Fig. 2, the data can appear noisy due to day-to-day fluctuations within the column. This variability can mask trends, and generally make a figure's interpretation difficult. A common technique for bringing out these trends and enhancing a figure's interpretation is to apply a rolling weighted mean, which is what we did in Figs. 2 and 5. Since this "smoothing" was only applied to long time series (i.e. Figs. 2 and 5), this comment was removed from section 4.0. A comment was added to the caption of both Figs. 2 and 5 regarding this smoothing. Again, we appreciate this comment as this would likely have led to confusion of many readers.

   A five-minute block average was applied to both datasets to allow direct comparison of the two data sets. Due to how the data were recorded, the two data sets do not have common time stamps (e.g. Pandora may report an SCD at 12:01:23 while the nearest neighbor for M07 may be at 12:03:11). To temporally align the two datasets a five-minute resampling was performed. This rolling average was done to all data sets and for all analyses. Wording has been changed to clarify this.

1. Section 4.1: Need to explicitly state that the R2 values are given in Table 3, to make it easier for the reader to find the numbers that support the result that the correlation increased with decreasing SZA. Might even be good to list a few R2 values for some of the SZA bins, since to this reviewer, the correlation for the 87.5-90 SZA bin looks strongest when looking at the plots in Fig. 3 (though this was not the bin with largest R2). A follow up question is whether the statistical significance of these correlations was tested, to determine if the correlations were statistically different from each other (at least for the bins containing SZAs less than 92.5); a direct comparison of correlation coefficients can be misleading.

   (a) Reference to table added.
   Yes, the R-squared values are significantly different at 95% confidence, except for two cases. Now specified in text.

1. Are R2 values available for the sub-correlations for each panel of Fig. 3? An example for at least one panel might be good, showing how the correlation decreased with lower SZA within that SZA bin (and by extension for the other SZA bins).

(a) No. Due to the nature of the intercomparison and the proposed validation application going to smaller bin ranges is not applicable. The statement within the text "Within each panel of Fig. 3 the data are color coded to correspond to the SZA range within each sub-panel and provide insight into how the short-term change in SZA influenced correlation. As an example, in panel a it is observed that data collected at higher SZA (red-shaded points) were further from the one-to-one line than data collected at lower SZA (blue-shaded points)" was intended to call out the fact that as the SZA changed, individual points moved either closer to or further away from the 1:1 line. The text has been updated to this to be better understood.

1. Why can the dependence on SZA not be separated from day-to-day chemical variability? I'm not sure what "day to day chemical variability" refers to, so this statement is confusing. Does this refer to the annual variability of the NO2 column, or daily variability of the column? There needs to be a justification for this statement. It would seem that the correlation's dependence on SZA is due to daily photochemistry (available sunlight for photochemical reactions involving NOx), as well as limitations of either instrument at high SZA. So to start the analysis presented in Fig. 3 could be extended, to investi- gate how the time series of the NO2 columns from both instruments within each SZA bin and over all SZA's compare, comparing to O3 column data, etc.

    (a) Agreed. This part of the sentence was confusing and has been deleted.

1. Section 4.2: Do the authors have a hypothesis for why the tailing behavior was limited to winter conditions? This would be good to state in the paper.

    (a) Yes. The "tailing" is only observed at low SCD values. Stratospheric NO_2 follows a seasonal cycle shown in Figure 2, with more NO_2 in the summer, less in the winter. Therefore, it makes sense that we see lower values in the wintertime. A reference to the seasonal cycle has been added to the text.

1. It's true that the R2 values remained high throughout most of the year, but it can be seen that R2 drops during the winter months for most SZA bins in Fig. 5, such as April-July 2015 bin for the 90-92.5 bin, and for the 80-85 bin. Is this just noise, or is this related to the trends observed in slope and SCD ratio for winter vs. summer? It needs some explanation, and this reviewer is not convinced that it can be said that there is no seasonal dependence seen in the correlation at this time.

    (a) We agree that there is fluctuation within the April-July bin for SZA between 90 and 92.5 degrees. However, this fluctuation is not indicative of a clear seasonal pattern in the coefficients of regression. We support that by failing to see the same behavior in the other SZA bins, and by seeing a similar behavior in December. Further, we would expect to see a similar pattern as shown in SCD or %diff panels if there were a seasonal dependence. At this time we cannot conclusively state there is a seasonal dependence in the R^2 values.

1. – Section 5: The second conclusions paragraph is a little confusing to read. Not quite sure what the message is about, particularly about the twilight retrievals. Some rewording should be all that is needed to make the message clearer.

    (a) Paragraph rewritten.

1. – When referencing parts of a figure, such as panel a in Fig. 3, use parentheses to encapsulate the letter to make it easier to distinguish for the reader (e.g.; Fig. 3 panel a Fig. panel (a)).

    (a) Changes made throughout

1. – Fig. 4 says "orrelation" in the plot titles rather than "Correlation"

    (a) We do not see an error in the titles of figure 4. Perhaps this was an error in the reviewer's file?

We appreciate the thorough review from referee #2. The manuscript has been updated to implement the recommendations as described below.

1. Page 1, line 1, Abstract: Add comma: 'In September 2014, a ...'

   (a) Comma added.

1. Page 1, line 8, Abstract: Change to '. . . to vertical column density (VCD) is '

   (a) "(VCD)" added to text

1. Page 2, line 1: delete 'the' before 'stratospheric'

   (a) Deleted.

