

Tracking ALMA System Temperature with Water Vapor Data at High Frequency

HAO HE,¹ WILLIAM R. F. DENT,² AND CHRISTINE WILSON¹

¹*McMaster University*

1280 Main St W, Hamilton, ON L8S 4L8, CAN

²*Joint ALMA Observatory*

Alonso de Córdova 3107, Vitacura, Santiago, Chile

Submitted to PASP

ABSTRACT

As the world-leading submillimeter telescope, the ALMA observatory is now putting more focus on high-frequency observations at Band 7 – 10 (frequencies from 300 – 890 GHz). However, high-frequency observations often suffer from rapid variations in atmospheric opacity that directly affects the measured system temperature T_{sys} . Current T_{sys} measurements use discrete atmosphere (ATM) calibrations done every few minutes, which usually involve 10 \sim 20 ATM calibrations within a <2 hour observation at high frequencies. In order to obtain more accurate flux measurements and reduce the number of atmospheric calibrations (Atm-cals), a new method to monitor T_{sys} continuously is proposed using existing data in the measurement set. In this work, we demonstrate the viability of using water vapor radiometer (WVR) data to track the T_{sys} continuously. We find a tight linear correlation between T_{sys} measured using the traditional method and T_{sys} extrapolated based on WVR data with scatter of 0.5% – 3%. Although the exact form of the linear relation varies among different data sets and spectral windows, we can use a small number of discrete T_{sys} measurements to fit the linear relation and use this heuristic relationship to derive T_{sys} every 10 seconds. Furthermore, we successfully reproduce the observed correlation using atmospheric transmission at microwave (ATM) modeling and demonstrate the viability of a more general method to directly derive the T_{sys} from the modeling. We apply the semi-continuous T_{sys} from heuristic fitting on a few data sets from Band 7 to Band 10 and compare the flux measured using these methods. We find the discrete and continuous T_{sys} methods give us consistent flux measurements with differences up to 5%. Furthermore, this method has significantly improved the flux accuracy of one dataset, which has large precipitable water vapor (PWV) fluctuation, from 10% to 0.7%.

Keywords: instrumentation: interferometers, atmospheric effects, techniques: interferometric, telescopes: ALMA

1. INTRODUCTION

1.1. Flux Calibration in ALMA

Calibration is the process by which the astronomer converts electronic signals from the telescope into meaningful astronomical data. Accurate calibration is crucial for the Atacama Large Millimeter/submillimeter Array (ALMA), as millimeter and submillimeter wavelength

radiation will be adversely affected by the atmosphere and the electronic signal path in a variety of ways, and the antennas will also be affected by the observing environment (Remjian et al. 2019, ALMA technical handbook, Chapter 10). One of the important calibration processes is the amplitude and flux calibration. The aim of this calibration is to convert the raw visibilities (and auto-correlations) from the correlator into brightness temperature or flux density by carefully tracking the instrumental and atmospheric variations and determining accurate conversion factors. Because of the large and rapidly varying opacity of water vapor, standard

calibration procedures are less accurate at submillimeter wavelengths. For the flux calibration, a well defined scientific goal can be elucidated and set as the requirement. In numerous meetings and discussions, the scientific community originally made clear its desire to reach 1% flux density accuracy (e.g. [Bachiller et al. 2003](#), report of the spring 2003 ASAC meeting), which means that we must be able to determine the overall flux density scale (and apply it to the visibilities and total power measurements) to 1% accuracy. In addition, the capability of achieving a dynamic range of 10000 or higher in ALMA images means that we must track the amplitude fluctuations to better than 1% ([Yun et al. 1998](#)). A later study by [Moreno & Guilloreau \(2002\)](#) showed that it is impractical to achieve 1% at submillimeter wavelengths, and so a requirement of 3% has been adopted for frequencies >300 GHz. The current achieved calibration accuracy for ALMA is 5% at the lower bands (100 GHz), 10% in mid-bands (200–400 GHz) and 20% in the higher bands (>400 GHz) ([Remjian et al. 2019](#)). This paper is part of the work being done to improve the overall flux calibration accuracy.

