ALMA Development Study Final Report: Improving the calibration of atmospheric spectral features in ALMA data

Todd Hunter (NRAO), Neil Phillips (JAO, ESO), Dominique Broguiére (IRAM), Justo Gonzalez (ESO)

May 2018

Abstract

On the 64-station baseline correlator (BLC), the atmospheric calibration of ALMA is currently performed while operating in the coarse spectral resolution (TDM) mode. While there were practical reasons behind this choice, including the lack of an easy way to compute the quantization correction (QC) in FDM mode and limited processing power in the machines running TelCal, the consequence on calibrated science data is that atmospheric line features, primarily ozone, are not fully removed. In this Development Study, we performed proof-of-concept tests both offline and at the telescope demonstrating that performing high-resolution atmospheric calibration would better remove these features, and identified the areas where improvements in ALMA software were needed to bring this calibration mode into routine usage. As a result, we requested and tested new infrastructure and speed improvements in the online software modules TelCal and DataCapturer. In the most recent on-sky tests, we recorded a science-like dataset with Tsys measured in four FDM spectral windows and 47 antennas. The new QC option in online TelCal succeeded and the dataset was successfully processed through the Cycle 5 pipeline. We describe the remaining steps needed to implement FDM QC and Tsys observations before they can be included as a standard observing mode for science observations (targeting Cycle 7), including cross calibration of the digitizers to the square law detectors (SQLDs), and validation of the accuracy obtained by the chosen QC formula.

In a parallel effort, we also developed a mechanism in the analysisUtils python package that can simulate higher resolution Tsys spectra by combining the existing low-resolution spectra with a high resolution atmospheric model. This functionality can potentially be used in the future to correct older (archival) data by inserting an extra stage in the ALMA pipeline after the standard Tsys caltable is generated. This software work revealed several issues with the atmospheric model code (ATM). Although the two minor bugs have been fixed, there are two outstanding issues remaining. First, ATM apparently does not distinguish between channel width and channel resolution, which are generally not the same for ALMA data (due to the correlator's default application of the Hann window function) unless a significant level of subsequent online channel averaging has also been employed. This effect must be accounted for when simulating high resolution Tsys spectra. It also means that TelCal is not optimally calculating Tsys spectra because the atmospheric model is used in this calculation, but the effect is probably minor for the non-DSB receivers since the contribution is weighted by the sideband gain ratio. Second, improvements to the layering in ATM are needed to better model the ozone line profiles in order to achieve a better match to ALMA observations and thereby reduce the residual features in calibrated data. Without these improvements, residual features will remain even after FDM Tsys measurements are implemented on the BLC. We summarize the required additions to ATM to make progress in this area.

1 Acronym glossary

ACA	Atacama Compact Array
ACAC	Atacama Compact Array Correlator
ALMA	Atacama Large Millimeter and submillimeter Array
AOS	ALMA Array Operations Site (location of ACA and BLC)
ARC	ALMA Regional Center (NA: NRAO-Charlottesville, Europe: ESO-Garching, East Asia: NAOJ-Mitaka)
ASDM	ALMA Science Data Model (raw data format)
ATM	Atmospheric Transmission at Microwaves (Pardo et al. 2001)
BB	BaseBand (a downconverted RF signal where the lowest frequency is near DC)
BLC	BaseLine Correlator (the ALMA 64-station correlator)
CARMA	Combined ARray for Millimeter Astronomy (defunct interferometer in California)
CAS	JIRA ticket system for CASA
CASA	Common Astronomy Software Applications (interferometry and single dish)
CDP	Correlator Data Processor (a cluster of CPUs)
CM	7-meter Mitsubishi Electric Corporation (Melco) ACA antenna
COMP	ALMA Computing group JIRA ticket system (pre-dates ICT)
CPU	Central Processing Unit
CSV	Commissioning and Science Verification
DA	12-meter Alcatel Corporation ALMA antenna
DC	Direct Current (response at zero frequency)
DSB	Double Side Band (a property of receivers)
DV	12-meter Vertex Corporation ALMA antenna
EOC	Extension of Capabilities
ESO	European Southern Observatory
FDM	Frequency Division Multiplexing (a BLC mode)
HITRAN	HIgh-resolution TRANsmission molecular absorption database
ICT	Integrated Computing Team, also a JIRA ticket series
IF	Intermediate Frequency
IRAM	Institut de Radio Astronomie Millimétrique
JAO	Joint ALMA Observatory (headquartered in Santiago, Chile)
JIRA	ticket tracking software (truncation of Gojira, the Japanese name for Godzilla)
LSB	Lower Side Band
NA	North America
NAASC	North American ALMA Science Center (the NA ARC)
NRAO	National Radio Astronomy Observatory
PI	Principle Investigator
PM	12-meter Mitsubishi Electric Corporation (Melco) ACA antenna
PWV	Precipitable Water Vapor
OSF	Operations Support Facility (ALMA mid-level activity site)
QA	Quality Assurance (including levels 0, 1, 2 and 3)
QC	Quantization Correction
RF	Radio Frequency
SB	Scheduling Block (XML instructions for an ALMA observation)
SCCB	Software Change Control Board, also a JIRA ticket series
SMA	SubMillimeter Array (operational interferometer on Maunakea, Hawai'i)
spw	spectral window (abbreviation used in CASA and ALMA)
SQLD	SQuare Law Detector
SSR	Science Software Requirements
TDM	Time Division Multiplexing (a BLC mode)
TelCal	(ALMA) Telescope Calibration software
TFINT	Test Facility INTerferometer (located at the OSF)
TMCDB	Telescope Monitor and Control DataBase
TOPO	Topocentric frequency frame (i.e. relative to ALMA AOS)
USB	Upper Side Band
WVR	Water Vapor Radiometer
XML	eXtensible Markup Language
XX or YY	Correlation product of two linearly polarized feeds (X or Y)

2 Introduction

Currently, the atmospheric calibration of ALMA observations on the 64-station baseline correlator (BLC) are performed in a coarse spectral resolution mode – Time Division Multiplexing (TDM) mode – which for the standard dual-polarization products provides 128 channels of width 15.625 MHz (Escoffier et al. 2007). Due to the default selection of the Hanning window applied in the correlator, the effective spectral resolution of ALMA data (that has not been spectrally pre-averaged) is 2.0 times the channel width (Hills 2012). In dual-polarization TDM mode, this resolution is 31.125 MHz, or $\approx$ 40 km/s in Band 6, which is insufficient to resolve the cores of atmospheric line features, primarily those of ozone. The reason behind this choice was twofold: **(1)** the quantization correction (QC) for ALMA's 3-bit digitizers is essential to obtain accurate calibration spectra (particularly as the input power levels change between the loads and the sky), and the full QC is currently available on the BLC only in TDM; **(2)** the manipulation of these spectra in the online software module TelCal (Broguiére et al. 2011) require much more processing time when high resolution observations are attempted in Frequency Division Multiplexing (FDM) mode spectra with (up to) 3840 channels per spectral window (spw). Due to the smaller number of antennas, item (2) is not an issue for the 16-station Atacama Compact Array (ACA) correlator (ACAC), and it has been acquiring full resolution Tsys since the start of Cycle 3. Furthermore, the ACAC was upgraded so that total power information is sent directly to the CDP cluster, which enabled the calculation of accurate quantization corrections on the integration timescale. However, bringing a similar capability to datasets obtained with the BLC requires additional work. There are two required ingredients: **(1)** the ability to apply the proper quantization correction to atmospheric calibration scan data recorded in FDM; **(2)** the ability to compute system temperature spectra at higher spectral resolution from these corrected data for (up to) 64 antennas.

3 Outline of the problem

ALMA atmospheric calibration (AtmCal) scans consist of sequential auto-correlation spectral measurements of the sky, the ambient temperature load, and the heated load contained in the ALMA Calibration Device mounted above the front-end (Casalta et al. 2008). As these data are being collected, TelCal computes the receiver temperature spectra (Trx), the sky temperature spectra (Tsky), and the system temperature spectra (Tsys), which it then stores into the ALMA science model (ASDM) dataset as the observation progresses. Lines of ozone (and other species) cause the autocorrelation spectrum and consequently the Tsys to spike upward over a narrow frequency range (Figure 1a), which is expected because their presence represents a real loss of sensitivity in those channels.

In CASA, the Tsys spectra are applied to the data to bring them onto a Kelvin scale and to flatten the amplitude vs. frequency. For high resolution spectral windows (FDM spws), CASA interpolates the Tsys spectrum to narrower channels as it applies the correction. However, because the lines are not fully resolved, after this Tsys calibration is applied, the subsequent antenna-based bandpass calibration solutions for the FDM spws still show significant features at the frequencies of the ozone lines (Figure 1b). In Cycle 0 data, this effect was primarily noticed as narrow residual dips ($\sim$5-20%). In more recent, higher S/N data, we sometimes also see broad emission wings in addition to the central dip, perhaps indicative of inadequacies of using the default atmospheric model parameters for ozone. In any case, these narrow features manifest in the bandpass solution because they are not removed by the coarse Tsys calibration. To the extent that these residuals are similar in the bandpass calibrator and the science target, one would imagine that they would be removed after bandpass calibration. The problem is that these residuals will depend on the elevation and there is no facility to correct for this dependence in CASA. Thus, even though the bandpass calibration applied to the bandpass calibrator looks reasonable (Figure 1c), the residual features propagate from the bandpass solution to the science target (Figure 1d) because the contrast between the atmospheric line and continuum emission is different. An exacerbating factor is the fact that the observatory-accepted observing modes typically do not observe Tsys on both the phase calibrator and the science target. When Tsys is available only on the phase calibrator, CASA will apply this same Tsys spectrum to the science target (since it will be the nearest target with a Tsys), and vice-versa. In this case, either the science target or phase calibrator will not have an equally-optimal Tsys spectrum applied to it as was applied to the bandpass calibrator, so any difference in sky conditions between these objects (i.e. due mostly to an elevation difference) can lead to a further residual feature.

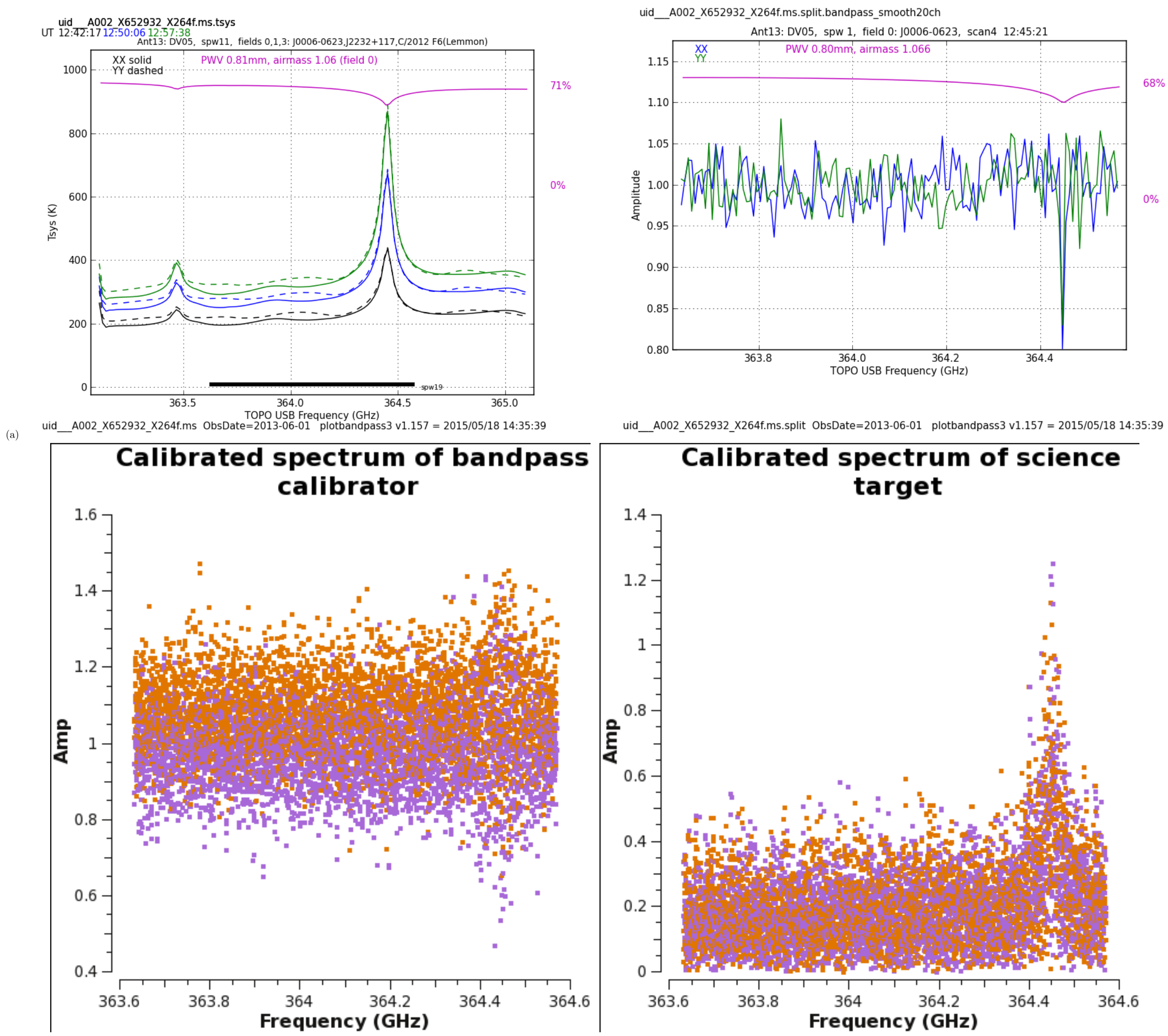

Figure 1: This figure demonstrates the quality of ALMA calibration of science datasets including strong ozone lines. This example comes from an observation on 2013-Jun-01 12:39 UT. **Top left panel)** TDM Tsys spectra recorded with the standard 31 MHz resolution on the same antenna (DV05) at 3 different times and elevations (corresponding to observations of the bandpass calibrator in black, flux calibrator in blue and science target in green) during a Band 7 science observation. The atmospheric transmission is shown in the magenta line. Note the effect of the strong ozone line at 364.45 GHz, and how the relative line strength increases at lower elevation. **Top right panel)** Bandpass solution solved (using 7.3 MHz pre-averaging) for the corresponding science FDM spw, solved after application of the (spectrally coarser) TDM Tsys spectrum. Much of the broad feature is successfully eliminated, but note the 20% drop at the line center due to the residual ozone feature. This feature could potentially be eliminated if the Tsys spectrum was acquired at higher spectral resolution. Note that the frequency span of the FDM bandpass solution is half that of the TDM Tsys spectrum. **Bottom left panel)** Calibrated spectrum of bandpass calibrator with 1 MHz effective resolution. It has been calibrated for Tsys and for bandpass (by itself), and averaged over the DV05 baselines and colorized by polarization. The result is relatively flat as expected, with simply higher noise at the frequency of the ozone line. **Bottom right panel)** Calibrated spectrum of a line-free science target, which shows a strong anomalous feature consistent with the shape of the ozone line, even though Tsys was independently measured on the science target and applied prior to the bandpass correction. The effect of the residual feature in the bandpass solution is large due to the large line strength and the large difference in elevation between the bandpass calibrator (70°) and science target (31°).

