Author response to Anonymous Referee #1

First, we would like to thank you for your kind words regarding the content and structure of the manuscript. We understand that more details would have been helpful in some cases and we have made an attempt to include revisions to the manuscript that directly respond to your comments and concerns. Our responses to your comments are highlighted in red in the text below.

Anonymous Referee #1

Received and published: 27 August 2018

GENERAL COMMENTS

This nice paper describes the level 0 and level 1 processing scheme for a new limb-sounding heterodyne spectrometer called SHOW. It proceeds to describe the first test campaign, where the instrument worked well enough such that first water vapour retrieval results could be gained that compare favourable to co-located radiosonde measurements.

The paper is well written and its scope fits very well into AMT. It describes the individual steps of the processing in a very approachable manner (commendable!), but leaves often small but important details unmentioned or implicit. A minor textual revision that ensures that all algorithms have been described with more rigour could turn this good paper into a great one. Specific notes below have been made below for the most important issues.

Further, there are several comments listed below, which must be addressed before I can recommend publication. I do not foresee any problem on the author's side to do so easily.

SPECIFIC COMMENTS page 23, Section 11.2:

I find the use of the term "degraded vertical resolution" with respect to the pointing uncertainty misleading. The vertical resolution of the measurements is primarily determined by the FOV/PSF of the instrument. In case that the line-of-sight cannot be actively managed, aircraft movements during a measurement also worsen the vertical resolution. Another factor in determining the vertical resolution of the final product (compared to the measurement) is also a potentially employed regularisation.

However, here these factor are intermingled with an uncertainty in absolute pointing, which seemingly stems from inaccurate pitch readings by the ER-2 attitude system. Further factors may also contribute here, e.g. thermal contraction of the instrument. Such a bias in the pointing would be assumed to be constant during the interferogram acquisition. Such a bias would generally cause a shift in derived features (in contrast to a smearing out) as well as likely quantitative errors due to the wrong assignment of measurements to tangent altitudes and corresponding problems with pressure and temperature assumptions.

From the variance in radiation, you properly estimate an estimate for the precision of the employed line-of-sight angle, which shouldn't enter the vertical resolution. In contrast, I miss an estimate for the accuracy of the line-of-sight angle, which obviously can be calibrated on ground, but may change slightly upon mounting the instrument or may change due to thermal effects. Also, can the wings move in relation to the main body of the aircraft depending on speed and prevalent winds? I also miss an

estimate for the intra-interferogram variation of the pitch angle, which might be derived from highly-resolved aircraft pitch measurements that would, in fact, worsen the vertical resolution of the measurement.

At the beginning of the SHOW ER-2 project and leading up to the ER-2 demonstration flights, there were several concerns regarding the stability of the viewing geometry due to the aircraft motions. These concerns were primarily focused on potential vibrations in the aircraft wings (through flexing) or inherent pitch variations in the aircraft body.   Clearly, mapping the observed lines of sight in the images to the corresponding georeferenced tangent altitudes requires us to account for these variations. After the demonstration fights, the concerns regarding the aircraft pitch variations were partially alleviated since we found that the aircraft attitude information (obtained from a sensor located in the fuselage) could be used in combination with the atmospheric signals to obtain georeferenced images. This is the mapping that is described in Section 11.2.   On the other hand, we found it difficult to quantify any possible "flexing" of the wings during the flight since the attitude sensor is located in the fusealage of the airplane. Quantifying this effect would require the use of a sensor in the wingpod to monitor the absolute pointing. Indeed, future versions of the instrument should make use of such a configuration to remove this uncertainty.

We agree that the use of the term "degraded vertical resolution" with respect to the pointing accuracy is not entirely appropriate. In order to address your concern, we have removed this wording and we have revised small parts of the manuscript to include a description of the uncertainty in the absolute pointing which was confirmed on the ground to 0.01 degrees. We separate this uncertainty from the instantaneous field of view (or PSF) of the instrument which sets the maximum possible vertical resolution that is possible for the measurements.

The vertical resolution of the SHOW measurements is now discussed in the following manner in the manuscript:

1. In section 9.3.2, we present the vertical field of view measurement that was performed on the ground in the lab at ABB Ltd. This provides us with a mapping of angle of incidence (or line of sight angle) to row number (Figure 4) and an estimate of the instantaneous field of view of the instrument (or point spread function). The instantaneous field of view of the instrument was measured to be roughly 0.0176 degrees. This sets the maximum vertical resolution at the limb that can be obtained with the instrument.

2. Aircraft movements that occur during the acquisition of an interferogram will also degrade the vertical resolution. This effect is negligible for the SHOW measurements, since the frame rate is high enough that the aircraft motions can be considered stable over a single acquisition(Figure 13 (a)).

3. In practice, the spacing of the retrieval grid and the type of and amount regularization also impacts the vertical resolution. In our case, very little smoothing is applied and the regularization has very little impact on the retrieved water vapour profile. This is illustrated by the averaging kernel in Figure 18 (b) which has widths that are equal to the 250 m retrieval grid.

The absolute pointing is now discussed in the manuscript with the following content included to address your concerns:

Page 25, Section 11.3

The section does a good job of giving a good overview of the employed technique but skimps on the detail. I assume that the computed correction factor for each spectral sample is of a multiplicative nature. While the method obviously works, it remains a question of how well this models the underlying failure and what error remains in the final data for proper error analysis.

The error may likely be assumed to be multiplicative in interferogram space due to the applied flat-fielding with potentially remaining higher-order effects (see Sect. 6). Such a multiplicative error in an interferogram sample will cause a sinusoidal error in the resulting spectrum. Due to the nature of the multiplicative error, I'd expect the influence of the pixels close to the ZOPD to be higher causing the error pattern to be of a more low-frequency nature and thus coloured noise.

So, I am looking for a physical motivation to the correction technique. If the effect is stable during this window, a point-wise multiplication in interferogram space should solve the issue, which would correspond to a convolution in spectral space.

Please motivate the correction technique and describe its application better to make clear its nature and implementation.

variability due to a less-than-perfect flat field correction. This flat field correction included variations due to optical effects as well as variations in the response of the pixels across the detector.

The SNR of the individual measurements is not an issue for the SHOW instrument – the SNR is high enough to minimize the impact of the noise associated with photon counting. For the SHOW measurements, we found that the primary source of noise in the interferogram samples is the inter-sample variability due to a less than ideal flat-field.

Indeed, the combined flat field correction and non-uniformity correction is multiplicative; however, the residual variation that remains after the application can be viewed as a perturbation. As noted in Section 6, this underlying failure leads to an effectively random relative pixel variation $\delta S/S$ across the interferogram row where S is the signal at the interferogram peak and δS is the error. In this case, the noise level on particular sample is not proportional to the source intensity and is fixed for a fixed source strength. It adds a small perturbation to each sample in a similar manner to digitization noise (see Davis (2003)).

This feature is discussed by Englert et al. (2006) where he notes that such a variation results in an SNR in the spectral samples of between $\sqrt{N/2}\ S/\delta S$ and $\sqrt{2/N}\ S/\delta S$ due to multiplex noise propagation for the case of a monochromatic line or continuum respectively. SHOW utilizes a relatively broad spectral feature and therefore suffers a loss in SNR closer to the continuum case.

While the pattern is effectively random over the several hundred interferogram samples, the inter-sample variability remains fixed for a fixed input intensity. If we take the same measurement hundreds of time with a stable source of fixed intensity, then we find that the amplitudes of inter-sample variations remains fixed in the interferograms, as well as, the spectra. The effect be viewed as adding and subtracting a fixed set of very small perturbations to the interferogram samples. The FFT of the resulting interferogram gives the desired spectrum plus a fixed (in wavenumber space) set of associated perturbations. Since the perturbations are small (< 5 %), a multiplicative factor can be isolated using the approach discussed in Section 11.3. The text in Section 6 and in Section 11.3 has been updated to make this physical understanding more clear.

Page 30, lines 1ff:

The agreement between the radiosonde and the SHOW retrieval results is indeed remarkable. However, the error bars due to noise errors are not applicable here, as noise is likely one of the smallest error sources of your retrieval compared to many of the instrumental uncertainties you (partially) corrected for and other inaccuracies by the forward model (e.g, line strength uncertainties, single/multiple-scattering) or other influencing factors (uncertainties in pressure, assumed aerosol profile, solar radiation variation, horizontal homogeneities along the line-of-sight, . . .).

The actual extent of the error bar is quite difficult to see on the plot. A logarithmic plot might be helpful here. It is certainly not expected that the SHOW retrieval results agree with any other measurement within error bars induced by noise alone. The retrieval paper (Langille 2018) has, e.g., a sensitivity towards aerosol of 1 ppm, which also fully explains discrepancies at higher altitudes. For lower altitudes, the larger variability of water vapour in the stratosphere may, as rightly noted, explain observed discrepancies.

The paper should take systematic errors identified by Langille (2018) into account for the comparison. The plots should make identifying the errors better.

Agreed, the main source of error is not associated with measurement noise and the error bars will be removed from the plot and text has been added to the Figure caption and in the text of the manuscript to explicitly discuss the uncertainties. The manuscript has been edited to relate the discrepancies in the plot back to the sensitivity study (Langille, et al., 2018).

MINOR REMARKS page 1, lines 33ff:

Amongst others, especially ACE-FTS has also significantly contributed to our knowledge about water vapour in the UTLS, partially due to its significantly higher resolution compared to the listed instruments. See for example doi:10.1029/2008JD009984. In contrast, it lacks in temporal and spatial coverage, obviously.

A reference to an ACP ACE papers has now been added to this paragraph.

page 6, lines 5: Reference of (Harlander, 1990) is not defined.

This has been changed to (Harlander, 1991) and (Harlander, 1992).

page 6, lines 32: Are the ranges of the integral correct? With the variable transformation from sigma to kappa, I'd expect a corresponding change of integration range. In addition, I suspect also a missing factor. This makes only sense under the assumption of an appropriately chosen passband filter, which is only introduced later.

You are correct. The range needed to be updated to reflect the change in variables. Thank you for picking up on that one.

page 9, line 24: Why is a nearest neighbour interpolation and not a linear one between the upper and lower neighbour used? Especially in case of only a few bad pixels, this should reduce the error introduced due to the increase of radiation with decreasing tangent point altitude.

For this iteration of the SHOW instrument we chose nearest neighbor interpolation since it appeared to to work reasonably well in the processing of lab measurements. We are in the fortunate scenario where the detector has very few bad pixels and in most cases, a single row has none or only one bad pixel. In this case, the nearest neighbor interpolation works fine. In future campaigns we plan to investigate other approaches to this procedure.

page 9, lines 30ff: It is not clear what correction is performed on the interferograms with respect to the changing aircraft pitch. I interpret it in the way that not the interferogram is changed, but that each interferogram is associated with a tangent point altitude determined from the current aircraft pitch.

In fact, this is not a correction, but rather a procedure that is applied to map the interferogram rows to tangent altitude at the limb. The manuscript has been edited to make this more clear.

page 9, lines 30ff: In what way is a possible bias between the LOS of SHOW and the given pitch angle of ER-2 determined? While the pitch variations of ER-2 can be readily applied, any systematic bias needs to be ruled out to properly assess the quality of the tangent altitudes.

There is the very real possibility of a systematic bias between the LOS of SHOW and the pitch angle of ER-2. As mentioned in the response to your "Specific Comment", the bias between the SHOW LOS and the ER-2 pitch was measured in the hangar to an accuracy of ±0.01 degrees. The manuscript has been updated to note this pointing accuracy.

page 10, lines 4ff: It is unclear, what correction is described here. Section 11.2 describes a means of estimating remaining uncertainties in pitch/LOS after taking ER-2 pitch data into account, but does not describe correcting these. Different effects that are difficult to differentiate may cause the variation in radiation (changing scenery/clouds as well as inaccurate or imprecise pitch values from ER-2) and may a correction in contrast to an error estimation difficult.

This is not really a correction but a mapping that maps interferogram rows to tangent altitude. The remaining variability is quantified by determining the variability on what appears to be a cloud feature. The text has been edited to make this clear. The remaining variability amounts to an uncertainty in the absolute pointing of the instrument of ±0.032 degrees.

page 10, lines 19: For a non-linear retrieval, the matrix K should be marked with an index i or as being dependent of the current iterate xi as K(xi) or F 0 (xi). with F 0 being the Jacobian of F.

This has been updated in the text.

page 12, lines 25ff: From the description, it is unclear what uniformity or non-linearity correction is performed. The sentence that no uniformity correction is done contradicts a statement of page 8 line 23.

The non-uniformity correction is performed as part of the flat-field. The text has been edited to remove the contradictory statement.

page 17, lines 3ff: One may also estimate the noise from the imaginary part of the spectrum after phasecorrection. You typically operate on magnitude spectra. Could one still compute the phase from calibration measurements and use this to both reduce and estimate the noise?

This type of correction is not feasible with the current configuration of the SHOW instrument due to the presence of the aliasing effect. As mentioned in Section 5, aliasing leads to a complicated phase relationship that is extremely difficult to isolate. This relationship changes with the

page 27, lines 8f: It is never spelled out precisely, how the correction is applied. It seems to be a multiplicative factor?

Yes, the correction factor is a multiplicative factor. The procedure that is used to isolate this correction and apply it to the data has been made more clear in the manuscript. Specific sections that have been edited are Section 6 and Section 11.3.

page 27, line 17f: There is a difference in scales between the two panels of Fig. 16. As such it is not apparent that the effect is minimal in wavenumber space. A delta plot would show that better in both instances. Especially the absorption feature at 1364.25 seems affected, but the l.h.s. might be affected by aliasing more. However also at 1365 is a visible different place with a large correction.

In this Figure, we were attempting to point out that the primary variability is in the vertical dimension. The statement saying the effect in wavenumber space is minimal has been removed.

page 29, line 10f: As shown in Fig. 18, the regularisation is insignificant and doesn't control or seemingly even influence the resolution.

This section has been edited to make it clear that the employed regularization does not significantly impact the vertical resolution.

page 29, line 10: It is unclear what "damping" refers to and how it relates to the regularisation. One typically uses a trust-region method like Levenberg-Marquardt which contains something like a dampening factor. Here it sounds like the dampening would affect the result of the retrieval, which weren't the case for Levenberg-Marquardt.

The damping term here does not influence the retrieval in the case that the retrieval has converged. The wording has been changed to state that "A small amount of damping is applied to ensure convergence of the retrieval"

TECHNICAL CORRECTIONS

page 10, line 14: The cite of Rodgers should be "described by Rodgers (2000)".

This has been edited in the text

page 10, line 19ff: The used mathematical notation is formatted differently in the equation as in the main text (cursive vs. regular and probably in a different font).

This has been edited in the text

page 17, Figure 7: Especially the "(c)" in image (c) is barely readable. Potentially a white rectangle behind the letter or a "path effect" of matplotlib would help here.

This has been edited in the text

page 18, Equation (9): There are two consecutive "+" in the formula.

This has been edited

page 24, Figure 13: Especially the "(b)" in image (b) is barely readable

This has been edited

Response to Anonymous Referee #2

Thank you very much for your comments on the manuscript. Please find our responses to your comments below highlighted in red.

Anonymous Referee #2

General Comments

The manuscript details water vapor measurements made with a Spatial Heterodyne Spectrometer (SHS) from an ER-2 high altitude aircraft and compares retrieved water vapor densities with coincident radiosonde measurements. My overall impression of the manuscript is quite positive. It provides sufficient context and the detailed descriptions provided make it quite easy to follow. The agreement between the SHOW and the radiosonde water vapor densities is remarkable, particularly given the complexity of the L0 to L1 processing required in the analysis of the SHOW data. Given the quality of the manuscript, my specific and minor comments below are to be considered as advisory to the authors and not requiring another detailed review.

Specific Comments

The manuscript details a series of steps required to remove instrumental effects from the level-0 data. The most complex of these stems from the combination of aliasing due to the passband of the interference filter spanning the Littrow wavenumber and a vertical fringe frequency tilt of the interferometer. The result is a row-dependent spectral modulation seen near left edges of Figures 7c and 7d and require that a detailed instrument model be in integral part of the retrieval. Although it's implicit in the discussion, I think it is important to point out that the difficulties and uncertainties associated with correcting for this effect could be eliminated by using a filter with passband shifted slightly to the red that blocks all light at wavelengths on the opposite side of Littrow. To simplify the analysis is there a plan to replace the interference filter for future flights of the instrument?

You are correct. This effect can be eliminated by changing the filter. Indeed, after the first ER-2 demonstration flights we obtained the financial resources to procure a new filter that is shifted so the Littrow wavelength sits at the edge of the filter passband. This ensures that there is negligible aliasing from the opposite side of Littrow. The filter was procured and installed in the Spring of 2018. The new configuration has been characterized in the lab and shows no signs of the aliasing effect. Several lines of text have been added to Section 5 of the manuscript to explicitly point out that this effect is entirely avoidable by design.

There are numerous places in the processing where a fitted high-order polynomial is subtracted from the data in an effort to assess the noise from the residual. It would be helpful to indicate the order of the polynomial used. Clearly with a very high order polynomial fit, some of the noise will be fit and the noise in the residual will be underestimated while fitting with a polynomial of too low an order will result in signal in the residual. How was the decision as to the order of polynomial used made?

In most cases, the high order polynomial was of order 8. However, we did not do a detailed analysis of the impact of lower or higher orders. For most of the cases, we are just using this approach to obtain a first order approximation of the noise in the measurements to check that photon-noise is the primary

Although not ultimately used in the analysis due to optical depth issues I found the description of the cloud artifacts evident in the lowest rows of Figure 12a and 12b somewhat lacking. If the entrance optics are anamorphic and aligned properly, they should completely defocus spatial information in the horizontal direction. Why then does the modulation in rows 0 to ~30 in Figure 12a tilted? Could this possibly be a fringe due to a spectral line very close to the Littrow wavenumber?

Thanks for pointing this out. After a second look at the images, it turns out that the primary effect is likely due to removal of the DC bias in the presence of nearly saturated pixels at the interferogram center burst. It is known that saturated pixels causes the pixels in the corresponding row to have non-ideal behavior. The manuscript has been edited to note that the cause of this effect is not completely clear; however, it is believed that the culprit is likely saturated pixels at the interferogram center burst combined with the DC bias removal. In any case, we do not use these rows in the retrieval and the effect does not appear in the rows above the cloud feature where the retrieval is performed.

It appears that amplitude spectra where used throughout in the analyses. If proper phase correction is performed the spectral information can be isolated to the real part of the Fourier Transform only thereby reducing the shot noise contribution from the imaginary part and reducing the shot noise contribution by roughly the square root of two. That said, I suspect that the dominant source of uncertainty is not shot noise so perhaps this improvement is not worth the substantial effort required.

Unfortunately, we could not perform a proper phase correction on the SHOW measurements due to the presence of aliasing in the system. We also did not have access to a tunable laser that could be used to perform this characterization. Such a characterization will be performed prior to the next SHOW measurement campaign using the new filter.