Currently there are two flux calibration strategies, the astronomical flux calibration and the direct instrumental amplitude calibration. The astronomical flux calibration uses an astronomical source with known flux and scales up the recorded amplitude based on that flux standard. This method requires the astronomical source to be bright and have stable flux. Currently planets are used as the primary flux standard while quasars are also used as an alternative flux standard due to their compact size and availability over the sky. However, the estimated accuracy of the planets’ fluxes is only about 10% ([Remjian et al. 2019](#), 10.4.7). Therefore, people are still searching for ideal flux calibrators with high accuracy, especially at high frequencies. In addition to the flux calibrator standard, the relative sky transmission on the calibrator and the science target, as well as the time variability of the transmissions will also affect the overall flux calibration. This is the standard method currently used by ALMA. On the other hand, for a stable system, one can directly translate the measured counts in total power into flux units using a direct instrumental amplitude calibration method. Both methods rely on accurate measurement of the sky opacity and tracking its variations during the observation. At millimeter wavelengths, the changes in atmospheric transparency will usually be very modest, under 1% over 10 minutes about 80% of the time. Since the same amount of water vapor results in much larger opacities in the submillimeter, the transparency fluctuations in the submillimeter over characteristic calibration time scales will be much

larger, typically several percent during median stability conditions and sometimes $> 10\%$.

For ALMA, both calibration methods require the precise measurement of the system temperature T_{sys} and complex gain G . T_{sys} represents the total thermal noise of the measurement. T_{sys} includes contributions from the sky, receiver, and system losses, with a large contribution coming from the sky temperature. Since ALMA is equipped with receivers of sufficiently low noise, the sky noise often dominates the total thermal noise. Therefore, it is necessary to track the changes in system temperatures caused by the fluctuations in the atmosphere. Current ALMA T_{sys} measurements use discrete atmosphere (ATM) calibrations done every few minutes with a cadence depending on the observing band. At low frequencies (< 300 GHz), we generally perform 2 or 3 T_{sys} measurements over a typical hour-long observation due to the assumed small variation in the atmosphere transmission. At high frequencies (> 300 GHz), due to the rapid opacity change in atmosphere, we generally perform 10 \sim 20 ATM calibrations per hour. So the time overheads just due to ATM calibration can become quite significant – up to 20-30% at the highest bands (9 and 10, at 600–900 GHz). Moreover the variations of T_{sys} on timescales faster than the ATM calibration interval are not tracked with this discrete ATM calibration method. Therefore, one of the major goals in high-frequency flux calibration is to track T_{sys} more closely while also reducing the time spent on discrete ATM calibrations.

1.2. T_{sys} Measurements in Flux Calibration

The system temperature (T_{sys}) is the fundamental parameter to determine the system sensitivity and the real flux of the source. T_{sys} includes various contributions, and can be written in a basic form as (adapted from [Mangum 2017](#))

$$T_{\text{sys}} = \frac{1}{\eta_f e^{-\tau_{\text{sky}}}} (T_{\text{rx}} + \eta_f T_{\text{sky}} + (1 - \eta_f) \times T_{\text{amb}}) \quad (1)$$

where

- T_{rx} is receiver temperature
- T_{sky} is sky temperature
- T_{amb} is ambient temperature where spillover is assumed to be terminated
- η_f is the forward efficiency. This is equal to the fraction of the antenna power pattern that is contained within the main beam and is currently assumed to be 0.95

- $e^{-\tau_{\text{sky}}}$ is the fractional transmission of the atmosphere, where τ_{sky} is equal to the atmospheric opacity along the target's line of sight.