4 Proposed solutions that motivated this Study

4.1 Observational approach for new data (Cycle 7 onward)

One obvious solution is to observe the AtmCal scans in the same correlator mode as the science data. This approach requires two major improvements to the online processing: (1) the computation and application of an accurate quantization correction (QC) for FDM data; (2) an increase in the capacity of online TelCal to process ∼15000 channels from 64 antennas in a reasonable amount of time (less than ∼1 minute).

4.1.1 Requirements for QC of FDM data

First, although an FDM QC is in the long-term plan of the BLC software, it is not high in the list of priorities primarily because (aside from enabling FDM Tsys) the benefits of a proper FDM QC are less pronounced to the user. Scheduling of this improvement will likely not be reached before the correlator upgrade development project[1]. Furthermore, if an FDM QC is someday implemented on the BLC, it will require extensive re-commissioning of the correlator software, a process which becomes more difficult since the Commissioning and Science Verification (CSV) period ended in March 2014, and the Extension Of Capabilities (EOC) period ended on October 1, 2015. Future commissioning and verification tests will fall to the Operations staff and the ALMA Regional Center (ARC) staff which have more limited resources for such work. However, an equivalent FDM QC could, in principle, be computed and applied entirely in TelCal. Moreover, for the purpose of FDM Tsys, it would only need to be applied to the AtmCal scans. The only requirement is access to a sufficiently rapid measurement of the baseband (BB) total power. We already have total power square law detectors (SQLDs) in the IF Processors which can be cross-calibrated with the digitizers, and these data can be written to the ASDM and thereby accessed by TelCal. It should be relatively easy to combine the uncorrected FDM data from the correlator with the BB powers to produce corrected FDM data within TelCal. Before he retired, Robert Lucas made some progress in the past trying this technique with TelCal and seemed confident in it (Lucas 2013). The fact that we derive similar BB-averaged Tsys and Trx from the BB detectors and from TDM data suggests that this technique should work. We thus decided to pursue implementation of the FDM QC in TelCal.

4.1.2 Requirements on TelCal

Second, online TelCal needs to be able to process FDM data which has a much larger number of channels per spw (up to 3840 vs. 128). As such, its performance will need to be improved, because currently the computation of Tsys for FDM data takes too long to complete in a reasonable amount of time, hence delaying subsequent observations. This requirement will remain even if the FDM QC is someday performed in the correlator. Although the calculations can in principle be run post-observation using offline TelCal, this change would further burden and disrupt the normal process of Quality Assurance Level 0 (QA0). Every FDM dataset would need to be extracted from the Archive, processed, and restored to the Archive before its quality (i.e. including Tsys) could be assessed for QA0. One could imagine observing both a normal TDM Tsys and an FDM Tsys, allowing QA0 to continue to immediately assess the current coarser Tsys, and only compute and apply the higher resolution Tsys at the QA2 stage (i.e. in CASA). But this compromise would still double the observing time required for atmospheric calibration. Clearly, the optimal goal for this approach is for the ASDM to contain full-resolution Tsys spectra at the conclusion of observations.

4.2 Theoretical approach for archival data

An alternative solution that would work for archival data (Cycles 0-6) would be to use our knowledge of the physics of atmospheric line profiles, as encoded in the ATM atmospheric model (Pardo et al. 2001). We could attempt to fit the observed TDM Tsys spectra by varying a few parameters of the atmospheric model tool (already in CASA), and then use those parameters to generate artificial high-resolution Tsys spectra to be applied to the data in CASA in place of the TDM Tsys. This approach requires development of a python task to perform the fitting and generation of the new Tsys calibration tables. Two such parameters are the tropospheric temperature lapse rate (default = -5.6 K/km) and the upper bound of the atmosphere (default = 48 km). Normally, these default values do a decent job of removing ozone

[1] We have communicated to R. Lacasse and R. Amestica the desire for future correlator to perform FDM QC

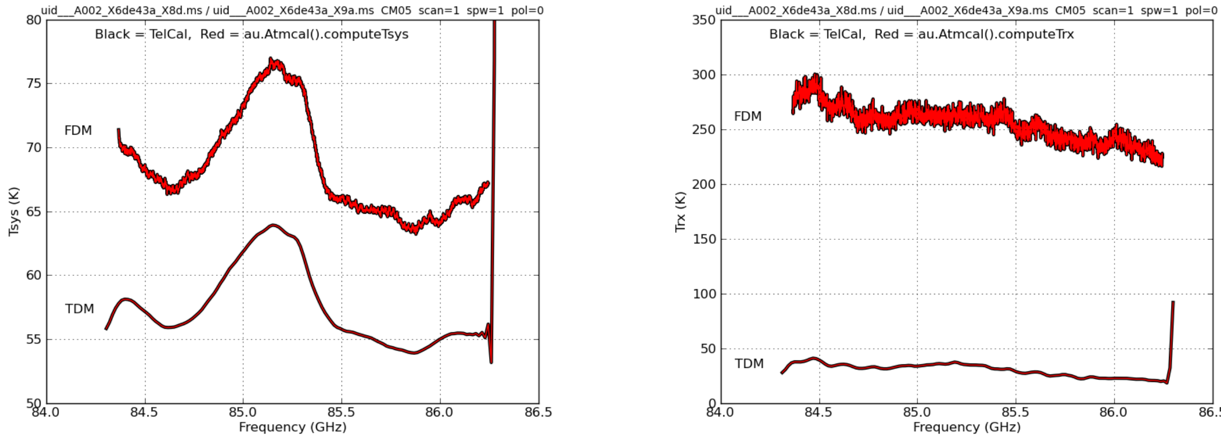

Figure 2: This figure demonstrates why Tsys measurements on the BLC currently must be performed in TDM. These datasets are from 2013-Sep-28 with 11 antennas **Left panel)** Tsys spectra obtained in FDM (upper traces) and TDM (lower traces), observed one right after the other in Band 3. The black curves are from the online TelCal calculation while the red curves are the offline calculation in analysisUtils, which agree perfectly. The DC offset between the two curves is due to the lack of quantization correction in FDM. **Right panel)** The same plots for the Trx spectra, which differ more from one another than the Tsys spectra, due to the lack of quantization correction. The FDM Trx is clearly wrong, given the known performance of the Band 3 receivers (and the fact that it is higher than the Tsys!). Note that the FDM spectra are by their nature narrower (1875 MHz) compared to TDM (2000 MHz). As such, they avoid the ringing edge effect from the digitizer channel 0 value of the baseband.

lines from TDM data. Increasing the lapse rate to -10 K/km causes the residual lines to appear more in emission after Tsys calibration, while decreasing it causes them to appear more in absorption. Also, by increasing the upper bound to 75 km, the profiles of the ozone lines get more peaked up, more similar to the high-resolution Tsys spectra measured in our FDM tests (see the following section). This should not normally affect Tsys measured in FDM in the case of a pure single sideband receiver. However, if the sideband rejection is not perfect (and it is not for ALMA receivers), then a correction is applied by TelCal in the expression for Tatm, which depends on the model profile in the signal band and on the image sideband ratio. So, this may affect the derived line shape in Tsys, even more if an error is made in the measurement of the image sideband gain ratio (Lucas 2013). Thus we may find that we need to adjust the upper bound to get the best performance from the online TelCal FDM Tsys calculation. In any case, the real FDM Tsys measurements that we acquire can be used to validate the modeling approach. If successful, then the model python task can in the future be added as a CASA task to correct the Tsys recorded in archival datasets.

5 Initial (pre-Study) tests of the observational FDM Tsys method

5.1 Observations

In September and October 2013, we tested the observational approach of measuring full FDM Tsys using the custom DelayCal_atmFDM.py script. For the first test (September 27), we recorded a scan on a bright quasar (called a DelayCal scan) with an initial AtmCal scan in TDM followed by a similar sequence in FDM, using 11 antennas in Band 3. Online TelCal processed both AtmCal scans to completion, requiring only 5 seconds on the TDM scan, but more than 2 minutes for the FDM scan. Because the FDM execution block ended before TelCal started to process the quasar scan, there were no DelayCal results from the quasar observation stored in that dataset (indicating that this method will not be viable without a reduction in the AtmCal calculation time). Nevertheless, a comparison of the Tsys spectra shows very similar shapes in FDM and TDM but with a significant (20%) DC offset (Figure 2). The offset is due to the lack of quantization correction in the FDM data.

For our second test, on October 22, 2013, we used Band 6 to record a similar pair of DelayCal scans first in TDM, next in FDM with 1875 MHz windows, and finally in FDM with 468.75 MHz windows. The tuning was chosen to include several ozone lines – the baseband center frequencies were 231.8, 233.9, 247.7 and 249.8 GHz. We observed quasars at two different elevations (J0522-364 at 75 degrees and J0854+201 at 32 degrees). In the resulting datasets, the SysCal.xml file was empty, which was not too surprising given the time required to calculated Tsys for 3840 channels exceeded the timeout (we were using software version ALMA64-9_1_3-B-2013-10-16-03-00-00). Therefore, we processed the FDM Tsys scans with offline TelCal. The linewidths of the ozone emission lines in the Tsys spectra are noticeably broader in the TDM data compared to the FDM data (Figure 3).

Next, the October data were calibrated as follows: online flags were applied, and 8 edge channels flagged in TDM corresponding to the roll-off of the anti-aliasing filter (as per normal manual and pipeline calibration procedure), followed by phase-only gaincal on the 'integration' timescale. Two bandpass solutions were then computed: one with the TDM Tsys and phase-only gain tables pre-applied, and one with the FDM Tsys and phase-only gain tables pre-applied. We then overlay these two solutions with the CASA task plotbandpass. Both the broad emission wing and narrow dip are eliminated (Figure 4). There are some minor broadband differences between these bandpass solutions, and the magnitude of the differences varies from antenna to antenna. But these differences are likely due to the fact that the FDM Tsys spectra have been constructed from data lacking a proper 3-bit quantization correction. Thus, we are seeing only the improvement due to the higher spectral resolution, and some anomalous broadband structure may be present. This improvement means that with a proper FDM measurement of Tsys on the bandpass calibrator and science target, no ozone line will remain in the bandpass solution and thus it cannot be transferred erroneously to the science target.

5.2 Initial test of a first order Quantization correction

For our third test, we used the datasets from the second observation to test an offline application of a 3-bit quantization correction, using the method of linear approximation described in § 7.3 of ALMA Memo 583 (Comoretto 2008). That is, only the linear term of the full three-term polynomial expression is used. We used the TDM spectra acquired closest in time to the FDM Tsys scan (8 minutes apart) as an estimate of the baseband total power, taking care to confirm that the IF attenuators had not changed between the scans, as that would violate the underlying assumption of constant power vs. time. We wrote and used a python class called "Atmcal" (in the publically-available analysisUtils package) to reproduce the calculations performed by TelCal to compute Trx and Tsys, which are described in Lucas & Corder (2005) and Lucas (2012). The result for Tsys in one baseband is shown in Figure 5. This baseband contains the 16(1,15)-16(0,16) ozone line at 231.281511 GHz[2]. In contrast to Figure 2, this measurement of FDM Tsys is now in excellent agreement in both shape and level with the normal TDM Tsys. This result provides great promise for our proposed study on how to implement and test this technique in online TelCal. Application of the FDM quantization correction to the quasar visibility data (in addition to the FDM Tsys correction) should also remove the small differences in the broadband shape of the bandpass solutions (Figure 4). Confirmation of this result is pending the development of script to modify the visibility data in CASA.

6 Progress achieved on deliverables during the Study

This Study included work in three areas described in the following subsections: hardware calibration and calibration infrastructure (§6.1), online software components including TelCal, DataCapturer and ATM (§6.2, 6.3, and 6.4) and offline software modules in analysisUtils (§6.5). The official period of work extended from March 2016 through September 2017.