1. Page 3, lines 13-18 & page 4, lines 2-10: Please add another paragraph to explain in more detail the processing procedure applied to retrieve the NO2 slant columns dis- cussed in the this paper; the same goes for the NIWA data processing and it would certainly be helpful to have a table summarizing the settings used for both instruments, e.g. wavelength interval used for the fit, which cross section were used, which polynomial, Ring, has an offset correction been applied, are the NO2 and O3 cross-sections Io corrected? This will help to understand the similarities and differences in the data processing algorithms applied to each of the data sets.

   (a) Additional text added, as well as a table containing retrieval information for both instruments and references to retrieval details.

1. Can you please provide some evidence that shows that the Pandora instrument used here has a high enough light throughput to make sensible measurements at SZA $\geq$90.

   (a) This was part of the intention of the work. Prior to this study, there were no studies conducted to conclusively determine whether Pandoras have the requisite light throughput to make twilight observations. Text was added to clarify.

1. Page 4, line 15: replace ')' with ','

   (a) Replaced

1. Page 4, line 21: should be 'similarly'

   (a) Changed

1. Page 4, lines 26/27: If these are SCDs please state so clearly in the text.

   (a) When the OMI NO2 product is introduced in the text it is now declared as VCD. The table caption was updated as well to indicate a VCD.

1. Page 5, line 9: '. . . the two sites are in phase . . ..'

   (a) Corrected

1. Page 7, Figure 2: panel (b): any comments re why the displayed ground-based data is clearly higher? Maybe I missed this somehow but this should be included in the caption and also be discussed in the manuscript text. Also state clearly in the first line of the caption if the OMI NO2 data product is SCD or VCD.

(a) As mentioned in the text, we have not inverted the surface instruments' SCD to a VCD. Figure 2, panel b presents the OMI VCD and the Pandora/M07 SCD. The three instruments are *not* plotted here to show quantitative agreement, rather it is to demonstrate that the surface instruments are in phase with the OMI seasonal cycle for stratospheric NO2. Clarification has been added to the text and the figure's caption.

1. Page 8, line 5: '. . . the correlations as seen in Fig. 3'
   (a) Corrected

1. Page 8, line 6: '. . . where it peaks at . . . Within each panel of Fig. 3 the data are color ... '
   (a) All changes implemented.

1. Page 8, line 6: It would also be helpful to refer here to Table 3 as well.
   (a) Reference to table added.

1. Page 8, line 13: '. . . instruments have strikingly . . .'
   (a) Verb tense changed

1. Page 9, Tables 3&4: Last column, close with ')' not ']'
   (a) It is common mathematical notation to use parentheses to indicate an exclusive range, and square brackets to indicate an inclusive range. Therefore, we chose to use square brackets to indicate the ranges specified in tables 3 & 4 are inclusive of the high SZA.

1. Page 10, Figure 3 caption: Change to '(i.e. red colors represent the upper SZA limit, blue colors represent the lower bounds for each individual panel).'
   (a) Hyphenation removed

1. Page 10, line 1: 'The decrease in correlation . . .'
   (a) Change implemented

1. Page 10, line 9: How about something like: 'To better understand the seasonal variability seen within the data sets, the data . . .'
   (a) Your wording provides better understanding of what we are doing. Change implemented.

1. Page 10, lines 10-11: This sentence needs to be improved, change to something like: 'Seasonal correlation plots were generated (Fig. 4); they show nearly identical behaviour to the aggregate (Fig. 3) with most of the tailing behaviour being attributed to winter conditions.'
   (a) Again, your wording allows quicker understanding of the text. Change implemented.

1. Page 11, line 11: Change to: '. . . at high SZAs . . .. or at very low NO2 . . .'
   (a) Change implemented.

1. Page 11, line 20-22: Can you please address if the observed seasonal dependence can cause issues when using Pandora instruments for satellite validation.

(a) Seasonal variability in satellite validation is a well known issue. Unfortunately, we cannot determine the seasonal dependence when comparing either Pandora or M07 to SAGE-III until after the observations are made. We can be certain that there will be a seasonal dependence on how well Pandora and M07 agree with SAGE-III, but that seasonality will have to be accounted for at a later time.

1. Page 11, line 24: delete '-' between 'NDACC' and 'standard'

   (a) Hyphen deleted.

1. Page 11, line 25: replace 'on' with 'using'

   (a) Change implemented.

1. Page 11, lines 26-27: Change to: 'We showed that the data obtained using the two instruments and retrieval algorithms were well . . ., and that the time of the year had just minimal impact on the comparison. ' However, didn't you just state in the paragraph above that there actually is a seasonal impact??

   (a) You are correct. It is clearly demonstrated within the manuscript that the seasonal impact was +/- 10%. The sentence should have said "minimal impact on correlation". Now corrected.

1. Page 11, line 30: The tailing effect should be explained when first mentioned.

   (a) Tailing effect now defined in the conclusions section.

1. Page 12, line1: The SZA range where the Pandora instrument may be useful for SAGE-III validation seems rather limited (around 90 deg, possibly as low as 85 deg) - can you please elaborate a bit on if that is realistic with re to known overpass information for suitable sites.

   (a) A table was added to show SZA statistics for all SAGE-III/ISS observations collected thus far. The SZA were calculated with respect to a potential surface instrument's viewing geometry. The table shows the SZA to be tightly grouped about 90 degrees.

1. Page 12, line 2: I don't understand the sentence: 'Lower SZAs may not . . ...'. Can you please explain what you mean here.

   (a) That sentence was superfluous and has been removed.