Note this equation is for single sideband (SSB) and sideband separating (2SB) receivers, which are used for ALMA Band 3 – 8 observation. For this configuration, the image sideband gain is assumed negligibly small. The Band 9 and 10 receivers are using the double sideband configuration and hence T_{sys} are calculated differently (Mangum 2017, Eq. 6). T_{sky} and T_{amb} can be further expressed as

$$\begin{aligned} T_{\text{sky}} &= T_{\text{atm}}(1 - e^{-\tau_{\text{sky}}}) \\ T_{\text{amb}} &\approx T_{\text{atm}} \end{aligned} \quad (2)$$

where T_{atm} is the representative atmosphere temperature. Note that $T_{\text{amb}} \approx T_{\text{atm}}$ is a reasonable approximation when the opacity originates close to the ground (e.g. due to water vapor). Therefore, by combining equation 1 and 2, T_{sys} can be calculated as

$$\begin{aligned} T_{\text{sys}} &= \frac{1}{\eta_{\text{f}} e^{-\tau_{\text{sky}}}} [T_{\text{rx}} + T_{\text{amb}}(1 - \eta_{\text{eff}} e^{-\tau_{\text{sky}}})] \\ &\approx \frac{1}{e^{-\tau_{\text{sky}}}} (T_{\text{rx}} + T_{\text{sky}}) \quad (\text{where } \eta_{\text{eff}} \sim 1) \end{aligned} \quad (3)$$

This equation suggests that the key parameters to measure T_{sys} are T_{rx} and T_{sky} (as τ_{sky} can be derived from T_{sky} using Eq. 2).

In ALMA flux calibration, the intensity of the observed source is directly proportional to T_{sys} by the following equation (e.g. Brogan 2018)

$$S_{\text{final}} \sim S_0 \times \sqrt{T_{\text{sys}}(i)T_{\text{sys}}(j)} \times \Gamma \quad (4)$$

where S_0 and S_{final} are the fluxes measured before and after the flux calibration. Γ is the antenna efficiency factor to convert K to Jy and i and j represent the two antennas forming the baseline. Note that ALMA uses an antenna-based calibration method to simplify the calibration process. Nearly all of the changes to the visibility function (e.g. atmosphere, system noise, amplitude changes, delay changes) can be decomposed into the two complex antenna-based gain factors associated with any baseline. This approach reduces the number of gain correction terms for an N -element array from $N(N-1)/2$ baselines to N antennas. In this case, T_{sys} is associated with each antenna and S_{final} is associated with each baseline.

T_{sys} also determines the achieved rms noise of the observation (e.g. Condon & Ransom 2016, a modified form of radiometer equation)

$$\text{rms} \approx \frac{cT_{\text{sys}}}{\sqrt{\Delta\nu t_{\text{int}}}} \quad (5)$$

where $\Delta\nu$ is the frequency bandwidth, t_{int} is the integration time of the observation and c includes the quantization and correlator efficiencies, and is typically 0.8-0.9 for ALMA. Therefore, higher T_{sys} means the data has a larger noise within one observation. In addition, for ALMA the weighting function used to combine visibility data is inversely proportional to T_{sys} as

$$\text{Weight} \propto \frac{1}{T_{\text{sys}}(i)T_{\text{sys}}(j)} \quad (6)$$

1.3. Traditional Method to Measure T_{sys}

As noted above, T_{sys} is highly dependent on the sky opacity (eq. 2). ALMA antennas use a two-load system for T_{sys} measurement in Band 3 and higher, which is different from the one-load system used for other radio telescopes. The two-load system in theory can achieve a T_{sys} measurement accuracy of 1%, which is significantly better than that of a one-load system ("chopper wheel") of 5% (Yun et al. 1998). For ALMA, T_{sys} is obtained from an atmospheric calibration (ATM-cal) scan where a hot load, ambient load and sky are consecutively placed in front of the feed using an Amplitude Calibration Device (ACD). Typically this process takes 30-40 seconds, including antenna slew time and overheads. At frequencies below about 400 GHz, where the system temperatures are stable, an ATM-cal scan is made every 10 to 20 minutes. However, at higher frequencies, and where the opacity is large and more variable, every scan on the astronomical target will have an associated ATM-cal measurement (Remjian et al. 2019), implying a cadence of ATM calibration as fast as once every 2-3 minutes. From the two-load system, we can also measure T_{rx} . In this case, the measured power can be expressed as (Mangum 2002)