6.1 Calibrating the SQLDs against the digitizers

6.1.1 Accuracy required

We computed the effect of relative gain errors and offset errors upon the derived Trx and Tsys values (Figure 6). The relative gain is the more important parameter to get right. We need ≤ 0.1 dB accuracy which is not difficult but better than we currently have on most IF processors. In contrast, typical offset

[2]Information from Splatalogue at `www.cv.nrao.edu/php/splat/advanced.php`

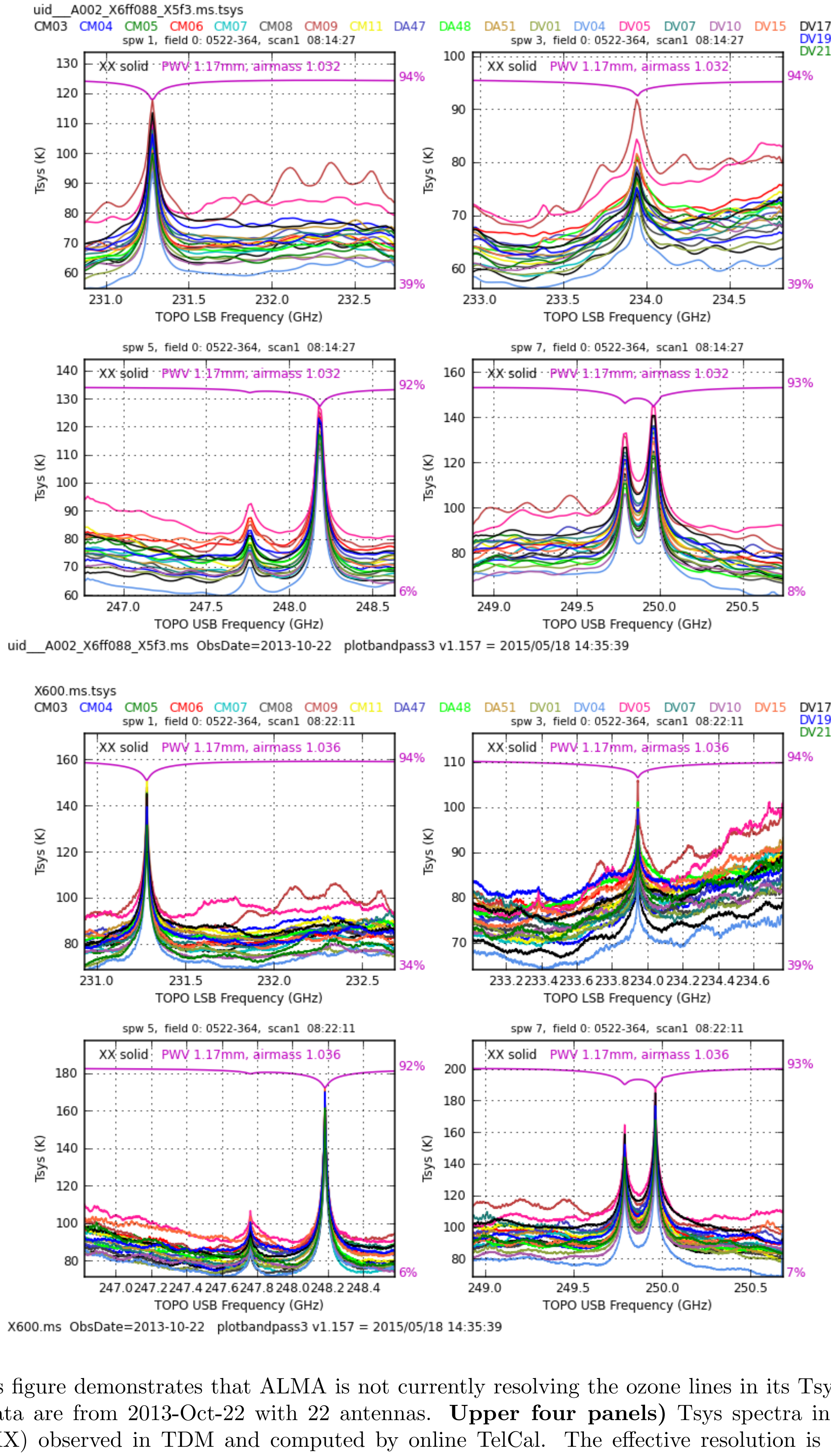

Figure 3: This figure demonstrates that ALMA is not currently resolving the ozone lines in its Tsys spectra on the BLC. These data are from 2013-Oct-22 with 22 antennas. **Upper four panels)** Tsys spectra in Band 6 for one polarization (XX) observed in TDM and computed by online TelCal. The effective resolution is 31.25 MHz. All antennas are overlaid. The magenta line is the atmospheric transmission curve. **Lower four panels)** Tsys spectra for the same tuning observed in 8 minutes later in FDM and computed by offline TelCal. Note the much narrower appearance of the ozone line peaks when the effective resolution is 0.9765625 MHz instead of 31.25 MHz.

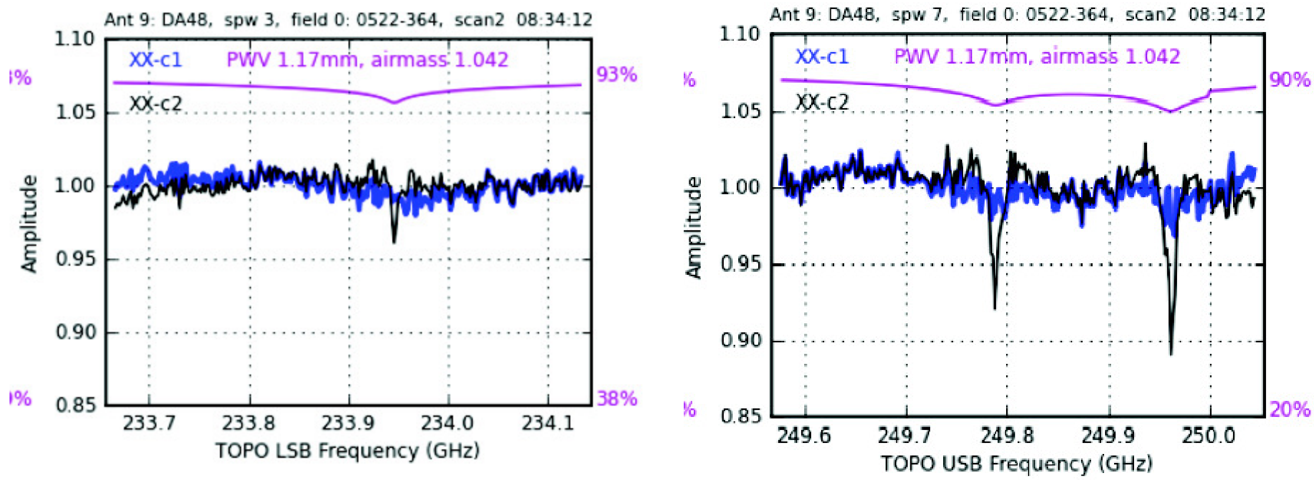

Figure 4: This figure demonstrates that measuring and applying higher resolution Tsys can improve the bandpass solutions obtained in CASA. These data are from 2013-Oct-22 with 22 antennas. In both panels, the black spectrum is the bandpass solution measured on quasar J0522-364 obtained after applying the normal TDM Tsys, while the blue spectrum is the bandpass solution obtained after applying the (experimental) FDM Tsys. The effect of the ozone lines has been almost completely eliminated. The minor residual differences on broad scales are likely due to the fact that the FDM Tsys spectrum was computed from data that lack a proper 3-bit quantization correction. The bandwidth of the spw is 468 MHz. The effective resolution of the data and the Tsys is 0.25 MHz, while the bandpass was solved using 0.5 MHz pre-averaging. The magenta line is the atmospheric transmission curve. The glitch at 250 GHz was due to a non-optimal configuration of the ATM code, which was fixed as a result of this study (see § 6.4.1).

errors of $\pm 10\mu$W produce less than 1% error in Trx and Tsys. The relative gain error can be estimated from any TDM subscan, by simply comparing the detector power with the autocorr value. The reference point is that +3.8 dBm == 2.3988 mW should give TDM autocorr of unity. Using this simple comparison requires that both the detector and digitizer offsets are reasonably accurate first. To reach that goal, the plan is to use subscans over a very wide range of powers to solve for all three independent parameters together (detector offset, digitizer offset, relative gain. Since the digitizer thresholds are being set to fixed values by the Correlator group, it should simply be a matter of adjusting the IF detector gains and offsets to compensate. For the IF Processors, it seems there are no objections to the Science group being able to update the calibrations as needed. A preliminary summary of this work was presented as a poster at SPIE meeting 9914 in Edinburgh in July 2016.

6.1.2 Moving the IF Processor calibration values to the TMCDB

To allow streamlining of routine updates of the IF Processor calibration parameters – primarily the detector offsets – we submitted an ICT ticket (ICT-8592) to move them from configuration files into the TMCDB, similarly to delay/pointing/focus models. There should be support added to tmcdb-hardware-explorer, and import/export model scripts added for bulk updates. History support should be included.

6.2 TelCal improvements

6.2.1 Speed improvements

In the original parallelization of the atmospheric calibration code in TelCal (May 2014), there was a limit of 4 threads in use. This limitation was removed by September 2017 (ICT-8009). Tests using a machine with processors that supported 7 threads yielded a factor of 1.53 improvement in speed. A further factor of 2-3 improvement was gained by adjusting the antenna altitude threshold that determines the number of absorption profiles to compute. Since some antennas are close together, there is little difference in the total path length of atmosphere they are viewing. Increasing the altitude threshold to 10m (equivalent to 0.05 K and 80 Pa using typical temperature and pressure lapse rates) yielded no discernible change in the Tsys results (ICT-8465). Finally, an option to pre-average the AtmCal data by powers of 2 (from 2 to

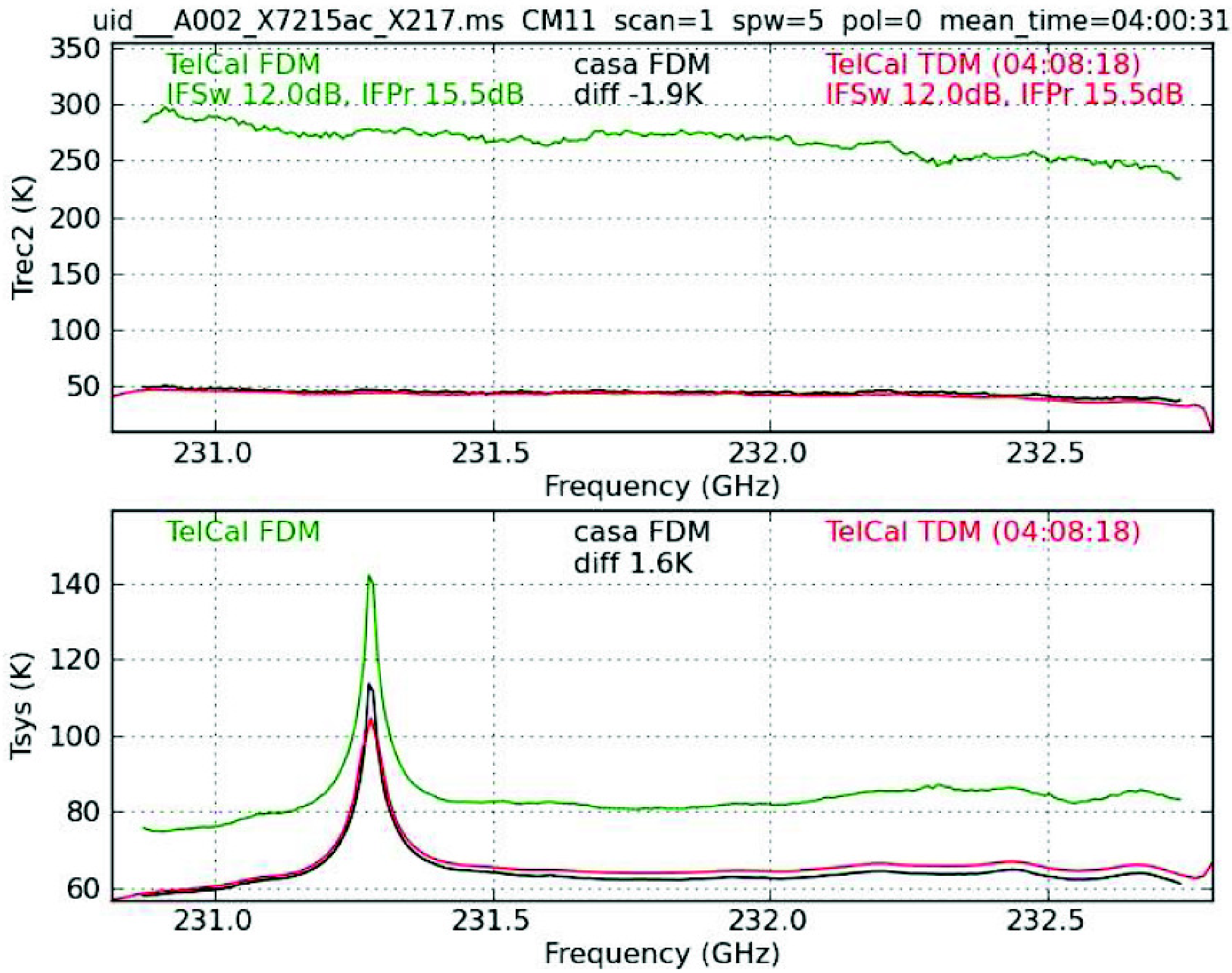

Figure 5: This figure demonstrates that by applying the linear approximation of the full 3-bit quantization correction allows us to compute FDM Trx and Tsys spectra (black curves) that agree in both shape and level with the normal TDM spectra (red curves). The upper panel is Trx and the lower panel is Tsys in one baseband. The green curves are produced from uncorrected FDM data, while the black curves are produced after applying the proposed quantization correction with a python script (using the same atmospheric model in CASA that is used by online TelCal). These data are from 2013-Nov-09 with 32 antennas.