Connected to the previous comment, the error bars shown for the SHOW measurement in Figure 19 appear to be quite small. Do these error bars include the systematic effects associated with uncertainties in the retrieval or are they simply an indication of the photon shot noise component of the noise? It would most illuminating, if the systematic uncertainties could be quantified. In general, I would suggest making a clearer distinction between statistical sources of noise (photo, dark, CCD read noise, etc.) and systematic sources of uncertainty (errors in the uniformity correction, uncertainties in the retrieval parameters, etc.). In various places in the manuscript all of these effects are referred to as "noise".

The error bars shown in this figure only include measurement noise. The error bars do not include systematic effects associated with uncertainties in the retrieval. In our sensitivity study (Langille et al (2018)) we examined some the sensitivities of the retrieval assuming an ideal instrument configuration. In the current work, the primary source of uncertainty is our ability to accurately capture the systematics associated with aliasing. We have made edits to the manuscript to make a clearer distinction between the various sources of noise.

To understand the instrument, it would be helpful to add a short section and perhaps a figure describing the anamorphic optics that feed the interferometer and the optics between the interferometer and the detector.

The anamporphic optics and imaging configuration are discussed in detail in the design paper (Langille et al., 2017).

Technical Corrections

Section 5 second paragraph, second sentence: Missing "to"

Corrected in the manuscript
Section 6 third paragraph: N in the SNR equations should be defined

Corrected in the manuscript

Equation 4: Define Ij and σj.

Corrected in the manuscript

Section 7, second sentence: Figure 3 reference should be Figure 2

Corrected in the manuscript

Section 9.4, end of first paragraph: LightMachinery should be more completely referenced (e.g. LightMachinery, Inc., Ottawa, Ont).

Corrected in the manuscript

Figure 9a: The black background makes it very difficult to read

Unfortunately, this is the only version of the flight path that was made available by AFRC.

Response to Anonymous Referee #3

Anonymous Referee #3

This paper describes the details of measuring and retrieving water vapor from the SHOW instrument on a high altitude aircraft, including some first results compared to a nearby radiosonde profile. The development of the SHOW instrument is an exciting capability for the research community. Most of the paper involves characterization of the instrument and measurement details on the aircraft and in the laboratory. Overall, the paper provides a comprehensive explanation of these aspects, although I am not an expert on any of these details and cannot provide critical comments. Hopefully these will come from other reviewers. The comparison of the SHOW retrieval and the single radiosonde water vapor profile (Fig. 19) looks quite reasonable, although I question why the uncertainties are so small in both measurements (see below). Overall the paper is clearly organized and well written, the figures are reasonable, and the topic is appropriate for AMT. I only have a few minor comments to contribute:

1) The uncertainties in Fig. 19 seem small to me. The radiosonde measurements (Vaisala RS41) probably makes accurate measurements (to a few %) in the upper troposphere down to 10-20 ppmv, but there are larger uncertainties for lower H2O amounts at higher altitudes. Where do the uncertainties (error bars) shown in Fig. 11 come from? The 1 ppmv uncertainty quoted here may be on the small side at upper levels. Likewise, the uncertainties in the SHOW retrieval look remarkably small, given all of the uncertainties and corrections discussed in the paper. I have to say this is a relatively minor point, given the quite good agreement in Fig. 19.

The uncertainties in Figure 19 only include an estimate of the measurement uncertainty calculated using the specifications for the Vaisala RS41. The raw measurement uncertainty on any given radiosonde sample is on the order of 1-2 ppm; however, these uncertainties are further reduced since they are transformed to the lower vertical resolution retrieval grid of the SHOW measurements. The radiosonde error bars do not include any estimates of the accuracy of the measurement. In the manuscript we point out that the accuracy of the measurements is a function of relative humidity and the temperature and that these accuracies vary from sensor to sensor. Similarly, the error bars on the SHOW water vapour measurements only include measurement noise. In our case, the errors due to measurement noise are quite small due to the high SNR of the measurements.

We agree that the error bars are too small to be useful in the Figures. We have removed from the figure and the text has been edited to make it clear that these uncertainties are not the limiting errors in the measurements.

2) Some definitions need to be included: N (page 8, line -4 and elsewhere), and DN, PRNU (p. 11, Table 2 and elsewhere).

The text has been edited to provide definitions of these quantities.

3) Figure 9a reproduces quite poorly, and no need to have a black background.

This figure was provided by AFRC and it not available with another background.

Spatial Heterodyne Observations of Water (SHOW) from a high altitude airplane: Characterization, performance and first results

Jeffery Langille[1], Daniel Letros[1], Adam Bourassa[1], Brian Solheim[1], Doug Degenstein[1], Fabien Dupont[2], Daniel Zawada[1]

[1]Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, S7N 5E2, Canada
[2]ABB Inc., Quebec city, G1P 0B2, Canada

Correspondence to: Jeffery Langille (jeff.langille@usask.ca)

Abstract: The Spatial Heterodyne Observations of Water instrument (SHOW) is a limb-sounding satellite prototype that utilizes the Spatial Heterodyne Spectroscopy Technique (SHS), operating in a limb-viewing configuration, to observe limb-scattered sunlight in a vibrational band of water vapour within a spectral window from 1363 nm to 1366 nm. The goal is to retrieve high vertical and horizontal resolution measurements of water vapour in the upper troposphere and lower stratosphere. The prototype instrument has been configured for observations from NASA's ER-2 high altitude airborne remote science airplane. Flying at a maximum altitude of ~21.34 km with a maximum speed of ~760 km /hr, the ER-2 provides a stable platform to simulate observations from a low-earth orbit satellite. Demonstration flights were performed from the ER-2 during an observation campaign from July 15 – July 22, 2017. In this paper, we present the laboratory characterization work and the Level 0 to Level 1 processing of flight data that was obtained during an engineering flight performed on July 18th, 2017. Water vapour profile retrievals are presented and compared to in situ radiosonde measurements made of the same approximate column of air. These measurements are used to validate the SHOW measurement concept and examine the sensitivity of the technique.

1   Introduction

Water vapour is an extremely important trace species in the upper troposphere and lower stratosphere (UTLS) region of Earth's atmosphere. Indeed, it is well known that the abundance and distribution of water vapour in the UTLS is strongly linked to climate processes (Gettleman et al., 2011; Sherwood et al., 2010). However, research over the past couple decades has indicated that the distribution of water vapour and its link to climate processes is not fully understood. For example, the impact of Stratosphere Troposphere Exchange (STE) and the formation of the Tropopause Inversion layer (TIL), as well as, the role of water vapour as mechanism for radiative feedback still requires detailed study (Randel and Jensen, 2010). The primary factor limiting the ability to perform a detailed study of these processes is the lack of accurate long term global measurements that have a high vertical resolution in the UTLS.

Remote sensing of water vapour can be achieved using many different techniques and platforms (space-based, balloon, in-situ, airplane or ground-based platforms). While each technique has its advantages, the best combination of vertical resolution and global coverage of trace species in the UTLS is achieved using a limb-viewing instrument operating from a low-earth orbit satellite. Indeed, the current measurement record of water vapour has been enriched with observations from limb viewing instruments such as the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) (Fischer et al., 2008; Milz et al., 2005; Stiller et al., 2009), the Microwave Limb Sounder (MLS) (Hurst et al., 2014; Hurst et al., 2016; Sun et al., 2016), and the Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY (SCIAMACHY) (Rozanov et al., 2011). The Atmospheric Chemistry Experiment (ACE) instrument, which performs measurements in the solar occultation mode has also made significant contributions

1

to this data record (Carleer et al., 2008). These instruments provide between 3 – 5 km vertical resolutions in the UTLS with $< \pm 1$ ppm error.

The Spatial Heterodyne Measurements of Water (SHOW) instrument is a prototype satellite instrument that is being developed in collaboration between the Canadian Space Agency (CSA), the University of Saskatchewan (USASK) and ABB Inc, to provide high spatial resolution measurements (< 500 m) of the vertical distribution of water vapour in the UTLS from a space borne platform (Langille et al., 2017; Langille et al., 2018). The instrument utilizes the Spatial Heterodyne Spectroscopy (SHS) technique (Connes, 1958; Harlander, 1991), operating in the limb-viewing configuration to observe limb scattered radiation within a vibrational band of water centered at 1364.5 nm. The prototype version of the instrument, discussed in this paper, is configured for measurements from NASA's ER2 high - altitude airborne science airplane. High altitude measurements (~21 km) from the ER-2 airplane provide a sub-orbital demonstration of the SHOW instrument technique. A vertically resolved image of the limb scattered water vapour absorption spectrum with an un-apodized spectral resolution of ~0.03 nm within a narrow spectral window from 1363 nm and 1366 nm is extracted from each limb image. These vertically resolved spectra are inverted using non-linear inversion techniques and a forward model of the measurement to retrieve the vertical distribution of water vapour.
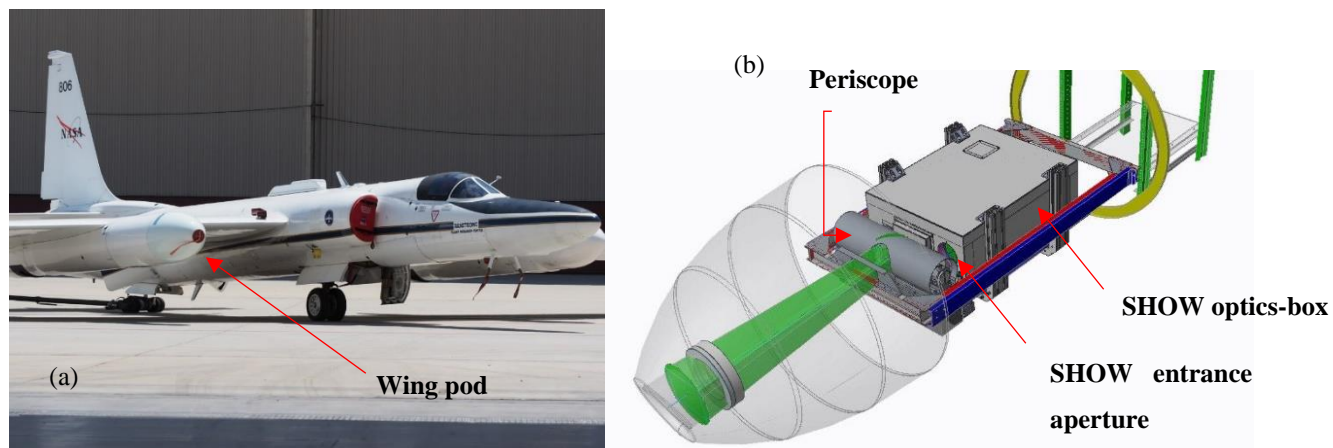
The SHOW measurement approach is unique in several ways. For example, it is the first demonstration of the measurement of atmospheric water vapour using the SHS technique; specifically, using limb scattered sunlight. In addition, the majority of the SHS instruments that have been developed, have been used to observe well-isolated emission features, such as in the case of SHIMMER (Harlander et al., 2002), and the DASH type instruments that are used to remotely detect motions in the upper atmosphere using well isolated airglow emissions (Englert et al., 2016; Englert et al., 2015). Therefore, SHOW is also one of only several SHS instruments (Englert et al., 2009) that has been developed to observe absorption features over a broader spectral range.

In two earlier publications, we presented the design of the prototype instrument (Langille et al, 2017) and demonstrated that the configuration and sensitivity of the instrument was suitable for airplane measurements of water vapour (Langille et al., 2018). This sensitivity study was performed assuming an ideal instrument configuration with realistic signal levels and Poisson noise added to the signals. A non-linear optimal estimation retrieval algorithm was developed to invert the spectral signatures and retrieve water vapour. Assuming the ideal configuration and an airplane altitude of 22 km, it was shown that SHOW is capable of providing vertically resolved measurements of water vapour with 1-ppm accuracy with 500 m – 1 km vertical resolution in the 8 km – 18 km altitude range.

In this paper, we focus on the Level 0 to Level 1 processing and characterization of the SHOW measurements that were obtained from NASA's ER-2 airplane during an engineering flight performed on July 18, 2017. During the flight, a radiosonde was launched from the Jet Propulsion Laboratory (JPL) facility located close to Table Mountain at the same time that SHOW observed the same approximate column of air. This provides a direct estimate of the in-situ water vapour abundance for the coincident SHOW measurements.

In practice, it was found that non-ideal instrument effects associated with the instrument configuration introduced systematic variations in the spectra that must be appropriately characterized prior to performing water vapour retrievals. In this paper, we present the work that was performed to characterize these effects. It is shown that knowledge of the instrument configuration can be utilized to develop an instrument model that is optimized to predict the systematic variations that are observed in the SHOW spectra. The approach is validated by using the measured in-situ water vapour abundance as input to a forward model to simulate the expected limb-scattered radiance profile for the coincident measurements. This radiance profile is then used as input to the SHOW instrument model to simulate the expected SHOW interferograms and spectra. The ability of the model to predict systematic

variability in the Level 1 spectra is examined. Preliminary water vapour retrievals are presented and compared to the in-situ measurements. These measurements are used to demonstrate the SHOW measurement technique and examine the performance of the instrument.

2   Spatial Heterodyne Observations of Water (SHOW) from NASA's ER-2 high - altitude airborne science airplane

The primary scientific goal of the SHOW instrument concept is the realization of high vertical resolution sampling (< 500 m) of the water vapour distribution in the UTLS region from a low earth orbit satellite with $< \pm1$ ppm accuracy. The SHOW instrument was originally developed as a laboratory prototype under the CSA Advanced Studies Program at York University (Langille et al.,

10   2017) and the instrument was further adapted through CSA's Space Technology Development Program (STDP) for demonstration from stratospheric balloon or high altitude airplane. The instrument was tested from a stratospheric balloon in 2014 (Dupont et al., 2014) and from 2015 to 2016, the instrument was configured for observations from NASA's ER-airplane (Dupont et al., 2016; Bourassa et al., 2016).

The SHOW prototype instrument is shown mounted inside the wing pod of ER-2 airplane in Figure 1. The instrument observes

15   the limb through a forward-looking window that is located at the front of the wing pod (see Figure 1 (a)). Limb-scattered sunlight enters the instrument through a periscope assembly that is mounted at the input of the optics box. This assembly consists of two mirrors that align the optical axis of the instrument with the wing pod window as shown in Figure 1(b). The aft optics are configured to provide a 4° vertical by 5.1° horizontal field of view of the limb and the instrument is tilted inside the wing pod to point the SHOW field of view downward by 3.23 degrees. The relative tilt between the airplane boresight and the ~~wingpod~~wing

20   pod combines to tilt the optical axis of the instrument downward 2.40 degrees from the airplane boresight.

The vertical sampling at the limb is defined by the airplane viewing geometry from the ER-2, and the configuration of the optical system and the SHS. In general, an SHS produces an interferogram image that corresponds to a set of overlapping Fizeau fringes with spatial frequencies that depend on the separation of the signal wavelength from the heterodyne (Littrow) wavelength. For SHOW, the SHS and aft optics are arranged so the interferograms are aligned horizontally. The fore-optics is configured to be

25   anamorphic (Langille et al., 2017) so the interferograms are imaged conjugate to the limb in the vertical dimension at the detector, whereas, each interferogram sample in the horizontal averages over the horizontal scene information. This ensures that the spatial information contained in the horizontal does not contaminate the spectral information recorded at a particular row.

3

Each SHOW measurement provides an interferogram image that has 295 vertical rows and 494 interferogram samples. The geographic location of the range of tangent altitudes within the field of view are determined from the airplane attitude information (altitude, heading, pitch, roll etc.). Variations in the pitch of the airplane during flight result in different interferogram rows observing different tangent altitudes during different parts of the flight. Post flight, the airplane attitude information is used to obtain a one-to-one mapping of each interferogram row to tangent altitude. For the measurements presented in this paper, the lowest observed tangent (set by the pitch of the airplane) varies from ~1 km to ~7km. Assuming zero pitch at an airplane altitude of 21.34 km and taking the radius of the Earth to be 6371 km, the SHOW viewing geometry (4 degrees vertical, tilted 2.40 degrees downward from the boresight) provides an altitude coverage from 2.5 km up to 21 km. Within this altitude range, SHOW obtains 295 vertical samples with an instantaneous angular resolution of 0.0176 degrees. This corresponds to an approximate tangent point vertical resolution of 20 m – 151 m, where the sample spacing~~sampling resolution~~ decreases with increasing altitude from ~2.5 km up to ~21 km.

During Level 0 to Level 1 processing, the raw interferogram signals are converted to calibrated interferograms and these calibrated interferograms are converted into corrected spectra. For each row in the image, a water vapour spectrum is obtained with an un-apodized spectral resolution of ~0.03 nm. A single processed SHOW data record includes the raw interferogram image, a vertically resolved image of the water vapour spectrum, as well as, housekeeping information and a mapping of each row in the image to tangent altitude at the limb. The SHOW retrieval approach, described in (Langille et al., 2018), implements a non-linear optimal estimation algorithm (Rodgers, 2000) to extract the vertical distribution of water vapour from the geo-referenced, vertically resolved water vapour spectra. The primary specifications of the SHOW prototype instrument are listed in Table 1. The reader is referred to (Langille et al., 2017; Langille et al., 2018) for more details regarding the design of the prototype instrument.

Instrument parameter	Specification
SHOW ER-2 altitude	**~21.34 km (70000 ft max)**
Airplane speed	**~760 km/hr (maximum at altitude)**
Field Of View	**4º vertical by 5.1º horizontal**
Instantaneous angular vertical resolution	**0.0176 degrees**
Mass	**222.68 lbs [101 kg]**
Power	**465 W (peak), 200 W (average)**
Dimensions	**(0.465 m × 1.32 m× 0.38 m)**
Operating Temperature *SHS temperature*	**-90°C (at 60 kft altitude) to +40°C +23°C**
Spectral Resolution (unapodized)	**~0.03 nm**
Spectral range	**1363 nm – 1366 nm**

Table 1 SHOW ER-2 instrument parameters

3   Data acquisition and storage

Control and operation of the instrument from the ground is achieved using the flight control software developed by ABB Inc and ground support equipment developed by USASK. Communication with the instrument is performed using the NASA Airborne Science Data and Telemetry system (NASDAT) and the airplane attitude information is also monitored and stored onboard the airplane. The control software and communication system allows for near real time download of images, as well as, the airplane attitude information (altitude, speed, pitch, roll, heading etc.) during the flight. The system was used for configuring the science modes (acquisition times, frame rates, FPA temperature etc.) and to monitor housekeeping data that is vital to the instrument survival; such as, PCB temperatures and SHS temperatures.

A 29 GB data storage unit is used to store the SHOW measurements onboard the instrument. The SHOW detector array is a Raptor OWL 640 camera with 512 x 640 15 um InGaAs pixels. Each acquired image is 650 kB and is stored along with the SHOW housekeeping data. For the ER-2 measurements, we use two primary modes of operation - a 1 HZ sampling mode with a 900 ms integration time and 0.5 Hz sampling mode that uses 1800 ms integration time. The two configurations are used in the case of higher and lower signals that correspond to forward scattering and back scattering signals respectively. Assuming the 1 Hz frame rate, the data storage capacity is met in approximately 12 hours.