$$\begin{aligned} P_{\text{hot}} &= K (T_{\text{rx}} + T_{\text{hot}}) \\ P_{\text{amb}} &= K (T_{\text{rx}} + T_{\text{amb}}) \\ P_{\text{sky}} &= K (T_{\text{rx}} + T_{\text{sky}}) \end{aligned} \quad (7)$$

where K is the gain to convert the temperature to the measured power. T_{hot} and T_{amb} are generally about 300 K and 270 K. Based on the equation above, we can express T_{rx} and T_{sky} as

$$\begin{aligned} T_{\text{rx}} &= \frac{T_{\text{hot}}P_{\text{amb}} - T_{\text{amb}}P_{\text{hot}}}{P_{\text{hot}} - P_{\text{amb}}} \\ &= \frac{T_{\text{hot}} - Y_1 T_{\text{amb}}}{Y_1 - 1} \\ T_{\text{sky}} &= \frac{P_{\text{sky}}T_{\text{amb}} - (P_{\text{amb}} - P_{\text{sky}})T_{\text{rx}}}{P_{\text{amb}}} \\ &= Y_2 T_{\text{amb}} - (1 - Y_2)T_{\text{rx}} \end{aligned} \quad (8)$$

where $Y_1 \equiv P_{\text{hot}}/P_{\text{amb}}$ and $Y_2 \equiv P_{\text{sky}}/P_{\text{amb}}$. Unlike the atmosphere, T_{rx} is relatively constant throughout the

observation. Measurements performed by ALMA show fluctuations of T_{rx} are generally smaller than 1% during normal sidereal tracking. With the measurement of T_{sky} , we can further derive the optical depth based on Eq. 2 and calculate the T_{sys} based on Eq. 3. In summary, the expressions for the key quantities to measure T_{sys} are

$$\begin{aligned} T_{\text{rx}} &= \frac{T_{\text{hot}} - Y_1 T_{\text{amb}}}{Y_1 - 1} \approx \text{const} \\ T_{\text{sky}} &= Y_2 T_{\text{amb}} - (1 - Y_2) T_{\text{rx}} \\ T_{\text{sys}} &\approx \frac{1}{e^{-\tau_{\text{sky}}}} (T_{\text{rx}} + T_{\text{sky}}) \end{aligned} \quad (9)$$

Therefore, during each ATM cal, we point the array to hot load, ambient load and sky to measure T_{rx} and T_{sky} and then calculate the T_{sys} at that time.

1.4. Candidate Data to Track the Continuous T_{sys}

As mentioned above, the current method takes extra time to obtain a spot measurement of T_{sys} every few minutes. If we want to continuously track T_{sys} , in theory there are 3 types of measurement data available from ALMA to achieve this goal: Water Vapor Radiometer (WVR) data, auto-correlation (AC) data or square law detector (SQLD) data. We will describe where these data arise, and the theory behind each method below. The advantages and disadvantages of each method are summarized in Table. 1.