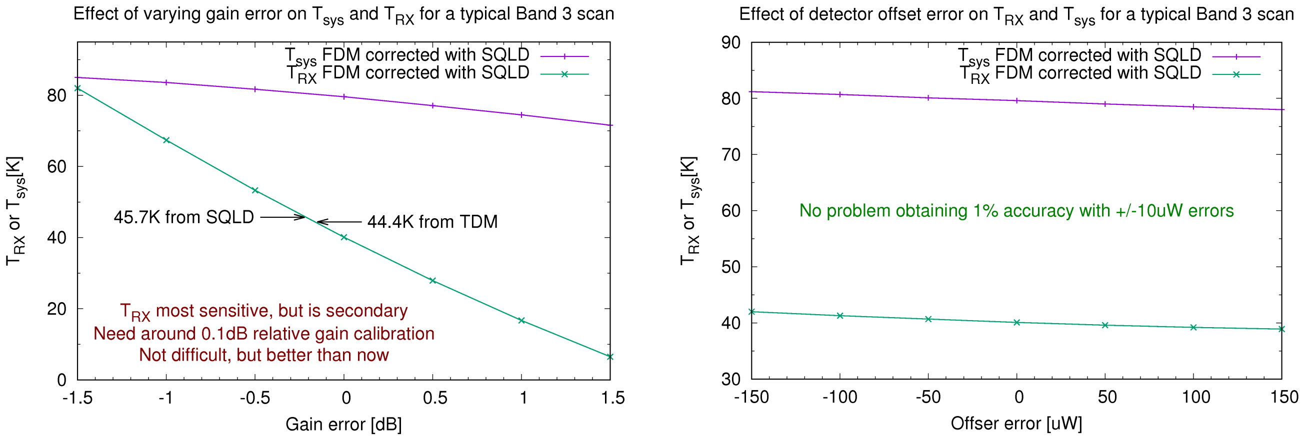

Figure 6: **Left panel:** Theoretical sensitivity of the derived Trx (green curve) and Tsys (red curve) to errors in the relative gain between the SQLD and digitizer. **Right panel:** Same for the relative offset between the SQLD and digitizer.

32) was added to support the possibility of using such averaging to speed up the calculations (ICT-9067). Tests showed a factor of 1.84 increase in speed when using a factor of 2 binning, which corresponds to the spectral resolution of the data when the online Hanning smoothing is active (by default in all ALMA data). Further improvements may be possible in the future if needed, such as using a common set of upper atmosphere layers for whole array and only computing the lower layers per antenna or group of antennas (ICT-10973).

6.2.2 Support for dual stream mode

By January 2017, the ability had been added to TelCal to receive data from two streams – the correlator and the SQLDs, in order to enable it to process them together to compute the QC (ICT-7499). The ability to invoke the dual stream mode at times other than a restart of TelCal still needs to be added.

6.2.3 Support for QC calculation

Furthermore, at the same time, TelCal's tc_atmosphere task was given the ability to apply the QC obtained from combining the SQLD and correlator data. The formula in use for QC needs to be further validated but it is giving more sensible values now for Tsys and Trx compared to the standard TDM values. The method used is:

- Average the TP integrations within each correlator integration (per baseband and polarization).
- Use equations 21 and 23 from ALMA Memo 583 to get the coefficients (a and b) from each averaged power value. (See Appendix A for code.)
- Loop over the autocorr spectral channels in the integration, replacing the values (v) by (a*v + b).

The new option (*quantcorrection=True*) was exercised on the three single-scan datasets of 2016-May-12 taken on the TFINT correlator with antenna DV22 at the OSF. They have spw bandwidths of 234, 117, and 59 MHz, respectively, and 3840 channels each. In Figure 7, we see that even without proper calibration of the total power detectors, the Tsky spectral data are brought into decent alignment with theory. We also see that using a higher maximum altitude in the ATM model would better reproduce the sharp peak of the line core. However, even with that modification, the predicted profile still differs in shape from the observed profile. That is, if one aligned the peaks, the wings of the line would be too broad. Qualitatively, this result is consistent with there being too much ozone in the model at high pressure (low altitude), suggesting that a redistribution of the ozone layering is required, not simply a change in the maximum altitude. This possibility is explored further in § 6.5.2 by comparing the ATM model result for atmType='tropical' (one of the six available approximate models in ATM and the default used by TelCal) and 'midLatitudeWinter'.

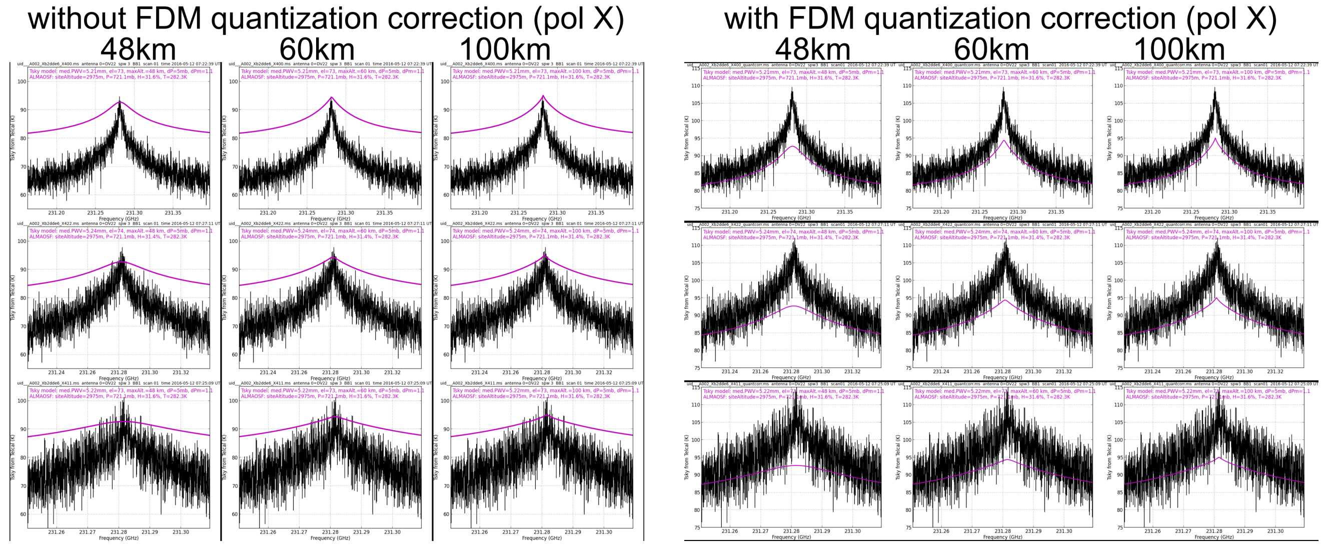

Figure 7: **Left panel:** Tsky spectra observed on 2016-May-12 with DV22 on TFINT from polarization XX (black) overlaid with predicted spectrum (magenta) with no QC applied and with three different maximum altitudes (columns). The top row is the 234 MHz spw (61 kHz channels), the middle row is the 117 MHz spw (31 kHz channels), and the bottom row is the 58 MHz window (15 kHz channels); **Right panel:** Same as left panel but with the new QC applied to the observed data. The mean level and shape are now better matched to the ATM model (as expected), but the central profile is still more sharply peaked than the model.

To confirm that the problem with the ozone profile is not somehow due to the lack of FDM QC on the TFINT data, or an error in our proposed FDM QC that we have applied it, we have similarly examined a Band 6 ACAC 7m dataset from 2015-Dec-29 observed at the high site for project 2015.1.00357.S. The ACA data have their normal (online) quantization correction applied. Figure 8 shows the Tsky spectra from two different antennas overlaid with the ATM model spectrum. The profile shape of the model does not match the observed profile, it is too broad relative to the height of the central peak, similar to Figure 7. There are also DC offsets on particular antennas. We should try to understand the origin of these offsets, but at least the distribution is centered near zero.

6.2.4 Other new features

As of July 2017, TelCal now has the ability to accept a vector profile for the temperature and pressure profile instead of constant lapse rates. We have not tested this in the context of Tsys calculations, but they could improve the match between the model and the observations.

6.3 DataCapturer improvements

The initial testing of FDM Tsys with many antennas revealed deficiencies in DataCapturer in handling the larger files associated with the large increase in Tsys channels. An unexpected latency in writing out the TelCal results to the database (close to the 1 minute timeout) led to the frequent (15%) loss of the subsequent WVR results that expired before they were written. This was fixed by Rachel Rosen in August 2017 (ICT-10614). After this improvement, some tables were still taking longer than 10 seconds, so a further improvement was made to the method used to write to the relational database, which was completed by the end of October 2017.

6.4 ATM improvements

During the course of this study, we identified several issues with the ALMA Atmospheric Transmission Model software written by Juan Pardo (Pardo et al. 2001).

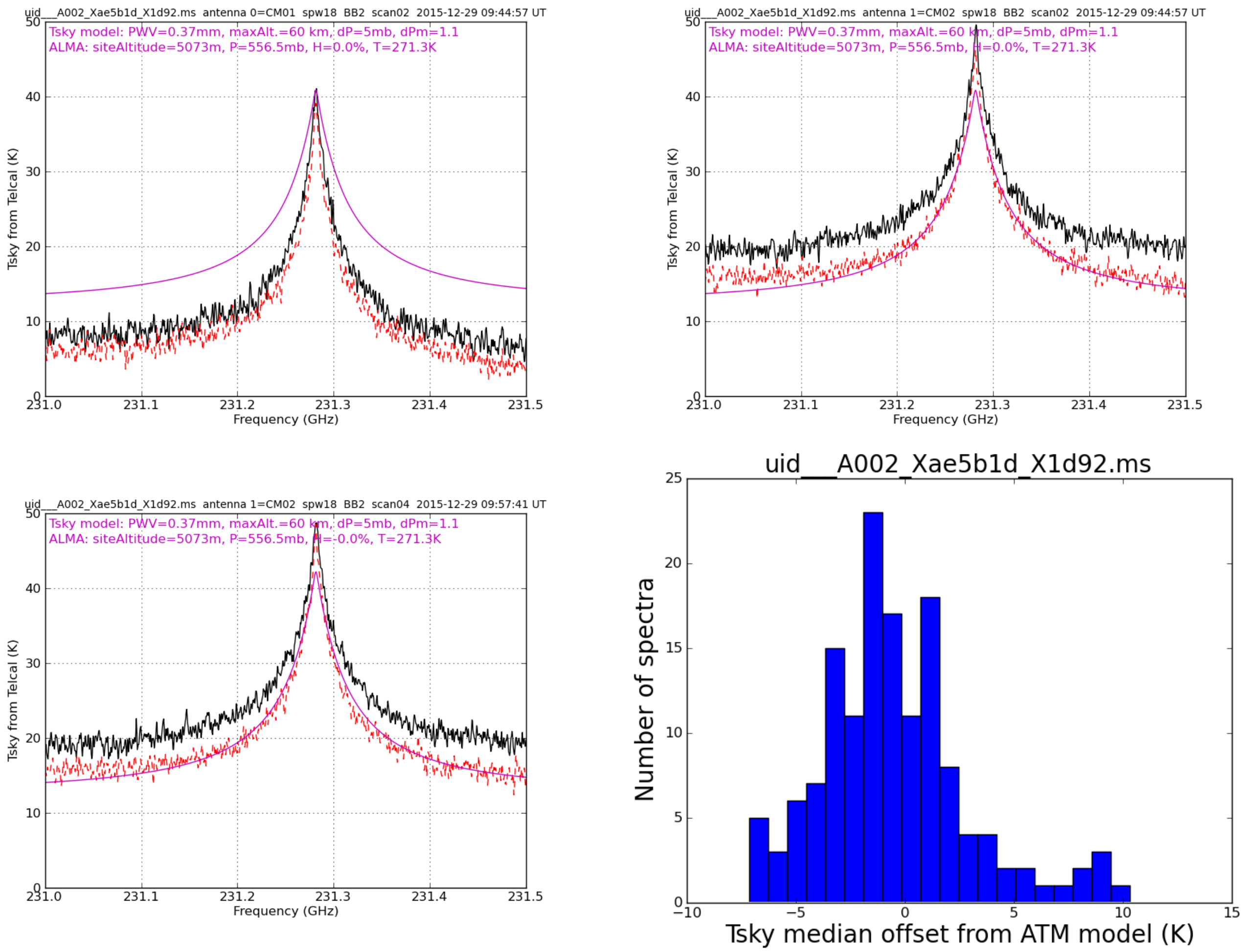

Figure 8: Sample Tsky spectra computed by TelCal (black=XX and red=YY) during observations with the ACAC on 2015-Dec-29 with nine 7m antennas (from science project 2015.1.00357.S) overlaid with predicted spectra (magenta) output from ATM (TebbSky). **Upper Left panel:** antenna CM01 scan 2; **Upper Right panel:** antenna CM02 scan 2; **Lower Left panel:** antenna CM02 scan 7. The DC level of CM02 pol YY is fairly well matched to the model, but CM01 has a large offset. More importantly, in both cases the discrepancy in the ATM profile (it being too wide for its corresponding peak) is similar to the OSF TFINT observation in Figure 7. **Lower Right panel:** A histogram of the DC offset between TelCal's Tsky and ATM model over all 144 spectra (9 antennas, 2 pols, 8 scans). The reason for the scatter has not yet been investigated but at least the center is near zero.

6.4.1 Minor bugs discovered (and fixed)

First, two minor bugs were found. There were small discontinuities in the output spectra, mostly at multiples of 2 GHz, which can be seen at 250 GHz in Figures 3 and 5. The problem was fixed by Juan by June 2017 (ICT-7997, CAS-9015). It was due to a failure to include enough lines of O_2 and other trace gases in the performance trade-off between speed and accuracy. The second bug was found by detailed examination of the code after noticing an asymmetry in the residual ozone feature (see Figure 10 in § 6.5.1). A typo was preventing the application of some of the rare ozone isotopologues and double application of others. This problem was fixed by March 2017 (ICT-7470, CAS-8448), but a brief follow-up test showed that it was not causing the asymmetry.

6.4.2 Lack of Hanning smoothing of the model

The other issue we very recently discovered affects all atmospheric lines and frequency ranges. Apparently, the model does not distinguish between channel width and channel resolution when defining a spectral window for model output (at least via the CASA interface). Since ALMA data use online Hanning smoothing by default, the model needs to be Hanning smoothed before using it in the calculation of Tsys by TelCal, except in the case of online channel averaging (which makes adjacent channels more independent). For more details, see § 6.5.2. We plan to request the addition of a resolution parameter to the spw definition, and that the output model spectra correctly take this into account. An ICT ticket will be written in the near future pending feedback from the developer.