Post flight, the raw data is downloaded and processed to generate a Level 0 netCDF file for each specific science mode chosen during the flight. The netCDF file stores the raw image files along with the measurement configuration, such as integration time, image size, UTC time and housekeeping data. The corresponding airplane attitude information is stored in a separate data file provided by NASA. The raw data is processed using the Level 0 to Level 1 processing chain to produce Level 1 calibrated interferograms and Level 1 spectra. The Level 1 data is stored as a netCDF file with each data record containing one minute of data.

4   Thermal Control

An important aspect of the design of the SHOW prototype is the thermal stability of the instrument. The two most temperature sensitive components are the narrowband filter and the SHS, both of which were designed for operation near room temperatures. Since the SHS is not thermally compensated, the heterodyne wavelength of the SHS shifts slightly with temperature. For the SHOW instrument this shift occurs at a rate of approximately ~0.06 nm per degree Celsius. On the other hand, the narrowband filter drifts at a rate of ~0.02 nm per degree Celsius. In practice, the shift in Littrow can be tracked during the flight by lining up (or fitting) the spectrum to the known water vapour absorption features. In any case, it is important to actively control the temperature of the SHOW optics box in order to minimize the extent of the thermal drift.

In the worst case, the inside of the wing pod will reach the ambient temperature where the extreme temperatures can easily reach 40 degrees Celsius on the tarmac and –90 degrees Celsius during flight. Therefore, the SHOW optics are housed within a thermally controlled environment (optics-box) designed and built by ABB Inc that can be operated in near vacuum conditions in the ambient temperature range from -90 degrees Celsius to 40 degrees Celsius. The temperature of the top and bottom plates of the optics-box are actively controlled using resistive heating elements. On the other hand, the SHOW detector is cooled to 0 degrees Celsius using a thermo-electric cooler. Heat generated by the detector is dissipated by thermally strapping the detector to the frame of the outer enclosure of the box, which is thermally connected to the frame of the wing-pod. Prior to flight, the instrument is purged with Nitrogen for over 24 hours to remove moisture and to achieve a stable operating environment. The active thermal control of the

optics-box is designed so that the SHS temperature stabilizes at approximately 23 degrees Celsius. The temperature of the optics-box, the wing-pod window, as well as, the SHOW periscope are monitored during the flight.

5 SHS spectroscopy

As noted in the introduction, non-ideal instrument effects associated with the instrument configuration result in systematic variations in the observed spectra that need to be characterized to perform water vapour retrievals. In order to understand and interpret these variations, it is important to review some key concepts regarding the SHS technique. The most rigorous theoretical background of SHS is provided by Harlander (1991̶0̶) and Harlander (1992̶1̶). Some of the more recent applications of the SHS technique for the remote sensing of the atmosphere can be found in (Englert, et al., 2015). The framework and notation that is used here to describe the SHS technique follows directly from work that was performed at the US naval laboratory (Englert et al., 2004; Englert et al., 2006).

Conceptually, the SHOW instrument operates in a similar manner to a traditional limb imaging FTS instrument such as MIPAS (Fischer et al., 2008); however, in this case, a limb image of the water vapour absorption spectrum is obtained from each frame without vertical scanning and without moving parts in the interferometer. The basic configuration of an SHS is similar to a Michelson interferometer that has its plane mirrors replaced by fixed, tilted diffraction gratings. Field-widening of the instrument is achieved by placing appropriately selected prisms in the arms of the interferometer. Collimated light enters the spectrometer and is incident on a beamsplitter that transmits and reflects (respectively) 50% of the input radiation down the two equal length arms of the interferometer. At the end of each arm, the radiation is incident on a diffraction grating that is tilted by the Littrow angle θ_L. In this configuration, a wavenumber dependent shear $\gamma(\sigma)$ is introduced between the wavefronts that exit the interferometer. An image of the associated wavenumber dependent Fizeau fringes are formed by imaging the grating onto a detector. The technique provides a heterodyned interferogram about the Littrow wavelength with a spatial frequency on the detector given by:

$$\kappa = 2\sigma sin(\gamma) \approx 4(\sigma - \sigma_0) \tan \theta_L \qquad (1)$$

, where the approximation assumes small γ, and σ_0 is the Littrow wavenumber for which $\gamma = 0$. Therefore, the spacing of the Fizeau fringes depends on the wavelength difference from the central heterodyne wavelength (or Littrow wavelength) and the zero spatial frequency fringe is generated by incident radiation at the Littrow wavelength.

For the SHS, each pixel essentially behaves as a separate detector, making the system extremely sensitive to variations in the intensity across the interferogram. Non-ideal flat field calibrations, bad/dead/hot pixels, and pixel-to-pixel non-uniformities (etc.) result in systematic variability in the interferograms and subsequently, errors in the retrieved spectra. The presence of phase distortions in the interferograms can also be important. In general, the basic form of an SHS interferogram, which includes these effects, can be described [see Englert et al., 2006] using Eq.2.

6

$$I(x,y) = \int_{4(-\sigma_o)\tan(\theta_L)\theta}^{\infty} B'(\kappa)\tau(\kappa)\{t_A^2(x,y) + t_B^2(x,y) + 2\varepsilon(x,\kappa)t_A(x,y)t_B(x,y) \cdot cos[2\pi\kappa x + \Theta(x,y,\kappa)]\}d\kappa$$

———— (2)

In this equation, the subscripts A and B denote the two arms of the interferometer, $B'(\kappa) = \left[B\left(\sigma_o + \frac{\kappa}{4\tan(\theta_L)}\right) + B\left(\sigma_o - \frac{\kappa}{4\tan(\theta_L)}\right) \right] \Big/ 4\tan(\theta_L)$ is the spectral radiance, $\tau(\kappa)$ is the filter transmission, $\varepsilon(x,\kappa)$ is the modulation efficiency of the interferometer, $t_A(x,y)$ and $t_B(x,y)$ are the transmission coefficients for each arm respectively and $\Theta(x,y,\kappa)$ is a phase distortion term. We have also substituted $\kappa = 4(\sigma - \sigma_o)\tan(\theta_L)$ where σ is the wavenumber and σ_o and θ_L are the Littrow wavenumber and Littrow angle respectively. For the SHOW instrument, each vertical row in the image is mapped one-to-one to a particular tangent altitude at the limb. Therefore, the interferogram equation is written with explicit dependence on the x and y position.

A robust theoretical framework has been developed by (Englert et al., 2004; Englert et al., 2006) that provides a template to correct the interferogram described by Eq.2. This includes the flat field correction and phase distortion. Once the flat field correction is performed, the corrected interferogram can be written as, Eq. 3 where we have substituted $B''(x,\kappa) = B'(\kappa)\tau(\kappa)\varepsilon(x,\kappa)$.

$$Ic = \int_{4(-\sigma_o)\tan(\theta_L)\theta}^{\infty} B''(x,\kappa)cos[2\pi\kappa x + \Theta(x,y,\kappa)]\,d\kappa \qquad ————$$

(3)

If $\Theta(x,y,\kappa) = 0$, the corrected interferogram has the familiar form of a typical FTS interferogram where κ has replaced σ as the variable. For an FTS, we strictly have $\sigma > 0$, the interferogram equation is easily symmetrized and a simple FFT of the interferogram returns the spectrum (Davis et al., 2001). On the other hand, if $\Theta(x,y,\kappa) \neq 0$, phase errors are introduced and a complex spectrum is returned by the FFT. In general, the phase errors are corrected using phase correction algorithms (Revercomb et al., 2003; Davis et al, 2001). For the SHS, one can perform an analogous correction, called a phase distortion correction, using the technique developed by (Englert et al., 2006). For the SHOW instrument, the ability to perform this correction is limited by aliasing effects, which we now describe.

Observe that an SHS cannot distinguish between signals that are generated by wavenumbers separated by the same amount on either side of Littrow ($\pm\kappa$). This results in aliasing of spectral information from one side of Littrow to the other. Generally, this effect is minimized by choosing a narrow band filter such that the Littrow wavenumber is close to the edge of the passband so we have: $\kappa > 0$. In this case, we can symmetrize Eq.3, the simple FFT recovers the spectrum and a phase distortion correction is feasible. The SHOW filter and SHS were both designed to operate at 20C in the laboratory and there was minimal aliasing at 20C (Langille et al., 2018), however, the ER-2 environment required an operating temperature of 23C and we did not have time or the financial resources to change the filter for the first flight . Consequently, signals with $\kappa < 0$ are transmitted through the passband. In our earlier study, we showed that the impact of these aliased signals on the SHOW retrievals associated with this effect is minimal when the phase term in Eq.2 is negligible ($i.e\ when\ \Theta(x,y,\kappa)\sim 0$).

7

In the prototype version of the SHOW instrument, the aliasing effect is complicated by the presence of a linear phase variation: $\Theta(\kappa, y) = 2\pi \left[\frac{\kappa}{4\tan(\theta_L)} + \sigma_o \right] y\alpha$ in the y direction. This phase variation was introduced by design, to allow for the characterization of the Littrow wavelength in the laboratory using well-isolated monochromatic calibration lines (see Section 9.5). The phase variation is generated by introducing a slight tilt α between the two gratings in the y-direction which has the effect that the Fizeau fringes associated with signals from either side of Littrow are rotated in opposite direction. In the case of a monochromatic line, once the side of Littrow has been identified, the location of Littrow can be experimentally determined using Eq.1 from knowledge of the observed spatial frequency and knowledge of the known calibration line position.

On the other hand, taking the FFT in the presence of the linear phase term, $\Theta(\sigma, y)$ in Eq.3 returns a complex spectrum. For a~~A~~liasing in the SHS, signals with ($\kappa < 0$) exist and introduce additional signals to the complex spectrum that do not exist in the context of the traditional FTS. These signals combine to form an "effective" complex spectrum that contains systematic variations which are difficult to isolate analytically. The best that one can do is directly measure the phase term $\Theta(\kappa, y)$ in the laboratory using a diode laser following the approach discussed by (Englert et al., 2004). In addition, the functional shape of the effective spectrum can be determined by taking measurements of a white light source in a vacuum to remove contamination from water vapour absorption features. Unfortunately, for the SHOW characterization work, we did not have access to a large enough vacuum tank to accommodate the instrument or a laser diode to perform the phase characterization. However, it is worth noting that since the ER-2 flights that are discussed in this paper, a new filter has been installed that is optimized to eliminate this aliasing effect. Therefore, this behavior is not expected to be an issue for future measurement campaigns.

In this paper, we attempt to use an instrument model to predict the aliasing~~is~~ behavior with the assumption that the phase term is linear and given by: $\Theta(\kappa, y) = 2\pi \left[\frac{\kappa}{4\tan(\theta_L)} + \sigma_o \right] y\alpha$. This instrument model is optimized using knowledge of the instrument configuration that is obtained through laboratory characterization work, as well as, flight measurements that are obtained while observing a column of air that is simultaneously sampled using a radiosonde to measure the in-situ water vapour abundance. This approach has several limitations. For example, the ability to adequately capture the systematic variations is limited by our knowledge of the true phase variation, as well as, the knowledge of the atmospheric state. In this paper, we attempt to isolate and characterize the remaining systematic behavior~~errors~~ and we examine the impact on the retrieval of water vapour with SHOW.

6   Noise **considerations**

The primary source of uncertainty~~noise~~ in the SHOW interferogram samples~~s~~ is associated with photon counting statistics and the inter-pixel variability of the detector response. Poisson statistics describe the measurement of the incident photon signal from the atmosphere (S), the thermally generated dark signal (D) and the readout noise (R). On the other hand, inter-pixel variations associated with incomplete flat field corrections (optical variations due to dust, aberrations etc.) and variations in the relative pixel response (known as photo-response non-uniformity (PRNU)(Ferrero et al., 2006; Schulz et al., 1995) across the detector introduce an inter-pixel variability (spatial noise) to the samples in the interferograms. This variability introduces uncertainty in the interferogram samples that is not described by photon counting statistics.

Generally speaking, flat field calibrations are performed to remove the intensity variations associated with optical effects and a non-uniformity correction is performed to minimize the impact of the relative pixel response. Ideally, the relative pixel response is corrected separately from the flat field variations by characterizing the PRNU prior to installing the camera in the instrument. For the SHOW prototype, this was not possible since the camera was installed by ABB prior to shipping the instrument for the

8

calibration work. ~~In this case, the flat field correction is performed in combination with the non-uniformity correction by obtaining white light flat field measurements at signal levels that closely match signal levels recorded during flight.~~ We characterize t~~T~~he approximate PRNU of the camera in the presence of optical effects ~~is characterized in the laboratory~~ by performing a two-point non-uniformity correction using white light flat field measurements. While this approach cannot be used to completely separate the flat field term from the non-uniformity correction term it does give a good indication of the approximate relative pixel response of the detector. For the measurements presented in this paper, the flat field correction is performed in combination with the non-uniformity correction by obtaining white light flat field measurements at signal levels that closely match signal levels recorded during flight. Therefore, each flat field frame corrects for the optically related variations as well as the inter-pixel response variations.

The SHOW flat field calibrations are performed using the flat field approach developed by (Englert et al., 2006) and described in Section 9.4. After applying the flat field corrections, the remaining inter-pixel variations are associated with the~~the~~ residual photo-response non-uniformity of the detector. Techniques have been developed to characterize these variations in the interferogram images in the laboratory, as well as, their influence on the spectral measurements. In the interferogram dimension, these variations add a relative error to the signal at each sample. Assuming a random relative pixel variation $\frac{\delta S}{S}$, this results in a~~n~~ SNR in the spectral measurements of between $\frac{S}{\delta s}\sqrt{N/2}$ and $\frac{S}{\delta s}\sqrt{2/N}$ (where N is the number of interferogram samples, S is the signal level at the interferogram center burst and δS is error in the signal) due to multiplex noise propagation for the case of a monochromatic line and a continuum respectively (Englert et al., 2006). For the SHOW instrument, we observe absorption features within a micro-window that is isolated with a 2 nm bandwidth filter. In this case, the multiplex noise associate~~d~~ with a random relative pixel response in the spectral measurements is closer to the continuum case.

Although the relative pixel response behaves quasi-randomly (across several hundred interferogram samples) in the interferogram domain, the relative response is a fixed pattern and the observed variability is fixed across the interferogram row for a fixed intensity level. The effect can be viewed as adding or subtracting a small fixed perturbation to each sample (a similar effect is observed with digitization noise in FTS (Davis (2001)). The FFT of this distribution produces an associated ~~Therefore, the observed noise~~ pattern ~~is also fixed~~ in spectral space that remains fixed ~~in wavenumber space~~ for the same signal level. This fact is used during the Level 0 to Level 1 data processing in order to characterize th~~this~~is effect~~rror~~. It is shown that a multiplicative secondary correction is feasible for the SHOW ER-2 airplane measurements that effectively minimizes the impact of this variability.

Assuming a negligible inter-pixel variability, the total uncertainty σ_j in the recorded interferogram signal (measured in D~~igital Numbers (DN))~~ is approximated by Eq.4 where g is the gain factor (electrons/DN) used to convert the measured DN to electrons and I_j is the signal at a particular sample.

$$\sigma_j = \sqrt{\left(\frac{I_j + D}{g} + R^2\right)} \tag{4}$$

9

In the ideal case, the dark signal and readout noise are small relative to the photon noise. We cool the camera to reduce the dark signal and calibrate the remaining signal by subtracting a dark frame obtained by averaging hundreds of measurements. The readout noise and noise associated with the relative pixel response are characterized in the laboratory and are shown to be negligible compared to the Poisson noise. The signal to noise ratio (SNR) in the spectral samples is estimated using Eq.5 (Brault, 1983).

5

$$SNR_\sigma = \sqrt{\frac{1}{N} \frac{B(\sigma)}{\overline{B_e}}} SNR_x \tag{5}$$

In this equation, N is the number of samples in the interferogram, SNR_x is the signal to noise ratio in the interferogram samples, $B(\sigma)$ is the spectral density at wavenumber, σ and $B_e = \frac{1}{2}[B(\sigma) + B(-\sigma)]$.

10 **7   Level 0 to Level 1 processing**

The SHOW Level 0 data consist of the raw SHS interferograms and noise estimates, as well as, the airplane attitude information (time, geographical location, speed, pitch, yaw etc). These raw data are processed in several stages as shown diagrammatically in Figure 2~~3~~. In the first stage, raw interferograms are corrected for dark signal and bad pixels (of which, there are very few) are removed by performing a nearest neighbor interpolation ~~in the vertical dimension.~~ Then, the flat field correction is~~and non-~~

15 ~~uniformity correction are~~ applied. Finally, the DC bias is removed from the interferogram signal in order to obtain the final Level 1A interferograms. In the second stage, corrected spectra are extracted from these interferograms. This consists of two primary steps. The first step is the application of Hanning apodization. The second step is the application of an FFT to each interferogram row to obtain the complex spectrum corresponding to each line of sight in the image. The amplitude of the complex spectrum is taken at each row to form a vertically resolved image of the water vapour absorption spectra.

20 As noted earlier, the airplane pitch varies as a function of time and this variation results in each row of the detector observing a range of tangent altitudes at the limb. As part of the Level 0 to Level 1 processing, the airplane attitude information is used to obtain a mapping of the interferogram rows in each image to the corresponding line of sight (LOS) at the limb.~~correct the interferograms for the "jitter" associated with the variation in the airplane pitch to generate a vector that contains corrected lines of sight (LOS) corresponding to each row in each image.~~ Therefore, the Level 1 data consists of the calibrated interferograms and

25 spectra, as well as, the geometrical information required to map each interferogram row in the image to a particular tangent altitude at the limb.