6.5 Offline Tsys simulation software

In a parallel effort, we produced Python software within the Atmcal class in the ALMA analysisUtils (au) suite[3] which generates synthetic high resolution Tsys spectra based on measured TDM Trx and Tsky spectra and the atmospheric model in CASA. It allows quantitative comparison with the new FDM Tsys observing technique as well as an avenue for improving older data. The main results are described in the following subsections.

6.5.1 Minimum required spectral resolution

We wrote the au.smoothTsys function to smooth existing FDM Tsys from Cycle 3+ science observations on the ACAC in order to compare calibrated target spectra using different levels of coarser Tsys spectra. In addition to smoothing, the function has a "decimate" option to reduce the number of channels by the specified factor in order to ensure that the behavior of the CASA applycal behavior is consistent for these tests. We searched for a Cycle 3 ACAC 7m dataset with spws covering ozone lines at high resolution, and with bright sources. We first used a Band 3 dataset with 110.84 GHz ozone line in a 2.0 GHz bandwidth spw with 1024 channels (channel width=1.953km/s, effective channel resolution=1.938MHz=5.16-5.88km/s) in execution block uid://A002/Xac5575/X8541 from project 2015.1.00804.S observed on 2015-Nov-01. We used smoothTsys to produce 512, 256 and 128-channel version of the Tsys spectra.

We ran the ALMA manual calibration script generator and compared the line residuals in the flux calibrator after the bandpass calibration stage having used these different Tsys spectra on both calibrators. The elevation separation between the bandpass calibrator and the flux calibrator was 18 degrees, from airmass 1.05 vs. 1.22 (Figure 9a). The baseline-averaged uv spectrum of flux calibrator Mars (Figure 9b) shows a significant line residual feature when calibrated with 128, 256 and 512 channels, but not 1024. This result suggests that < 5 km/s resolution is sufficient to remove ozone line cores. However, the non-ideal use of the un-Hanning-smoothed ATM model by TelCal (see § 6.4.2) may be affecting the results. Also, other stronger ozone lines may produce larger residuals. We next examined a Band 6 dataset from project 2015.1.00357.S with execution block uid://A002/Xae5b1d/X1d92 observed 2015-December-29. It has the 231.3 GHz ozone line (§ firstQCtest) in a 4096-channel 2000 MHz window and the flux calibrator is Mars. Since no online pre-averaging was performed, the effective resolution is 0.9765625 MHz, i.e. twice the channel width due to the default online Hanning smoothing. In this case (Figure 10), the narrow feature becomes apparent as soon as the channel width (and effective resolution) is increased to 2 MHz (2.6 km/s) across 1024 channels. This result means that Tsys data should be taken with a spectral resolution better than $\approx$1.5 km/s (i.e. < 1.1 MHz at $f = 220$ GHz). Therefore, the channel width should

[3]available at `https://casaguides.nrao.edu/index.php/Analysis_Utilities`

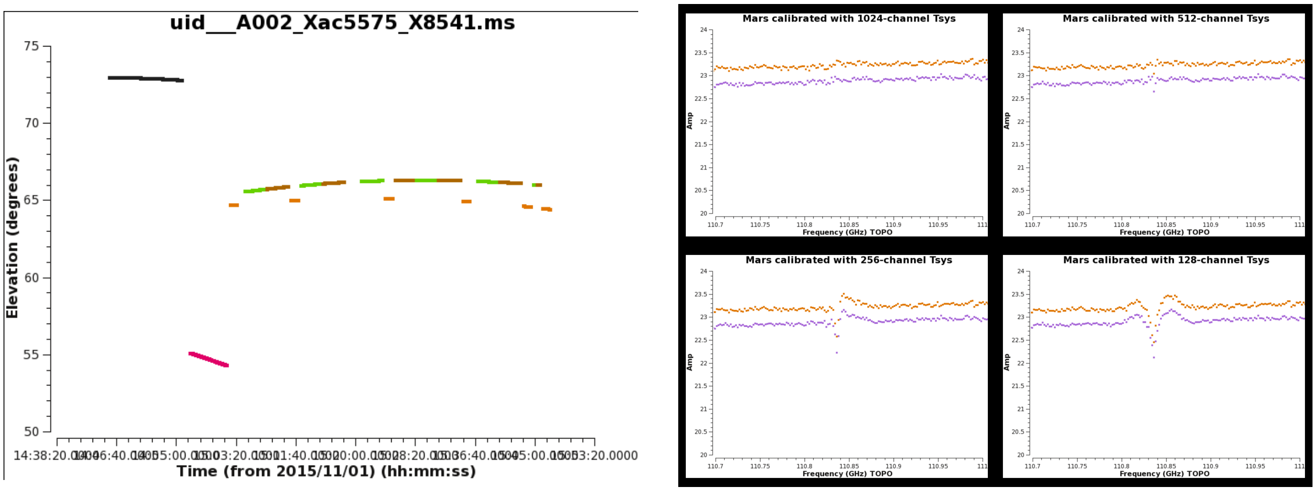

Figure 9: Band 3 Cycle 3 ACAC science observation with FDM Tsys. **Left panel:** Elevation of the targets vs. time: bandpass calibrator (black), flux calibrator (red = Mars), phase calibrator (orange) and science targets (green and brown); **Right panel:** Bandpass-calibrated spectrum of Mars using the four different resolution Tsys spectra generated by smoothing the observed FDM Tsys spectrum by 1, 2, 4 and 8 channels. The progression shows how the residual feature increases in strength as you apply coarser Tsys spectra.

be $< 0.0024 f_{\mathrm{GHz}}$. The required number of channels is listed in Table 1 for the dual-polarization case. However, these values should be considered preliminary until the other issues raised during this Study have been addressed.

Table 1: Minimum requirements for Tsys spectral setups[4] to recover ozone line profiles (preliminary!)

Frequency	Resolution	Minimum number of channels for spw bandwidth (MHz)[5]					
(GHz)	(MHz)	1875-2000	937.5	468.75	234.875	117.1875	58.59375
110	0.55	3840	1920	960	480	240	120
220	1.1	1920	960	480	240	120	
440	2.2	960	480	240	120		
880	4.4	480	240	120			

6.5.2 Generating high resolution Tsys spectra by modeling

The analysisUtils package contains python functions for manipulating ALMA metadata and binary data for commissioning, debugging, and quality assessment purposes. The class `Atmcal` was written about 5 years ago to reproduce the Trx and Tsys calculations from TelCal into python for various test purposes, and was confirmed to reproduce the results of TelCal. During this Development Study, we added a method to this class called generateHighResTsysTable, which accepts the original low resolution Tsys table and its parent measurement set as input and generates a new Tsys table at full resolution. For each combination of antenna, scan, spw, and polarization, it calls another new method simulateHighResTsys. A full list of parameters for both methods are shown in Appendix B. The latter method performs the following sequence:

1. Compute low-resolution spectra of the loads and sky from the autocorrelation data
2. Compute low-resolution Trx and Tsys spectra

[4] for dual-polarization spws
[5] Assuming online Hanning smoothing

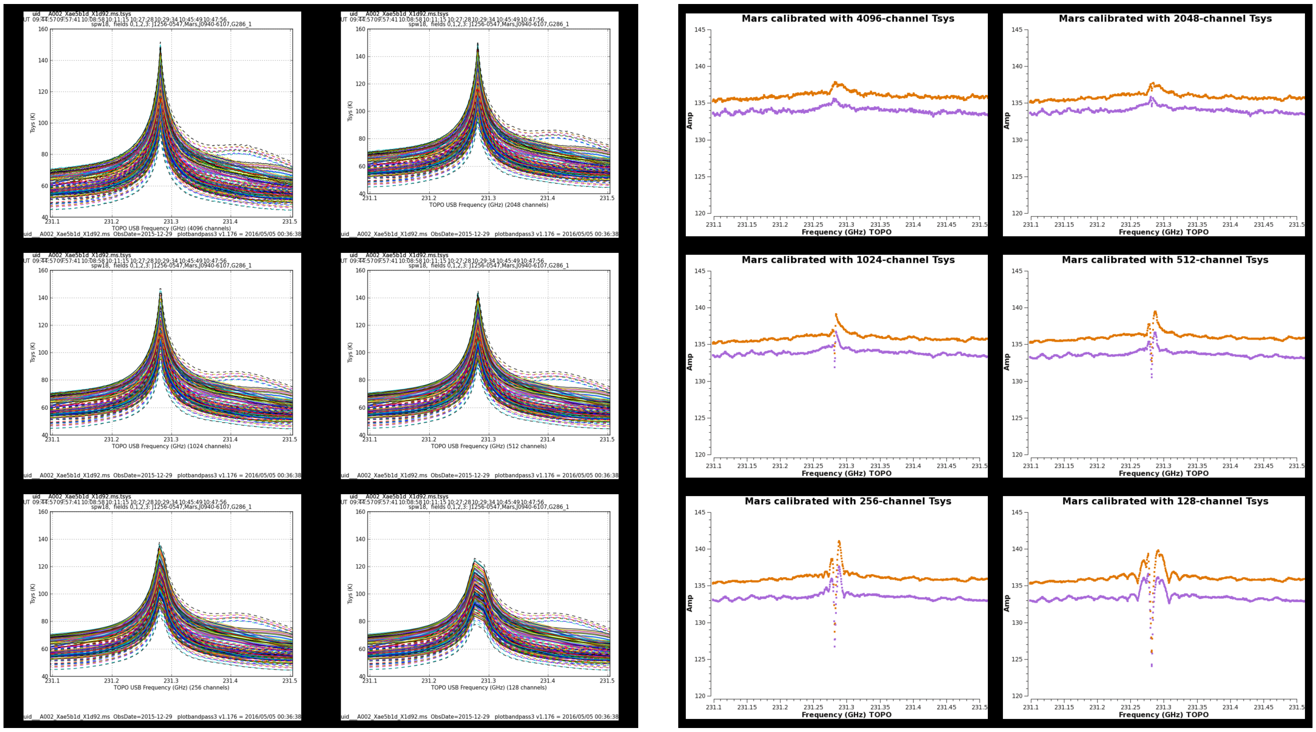

Figure 10: Band 6 Cycle 3 science observations. **Left panel:** Tsys spectra of all antennas overlaid showing the strong ozone line at 231.3 GHz, smoothed to different resolutions by factors of 2 from the original 4096 channels down to 128 channels; **Right panel:** Bandpass-calibrated spectrum of Mars having applied the different Tsys solutions with different effective resolutions.

3. Resample the spectra on the loads to the high-resolution science spw
4. Compute a theoretical high-resolution model of the sky spectrum using ATM
5. Combine the low-res sky spectrum with the high-res ATM model into a simulated high-res sky spectrum
6. Compute high-resolution Trx and Tsys spectra

Clearly, Step 5 is the tricky part. We must first assume that there are no spectral features due to the hardware that are narrower than the TDM resolution, but this is a fairly safe assumption (excluding the presence of pathological amplifier oscillations or birdies). The goal then is to replace the atmospheric features observed at low resolution with their true high resolution profiles, without biasing the mean continuum level of the low resolution observation. A few methods were tried, but what seems to work is to convert the observed sky spectrum to Tsky, then add the high-resolution Tsky spectrum calculated with ATM, then subtract the low-resolution Tsky spectrum calculated with ATM. This imparts the true high-resolution line profile onto the observed spectrum so that the re-solved Tsys will then show a more accurate (sharper) profile. Note that both ATM spectra must be Hanning smoothed when necessary (see § 6.4.2), a fact which we discovered while trying to produce pleasing results. Figure 11 shows the importance of applying Hanning smoothing to the model before re-computing the Tsky and Tsys spectra. Without it, non-sensical profiles are obtained from this simulation method. To achieve the smoothing, we used the pre-existing function analysisUtils.casaHanning which performs the standard triangular weighted smoothing in the spectral domain ($T_i' = 0.25 * T_{i-1} + 0.5 * T_i + 0.25 * T_{t+1}$) that mimics the correlator's application of the Hanning window in the time domain. In the analysisUtils implementation, we use the padOutput=True option, in which edge channels are retained, and are weighted in a 2:1 ratio with the adjacent channel. Of course, ATM could create spectra that are slightly wider than requested by the user, perform the smoothing, and then trim the edges. Ideally, it would perform the same steps as the correlator in the same order: Hanning smoothing followed by (in some cases) channel averaging.

To further test the hypothesis that the ATM spectra need to be Hanning smoothed, we examined directly the Tsky spectra created by TelCal for the Band 6 TDM dataset. To reduce vagaries in the

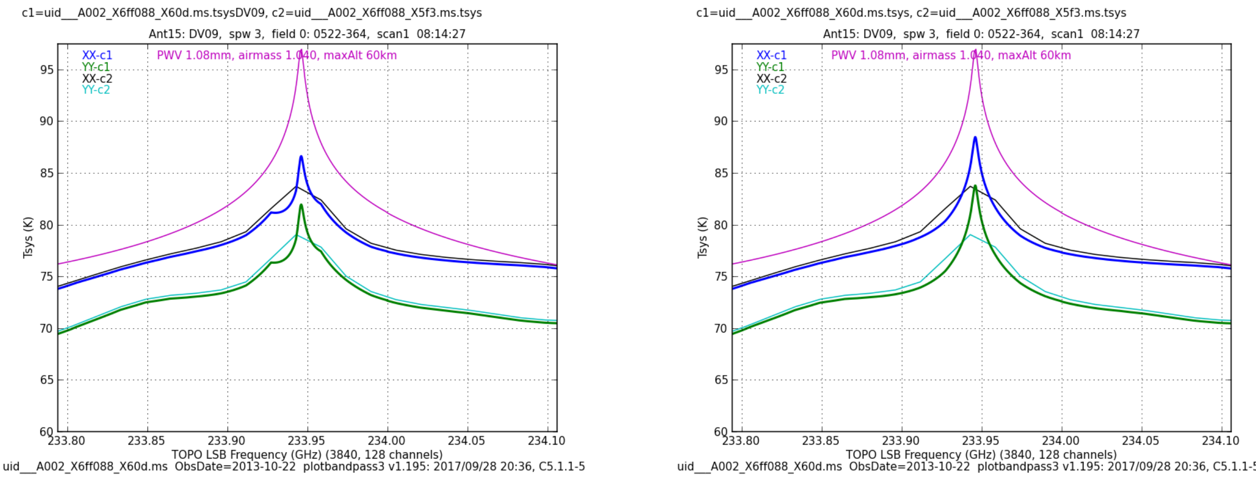

Figure 11: Comparison of simulated Band 6 Tsys spectra obtained by adding the high-resolution ATM model Tsky spectrum to the observed Tsky spectrum, subtracting the low-resolution ATM model Tsky spectrum and re-solving for Tsys. **Left panel:** Result with no Hanning smoothing applied to the ATM models before usage. The black and cyan lines are the XX and YY polarization of the TelCal TDM Tsys, while the blue and green lines are the simulated FDM Tsys. The magenta line is the full-resolution ATM model. **Right panel:** Same thing but with Hanning smoothing applied to the ATM models before they were used to create the blue and green lines. It is the same dataset as Figure 5. The non-physical feature of the left panel demonstrates the need to Hanning smooth the ATM model whenever the spw did not use online channel averaging. This conclusion should also apply to TelCal's usage of the ATM model.

profile between antennas, we computed the median spectrum over 20 good antennas, shown in Figure 12. We then overlay the raw ATM model in the left column, and the Hanning smoothed ATM model in the middle column. First of all, in both cases, we see that the low-level line wings are too broad. In spw 7 (bottom row), this is easily seen in the valley between the two lines – the model valley is much higher than the observed valley. Second, there is little difference in the FWHM before or after smoothing, mainly because the lines are significantly wider than the resolution. Third, we see, particularly in spw 3 (second row), that the unsmoothed ATM model is sharper than the observed spectra while the smoothed model is more similar. Therefore, we conclude two things: (1) Hanning smoothing is formally necessary, but (2) the layering in the ATM model is the larger culprit in the failure to reproduce the total line profiles.

6.5.3 Initial exploration of ozone layering

To explore the layering of ozone in ATM, we note that there are six ozone profiles defined in ATMProfile.cpp (see Appendix C), and ALMA TelCal is using the one labeled 'tropical'. The concentration profiles are plotted in Figure 13. Interestingly, this one shows higher ozone abundance at low altitude compared to the mid-latitude models, and lower abundance at high altitudes, suggesting that using the mid-latitude models may provide a better match to ALMA data. However, as the third column in Figure 12 shows, the mid-latitude winter model provides a worse match to the observations, but the magnitude of the changes are similar to what is needed, but in the opposite direction. Apparently, what is needed is a profile that lies below (i.e. to the left of) the 'tropical' profile between 15-25 km rather than between 25-40 km, in contrast to what the mid-latitude profiles provide. Clearly, further research is needed. A comparison with the ozone profiles used in Scott Paine's "am" model (Paine 2017) of the Chajnantor atmosphere would likely be fruitful[6]. In section 8.1, we describe the required additions to ATM that would allow this research.

In any case, an example of applying the results of a simulated FDM Tsys spectrum to PI astronomy data are shown in Figure 14. Here we compare the calibrated spectrum of a quasar after applying the observed TDM Tsys vs. after applying the simulated FDM Tsys. The weak lines are completely removed (spw 3). Stronger lines no longer have positive shoulders, but they still have negative narrow core features. This residual may be consistent with the non-ideal modeling of ozone that we have discussed.

[6] https://www.cfa.harvard.edu/~spaine/am/

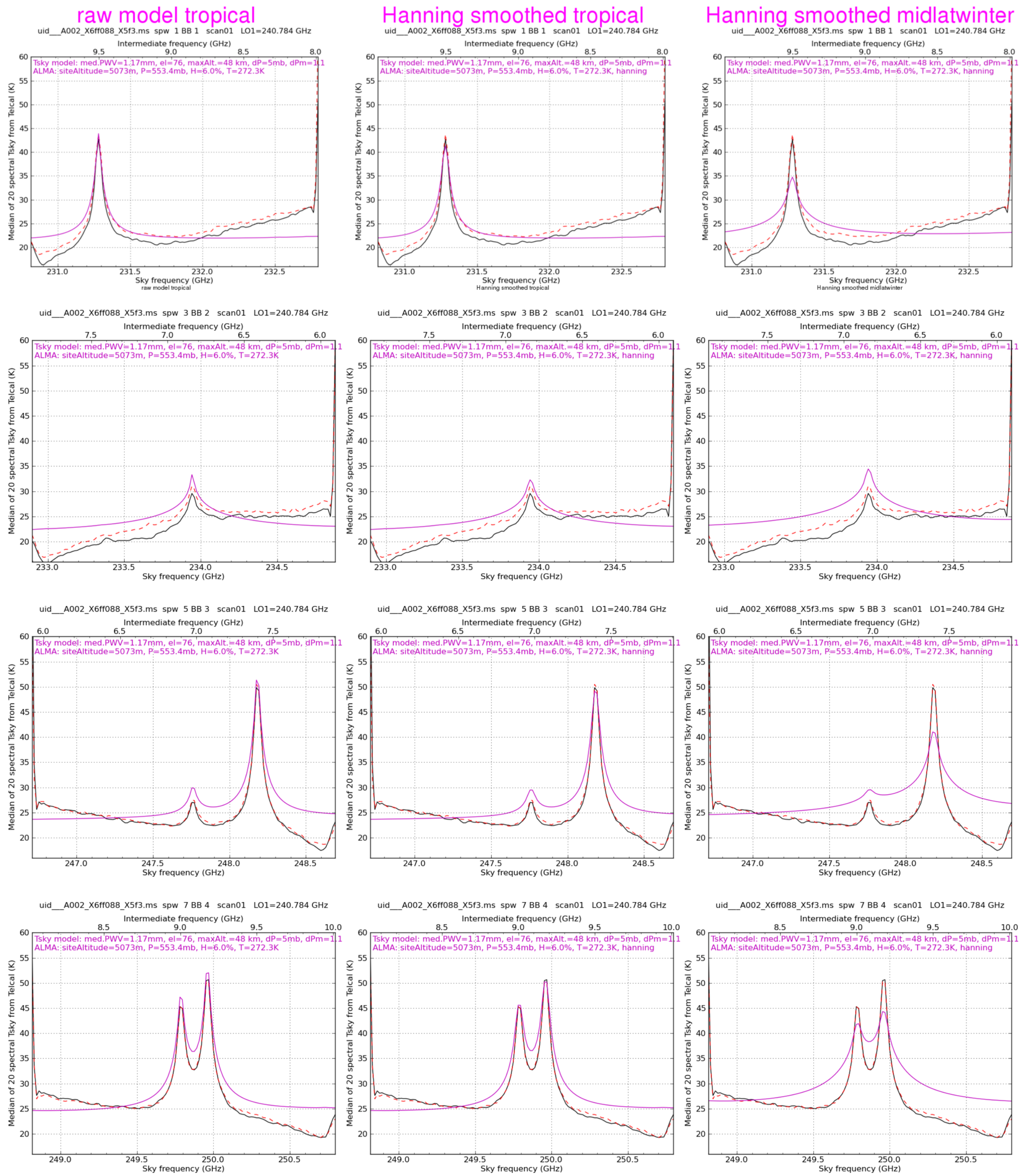

Figure 12: Median over 20 good antennas of the Band 6 TDM Tsky spectra created by online TelCal (red dashed line=YY polarization and black solid line=XX polarization) in the AOS execution uid://A002/X6ff088/X5f3 from 2013-Oct-22. **Left column:** with ATM tropical model overlaid (magenta); **Center Column:** with Hanning smoothed version of ATM tropical model overlaid. Rows correspond to spws 1, 3, 5 and 7. While the result for spw 3 indicates that the model profile is too sharp without Hanning smoothing, the bigger problem is that all the model profiles are too wide in the line wings, suggesting inaccurate layering in the model. **Right Column:** with Hanning smoothed version of the ATM mid-latitude winter model overlaid. Clearly, this model provides a worse match. Examining the profiles in Figure 13, it seems to be the portion below 25 km that is the most important to reduce, and the tropical model is already the lowest model available.

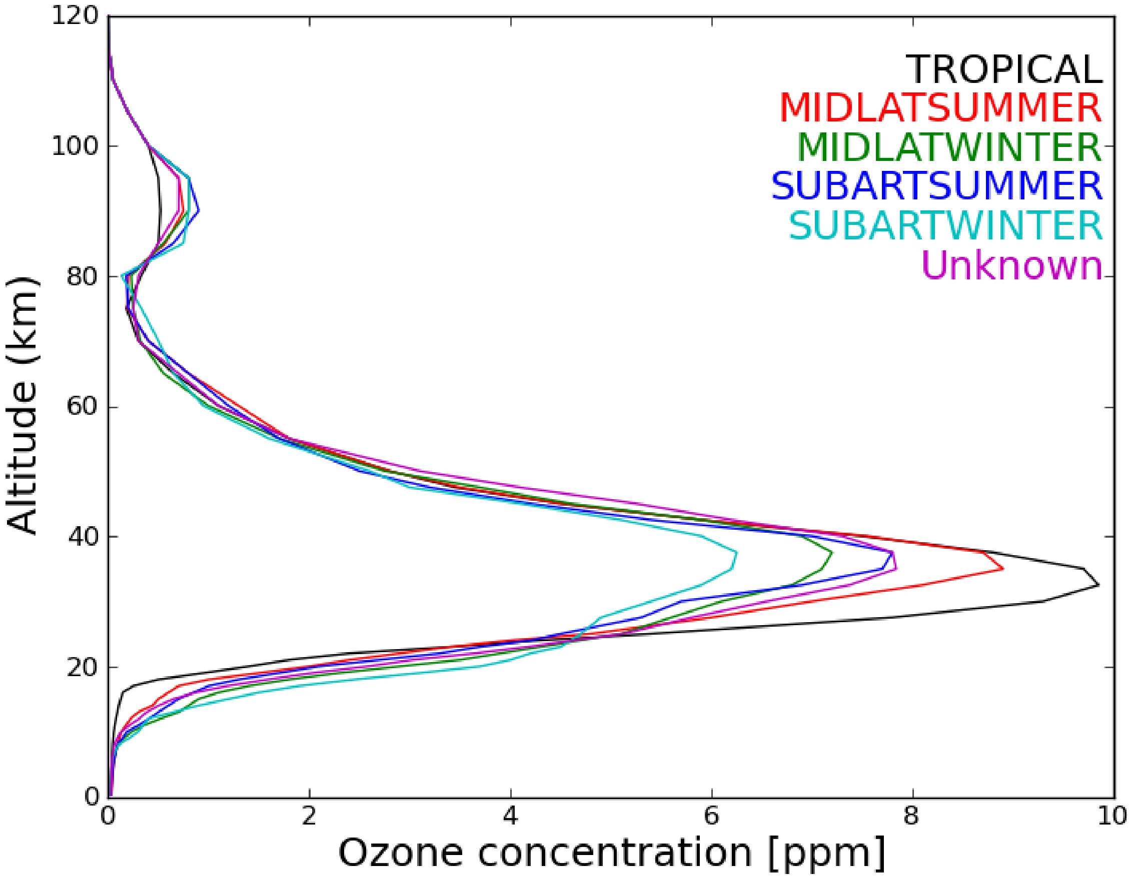

Figure 13: Plot of the ozone profiles defined in the ATM model, in ATMProfile.cpp. For the values and plotting script, see Appendix C. The units are not specified.

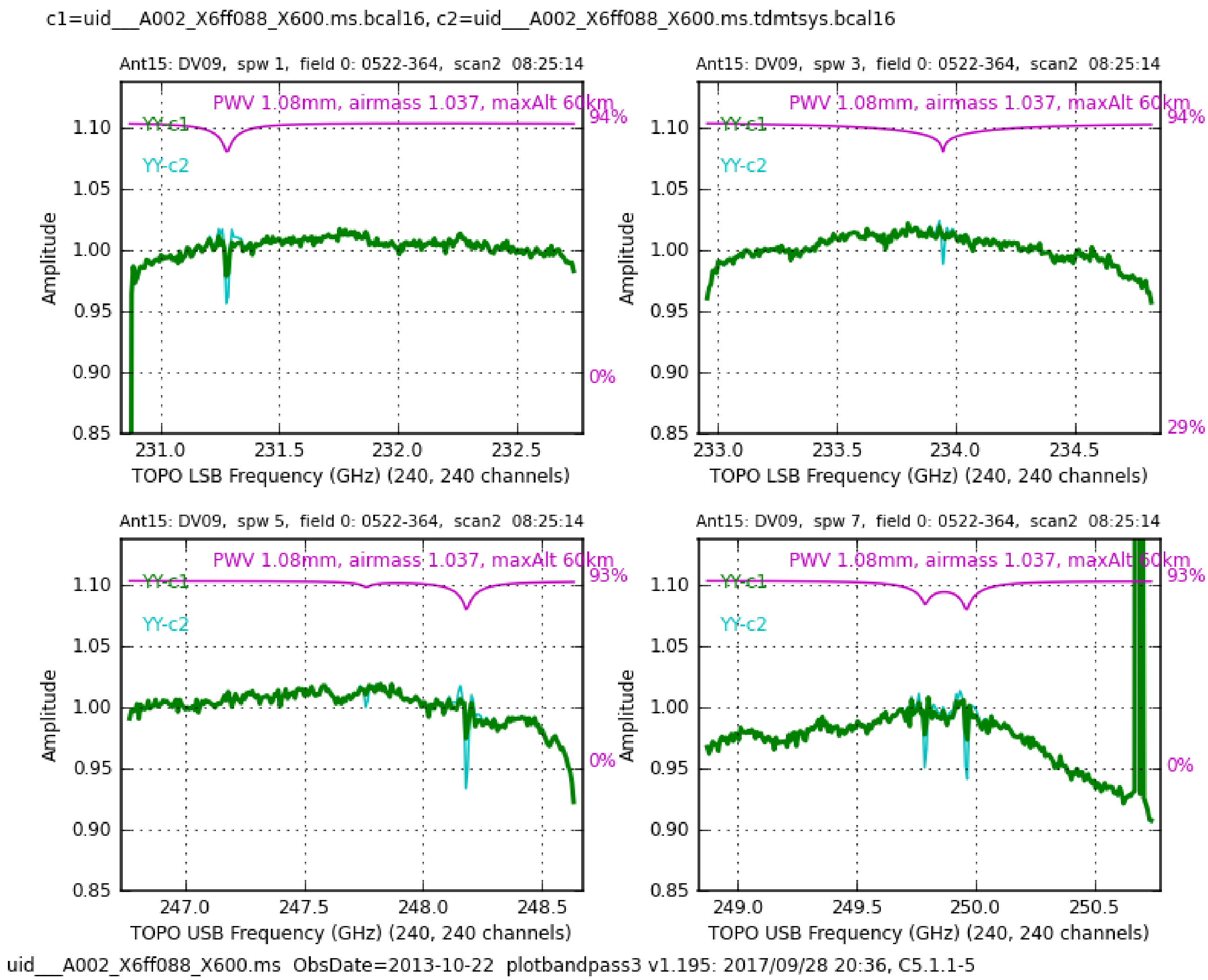

Figure 14: This figure compares the calibrated Band 6 spectra of a quasar on DV09 (YY) after applying the observed TDM Tsys (green profile) vs. after applying the simulated FDM Tsys (cyan profile). It is the same dataset as Figure 5.