The pitch variation of the airplane is also used to obtain a secondary correction that can be applied to minimize the impact of the systematic variations in the spectra associated with the residual relative pixel response in the interferogram signals. In practice, over a short period of time, it can be assumed that the radiance level at a particular altitude remains constant. Therefore, as the

30 airplane pitches up and down, the signal level crosses different rows and this variability can be tracked in the spectral images. This allows this detector related systematic variability to be characterized and corrected in the measurements. This correction, discussed in detail in section 11.2, produces Level 1C corrected spectra.
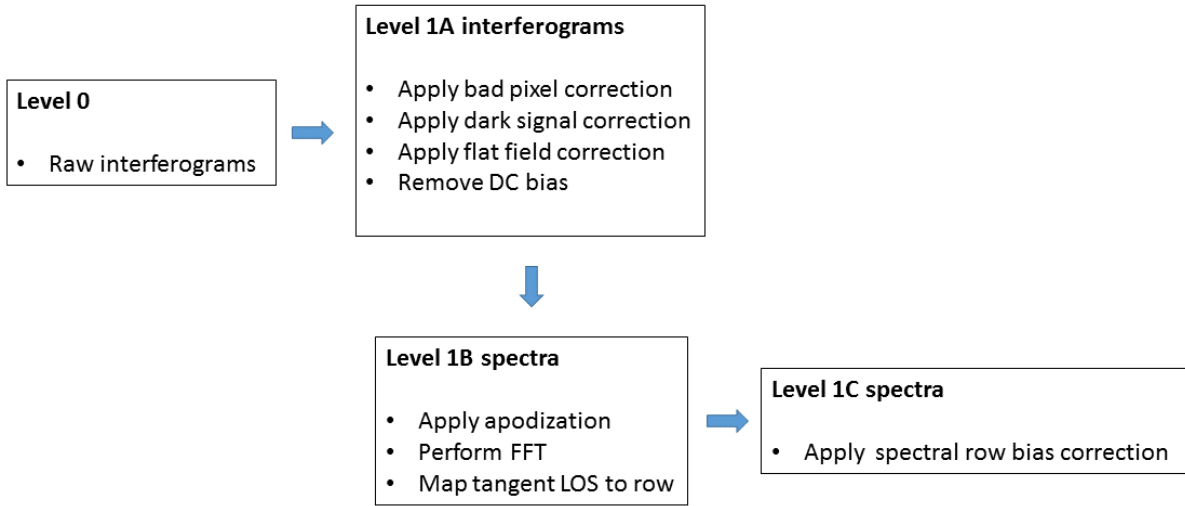
Figure 2 SHOW Level 0 to Level 1 processing chain

8　Retrieval approach

The SHOW retrieval algorithm utilized to process the flight data builds on the SHOW sensitivity study presented in a previous
publication (Langille et al., 2018). The approach utilizes the non-linear optimal estimation formalism described by ~~in~~ (Rodgers~~(,~~
2000) to extract water vapour profiles from the vertically resolved spectral measurements. The retrieval performs an iterative step
given by Eq.6 in order to find the next best estimate of the state vector, x . In this case, x is the retrieved water vapour profile as
the logarithm of number density.

$$x_{i+1} = x_i + \left(K_i^T S_y^{-1} K_i + R + \lambda I\right)^{-1} \{K_i^T S_y^{-1}[y - F(x_i)] - R(x_i - x_a)]\} \tag{6}$$

In this equation, y is the measurement, i is the iteration number, x_a is the apriori state vector, F is the forward model, K is the
Jacobian matrix, S_y is the measurement covariance matrix, λ is a Levenberg-Marquart damping term, and R is a second order
regularization matrix that is used to constrain the vertical resolution of the retrieval. The measurement vector, y, contains the
logarithm of vertically resolved spectra that are normalized by their spectral mean. The forward model, F, is composed of the
SASKTRAN radiative transfer model (Bourassa et al, 2007; Zawada et al., 2015; Zawada et al., 2017) and the SHOW instrument
model (Langille et al., 2018). SASKTRAN is used to produce high resolution, forward modelled radiances and the SHOW
instrument model is used to simulate SHOW measurements $F(x_i)$ from these modelled radiances.

The retrieval of the water vapour profile requires a forward model that accurately predicts the instrument behavior. For a typical
FTS instrument, this is achieved by convolving the instrument line shape (ILS) with a high-resolution forward modelled spectra.
Generally, the ILS is obtained by taking measurements with a quasi-monochromatic calibration source. For the SHOW instrument,
the convolution with the ILS at each row in the instrument produces a spectrum that does not capture the systematic variations
associated with the aliasing effect discussed in Section 5. In the current work, we have performed a detailed instrument
characterization to obtain appropriate calibrations that are utilized during Level 0 to Level 1 processing and we optimized an
instrument model to accurately predict the systematic variations in the observed spectra. In Section 12, we utilize this model in

11

the retrieval to extract water vapour profiles from measurements that were taken during a SHOW ER-2 engineering flight on July 18 th, 2017.

9 Laboratory Characterization

9.1 Overview

5   The SHOW instrument consists of several primary components. This includes, the field widened SHS, the optical system (anamorphic fore optics and exit imaging optics), and the FPA. The configuration of these components introduces variations in the interferogram signals that must be characterized and calibrated in order to obtain the highest quality measurement. As noted in Section 5 and Section 6, treating each pixel in each interferogram row as a separate detector requires knowledge of the behavior of each pixel in the array. Converting the raw Level 0 interferograms to corrected Level 1 spectra requires detailed knowledge of

10   the SHS configuration and the imaging quality of the instrument. In addition, knowledge of the vertical field of view is essential to mapping the rows of the detector to the tangent altitudes at the limb. We also utilize the knowledge of the instrument configuration to develop the optimized instrument model that is applied to predict the systematic effects associated with aliasing.

Detector parameter	Value
Bad pixels	46
Dark signal (1800 ms exposure)	161 DN*
Dark noise (single frame)	2-5 DN *
Dark noise (calibration frame)	< 0.5 DN
Bias	1974 DN *
Gain	45.7 e / DN *
Readout Noise	165.7 e (3.62 DN)
Non-linearity	< 1%
PRNU	< 1.2 %

15   **Table 2 SHOW OWL 640 detector parameters characterized in the laboratory. Values with a * are representative of the average behavior across the array**

9.2 Detector characterization

The FPA utilized in the SHOW-ER2 instrument is a commercial off the shelf (COTS) OWL 640 InGaAs detector manufactured

20   by Raptor Photonics. The FPA has 640 x 512 pixels covering an area of 0.74 cm^2, 14 bit ADC, with a > 80 % quantum efficiency in the SHOW passband. The temperature of the FPA is controlled to 0 degrees Celsius using a built in thermoelectric cooler in order to reduce the dark signal. As noted earlier, bad/hot pixels, Poisson noise and the variation in the relative pixel response in the interferogram samples results in noise that propagates into the spectra. These parameters have been characterized in the laboratory and are summarized in Table 2.

12

A total of 46 bad/defective pixels were identified within the imaged field of view (FOV) of the SHOW FPA. A handful of pixels around the edge of the FPA also exhibit odd behavior and have been identified as bad pixels. During data processing, a bad pixel map is utilized to identify these pixels and we interpolate over these pixels using nearest neighbor interpolation. To correct for dark signal we obtain average dark calibration frames for a range of exposure times by averaging several hundred dark measurements at each setting. The observed dark signal was found to vary non-linearly in the SHOW operating range using less than 2-second exposures. For the 0.5 Hz acquisition setting, the mean bias level of the detector was found to be approximately 1974 DN and the average dark signal across the frame using the 1800 ms exposure setting was found to be ~161 DN. The measurement noise at a single pixel in the average dark frames was found to be < 0.5 DN. An average dark frame is subtracted from each flight measurement obtained with the same exposure time.

Propagation of error in the signal measurement using Eq.4 requires knowledge of the gain, readout noise error and the magnitude of the relative pixel response. The approximate gain (electron/DN) and readout noiseerror (R), calculated from an experimentally determined photon transfer curve, are 45.7 electrons/ DN and 165.7 electrons respectively. These values match closely with the values specified by the manufacturer of 39.67 electrons/DN and 150 electrons. Therefore, the readout noise is estimated to be 3.62 DN (~ 4 DN). The detector was also found to exhibit a slight non-linear response to incident radiation. This non-linearity is on the order of 1 % across the full dynamic range of the detector. In practice, it was found that the non-linearity changed slightly with the detector operating settings (PCB temperature, FPA temperature etc.).This made characterizing the effect difficult; therefore, the non-linearity is not corrected in the calibration or flight measurements.

The relative pixel response was characterized in the laboratory by performing a two-point non-uniformity correction using flat field measurements. This approach inherently includes some variations from optical effects and iIt was found that the PRNU associated with an uncorrected frame results in a ~1.2 % relative variation in the interferogram samples. In addition, tThe presence of the 1% non-linearity makes performing separating the a non-uniformity correction from the flat-field correction difficult, since the extent of the linearity was found to change with the level of incident radiation, as well as, exposure time. In the current work, the effect of the relative pixel response is minimized by obtaining flat field measurements using intensity levels that closely match intensity level of the observations. Instead of performing the non-uniformity correction, we correct the flight measurements using the flat field correction that best corrects for the relative pixel response variations at the observed intensity.

9.3   Imaging system characterization

The 4 degree vertical field of view by 5.1 degree horizontal field of view of the limb observed by SHOW is collimated at the entrance aperture and is passed through the SHS spectrometer using the anamorphic entrance optics. The optics are arranged to produce a vertically resolved image of the limb at the grating location, whereas, the horizontal scene is averaged at the grating location. At the output, the detector and exit optics are mounted on a common holder that is aligned with the monolithic SHS. The aft imaging optics are arranged to form an image of the SHS interferograms conjugate to a vertically resolved image of the limb at the detector. This is achieved by imaging the grating at the detector location. Therefore, the magnification of the exit optics determines the scale factor relating the observed spatial frequency on the detector to the spatial scale given by Eq.1. This parameter is used to define the bin spacing between the interferogram samples and ultimately the frequency scale of the observed spectra. In addition, this configuration establishes a one- to-one mapping between interferogram row and tangent altitude at the limb. Accurate knowledge of these parameters is required for Level 0 to Level 1 processing.

9.3.1 Exit optics magnification

An example raw image obtained by illuminating the entrance aperture with white light from an 8" LabSphere integrating sphere is shown in Figure 3 (a). This image corresponds to the full frame provided by the (512 pixel x 640 pixel) SHOW detector. The contrast has been adjusted in Figure 3 (b) to highlight the SHOW FOV defined by the image of the aperture at the grating (red) and the edge of the grating illuminated by scattered light (blue). The magnification of the exit optics is calculated by measuring image size of the grating (blue) on the FPA and taking the ratio of the known grating height to the measured grating height. For the SHOW ER-2 configuration, the magnification was measured to be 0.22.

Figure 3 An example dark corrected white light image (a) and the same image with a color scale that shows the SHOW field of view in red and the edge of the grating in light blue (b)

9.3.2 Field of view

Scattered sunlight from the limb is collimated at the input to the instrument, therefore, a particular off-axis angle illuminates a single row on the SHOW detector. As the off-axis angle is varied, the illuminated row also changes. Knowledge of this mapping is critical in order to determine accurate lines of sight that are used to obtain accurately georeferenced spectra. The SHOW field of view only fills a portion of the detector frame – called the SHOW FOV from pixel 197 to 491 in the vertical and 9 to 502 in the horizontal. The outer edges of the red region in Figure 3 (b) have been removed from the FOV to ensure no edge effects are present. Later in the paper, only this FOV is shown, rather than the full frame that is shown in Figure 3.

This vertical field of view was characterized in the laboratory at ABB Inc. by directing a well-collimated beam into the instrument entrance at different incident angles. For each incident angle, the position of the illuminated row was determined. A plot of the illuminated row number versus incident angle is shown in Figure 4. A LMS linear fit was performed to determine the slope of $m = -79.968 \ pixels/deg$. The R^2 fitting coefficient has a value of 0.9996. From these measurements, the maximum angular vertical resolution of the measurement is ~0.0126 degrees. In practice, the FWHM of the imaged row was found to be approximately 1.4 pixels. Having 0.0126 degrees/pixel, this gives an instantaneous FOV for each sample of 0.0176 degrees. If we take the radius of the Earth to be 6371 km and assume the airplane to have zero pitch and an altitude of 21.343 km, this angular resolution converts to a tangent point vertical resolution of approximately 151 m at 2.5 km and 20 m at 21 km.
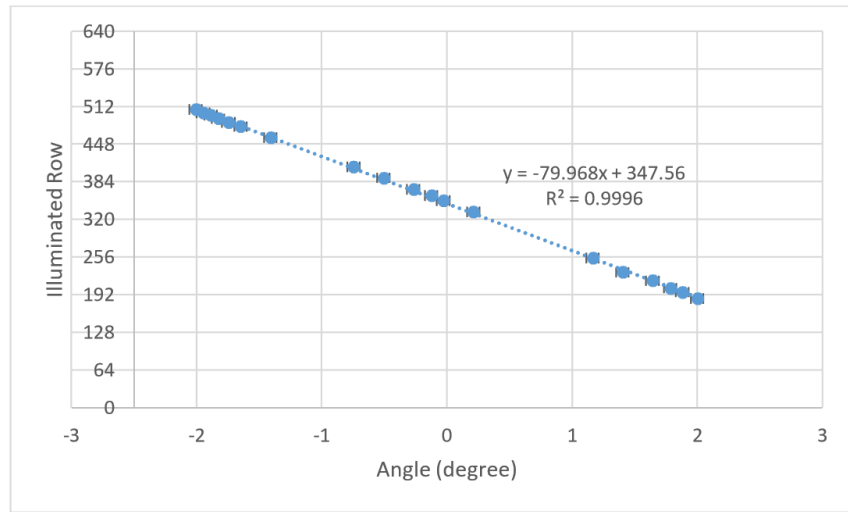
14

Figure 4 SHOW measured vertical FOV

9.4   Flat field calibration

The flat field calibration is performed using the approach developed by (Englert et al., 2006) and takes into account the flat field variations from the optics, as well as, slight differences between the two arms of the interferometer. First, the entrance aperture of the instrument is uniformly illuminated with white light. Then, separate images are obtained with each arm of the interferometer (arm A and arm B) blocked using an opaque material. Hundreds of measurements are obtained, dark corrected and then averaged to minimize the Poisson noise. Taking these images to be I_A and I_B respectively, the flat field term associated with optical effects is given by Eq.7.

$$FF1 = I_A + I_B \qquad\qquad (7)$$

The correction term associated with relative differences between the two arms of the interferometer is given by Eq. 8.

$$FF2 = \frac{2\sqrt{(I_A \cdot I_B)}}{(I_A + I_B)} \qquad\qquad (8)$$

Corrected interferogram images are obtained by dividing the dark corrected SHOW interferogram images by Eq.7, subtracting off the mean DC signal at each row and then dividing by the image generated using Eq.8.   If the FF2 term is close to 1, it can be neglected. In this case, the interferograms are corrected by dividing the interferograms by $\frac{FF1}{FF1}$ and then subtracting off the mean DC signal level at each row in the image.

Example FF1 and FF2 correction images (normalized) that were obtained in the laboratory using an 8" aperture LabSphere integrating sphere are shown in Figure 5 (a) and Figure 5 (b) respectively. Optical variations of ± 5 % are observed in the FF1 image and the presence of dust and vignetting in the optical system is also apparent. The FF2 image is close to 1 and the variations are on the order of 0.1 %. Therefore, the SHOW SHS can be assumed to be well-balanced and the dominant correction term is due

15

to optical flat field variations. This demonstrates the high quality of the optical system that was designed and manufactured by ABB Inc. and York University (Center for Research in Earth and Space Science), as well as, the interferometer that was designed and fabricated by Light Machinery, Inc., Ottawa, Ont.

For the SHOW data processing, we take the FF2 term to be equal to 1 and just apply the FF1 term which corrects for the optical variations. This also corrects for the inter-pixel variations associated with the relative pixel response since the flat field images are obtained without performing a non-uniformity correction. However, the measured relativenon-uniform pixel response of the detector depends on the intensity of the incident radiation. Therefore, we obtain flat field correction images within the range of the intensity levels that are expected during flight. The flat field correction terms shown in Figure 5 were obtained with roughly quarter of the full well and an exposure time of 250 ms. The flight measurements are corrected using the calibration frames that are obtained at the closest matching intensity level. This ensures that the flat field correction also appropriately corrects for the relative pixel response.
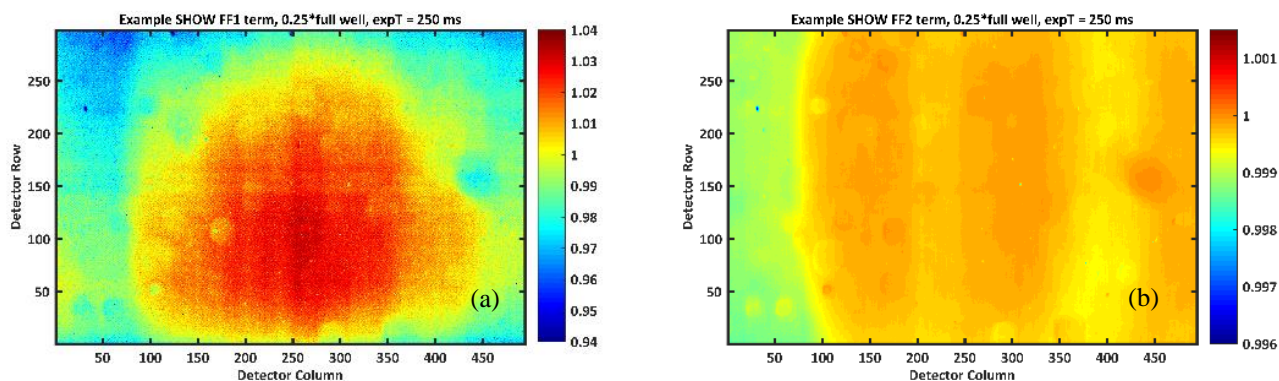

Figure 5 Example SHOW flat field correction images obtained using a 250 ms exposure time to obtain a quarter full well signal. The FF1 term is shown in (a) and the FF2 term is shown in (b).

9.5   SHS characterization

The primary parameters that are used to characterize the configuration of the SHOW SHS system are the Littrow wavelength and the spectral resolution. The SHS is designed to have an unapodized spectral resolution of: $\delta\lambda_{resolution} \simeq 0.02$ nm and a Littrow wavelength of 1363 nm (in air) at 20 degrees Celsius (Langille et al., 2017). This spectral resolution was calculated assuming that the full grating width is utilized. For the current version of the SHOW instrument, the imaged portion of the grating cross-section is slightly smaller - roughly 3.37 centimeters (494 pixels x 15um/pixel / 0.22). This gives a maximum theoretical resolution of ~ 0.03 nm where the maximum resolving power of the instrument is given by (Harlander, 1991): R= $4W_g\sigma\sin(\theta_L) =$ $4\left(\frac{3.37}{\cos(28.5)}\right)(7336.8\ cm^{-1})\sin(28.5) = 53698$. In addition, the SHOW interferograms are apodized using Hanning apodization which results in a factor of ~1.8 increase in the spectral resolution (Davis et al., 2001). It is also known that the Littrow wavelength drifts in temperature by roughly 0.06 nm/ deg Celsius, and the flight instrument is operated at 23 degrees Celsius. Therefore, it was anticipated that the Littrow wavelength is higher than the design value by close to 0.18 nm.

Characterization of the SHS configuration was performed in the laboratory by uniformly illuminating the entrance aperture of SHOW with light from a Krypton calibration lamp. The Krypton lamp contains a well-isolated line at 1363.422 nm (in air) that lies within the SHOW passband. Measurements with the Krypton lamp were used to measure the Littrow wavelength at the operating temperature and determine the spectral resolution of the instrument. Measurements of a white light source were utilized

to characterize the full instrument. The instrument was purged with dry nitrogen for 24 hours prior to the characterization work in order to ensure a stable operating environment and an SHS temperature of 23 degrees Celsius.

An example image of the Krypton fringes obtained with SHOW is shown in Figure 6 (a). The rotation of the fringes is due to the presence of a slight tilt in the y-direction between the two gratings discussed in Section 2. The counter-clockwise rotation indicates that the Krypton line is located to the long wavelength side of the Littrow wavelength. The Littrow wavelength was determined by performing a non-linear least-mean squares cosine fitting to an interferogram row to determine the spatial frequency of the observed fringes. The spatial frequency of fringes measured on the detector surface is given by Eq.1, therefore the Littrow wavelength is calculated from the observed spatial frequency and knowledge of the target wavelength. Using this approach, the spatial frequency of the observed fringes was found to be 1.98 fringes/cm and the Littrow wavelength was determined to be $\lambda_L =$ 1363.62 nm (in vacuum – calculated using the Ciddor equation) and 1363.25 nm (in air). This is close to the expected increase in Littrow based on the design of the instrument.

The Krypton spectrum, obtained by taking an FFT of an interferogram row close to the center of the image is shown in Figure 6 (b) in blue. The theoretical spectrum is obtained by taking the FFT of a simulated interferogram that is modelled by assuming a perfectly monochromatic line at 1363.422 nm is shown in red. Hanning apodization was applied to both interferograms prior to taking the FFT. The spectral resolution was estimated by determining the FWHM from a Gaussian fitting to the spectral line. The FWHM of the measured spectrum is 0.0516 nm and the FWHM of the theoretical line is 0.0451 nm. The two values differ by roughly 0.0065 nm and the difference is likely due to slight distortions in the spectrum that are introduced by the imaging optics.
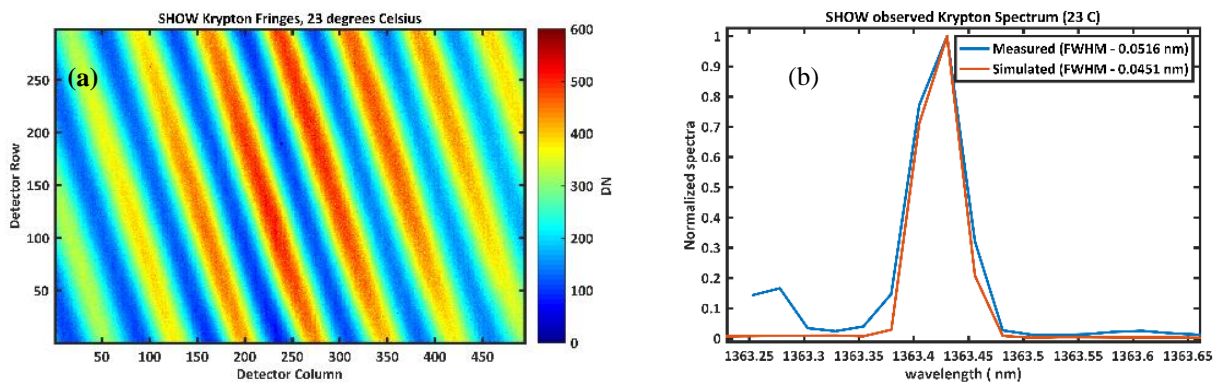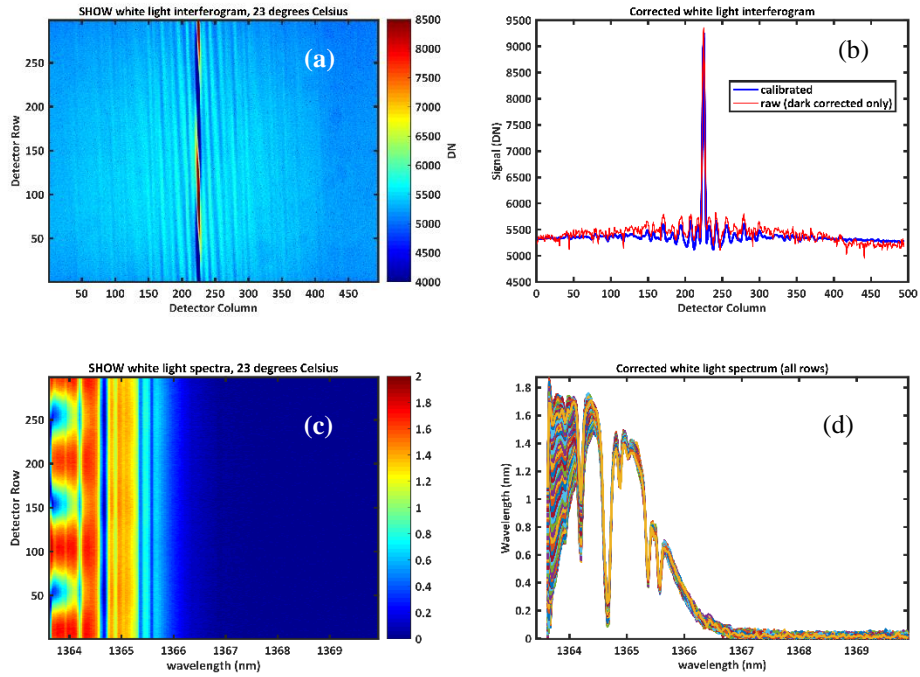

Figure 6 Krypton fringes observed with the SHOW SHS instrument at 23 degrees Celsius (a) and the spectrum obtained by taking the FFT of the interferogram at row 150 (b). The measured spectra is shown in blue and the simulated spectrum is shown in red.

9.6   Performance

The performance of the SHOW flight instrument was characterized in the laboratory by uniformly illuminating the instrument with white light from an 8 " aperture LabSphere integrating sphere. One hundred and fifty interferogram images were obtained and then processed using the Level 0 to Level 1 processing chain to obtain calibrated interferograms and corrected spectra. An example calibrated interferogram image is shown in Figure 7 (a) and an example raw and corrected interferogram row is shown in Figure 7 (b). The corresponding spectral image formed by taking the FFT of each row of the interferogram image is shown in Figure 7 (c). All rows of the spectral image are plotted in Figure 7 (d). The SHOW filter is a narrow band filter (see Langille et al., 2018) with a 2 nm bandwidth and a peak transmission of 0.77 that is centered at 1364.52 nm at 23 degrees Celsius. The shape of the filter passband is clear in Figures 7(c-d) and the sharp cut off in the spectra on the left hand side corresponds to the location of the Littrow wavelength at 1363.62 nm (in vacuum).

17

The goal here is to examine the quality of the corrected interferograms and spectra in order to estimate the remaining variability in the images that is not associated with photon noise. It is difficult to characterize the noise in each individual interferogram and spectral row since samples within each row contains variations associated with the source. However, the vertical dimension of the interferogram and spectral image is expected to be smooth since the illumination within the field of view is uniform and the optical configuration is anamorphic. Therefore, we examine the inter-sample variability in each vertical column in order to characterize the noise.

Figure 7 Example corrected white light interferogram (a) and a row cut taken at row 201 (b). The dark/bad pixel corrected signal is shown in red and the dark/bad pixel and flat field corrected signal is shown in blue. Example spectral image obtained by taking an FFT of each row in the interferogram image (c) and all rows of the spectral image plotted (d)

Ideally, we want to operate the detector with the minimum amount of uncertainty in the measurements noise. Taking the variance on any particular interferogram sample to be given by Eq.4, we estimate the noise floor by averaging the images to reduce the photon noise. For an average signal level of ~ 5300 DN and averaging 192 images, we find the noise floor for the samples to be roughly $\sigma = 10.7$ DN. Taking into account photon noise, dark noise, and readout noise, we estimate the observed variability noise associated with the residual relative pixel response (using Eq.4 to estimate the variances from the known sources) to be approximately 0.2 %. This is an order of magnitude smaller than the PRNU of an uncorrected image (~1.2 %) and demonstrates the quality of the combined flat field calibration and non-uniformity correction.

On the other hand, the observed variations in the spectral image are is slightly more difficult to quantify. We expect that each row in the image will record the same spectrum; however, this is not what is observed. Instead, we observe a strong modulation of the spectra in the vertical dimension. This variation is systematic and is due to the presence of a small relative tilt between the two gratings in the SHS combined with aliasing of spectral information from the left side of Littrow into the right hand side (see Section

18

5). This systematic variation is characterized in the next section by optimizing an instrument model that predicts the effect. Here we focus on the propagation of error from the interferogram samples to the spectra.

Since each row of the interferogram should produce the same spectrum, the presence of the inter-sample variability in the spectra is determined at each wavelength (column) by subtracting off a high order polynomial (n= 8) from each column cut to remove the modulation and then determining the standard deviation of the subsequent distribution. Following this approach, we determine the inter-sample variability in the vertical dimension to be roughly 3 %. This variability~~noise~~ is primarily associated with the propagation of multiplex noise due to the residual relative pixel response in the interferogram samples; therefore, the spatial variation is fixed within any particular image (see Section 6). In Section 11.3, we show that the pitch of the airplane and the subsequent changes in intensity level across the rows, can be used to further correct this variation.

9.7   Model optimization

The aliasing effect that is shown Figure 7 (c-d) is extremely problematic for the retrieval of the vertical distribution of water vapour. The observed variation modulates the true spectrum and the amplitude of the effect depends on the relative position of the filter center peak and the Littrow wavelength, as well as, the amount of absorption due to water vapour. In order to implement the retrieval approach described in Section 8, a forward model is constructed that captures these perturbations. The model takes a simulated high-resolution spectrum $B(\sigma)$ that is simulated using SASKTRAN and constructs the interferogram image given by Eq.9.

$$I(x,y) = \int_0^\infty \tau(\sigma)B\,(\sigma)[1 + \text{+}\cos[2\pi\big(4(\sigma - \sigma_0)\big(x - x_{shift}\big)\tan(\theta_L) + \alpha\sigma(y - y_{shift})\big)]]d\sigma \qquad (9)$$

This equation follows directly from Eq. 3, where we have assumed the interferogram has been dark and flat field corrected and we have included a linear phase variation in the y-direction associated with the grating cross tilt, as well as, a relative shift to both the x and y dimensions. These relative shifts account for a lateral detector misalignment or a shift in the relative axis of rotation between the two gratings. Each row of the modelled and measured spectral images is normalized by the mean of the full spectrum at each row to ensure the spectra are on the same scale.

As a first step prior to performing the optimization, the Littrow wavelength ~~(σ_o)~~ was determined by aligning the observed water absorption features in the white light spectra with known features in the high-resolution simulated spectrum. In order to model the correct amount of absorption, we need to estimate the water vapour abundance, as well as the path length. For this calculation, we assumed standard temperature and pressure and a relative humidity of 50 %. The path length was then manually adjusted in the model until the depth of the absorption features were closely matched.

The instrument model was then optimized by iteratively adjusting the center peak of the filter, the value of the grating rotation α, and the relative shifts in both the x and y positions on the grating. After each iteration, the spectral agreement between the model and the SHOW white light measurements was checked by calculating the average absolute difference in the spectra signal, as well as, the standard deviation of the difference in the systematic variations introduced by the aliasing effect. The parameters used on

19

the iteration which produces the best agreement between the model and SHOW is then taken to be the best representation of the true configuration of the instrument for this measurement.

Figure 8 (a) shows the final agreement between the instrument model and the SHOW white light laboratory measurements as a result of this process. The normalized interferograms are shown in the top panel and the normalized spectra are shown in the bottom panel. Here, the image has been rotated so the lower rows correspond to the lower altitudes when compared to the flight measurements. For this set of measurements, the Littrow wavelength was found to be 1363.62 nm (in vacuum), the filter center was found to be 1364.52 nm, the grating tilt was found to be -5.108 x 10 $^{-5}$ radians, and the x, y offset was found to be (15.77 pixels, -15.52 pixels) respectively. Two rows corresponding to the best and worst agreement are shown in Figure 8 (b) and Figure 8 (c). The blue line in both of these figures shows the forward modelled high-resolution water vapour spectrum. There is reasonable agreement between the model and measurement and the model appears to capture the shape of the systematic variations associated with aliasing. Unfortunately, it is difficult to quantify the differences between the spectra since the amount of water vapour in the laboratory (and our lack of knowledge of it) has a strong influence on the amplitude of the variations; especially close to the absorption features. We will come back to the optimization in the case of flight data where we have more information regarding the water vapour abundance via the in-situ radiosonde measurements.
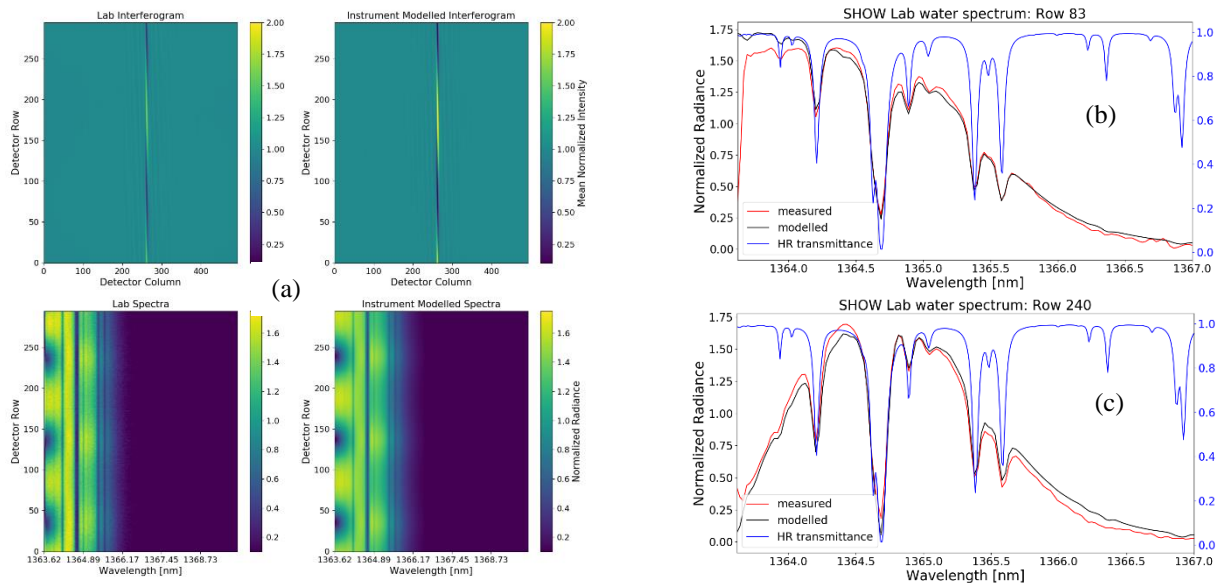

Figure 8 Instrument model optimization using laboratory data. Comparison between the modelled and measured interferograms and spectral images (a). Example spectral row cuts at row 83 (b) and at row 240 in (c). The measured spectrum is shown in red and the modelled spectrum is shown in black. The normalized high-resolution spectrum that was utilized in the instrument model is shown in blue.

10   SHOW demonstration flight on July 18th, 2017

On July 18th, SHOW flew an engineering flight on the ER-2 airplane from the Air Force Flight Research Center (AFRC) located near Palm Dale, California in order to test several aspects related to the performance of the instrument. The flight track of the ER-2 airplane for the entire engineering flight is shown in Figure 9 (a) overlaid on top of a map showing the geographic location of the aircraft above California. The airplane attitude information during the flight is shown in Figure 9 (b) where the recorded airplane pitch, roll, heading, height, latitude and longitude are shown as a function of time. Portions of stable flight correspond to regions where the pitch and roll are relatively close to zero; although small variations in the pitch are always recorded – even for the most
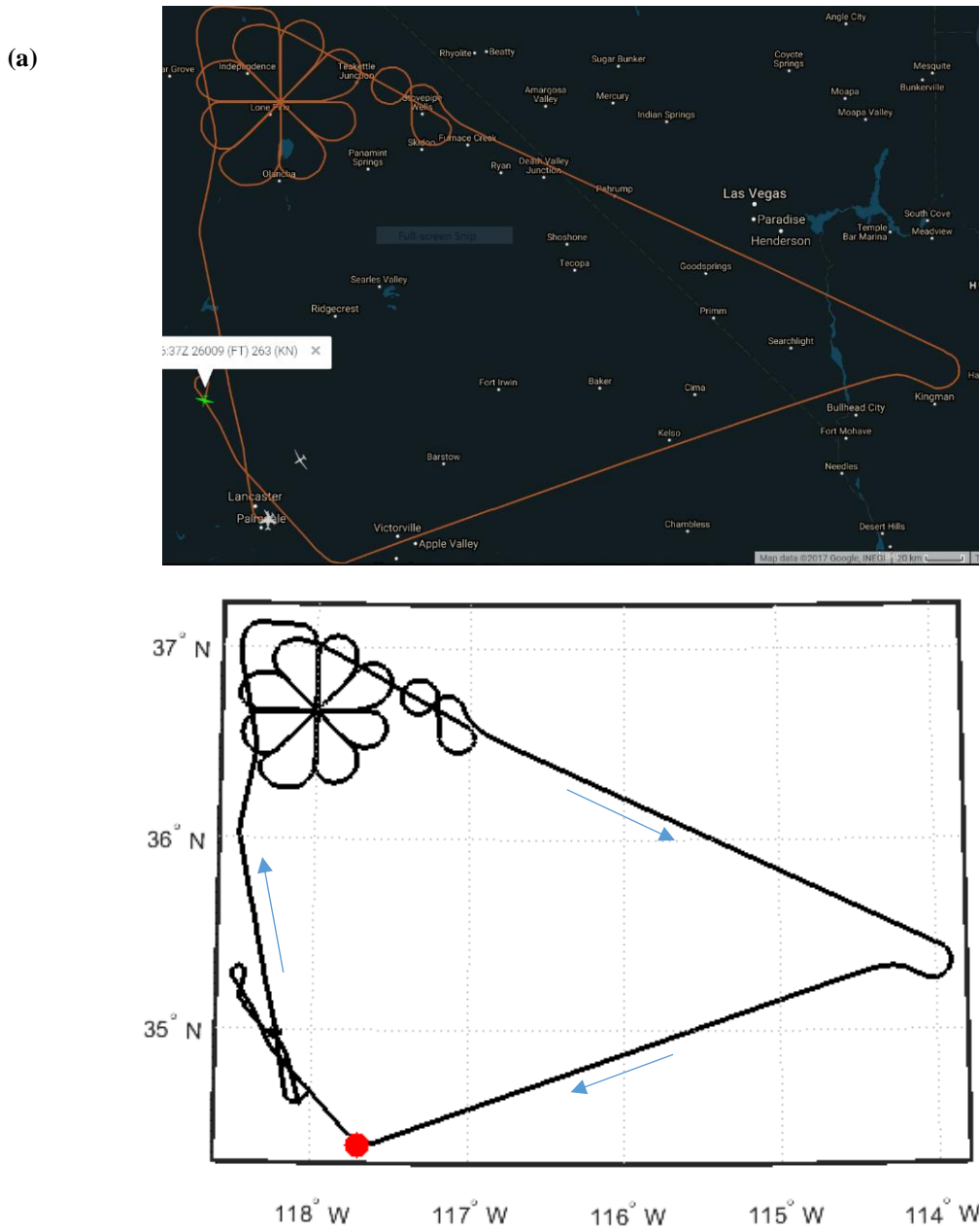
20

stable portions of the flight. In practice, this airplane attitude information is used to map the interferogram rows to tangent altitude at the limb for each frame obtained with the SHOW instrument.