7 Current status on the telescope (January 2018)

Following the correction of a memory leak in TelCal (ICT-11091 / SCCB-781) and improvements in writing to the relational database by DataCapturer (ICT-10980), it became feasible to attempt testing of the system load implications of observing with FDM Tsys. On 2017-December-01, we made some minimal tests with an array of 47 antennas on the BLC using the maximum possible number of spectral channels (4*3840 = 15360, ignoring the additional factor of 2 enabled by 90-degree Walsh switching on the DSB receivers). We ran some usual manual execution tests, and then ran a science-like Scheduling Block (SB) with a modified version of the Science Software Requirements (SSR) code to generate an AUTO_ONLY FDM spectral setup as we do for ACA observations (to avoid storing unnecessary cross-correlation data). The science-like Scheduling block (SB) execution is uid://A002/Xc75eba/X5b78, observed 2017-12-01T01:41:42-2017-12-01T02:15:10 in project code 2010.2.99003.CSV. TelCal processing time was around 70-80 seconds per scan, which is acceptable. That was with 8 threads (800% CPU utilization was seen in "top"). The processing did not disrupt operation of the ACA array running in parallel, or other calibrations in the same execution. However, we could not enable the receiving of SQLD data to do the quantization correction because we could not restart TelCal engine without interrupting the concurrent PI observing on the ACAC. We opened ICT-11578 to request the enabling of the total power data reception (dual stream) by default, to facilitate such shared test/operations system use. There were also timeouts getting the results from the TelCal_PUBLISHER to DataCapturer (such that the SysCal.xml file was not written) due to an insufficient Java heap size of the DataCapturer container that receives TelCal results (ICT-11594). Once this issue was fixed, we obtained new 30-minute, 47-antenna dataset on 2018-January-09 (uid://A002/Xc8ed16/X815a) for which the new QC option in TelCal worked successfully. Despite two minor display issues in TelCalSpy and quicklook (PRTSPR-30765 and PRTSPR-30766), the data were stored and the execution was successfully processed through the Cycle 5 pipeline (the weblog is attached to SCIREQ-1213).

In addition to testing the online performance with larger datasets, further tests of the proposed QC are needed to:

- Confirm that sufficient accuracy can be achieved with best-case digitizer vs. detector calibration (previous attempts using the current implementation in TelCal showed differences around 1-2% compared to TDM, the origin of which was not yet determined).
- Develop the digitizer vs. detector calibration procedure into something usable in production (currently just using some hacks for testing).
- Determine what error distributions can be achieved in practice.

8 Future work

8.1 Plans for development during Cycle 6

While FDM Tsys on the BLC will not yet likely be offered in Cycle 6, some other improvements related to Tsys calculations may available by then. Accounting for the effective resolution of ALMA spectra when producing ATM model spectra should be implemented in TelCal's usage of ATM. Also, a better choice of atmosphere layers in ATM should be pursued in order to better model the ozone lines (see § 6.2.3 and § 6.5.3). To enable this research, the list of required improvements to ATM are:

- Expose the setLayerO3 method to allow experimentation with different ozone profiles, and examine seasonal variation from archival data. Although changes in stratospheric ozone are slow, mesospheric ozone goes through a diurnal photochemical cycle with concentration varying by about a factor of 2. Without accounting for this effect, the center $\sim$1 MHz of the line will show excess residuals relative to the wider parts of the line. See Fig. 7.3 of Paine (2017) for an example.
- Add an input for the air pressure broadening coefficient ($\gamma_{\rm air}$) to allow values different from the approximate value of 2.5 MHz mb^{-1} that is hardcoded. But it would be better if ATM would simply use the tabulated HITRAN values for each line, which vary by about 20% between lines (Rothman et al. 2013). For example, the median of the 49 O_3 lines between 300-317 GHz is 0.0749 cm^{-1} atm^{-1} = 2.216 MHz mb^{-1}, but the values range from 2.01 to 2.55 MHz mb^{-1}.
- Implement speed improvements (if necessary) for using higher altitude maxima ($>$ 48 km). Perhaps fewer layers could be used if the "linear-in-tau" method of Clough et al.(1992) was implemented?

In addition, at somewhat lower priority, the ability to control the number of lines included in the model (ICT-9489) would allow offline testing to determine if any significant improvements can be obtained in this area.

To ease the testing of these features and improvements, we will try to create an expert mode option for the StandardInterferometry method that can be available in the Cycle-6 software to enable observations with FDM Tsys. With this capability, test executions can be efficiently performed from SBs rather than resorting to manual mode.

8.2 Cycle 7 and beyond

Once the implementation of FDM Tsys is successful, targeted for Cycle 7, we should consider an additional conceptual improvement in Tsys. While ALMA's implementation of Tsys consists of snapshots in time, most current and previous millimeter observatories (SMA, CARMA, etc.) continuously update(d) their Tsys measurement (albeit a non-spectral or very coarse spectral measurement) by using the current measurement of sky power along with the most recent measurements of power from the calibration load(s). Such a behavior could be imagined for ALMA if online TelCal was modified to compute a system temperature for every scan, possibly by using the 'ignoreloads' option in tc_atmosphere (COMP-6518), which has never been used. This option will combine the sky data from the current scan with the load spectra from the most recent AtmCal scan. Then the CASA "applycal" engine would no longer need to interpolate the Tsys correction between AtmCal measurements, resulting in more accurate flux scales for the calibrators and science targets. However, the calculation overhead for implementing such a feature online would need to be assessed. Also, it was implemented before the two sets of attenuation settings was adopted (one for AtmCal scans, and one for science scans). Thus the change in attenuation would need to be accounted for between the AtmCal subscan on the sky and the subsequent science subscans on the sky.

Further refinement in Tsys calibration that goes beyond using an improved *static* model in ATM may require the use of the recently-added ability to enter different atmospheric temperature and pressure profiles (ICT-7787), instead of the constant lapse rate (see § 6.2.4). One could imagine using *dynamic* profiles from the oxygen sounder, but this would require more reliable operation of that device.

Acknowledgements

We thank Scott Paine for valuable comments on this document.

References

- Broguiére, D., Lucas, R., Pardo, J., & Roche, J.-C., 2011, ADASS XX, ASP Conference Proceedings, Vol. 442, p.277
- Casalta, J. M., Molins, A., Bassas, M., et al. 2008, Proc. SPIE, 7018, 701838
- Clough, S. A., Iacono, M.J., & Moncet, J. 1992, "Line-by-Line Calculations of Atmospheric Fluxes and Cooling Rates: Application to Water Vapor," J. Geophys. Res. 97:15761
- Comoretto, G. 2008, ALMA Memo 583, "Algorithms and formulas for hybrid correlator data correction", `http://library.nrao.edu/public/memos/alma/main/memo583.pdf`
- Escoffier, R., Comoretto, G., Webber, J. C., Baudry, A., et al. 2007, A&A, 462, 801
- Hills, R. 2012-05-1, "Note on Spectral Response", version 2, `https://safe.nrao.edu/wiki/pub/Main/ALMAWindowFunctions/Note_on_Spectral_Response_V2.pdf`
- Lucas, R. 2013, private communication
- Lucas, R. 2012-06-08, oral presentation at JAO, "Temperature Scale Calibration", `https://safe.nrao.edu/wiki/pub/Main/ToddHunter/CalibrationSlides.pdf`
- Lucas, R. & Corder, S. 2005-07-27 memo, "Dual Load Amplitude Calibration in ALMA", `https://safe.nrao.edu/wiki/pub/Main/ToddHunter/AmpliCalMemo.pdf`
- Paine, S. 2017, The *am* atmospheric model manual, `https://zenodo.org/record/438726/files/am-manual-9.2.pdf`

- Pardo, J., Cernicharo, J., & Serabyn, E. 2001, IEEE Trans. on Antennas & Propagation, 49, 1683
- Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, "The HITRAN2012 Molecular Spectroscopic Database", JQSRT, 130, 4

Appendices

A Example code demonstrating a 3-bit quantization correction

Code written by N. Phillips and sent to D. Broguiére.

```c
/*
 * Routines related specifically to quantisation correction, including
 * using SQLD data as input.
 */

#include <math.h>

/*
 * This is the BB power used to calibrate the digitisers. It is a value from
 * calibration procedures, although assumes that the procedure really does use
 * this exact power, and that the digitiser calibration algorithm really does
 * optimise for optimum quantisation and doesn't introduce systematic errors
 * (this may be optimistic as of 2016...).
 */
#define ALMA_3BIT_OPTIMUM_BB_POWER_dBm (2.4)
/* the digitisers are optimised for 3-bit, so the optimum power for 2-bit
 * scales from that. ALMA Memo 583. states that the voltage scaling is
 * 2.01/1.70 = 1.455 dB */
#define ALMA_2BIT_OPTIMUM_BB_POWER_dBm (1.455+ALMA_3BIT_OPTIMUM_BB_POWER_dBm)


struct quantisationCorrectionLinearCoeffs {
  double a;
  double b;
};


/*
 * 2-bit / 4-level zero lag from analog signal voltage Gaussian sigma
 * in units of digitisation threshold (ALMA Memo 583 eqn. 3)
 */
static inline double R4ZeroLag(double sigma)
{
  return 9.0 - 8.0*erf(1.0 / (M_SQRT2 * sigma));
}
/* same in units of 3-bit digitiser threshold assuming the 2 MSBs are used */
static inline double R4ZeroLagFromSigma3bit(double sigma3bit)
{
  return R4ZeroLag(0.5 * sigma3bit);
}


/*
 * 3-bit / 8-level zero lag from analog signal voltage Gaussian sigma
 * in units of digitiser threshold (ALMA Memo 583 en. 4 with N=8)
 */
static inline double R8ZeroLag(double sigma)
{
  double sqrt2sigmaInv = 1.0 / (M_SQRT2 * sigma);
  return 49.0 - 8.0*erf(sqrt2sigmaInv) - 16.0*erf(2.0*sqrt2sigmaInv) - 24.0*erf(3.0*sqrt2sigmaInv);
}

/*
 * The SQLD data is stored in Watts (linear), as was defined in ICT-898.
```

```
 * This is the conversion from that to voltage distribution sigma in
 * units of the digitiser level (here assuming sigma for
 * ALMA_3BIT_OPTIMUM_BB_POWER_dBm corresponds to 1.706 times the threshold).
 */
static inline double dBmFromWatts(double p)
{
  return 10.0*log10(1000.0*p);
}
static inline double WattsFromdBm(double pdBm)
{
  return exp10((1.0/10.0) * (pdBm-30.0));
}
static inline double sigmaFromBBPower_dBm(double pdBm)
{
  return 1.706 * exp10((1.0/20.0) * (pdBm - ALMA_3BIT_OPTIMUM_BB_POWER_dBm));
}
static inline double sigmaFromBBPowerWatts(double p)
{
  return sigmaFromBBPower_dBm(dBmFromWatts(p));
}




/*
 * This is Memo 583 eqn. 21 for the auto case. It evaluates to the value 0.2698
 * stated in section 7.4 for the optimal quantisation case.
 */
static inline double autoCorrQuantisationGain3bit(double sigma)
{
  double sigma2 = sigma*sigma;
  double inv2sigma2 = 1.0 / (2.0*sigma2);
  double x = 1.0 + 2.0*(exp(-1.0*inv2sigma2) + exp(-4.0*inv2sigma2) + exp(-9.0*
    inv2sigma2));
  return M_PI * 0.5 * sigma2 / (x * x);
}
/*
 * This is Memo 583 eqn. 23.
 */
static inline double autoCorrQuantisationOffset3bit(double a, double sigma)
{
  return a*R8ZeroLag(sigma) - sigma*sigma;
}
/* This function does the lot */
static inline void autoCorrQuantisationCoeffs3bitFromPowerWatts(struct
    quantisationCorrectionLinearCoeffs *coeffs, double powerWatts)
{
  double sigma = sigmaFromBBPowerWatts(powerWatts);
  coeffs->a = autoCorrQuantisationGain3bit(sigma);
  coeffs->b = coeffs->a*R8ZeroLag(sigma) - sigma*sigma;
}
```

```
#include <stdio.h>
#include <string.h>
#include <stdlib.h>

/* includes math.h */
#include "quant.h"


int main(int argc, const char **argv)
{
```

```c
  double v, p, v_corrected;
  struct quantisationCorrectionLinearCoeffs q;
  char *endptr;

        if(argc < 3) {
                fprintf(stderr, "usage: %s autocorr_value BB_power_dBm\n", argv[0]);
                return 1;
        }

  v = strtod(argv[1], &endptr);
  if(endptr == argv[1] || *endptr != '\0') {
                fprintf(stderr, "autocorr value \"%s\" didn't parse as float\n", argv[1]);
                return 1;
        }

  p = strtod(argv[2], &endptr);
  if(endptr == argv[2] || *endptr != '\0') {
                fprintf(stderr, "BB power \"%s\" (Watts) didn't parse as float\n", argv[2]);
                return 1;
        }
  /*p = WattsFromdBm(pdBm);*/

  printf("Using autocorr value: %.6f, BB power: %.6f Watts\n", v, p);

  autoCorrQuantisationCoeffs3bitFromPowerWatts(&q, p);

  printf("Computed quantisation correction coefficients: a=%.6f, b=%.6f\n", q.a, q.b);

  v_corrected = q.a*v - q.b;

  printf("Corrected autocorr: %.6f (c.f. original %.6f)\n", v_corrected, v);

  return 0;
}
```