Prior to the flight, the instrument was mounted in the wing pod (detached from the airplane) and was purged with nitrogen for over 24 hours in order to remove moisture and to achieve a stable thermal environment (23 degrees Celsius) within the instrument optics box. About an hour before the flight, the instrument was turned off and wheeled out to the tarmac and the wing pod was mounted to the airplane. The temperature on the tarmac was roughly 25 degrees Celsius during this time and the SHOW instrument and thermal control was powered on just before takeoff. The recorded instrument temperatures inside the optics box (SHS and enclosure) and inside the wing pod (periscope and wind pod window) during the complete engineering flight are shown in Figure 10(a) and Figure 10 (b) respectively.

(a)

(b)

Figure 9 ER-2 flight track where the red dot indicates the location of the JPL Table Mountain facility (a) and the corresponding flight attitude information (b) for the engineering flight performed on July 18, 2017. The blue arrows indicate the forward direction of the airplane during the flight.

We can see that the temperature inside the wing pod at the beginning of the flight is relatively warm, corresponding to the warm morning temperatures on the tarmac at takeoff. Once the airplane reached its cruising altitude, the temperature dropped steadily and reached close to – 40 degrees Celsius at the wing pod window and close to – 20 degrees Celsius at the SHOW periscope. During this same period, the temperature inside the optics-box initially increased to 24.5 degrees Celsius and then steadily decreased towards 22 degrees Celsius.

The thermal control is designed so that the heaters on the top and bottom of the optics-box keep the top and bottom center area of the SHS close to 23°C; however, we observed that the SHS temperature settles closer to 24.3 degrees Celsius once the optics-box temperature begins to stabilize. Since the thermal time constant of the SHS is quiteiet long, we suspect that the SHS was still stabilizing after the large rapid increase in the temperature inside the optics box at the beginning of the flight. We believe that the SHS temperature would have begun to decrease towards 23 degrees Celsius over a longer period. In any case, the temperature of the SHS varies by roughly 1.5 degrees Celsius during the flight and the temperature of the SHS remained relatively stable for the last hour of the flight. In practice, the Littrow wavelength is deduced by lining up the known water vapour absorption features with the high resolution forward modelled spectrum during data processing.
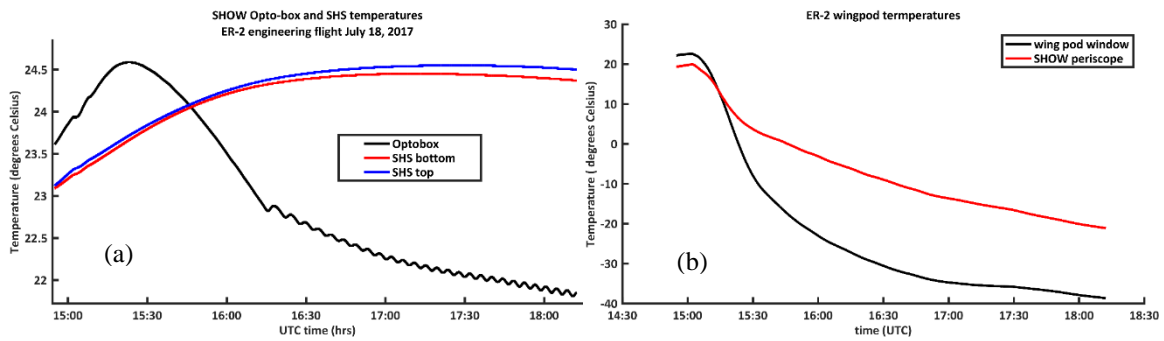
22

Figure 10 SHOW SHS and opto-box temperatures during the course of the flight (a) and the ER-2 wing pod temperatures recorded at the wind pod window (black) and the SHOW periscope (red)

Following from Figure 9 (a-b), the ER-2 airplane took off from AFRC at roughly 15:02 UTC and flew northward where the pilot performed a set of flight patterns designed to examine the sensitivity of the instrument to the observed scattering angle. Afterwards, the airplane flew towards the southeast and then turned to fly in the direction towards Table Mountain (identified by the red dot in Figure 9 (a)). During this time, SHOW was configured to continuously record images with integration times of 1800 ms at 0.5 Hz sampling. At 17:59 UTC, a radiosonde (model Vaisala RS41) was launched from the JPL facility located close to Table Mountain to measure relative humidity, temperature and pressure as a function of altitude. The in-situ water vapour abundance was deduced from these measurements using the Hyland and Wexler formulation (Hyland and Wexler, 1983). Roughly 12 minutes of coincident measurements (over 300 samples) of the same approximate column of air were obtained. The in-situ water vapour abundance that was measured by the radiosonde is shown in Figure 11. The abundance increases from 5 ppm at 20 km to roughly 35 ppm at 13 km. In this Figure, the error bars only show the measurement uncertainties. Systematic errors are not included. .—In general, the uncertainty on the radiosonde measurements is between~~less than~~ ± 1 – 3 ppm.

In the remainder of this paper, we focus on the set of measurements performed between 17:59 UTC and 18:09 UTC (bracketed by the red vertical bars in Figure 9 (b)). During this period, the airplane remained stable with pitch variations of $< \pm 0.5$ degrees and a mean sea level altitude of ~20.85 km. The temperature of the SHS was also stable during this time. According to the radiosonde measurements, the water vapour abundance increases rapidly below 13 km (not shown in the Figure), therefore, the optical thickness becomes exceeding large making it difficult to observe the tangent altitudes below that level. Therefore, the useable field of view within this range is from 13 km to 21 km, corresponding to roughly 42 % of the SHOW limb image.
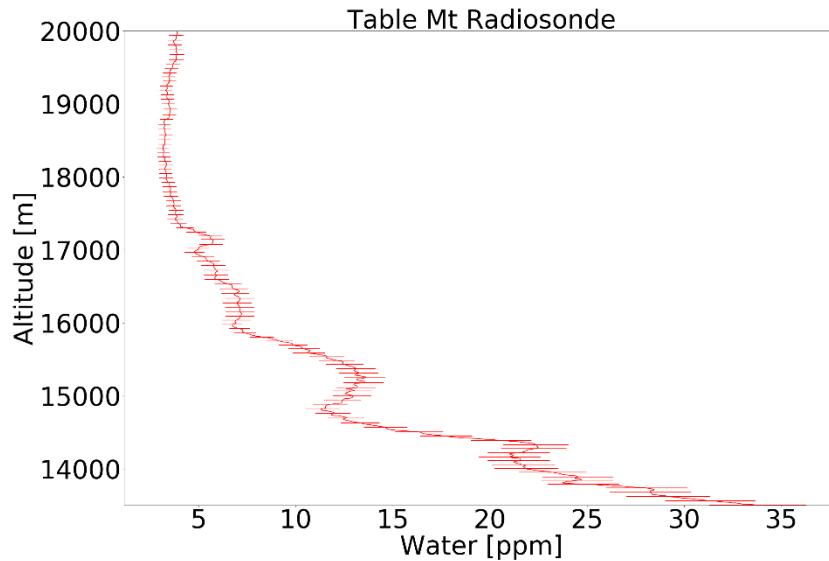
23

Figure 11 Measured in-situ water vapour above Table Mountain using the JPL Vaisala RS41 radiosonde measurements.

11 Flight data processing

11.1 Level 1A interferograms

5   Each image obtained by SHOW during the Table Mountain portion of the flight is corrected using the Level 0 to Level 1 processing chain illustrated in Figure 2. An example of the correction of a raw interferogram is shown in Figure 12. The raw interferogram is shown in Figure 12 (a) and the fully corrected interferogram is shown in Figure 12 (b). In these figures, the SHOW image has been rotated so the bottom part of the image corresponds to lower tangent altitudes. Figure 12 (c) shows the application of the correction for the interferogram row – 150. Here the DC bias has been removed by subtracting off the mean of each interferogram

10   row in order to isolate the modulated component of the interferogram. These interferograms represent the SHOW Level 1A data product.

In this particular measurement, the presence of clouds is seen in both the raw, as well as, the fully corrected modulated component. Observe the observed horizontal variability in lower part of the non-modulated component shown in Figure 12 (b). Ideally, the anamorphic optics averages over the horizontal variability in the scene radiance and the flat field correction assumes a uniform

15   illumination of the input aperture. Here we observe that the variability associated with the cloud appears to be slightly different on either side of the interferogram center burst. While not completely understood, the effect at these altitudes can partially be explained as a combination of non-uniform aperture illumination (which results in a poor flat field correction) and non-ideal pixel response due to saturation effects at the interferogram center burst. As the pixels approach saturation their response is no longer linear and the subtraction of the DC bias in the presence of this effect, as well as, non-uniform illumination will result in such variability. In

20   the case of the Table Mountain measurements, the cloud effects were confined to a small region at altitudes below the lower altitude cutoff of the retrieval at 13.5 km. Clouds also have a large optical depth, making it difficult to see the tangent point at the limb. Therefore, we do not expect to be able to perform retrievals on this particular measurement below 13.5 km.

In this particular measurement, the presence of clouds is seen in both the raw, as well as, the fully corrected modulated component. There are several primary effects on the SHOW measurements associated with the presence of clouds. The most obvious is the

In some cases, scattering from clouds saturates pixels at the interferogram center burst, resulting in corrupted pixels throughout the image. The presence of the saturation effect associated with clouds is detected by observing a small number of pixels off to the edge of the detector, far away from the SHOW FOV, as well as, the average signal across the FOV. This region also acts as a dark current and bias monitor during the flight. In practice, it was found that spikes in the number of saturated pixels correlate with spikes in both the FOV average and the off-image average. For the measurements presented in this paper, images that have strong cloud effects are identified and removed from the analysis. In the case of the Table Mountain measurements, the cloud effects were confined to a small region at altitudes below the lower altitude cutoff of the retrieval at 13.5 km.
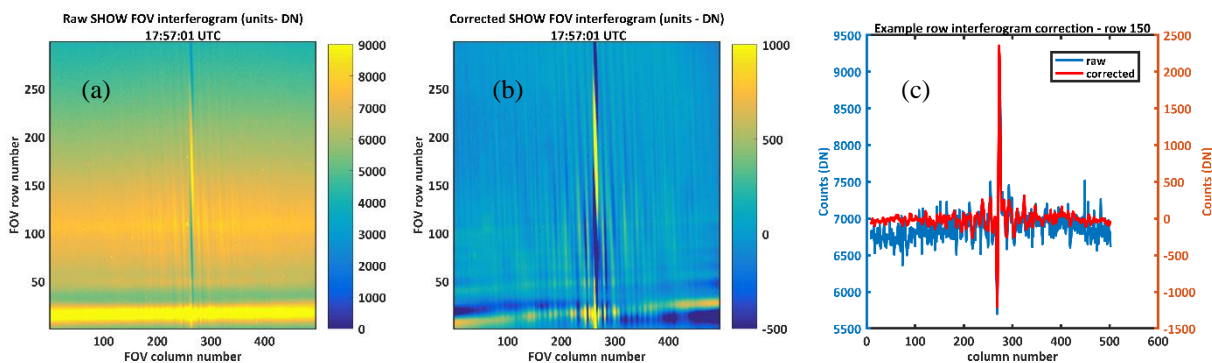

Figure 12 Example conversion of raw to Level 1A interferograms from raw images (a) to corrected interferograms (b). An example row cut at row 150 is shown in (c).

The quality of the corrected interferogram has been characterized by examining the observed variability in column cuts of the image using the same approach that was used in Section 9.6. We understand that this approach is only approximate, since the variance of the measurement varies as a function of signal level.   From Figure 12 (a), we can see that the raw interferogram signal varies from ~4000 DN at high altitudes (higher row number) up to ~ 8000 DN at lower altitudes (lower row number). The dark and bias corrected interferogram varies from 2000 DN up to 6000 DN.
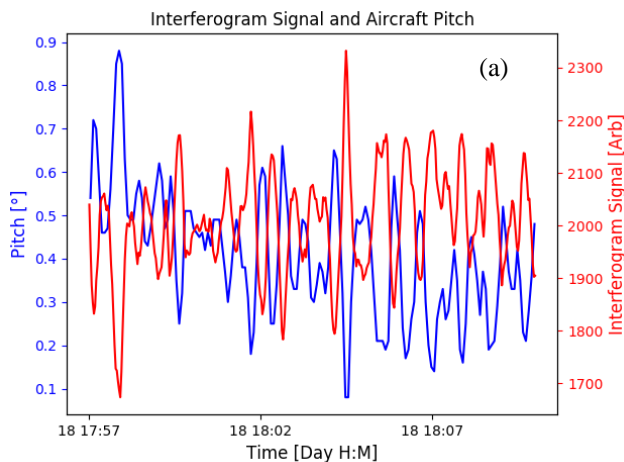
We estimate the measured variability in the image by subtracting a high order polynomial (n = 8) from each column in the interferogram to remove the atmospheric radiance profile and then calculating the standard deviation of the difference between the fit and the measured interferogram signal. The measured variability is roughly ± 23.48 DN. Assuming a mean signal level of 3700 DN, dark noise of 0.5 DN and a readout noise of 3.62 DN (see Table 2) we estimate the noise associated with the residual relative pixel response to be roughly 0.6%. This value is approximately 3 times higher than the residualrandom relative pixel response that was characterized in the laboratory. The increase is likely due to the fact that the flat field correction is chosen to match the mean intensity level and will not completely correct for the relative pixel response at the different intensities.

25

11.2   Airplane-pitch correction

The retrieval algorithm described in Section 8 utilizes a forward model which requires accurate information regarding the viewing geometry of the instrument. One must map the observed interferogram rows to the line of sight angles which correspond to specific tangent altitudes at the limb. This mapping is set by the absolute pointing, the field of view of the instrument and the number of rows in the image. On one hand, the maximum vertical resolution of the measurements is fixed by the instantaneous field of view (~0.0176 degrees) as described in Section 9.3.2. On the other hand, the absolute pointing of the instrument is set by the mechanical mounting of the instrument in the wing pod, the potential flexing of the wings during flight and the inherent pitch variations of the airplane. Pitch variations result in the same interferogram row observing a small range of different tangent altitudes as the airplane pitches up and down during the flight. In this case, we use the airplane attitude information to map the interferogram rows to tangent altitude at the limb. The instrument mounting in the wing pod was designed within the mechanical tolerances to point the field of view downward by 2.40 degrees from the aircraft boresight. Prior to the flights, the absolute pointing of the instrument was confirmed to within ±0.01 degrees. Unfortunately, we did not have the ability to monitor the flexing of the aircraft wings during flight.

As noted earlier, the field of view of the instrument at the limb is mapped one-to-one to a particular row on the detector. Therefore, the same interferogram row observes a range of different tangent altitudes as the airplane pitches up and down during the flight. The retrieval algorithm described in Section 8 utilizes a forward model which requires accurate information regarding the viewing geometry of the instrument. In the worst case, large errors in the knowledge of the viewing geometry will result in a forward model that does not accurately simulate the radiative transfer problem for this volume of the atmosphere. In the case of smaller errors, the result is a degradation of the vertical resolution of the measurement for any particular frame.

The ER-2 airplane pitch variability and its impact on the SHOW measurements during the Engineering flight on July 18, 2017 is demonstrated in Figure 13 (a) where we have plotted the signal measured by a single pixel as a function of time along with the ER-2 pitch variation. It is very clear that a strong correlation exists between the temporal variability of the interferogram signal and the pitch variation. In this particular case, the interferogram signal is anti-correlated with the pitch variation, corresponding to the location of detector pixel [150, 205] (see Figure 12 (a)) in the SHOW FOV, where we have a lower atmospheric signal above and a higher atmospheric signal below the corresponding tangent altitude. It is also clear that the pitch variation is slow relative to the 0.5 Hz sampling cadence of the SHOW measurements.
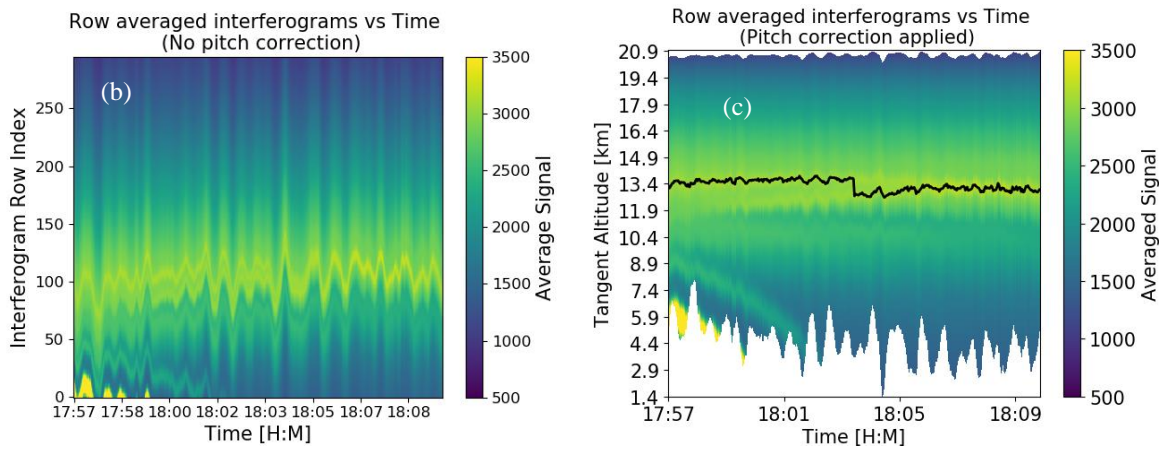
Figure 13 Example of the correction for the airplane pitch variation. The interferogram signal at a single pixel is plotted with the pitch variation as a function of time is shown in (a). The average signal at each interferogram row plotted as a function of row number and stacked as a function of time is shown in (b). The average signal at each interferogram row corrected for pitch variation and plotted as a function of tangent altitude and stacked as a function of time is shown in (c) and the solid black line is described in the text.

The pitch variation is further demonstrated in Figure 13 (b), where the mean signal level of each interferogram row is plotted as a function of altitude and then stacked as a function of time. For each row in each frame, the signal at each vertical bin corresponds to the DC component of the corresponding interferogram row. Therefore, the vertical variation of these samples in each frame provides a rough approximation of the shape of the integrated atmospheric radiance profile. In the ideal case, without pitch variation, we expect this profile to be relatively smooth, both temporally and spatially over this short 12-minute window. For the SHOW ER-2 measurements, we see that the airplane pitch variation shows up as vertical bands that correspond to the detector rows scanning up and down the atmospheric radiance profile. We also observe the presence of several cloud features below row 100 with the dominant cloud feature sitting between row 80 and row 120 and persisting through the full observation period. The impact of the pitch variation is the most obvious at the location of the clouds since the gradient in the limb radiance is largest at the cloud feature.

Since the pitch variation is slow relative to the 0.5 Hz sampling cadence, we correct for the pitch variation by using the airplane attitude information to map each interferogram row to the corresponding tangent altitude at the limb as shown in Figure 13 (c). This removes a significant portion of the variability; however, a residual temporal variation on the order of 1% of the mean signal level remains. The primary impact of the presence of this variability is ~~an~~the uncertainty in the absolute pointing of an individual frame. ~~reduction in vertical resolution associated with the error in the pointing knowledge~~. This degradation is quantified by examining the vertical perturbations in the two dominant thin cloud features located at mean altitudes of 13.6 km and 13.1 km respectively. The black line shows the location of the peak signal associated with these features. The standard deviation from the mean altitude level is found to be 124.45 m for the feature at 13.6 km and 175.03 m for the feature at 13.1 km. Assuming a mean ER-2 altitude of 20.85 km, these perturbations correspond to an angular uncertainty in the pointing of 0.023 degrees and 0.032 degrees for these two cases. Taking the maximum of these perturbations to be representative of the uncertainty in the ~~-~~ pointing ~~error~~, we estimate the uncertainty in the absolute pointing associated with this mapping ~~g knowledge~~ to be ~0.032 degrees. ~~This is larger than the, 0.0176 degree, instantaneous tangent point angular resolution provided by the instrument. Therefore, we find the ER-2 measurements obtained during the engineering flight to have a slightly degraded vertical resolution.~~

27

11.3   Level 1B spectra

SHOW Level 1B spectra are generated from the Level 1A interferograms by applying Hanning apodization to each of the interferogram rows, performing an FFT and calculating the amplitude spectrum. Each row of the spectrum is mapped to a tangent altitude at the limb using the pitch correction described in the previous section. Therefore, a single record consists of the Level 1B spectrum and a mapping of spectral row to the tangent altitude at the limb. An example Level 1B spectrum obtained by applying this procedure to the interferogram image obtained at 17:59:05 UTC is shown in Figure 14 (a). The wavelength registration is set using knowledge of the instrument configuration and by manually adjusting Littrow wavelength to line up the absorption features with a high-resolution forward modelled spectrum. For this particular measurement, the Littrow wavelength was found to be 1363.76 nm. Over the course of the 12-minute Table Mountain portion of the flight, the Littrow wavelength was not found to vary significantly as a function of time.
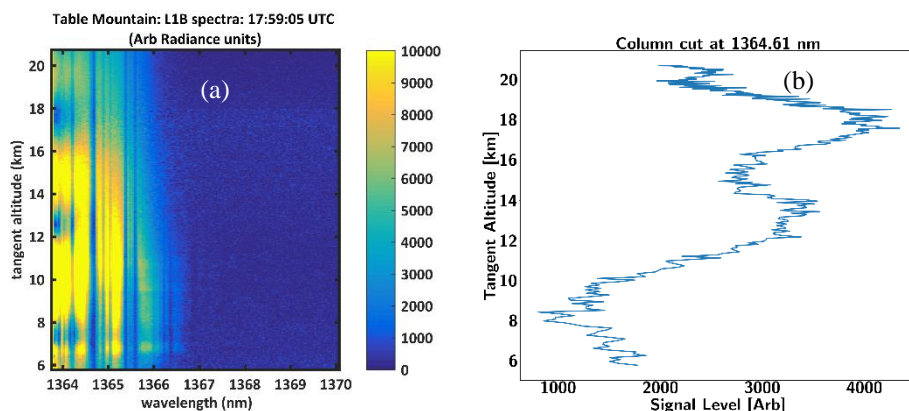

Figure 14 Example Level 1B spectral image (a) and a column cut taken at a wavelength of 1364.61 nm (b)

Figure 14 (b) shows a column cut at a wavelength of 1364.61 nm. Several sources of variability are apparent in the spectral samples. First, note the natural variability associated with the limb scattered radiance profile. The shape of this profile depends heavily on the atmospheric state (aerosol content, the water vapour abundance and clouds etc.). This particular wavelength sits in the wing of a strong absorption feature that is centered near 1364.66 nm. In this case, the radiance profile is modulated by the systematic variations in the spectrum introduced due to aliasing effects. This effect was anticipated earlier in the paper and is characterized in the next section. Here we focus on characterizing the signal-to-noise ratio of the corrected spectral samples.

The two dominant sources of noise in the spectral measurements are Poisson noise, as well as, a systematic inter-sample variability, both of which are introduced by the propagation of error into the spectrum from the interferogram samples (see Section 6). The systematic inter-sample variability in the spectral measurements is minimized by utilizing the temporal series of samples obtained while the airplane pitches up and down to obtain a secondary correction factor that is applied to the spectral measurements. In this case, the noise on individual spectral samples is primarily due to Poisson statistics and is estimated The noise in the interferogram samples is calculated using Eq.4 and the associated noise in the spectral domain is calculated using Eq.5. However, as we shall now show, the impact of the relative pixel response on the spectral measurements can be minimized by utilizing the temporal series of samples obtained while the airplane pitches up and down to obtain a secondary calibration that can be applied to the spectral measurements. In this case, the noise on individual spectral samples is primarily due to Poisson statistics.

28

To characterize the systematic inter-sample variabilityFor this characterization, we assume that the shape of the true atmospheric profile is smooth and constant. This is a justified assumption given the stability of the observed integrated intensity profile shown in Figure 13 (c). Therefore, as the airplane pitches up and down, multiple rows in the SHOW FOV observe the same atmospheric signal. For example, Figure 15 (a) shows the raw non-normalized L1B spectral signal level at a wavelength of 1364.81 nm for each detector row that observed the tangent altitude at 13.122 km (shown in black) within the 12 minute observation window. The black data points form clusters at each detector row that are approximately randomly distributed. This distribution includes Poisson noise, as well as, noise due to the temporal variability in the signal level associated with natural variability and the absolute uncertaintyerrors in the pitch correction of ± 0.032 degrees (see Section 11.2). The error bars show only the Poisson distributed spectral noise.
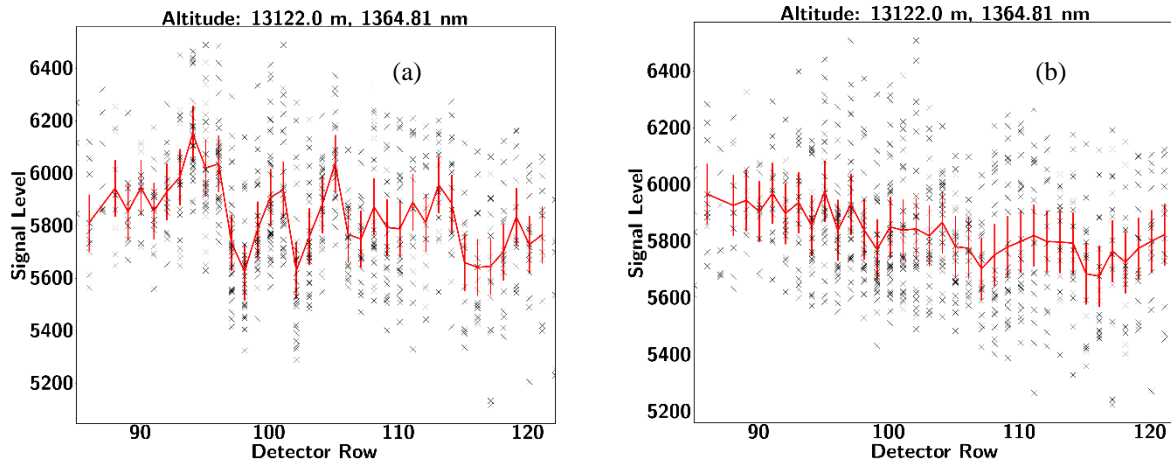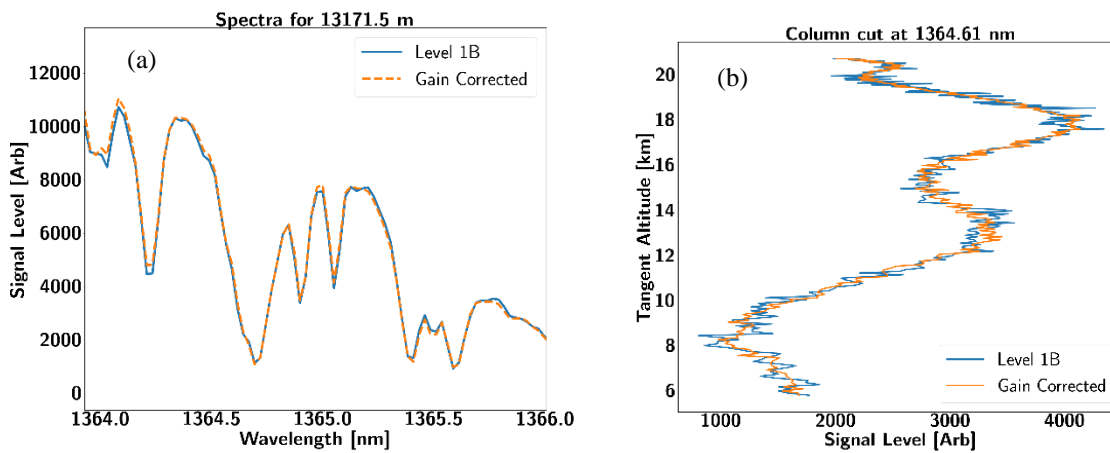

Figure 15 Observed systematic inter-sample variability in the measured spectra for a single wavelength (1364.81 nm) and altitude (13.122 km) (a). The same information is plotted in (b) after the secondary correction is applied. The error bars denote Poisson noise and the average SNR of the measurements for this set of samples is 53.83.

Ideally, each row will register the same radiance; however, we observe that each cluster exhibits a slight offset for each detector row. In this case, tThese offsets were anticipated in Section 6 and are due to the propagation of error associated with the relative pixel response across the interferogram rows. For these particular measurements, the systematic inter-sample variability is of the same magnitude as the combined temporal variability and Poisson noise.

TThe observed systematic biases are isolated in several steps. First, each of the sample clusters is averaged to minimize the random noise (shown in red) and clusters that have fewer than ten samples are discarded from further analysis. We then normalize the averaged values by the mean signal level across all of the detector rows that observed this altitude in order to isolate the deviations relative to the mean. This data is then merged with the normalized averaged clusters from all the other tangent altitudes in order to isolate the inter-row biases for this particular wavelength. In order to remove the larger-scale variation associated with aliasing, we divide out a high order polynomial from the resulting data. This procedure is repeated for all of the measured wavelengths to obtain a correction for each spectral column in the image. Therefore, we form a secondary spectral image multiplicative correction factor that can be applied to the Level 1B spectral images to minimize the systematic inter-sample variability.

29

The results of applying this correction to the L1 B spectral images is demonstrated in Figure 15 (b) where we plot the same data as shown in Figure 15 (a), this time with the correction applied to the spectral measurements. In this case, the amplitude of inter-row sample variability has been reduced by a factor of roughly 5 and this variability is now significantly smaller than the noise due to Poisson statistics. Assuming only Poisson noise, the average SNR of this set of samples is 53.83.

5   Each Level 1B spectral image obtained during the Table Mountain portion of the flight was corrected using this approach in order to obtain Level 1C spectra. An example of the impact of the secondary correction on an individual spectral row and an individual spectral column within a single image is shown in Figure 16 (a) and Figure 16 (b) respectively. ~~While the correction has a minimal impact on the wavelength dimension, it has a significant impact on the vertical variability.~~ This correction minimizes the impact of the systematic inter-sample variability; therefore, the noise on the corrected spectral samples is taken to be Poisson distributed

10   (Eq. 5) and this noise level is included in the retrieval algorithm in order to compute the associated measurement error of the retrieved vertical water vapour distribution.

Figure 16 Correction of the L1B spectra. Variability before and after the secondary correction is applied for an example row spectrum
15   **at a fixed altitude (a) and for a single wavelength column versus altitude (b)**

11.4   Instrument model optimization

The systematic perturbations to the radiance profile associated with aliasing are characterized by applying the optimization
20   approach that was used in Section 9.7. However, for the Table Mountain flight measurements, we use in-situ water vapour measurements as the input to the forward model. Knowledge of the in-situ water vapour abundance provides a better representation of the true atmospheric state and allows for a better quality optimization to be performed. Forward modelled interferograms and spectra are generated and compared to the flight measurements and the instrument model is iteratively adjusted to identify the instrument configuration that best predicts the measurements.

25   The results from this optimization are shown in Figure 17. The optimized configuration corresponds to a Littrow wavelength of 1363.76 nm (in vacuum), a grating tilt of -5.085 x 10 $^{-5}$ radians, a filter shift - 0.048 nm, and x, y offsets of (17.27 pixels, -18.58 pixels) respectively. Two spectral rows are shown in Figure 17 (b) and Figure 17 (c) corresponding to the tangent altitudes of

30

9.878 km and 19.993 km respectively. There is general agreement between the forward modelled spectra and the measured spectra; however, the best agreement is observed from 1364.52 nm to 1365.67 nm. The presence of clouds below 13.5 km also makes the retrieval of water vapour below this altitude difficult. Therefore, the best region to target for a SHOW retrieval is from 1364.5 nm to 1365.67 nm and from 13.5 km up to 18 km. In this window, the observed relative variability between the measured and forward modelled spectra is less than ± 5 %.
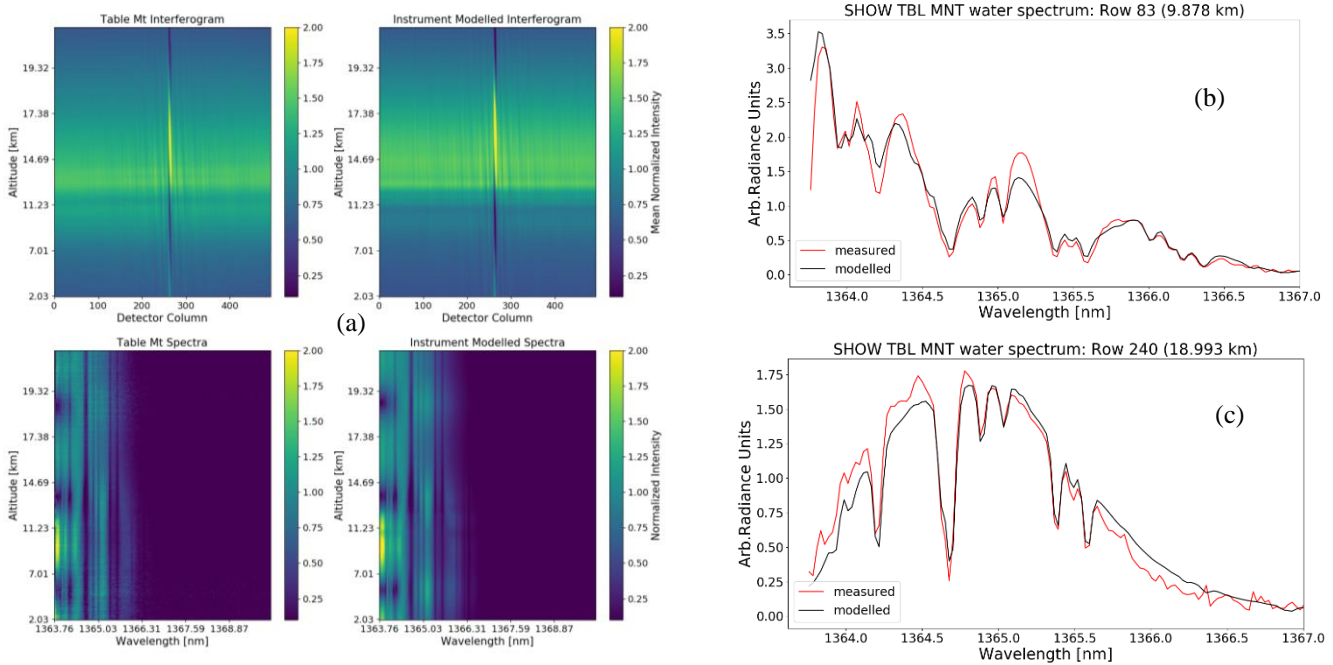

Figure 17 Instrument model optimization using flight data. Comparison between the modelled and measured interferograms and spectral images (a). Example spectral row cuts at row 83 (b) and at row 240 in (c). The measured spectrum is shown in red and the modelled spectrum is shown in black.
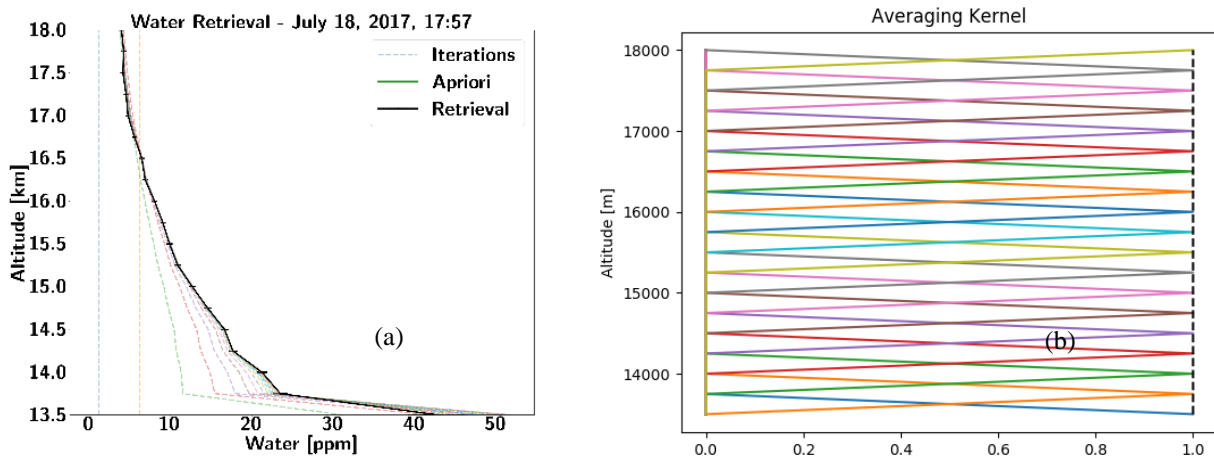
12   Example water vapour retrieval

For the water vapour retrieval, we focus on a 10-image measurement window from 17~~5~~:57:29 UTC to 17~~5~~:57:49 UTC and average the frames to increase the SNR of the measurements. During this time, SHOW has good coincidence with the radiosonde measurements and the aircraft is pitched slightly upward above the cloud features that are present at the beginning of the Table Mountain portion of the flight. The retrieval is performed using the approach discussed in Section 8. For the retrieval, we take the upper and lower altitude cutoff to be 18 km and 13.5 km respectively. The upper cutoff was chosen to be a few kilometers below the airplane altitude. Above this altitude, the region is optically thin and produces weak spectral signatures that reduce the accuracy of the retrieval. The lower cutoff was chosen to be above the region with clouds and above the region where the atmosphere is too optically thick to observe the tangent point.

The SHOW measurements utilize the full spatial resolution that is provided by the vertical sampling within the imaged FOV. This provides an instantaneous vertical resolution of 0.0176 degrees and an associated tangent point resolution of 151 m at 2.5 km and 20 m at 21 km assuming an airplane altitude of 21.34 km (see Section 9.3.2). ~~However, i~~It was shown in Section 11.2~~, that there are that the instantaneous angular vertical resolution is degraded to 0.032 degrees~~ also potential uncertainties in the absolute pointing of the instrument. These include uncertainties associated with the mechanical mounting in the wing pod (±0.01 degrees)

31

and a residual pointing uncertainty of ±0.032 degrees that remains after mapping the interferogram rows to tangent altitude at the limb using the aircraft attitude information. Since we average 10 images, the residual pointing uncertainty is reduced to approximately ±0.01 degrees. For the demonstration measurements, the retrievals are performed in the presence of these uncertainties in the absolute pointing of the instrument.