B Parameters to Atmcal method generateHighResTsysTable

```
    def generateHighResTsysTable(self, oldcaltable='', newcaltable='',
                                 maxRows=-1, maxScans=-1, verbose=False,
                                 showplots=False, keepLowRes=True,
                                 simulateTDMTsysScan=-1, showplotLoads=False,
                                 writeNewLowResSpectrum=False, antenna='',
                                 spw='', filterOrder=8, showpoints=True,
                                 separateFigures=False, scan='', maxAltitude=60,
                                 s=0, mode='difference',
                                 highresEB='', pwv=None):
        """
        Takes an existing TDM Tsys table for the active measurement set that
        contains FDM science data scans, and builds a high resolution version
        of the Tsys spectra using the atmospheric model and an upsampled
        version of the Trx spectra, producing one resampled Tsys spw
        per science spw.
        oldcaltable: name of existing TDM Tsys table to read
        newcaltable: name for new FDM Tsys table to write
        maxRows: if > 0, then limit the number of rows processed
        maxScans: if > 0, then limit the number of scans processed
        keepLowRes: if True, then also copy the low resolution Tsys solutions
            to the new table
        simulateTDMTsysScan: if > 0, then smooth the high-res Tsys scan data
            to this many channels before using it
        writeNewLowResSpectrum: if True, and simulateTDMTsysScan<=0, then instead
           of copying the existing low-res spectrum from the old table to the new,
           write the newly-calculated low-res spectrum;   if simulateTDMTsysScan >0,
           then write another new cal table containing the simulated TDM Tsys result
        antenna: if specified, then restrict new caltable to this list of
            antennas (python list of IDs, or comma-delimited string of names)
        scan: if specified, then restrict to these scans (python list of integers,
             single integer, or comma-delimited string)
        showplots: passed to simulateHighResTsys
        s: smoothing parameter passed to scipy.interpolate.UnivariateSpline when
           up-interpolating to high-res Tsys again
        mode: 'model', 'data', 'ratio', or 'difference' (passed to simulateHighResTsys)
        maxAltitude: of the atmosphere, in km
        highresEB: if the high-resolution data is in a different EB from the TDM Tsys
          scan, then set this parameter to the former
        pwv: if specified, then use this value to override what was in the ASDM
        Note: the ATM model spectrum is Hanning smoothed if self.hanningSmoothed[spw] == True,
        which is set automatically upon initialization of the Atmcal class.  To override
        this feature, one can set that dictionary entry to True or False before calling this function.
        """

def simulateHighResTsys(self, scan, antenna, pol, spw, fdmspw,
                        asdm=None, etaF=0.98, lo1=None,
                        dataFraction=[0.0, 1.0], parentms=None, verbose=False,
                        siteAltitude_m=5000, altscan=None, ignoreFlags=False,
                        tdmscan=None, tdmdataset=None, showplot=True,
                        showplotLoads=False, atmosphereLowRes=None,
                        atmosphereHighRes=None,
                        showpoints=False, simulateTDMTsysScan=-1,
```

```
                               plotedge=10, listpeak=False, filterOrder=8,
                               separateFigures=False, maxAltitude=60, s=0,
                               mode='difference', highresEB='',
                               newFDMFrequencyAxis=None, pwv=None):
        """
        For a specified scan, antenna, pol, spw in the current caltable,
        takes the low-resolution Trx spectrum and predicts a high resolution
        Trx and Tsys using the atmospheric model for this spw.  Trims resulting
        spectra to match the extent of the FDM spectrum.
        scan: integer or string
        antenna: integer ID, string ID or string name
        pol: 0 or 1, integer or string
        spw: Tsys spw (integer or string integer)
        fdmspw: science spw (integer or string integer), only the channel
                width is used
        simulateTDMTsysScan: if > 0, then smooth the high-res Tsys scan data
                to this many channels
        plotedge: skip this many edge channels when plotting Tsys
        separateFigures: if True, then open a new gui for each figure
        s: smoothing parameter passed to scipy.interpolate.UnivariateSpline when
          up-interpolating to high-res Tsys again
        mode: 'model', 'data', 'ratio', or 'difference'
        maxAltitude: of the atmosphere, in km
        highresEB: if the high-resolution data is in a different EB from the TDM
        Tsys scan, then set this parameter to the former
        newFDMFrequencyAxis: use this spectral grid to compute Tsys (e.g. from the
    proto-caltable)
        pwv: if specified, then use this value to override what was in the ASDM
        filterOrder: only used in simulating a TDM Tsys scan from a higher
   resolution one (e.g. ACA)
        Returns: trec, gain, tsky, freqHz, tcal, tsys, atmosphereLowRes,
   atmosphereHighRes, tsysLowRes
                  1D,   1D,   1D,    1D,    2D,  2D,   dictionary, dictionary, 2D
        Note: the ATM model spectrum is Hanning smoothed if
        self.hanningSmoothed[spw] == True, which is set automatically
        upon initialization of the Atmcal class.  To override
        this feature, one can set that dictionary entry to True or
        False before calling this function.
        """
```

C Ozone profiles in ATM

```python
import pylab as pl
import numpy as np
# The following comes from ATMProfile.cpp in CASA code/atmosphere/ATM
ozone = [ [ 2.869E-02, 3.150E-02, 3.342E-02, 3.504E-02, 3.561E-02, 3.767E-02,
            3.989E-02, 4.223E-02, 4.471E-02, 5.000E-02, 5.595E-02, 6.613E-02,
            7.815E-02, 9.289E-02, 1.050E-01, 1.256E-01, 1.444E-01, 2.500E-01,
            5.000E-01, 9.500E-01, 1.400E+00, 1.800E+00, 2.400E+00, 3.400E+00,
            4.300E+00, 5.400E+00, 7.800E+00, 9.300E+00, 9.850E+00, 9.700E+00,
            8.800E+00, 7.500E+00, 5.900E+00, 4.500E+00, 3.450E+00, 2.800E+00,
            1.800E+00, 1.100E+00, 6.500E-01, 3.000E-01, 1.800E-01, 3.300E-01,
            5.000E-01, 5.200E-01, 5.000E-01, 4.000E-01, 2.000E-01, 5.000E-02,
            5.000E-03, 5.000E-04 ],
          [ 3.017E-02, 3.337E-02, 3.694E-02, 4.222E-02, 4.821E-02, 5.512E-02,
            6.408E-02, 7.764E-02, 9.126E-02, 1.111E-01, 1.304E-01, 1.793E-01,
            2.230E-01, 3.000E-01, 4.400E-01, 5.000E-01, 6.000E-01, 7.000E-01,
            1.000E+00, 1.500E+00, 2.000E+00, 2.400E+00, 2.900E+00, 3.400E+00,
            4.000E+00, 4.800E+00, 6.000E+00, 7.000E+00, 8.100E+00, 8.900E+00,
            8.700E+00, 7.550E+00, 5.900E+00, 4.500E+00, 3.500E+00, 2.800E+00,
            1.800E+00, 1.300E+00, 8.000E-01, 4.000E-01, 1.900E-01, 2.000E-01,
            5.700E-01, 7.500E-01, 7.000E-01, 4.000E-01, 2.000E-01, 5.000E-02,
            5.000E-03, 5.000E-04 ],
          [ 2.778E-02, 2.800E-02, 2.849E-02, 3.200E-02, 3.567E-02, 4.720E-02,
            5.837E-02, 7.891E-02, 1.039E-01, 1.567E-01, 2.370E-01, 3.624E-01,
            5.232E-01, 7.036E-01, 8.000E-01, 9.000E-01, 1.100E+00, 1.400E+00,
            1.800E+00, 2.300E+00, 2.900E+00, 3.500E+00, 3.900E+00, 4.300E+00,
            4.700E+00, 5.100E+00, 5.600E+00, 6.100E+00, 6.800E+00, 7.100E+00,
            7.200E+00, 6.900E+00, 5.900E+00, 4.600E+00, 3.700E+00, 2.750E+00,
            1.700E+00, 1.000E-00, 5.500E-01, 3.200E-01, 2.500E-01, 2.300E-01,
            5.500E-01, 8.000E-01, 8.000E-01, 4.000E-01, 2.000E-01, 5.000E-02,
            5.000E-03, 5.000E-04 ],
          [ 2.412E-02, 2.940E-02, 3.379E-02, 3.887E-02, 4.478E-02, 5.328E-02,
            6.564E-02, 7.738E-02, 9.114E-02, 1.420E-01, 1.890E-01, 3.050E-01,
            4.100E-01, 5.000E-01, 6.000E-01, 7.000E-01, 8.500E-01, 1.000E+00,
            1.300E+00, 1.700E+00, 2.100E+00, 2.700E+00, 3.300E+00, 3.700E+00,
            4.200E+00, 4.500E+00, 5.300E+00, 5.700E+00, 6.900E+00, 7.700E+00,
            7.800E+00, 7.000E+00, 5.400E+00, 4.200E+00, 3.200E+00, 2.500E+00,
            1.700E+00, 1.200E+00, 8.000E-01, 4.000E-01, 2.000E-01, 1.800E-01,
            6.500E-01, 9.000E-01, 8.000E-01, 4.000E-01, 2.000E-01, 5.000E-02,
            5.000E-03, 5.000E-04 ],
          [ 1.802E-02, 2.072E-02, 2.336E-02, 2.767E-02, 3.253E-02, 3.801E-02,
            4.446E-02, 7.252E-02, 1.040E-01, 2.100E-01, 3.000E-01, 3.500E-01,
            4.000E-01, 6.500E-01, 9.000E-01, 1.200E+00, 1.500E+00, 1.900E+00,
            2.450E+00, 3.100E+00, 3.700E+00, 4.000E+00, 4.200E+00, 4.500E+00,
            4.600E+00, 4.700E+00, 4.900E+00, 5.400E+00, 5.900E+00, 6.200E+00,
            6.250E+00, 5.900E+00, 5.100E+00, 4.100E+00, 3.000E+00, 2.600E+00,
            1.600E+00, 9.500E-01, 6.500E-01, 5.000E-01, 3.300E-01, 1.300E-01,
            7.500E-01, 8.000E-01, 8.000E-01, 4.000E-01, 2.000E-01, 5.000E-02,
            5.000E-03, 5.000E-04 ],
          [ 2.660E-02, 2.931E-02, 3.237E-02, 3.318E-02, 3.387E-02, 3.768E-02,
            4.112E-02, 5.009E-02, 5.966E-02, 9.168E-02, 1.313E-01, 2.149E-01,
            3.095E-01, 3.846E-01, 5.030E-01, 6.505E-01, 8.701E-01, 1.187E+00,
            1.587E+00, 2.030E+00, 2.579E+00, 3.028E+00, 3.647E+00, 4.168E+00,
            4.627E+00, 5.118E+00, 5.803E+00, 6.553E+00, 7.373E+00, 7.837E+00,
            7.800E+00, 7.300E+00, 6.200E+00, 5.250E+00, 4.100E+00, 3.100E+00,
            1.800E+00, 1.100E+00, 7.000E-01, 3.000E-01, 2.500E-01, 3.000E-01,
            5.000E-01, 7.000E-01, 7.000E-01, 4.000E-01, 2.000E-01, 5.000E-02,
            5.000E-03, 5.000E-04 ] ]
alt = [ 0.0, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0,
        13.0, 14.0, 15.0, 16.0, 17.0, 18.0, 19.0, 20.0, 21.0, 22.0, 23.0,
```

```
        24.0, 25.0, 27.5, 30.0, 32.5, 35.0, 37.5, 40.0, 42.5, 45.0, 47.5,
        50.0, 55.0, 60.0, 65.0, 70.0, 75.0, 80.0, 85.0, 90.0, 95.0, 100.0,
        105.0, 110.0, 115.0, 120.0 ]
pl.clf()
desc = pl.subplot(111)
names = [ "TROPICAL", "MIDLATSUMMER", "MIDLATWINTER","SUBARTSUMMER", "SUBARTWINTER
    ","Unknown"]
nTypes, nLayers = np.shape(ozone)
colors=['k','r','g','b','c','m']
size=18
for i in range(nTypes):
    pl.plot(ozone[i],alt,'-',color=colors[i])
    pl.text(0.99,0.95-0.05*i,names[i],color=colors[i], ha='right', va='top',
            transform=desc.transAxes,size=size)
pl.ylabel('Altitude (km)',size=size)
pl.xlabel('Ozone concentration [ppm]',size=size)
pl.draw()
pl.savefig('ozoneProfile.png',bbox_inches='tight')
```