5 ~~due to uncertainty in the pointing. Therefore, within the target retrieval range, we estimate the tangent point vertical resolution to be roughly 171 m at 13.5 km and roughly 107 m at 18 km. We perform the retrieval on a 250 m sampling grid that is larger than this uncertainty and we allow the regularization to control the vertical resolution. A small amount of damping is also applied to constrain the retrieval.~~

An example retrieval that was obtained by averaging the 10 successive coincident measurements is shown in Figure 18 (a). Here,

10 we perform the retrieval on a 250 m sampling grid that is larger than the instantaneous vertical resolution of the measurements. We include regularization; however, very little smoothing is applied and this term has a negligible impact on the vertical resolution. In addition, a small amount of damping is applied to ensure convergence of the retrieval. -The figure shows 15 iterations; however, the retrieval converges in approximately 10 iterations. At each iteration, we produce a high-resolution forward modelled radiance for each of the 10 viewing geometries. Averaging is performed on the georeferenced forward modelled images and the

15 georeferenced measurement images using the exact same procedure. The associated averaging Kernel is shown in Figure 18 (b). Very little smoothing has been applied; therefore, the approximate vertical resolution of the measurements, determined from the full width half maximum of the averaging Kernel, is closely matched to the 250 m spacing of the retrieval grid. The measurement error due to Poisson noise is less than ± 1ppm for all altitudes.

Figure 18 Example SHOW water vapour retrieval (a) and the associated averaging Kernel (b).

The retrieved water vapour profile is plotted along with the measured in-situ water vapour profile in Figure 19 (a). Since the in-situ profile is measured on a finer grid than the retrieval, the radiosonde data is transformed to the 250 m retrieval grid using the

25 method that is suggested by (Rodgers, 2000). In effect, this appropriately smooths the radiosonde measurements down to the vertical resolution of the retrieval grid. The measurement uncertainties for both instruments are less than 1 ppm and are not shown in the Figure.~~The measurement uncertainties for both instruments is less than 1 ppm and these uncertainties are not shown in the~~

32

Figure. There is good agreement between the overall shape of the retrieved water vapour profile (shown in red) and the smoothed in-situ water profile (shown in blue) with the best agreement occurring at the high altitudes. The ppm difference between the radiosonde and SHOW measurement is shown in Figure 19 (b). At higher altitudes, the difference is on the order of 1-2 ppm and the two measurements match quite well~~overlap within the error bars~~. Such discrepancies are anticipated due to the sensitivities to aerosols that are described in our earlier sensitivity study (Langille et al., 2018). For example, differences in the knowledge of the true aerosol profile utilized in the forward model can result in systematic uncertainties on the order of 1 ppm. However, it is clear that the SHOW measurements do not capture some of the small-scale variability that is observed with the radiosonde below 16 km where the difference is between 2-5 ppm. This type of difference is not uncommon when comparing limb sounding water vapour measurements with radiosondes (Stiller et al., 2009).
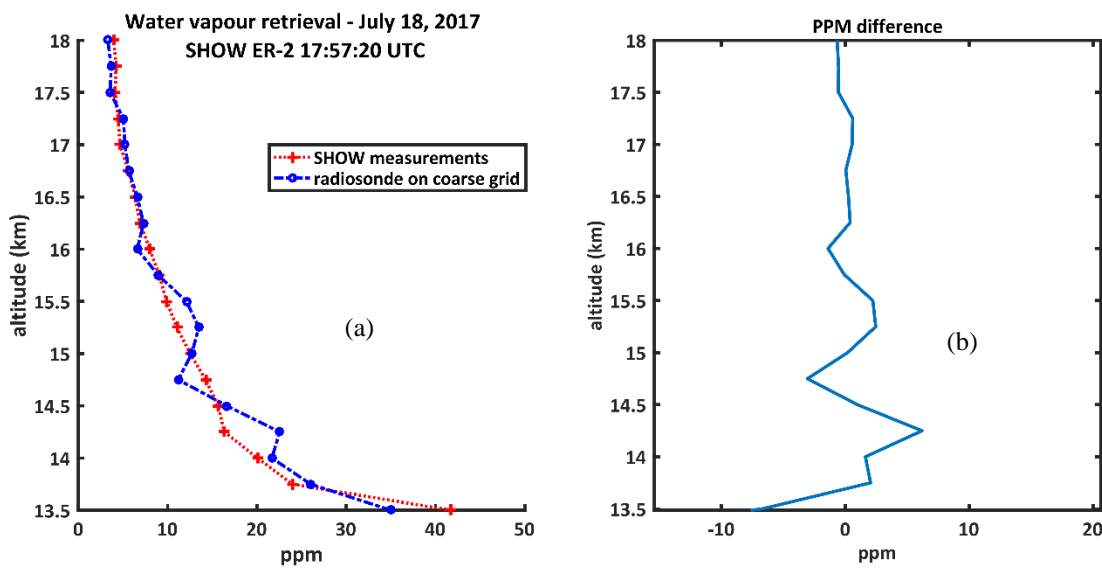

Figure 19 SHOW water vapour retrieval plotted with the radiosonde measurements transformed to the coarse retrieval grid (a) and the ppm difference between the radiosonde and SHOW measurements.

There are several potential reasons for differences between the Vasaila RS41 radiosonde and SHOW. The most likely is differences in the viewing geometry between the two instruments, coupled with the fact that both instruments have unknown accuracies. The radiosonde samples the spatial and temporal variability of the water vapour field as a function of altitude, whereas SHOW measures the 1-dimensional water vapour abundance, which is heavily weighted to the tangent point. It is known that radiosonde accuracies vary as a function of relative humidity and temperature and from sensor to sensor (Miloshevich et al., 2006). The Vaisala model RS41, utilized in this paper, has been shown to have improved accuracy over earlier models; however, errors between sensors on the order of 2-5 % have been recorded (Jensen et al., 2016). On the other hand, incomplete knowledge regarding the true instrument configuration (see Section 5) and the lack of knowledge of the true state of the atmosphere limits our ability to perform a more detailed characterization of the systematic effects associated with aliasing. In addition, uncertainties~~error~~ in the absolute pointing knowledge of SHOW measurements leads to potential uncertainties in the forward modelled radiances that are~~a smearing of vertical information that is~~difficult to characterize when coupled with the aliasing effect. Therefore, more work is required to fully examine the expected biases between these two types of measurement.

33

13 Conclusion

The SHOW prototype instrument has performed successful demonstration flights from NASA's ER-2 airplane. In this paper, we have presented the characterization and the Level 0 to Level 1 processing of flight measurements that were obtained with SHOW during an engineering flight that was performed on July 18th, 2017. Extracting water vapour profiles from the SHOW measurements required a significant amount of calibration and characterization work and the development of an instrument model that is optimized to capture systematic variations that are observed in the measured spectra.

We have applied this approach to the SHOW measurements that were obtained during a period of stable flight where the SHOW instrument observed the same approximate column of air as a radiosonde that was launched from a JPL facility nearby Table Mountain. These coincident measurements were compared to the SHOW measurements and were found to agree to within 1-5 ppm from 13.5 km – 18 km. The work presented in this paper provides initial validation of the SHOW measurement technique and demonstrates that high vertical resolution (< 500 m) measurements with $< \pm 1ppm$ accuracy are feasible using this approach.

References

1 Bourassa, A.E., Degenstein, D.A. and E.J. Llewellyn, SASKTRAN: A spherical geometry radiative transfer code for efficient estimation of limb scattered sunlight, Journal of Quantitative Spectroscopy and Radiative Transfer, Volume 109, Issue 1, January 2008, Pages 52-73, ISSN 0022-4073, http://dx.doi.org/10.1016/j.jqsrt.2007.07.007.

2 Bourassa, A.E., Langille, J., Solheim, B., Degenstein, D. and Fabien Dupont, "The Spatial Heterodyne Observations of Water (SHOW) Instrument for High Resolution Profiling in the Upper Troposphere and Lower Stratosphere," in Light, Energy and the Environment, OSA Technical Digest (online) (Optical0 Society of America, 2016), paper FM3E.1.

3 Rozanov, A., Weigel, K., Bovensmann, H., Dhomse, S., Eichmann, K.-U., Kivi, R., Rozanov, V., Vömel, H., Weber, M., and Burrows, J. P.: Retrieval of water vapour vertical distributions in the upper troposphere and the lower stratosphere from SCIAMACHY limb measurements, Atmos. Meas. Tech., 4, 933-954, https://doi.org/10.5194/amt-4-933-2011, 2011.

4 Ferrero, A., Campos, L., and Alicia Pons, "Correction of photoresponse nonuniformity for matrix detectors based on prior compensation for their nonlinear behavior," Appl. Opt.45, 2422-2427 (2006)

5 Gettelman, A., P. Hoor, L. L. Pan, W. J. Randel, M. I. Hegglin, and T. Birner (2011), The extratropical upper troposphere and lower stratosphere, Rev. Geophys., 49

6 Rodgers, C., Inverse Methods for Atmospheric Sounding: Theory and Practice, 2000

7 Englert, C.R., Harlander, J.M., Cardon, J.G and Fred L. Roesler, "Correction of phase distortion in spatial heterodyne spectroscopy," Appl. Opt. 43, 6680-6687 (2004)

8 Englert, C.R., and J.M Harlander, "Flatfielding in spatial heterodyne spectroscopy," Appl. Opt. 45, 4583-4590 (2006)

9 Englert, C.R., Harlander, J.M., Brown, C.M. and Marr, K.D, "Spatial heterodyne spectroscopy at the Naval Research Laboratory," Appl. Opt. 54, F158-F163 (2015)

10 Englert, C.R., Babcock, D.D and J. M. Harlander, "Spatial heterodyne spectroscopy for long-wave infrared: First measurements of broadband spectra," Opt. Eng. 48, 105602 (2009)

11 Zawada, D. J., Dueck, S. R., Rieger, L. A., Bourassa, A. E., Lloyd, N. D., and Degenstein, D. A.: High-resolution and Monte Carlo additions to the SASKTRAN radiative transfer model, Atmos. Meas. Tech., 8, 2609-2623, https://doi.org/10.5194/amt-8-2609-2015, 2015.

12 Zawada, D. J., Rieger, L. A., Bourassa, A. E., and Degenstein, D. A.: Tomographic retrievals of ozone with the OMPS Limb Profiler: algorithm description and preliminary results, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-236, in review, 2017.

13 Marquardt, D.W. "An algorithm for least-squares estimation of nonlinear parameters." Journal of the society for Industrial and Applied Mathematics 11.2 (1963): 431-441.

14 Talbot D. et al. (2008) Development of a New Canadian Operational Air Quality Forecast Model. In: Borrego C., Miranda A.I. Air Pollution Modeling and Its Application XIX. NATO Science for Peace and Security Series. Environmental Security. Springer, Dordrecht, Infrared Physics & Technology,Volume 36, Issue 4,1995,Pages 763-777,ISSN 1350-4495, https://doi.org/10.1016/1350-4495(94)00002-3.

15 Revercomb, H.E., Buijs, H., Howell, H.B., LaPorte, D.D., Smith, W.L and Sromovsky, L.A, "Radiometric calibration of IR Fourier transform spectrometers: solution to a problem with the High-Resolution Interferometer Sounder," Appl. Opt. 27, 3210-3218 (1988)

16 Hurst, D. F., A. Lambert, W. G. Read, S. M. Davis, K. H. Rosenlof, E. G. Hall, A. F. Jordan, and S. J. Oltmans (2014), Validation of Aura Microwave Limb Sounder stratospheric water vapour measurements by the NOAA frost point hygrometer, J. Geophys. Res. Atmos., 119, 1612–1625, doi:10.1002/2013JD020757.

17 Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H. B., Rosenlof, K. H., Davis, S. M., Hall, E. G., Jordan, A. F., and Oltmans, S. J.: Recent divergences in stratospheric water vapour measurements by frost point hygrometers and the Aura Microwave Limb Sounder, Atmos. Meas. Tech., 9, 4447-4457, https://doi.org/10.5194/amt-9-4447-2016, 2016.

18 Dupont, F., Grandmont, F., Solheim, B., Bourassa, A., Degenstein, D., Lloyd, N. and R. Cooney, "Spatial Heterodyne Spectrometer for Observation of Water for a Balloon Flight: Overview of the Instrument & Preliminary Flight Data Results," in Fourier Transform Spectroscopy and Hyperspectral Imaging and Sounding of the Environment, OSA Technical Digest (online) (Optical Society of America, 2015), paper FW4A.5.

19 Dupont, F., Grandmont, F., Paradis, S., Jacques, L., Bourassa, A.E., Solheim, B., Degenstein, D., Walsh, K., and M. Kapitzke, "Spatial Heterodyne Spectrometer for Observation of Water (SHOW) aboard the ER-2 Airplane: Overview of the Instrument," in Light, Energy and the Environment, OSA Technical Digest (online) (Optical Society of America, 2016), paper FM3E.2.

20 Stiller, G. P., Kiefer, M., Eckert, E., von Clarmann, T., Kellmann, S., García-Comas, M., Funke, B., Leblanc, T., Fetzer, E., Froidevaux, L., Gomez, M., Hall, E., Hurst, D., Jordan, A., Kämpfer, N., Lambert, A., McDermid, I. S., McGee, T., Miloshevich, L., Nedoluha, G., Read, W., Schneider, M., Schwartz, M., Straub, C., Toon, G., Twigg, L. W., Walker, K., and Whiteman, D. N.: Validation of MIPAS IMK/IAA temperature, water vapour, and ozone profiles with MOHAVE-2009 campaign measurements, Atmos. Meas. Tech., 5, 289-320, https://doi.org/10.5194/amt-5-289-2012, 2012.

21 Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151-2188, https://doi.org/10.5194/acp-8-2151-2008, 2008.

22 Langille, J.A., Solheim, B., Bourassa, A., Degenstein, S., Brown, S., and G.G Shepherd, "Measurement of water vapour using an imaging field-widened spatial heterodyne spectrometer," Appl. Opt. 56, 4297-4308 (2017)

23 Langille J.A., Letros, D., Zawada, D., Bourassa, A., D. Degenstein, B Solheim, Spatial Heterodyne Observations of Water (SHOW) vapour in the upper troposphere and lower stratosphere from a high altitude airplane: Modelling and sensitivity analysis, Journal of Quantitative Spectroscopy and Radiative Transfer,Volume 209,2018,Pages 137-149,ISSN 0022-4073, https://doi.org/10.1016/j.jqsrt.2018.01.026.

24 Harlander, J. M. (1991). Spatial heterodyne spectroscopy: Interferometric performance at any wavelength without scanning.

25 J. Harlander, R. J. Reynolds, and F. L. Roesler, "Spatial heterodyne spectroscopy for the exploration of diffuse interstellar emission lines at far-ultraviolet wavelengths," Astrophys. J. 396, 730–740 (1992).

26

27 Harlander, J.M., Roesler, F.L., Cardon, J.G., Englert, C.R and Robert R. Conway, "Shimmer: a spatial heterodyne spectrometer for remote sensing of Earth' middle atmosphere," Appl. Opt. 41, 1343-1352 (2002)

28 Brault., J.W.:Fourier transform spectrometry, High Resolution in Astronomy, Proceedings of the Fifteenth advanced course of the Swiss society of Astronomy and Astrophysics (1985)

29 Rothman, L. S., I. E. Gordon, A. Barbe, D. C. Benner, P. F. Bernath, M. Birk, V. Boudon, L. R. Brown, A. Campargue, J.-P. Champion, K. Chance, L. H. Coudert, V. Dana, V. M. Devi, S. Fally, J.-M. Flaud, R. R. Gamache, A. Goldman, D. Jacquemart, I. Kleiner, N. Lacome, W. J. Lafferty, J.-Y. Mandin, S. Massie, S. N. Mikhailenko, C. E. Miller, N. Moazzen-Ahmadi, O. V. Naumenko, A. V. Nikitin, J. Orphal, V. I. Perevalov, A. Perrin, A. Predoi-Cross, C. P. Rinsland, M. Rotger, M. S̆imečková, M. A. H. Smith, K. Sung, S. A. Tashkun, J. Tennyson, R. A. Toth, A. C. Vandaele, and J. V. Auwera (2009), The HITRAN 2008 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 110, 533-572, doi:10.1016/j.jqsrt.2009.02.013.

30 Carleer, M. R., Boone, C. D., Walker, K. A., Bernath, P. F., Strong, K., Sica, R. J., Randall, C. E., Vömel, H., Kar, J., Höpfner, M., Milz, M., von Clarmann, T., Kivi, R., Valverde-Canossa, J., Sioris, C. E., Izawa, M. R. M., Dupuy, E., McElroy, C. T., Drummond, J. R., Nowlan, C. R., Zou, J., Nichitiu, F., Lossow, S., Urban, J., Murtagh, D., and Dufour, D. G.: Validation of water vapour profiles from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys. Discuss., 8, 4499-4559, https://doi.org/10.5194/acpd-8-4499-2008, 2008.

31 Miloshevich, L. M., Vomel, H., Whiteman, D. N., & Leblanc, T. (2009). Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements. Journal of Geophysical Research, 114, D11305. https://doi.org/10.1029/2008JD011565

32 Jensen, M. P., Holdridge, D. J., Survo, P., Lehtinen, R., Baxter, S., Toto, T., and Johnson, K. L.: Comparison of Vaisala radiosondes RS41 and RS92 at the ARM Southern Great Plains site, Atmos. Meas. Tech., 9, 3115-3129, https://doi.org/10.5194/amt-9-3115-2016, 2016.

33   Milz, M., et al. (2005), Water vapour distributions measured with the Michelson Interferometer for Passive Atmospheric Sounding on board Envisat (MIPAS/Envisat), J. Geophys. Res., 110, D24307, doi:10.1029/2005JD005973.

34   Schulz, M., and Caldwell, L., Nonuniformity correction and correctability of infrared focal plane arrays,

35   Sherwood, S. C., R. Roca, T. M. Weckwerth, and N. G. Andronova (2010), Tropospheric water vapour, convection, and climate,Rev. Geophys., 48,

36   Connes, P. (1958). Spectromètre interférential à selection par l'amplitude de modulation. Journal de Physique et le Radium, 19, 215–222. doi: 10.1051/jphysrad:01958001903021500

37   Sumner P. Davis, Mark C. Abrams and James W. Brault, Introduction, In Fourier Transform Spectrometry, Academic Press, San Diego, 2001

38   Randel, W.J. and Jensen, E.J, Physical processes in the tropical tropopause layer and their roles in a changing climate", Nature Geoscience, 6,169–176 (2013)

39   Sun, Y., Chen, Q., Gui, K., Dong, F., Feng, X., and Long, Q., "Characteristics of Water Vapour in the UTLS over the Tibetan Plateau Based on AURA/MLS Observations," Advances in Meteorology, vol. 2017, Article ID 3504254, 13 pages, 2017. https://doi.org/10.1155/2017/3504254.