The WVR data are used by ALMA to track the optical depth of the water vapor along the line of sight to each antenna, and hence are used to correct for the resulting effective pathlength and delay errors. The WVRs do Dicke switching and have internal calibrated loads, so the output from each WVR is the calibrated sky temperature (T_{WVR}) at 4 frequencies (184.19, 185.25, 186.485 and 188.51 GHz respectively; Hills 2004) around the 183 GHz water line taken every 1.152 seconds (Remjian et al. 2019, Section A.6). By comparing T_{WVR} with T_{amb} , we can calculate the precipitable water vapour (PWV), which is proportional to the atmospheric opacity caused by the water absorption. Since T_{sky} at our observing frequency (Eq. 9) and T_{WVR} are tracking sky temperatures at different frequencies, we would expect

$$\begin{aligned} \tau_{\text{sky}} &= C \times \tau_{183\text{GHz}} + \tau_{\text{dry}} \\ T_{\text{sky}} &= C \times T_{\text{WVR}} + T_{\text{sky,dry}} \quad (\tau_{\text{sky}}, \tau_{183\text{GHz}} \ll 1) \end{aligned} \quad (10)$$

where τ_{sky} is the sky opacity at the observed frequencies, $\tau_{183\text{GHz}}$ is the optical depth for the water vapor line, τ_{dry} is the optical depth for the dry component, $T_{\text{sky,dry}}$ is the sky temperature for the dry component and C is the constant. The overall sky opacity includes contributions the wet component (H_2O lines which are relatively wide due to pressure broadening), and from the

dry component (mostly due to relatively narrow lines of O_3 , but also including a continuum component as well as other molecules). If optical depth is small enough, we would expect the observed temperature is proportional to the optical depth and hence the proportion relation between the two optical depths holds also for the two measured temperatures. Since the major change in T_{sys} is caused by the variation in T_{sky} , we would expect that T_{WVR} is tracking T_{sys} . The major advantage of using the WVR data to trace T_{sys} is that the radiometer is constantly monitoring the sky and internally calibrating itself. Therefore, we can extrapolate T_{sys} throughout the entire observation based on the WVR data. Furthermore, since T_{WVR} is internally calibrating and tracking the sky variation, it does not suffer from the internal electronic drift or small changes in system gain, which can affect the measured values of an uncalibrated signal (see Table 1)

Alternatively, we would expect T_{sys} is tightly correlated with the total power signal received by each antenna. To be more precise, the total power signal should be directly proportional to $(T_{\text{rx}} + T_{\text{sky}})$, which can be used to calculate T_{sys} given the optical depth τ_{sky} using Eq. 9 (see detailed discussion in Section 4). The total power signal received by each antenna is measured by a square-law detector built into the ALMA signal path, whose data is also recorded in the datasets. Additionally, the autocorrelation data recorded in the measurement set should also give us the total signal received by each antenna. We would expect

$$T_{\text{rx}} + T_{\text{sky}} \propto P_{\text{AC}} \propto P_{\text{SQLD}} \quad (11)$$

where P_{AC} and P_{SQLD} are the power of auto-correlation data and SQLD data read from the measurement set, respectively. If τ_{sky} is small, we would expect direct proportionality between T_{sys} and the total power received which could help us derive continuous T_{sys} . In addition, both AC and SQLD data cover the same frequency rangels as the actual observed science data so we do not need to assume atmosphere variation has the same effect on data at different wavelengths (the ratio proportionality in Eq. 10).

In Fig. 1, we plot the correlation between AC and WVR and SQLD data. We can see the AC and SQLD data follow a tighter linear correlation. These two types of data are expected to be equivalent and thus should follow a proportional correlation. We see an offset from direct proportionality between AC and SQLD data in this observation as no linearity correction for the effect of the 3 bit samplers is applied in this correlator mode, and there can be residual DC offsets in the SQLD data. On the other hand, we can see that the WVR and AC

Table 1. Summary of different methods to track T_{sys}

	WVR ¹	AC & SQLD ²
Advantage	1. The data is continuously calibrated to measure T_{WVR} ^a 2. Does not suffer from internal electronic gain drift ^b	1. Directly proportional to T_{sys} when τ_{sky} is small 2. Has the same frequency coverage as the science target
Disadvantage	1. Has different frequency coverage as the science target 2. Directly tracks T_{sky} not T_{sys}	1. The data is not calibrated. 2. Suffers from the electronic gain drift ^b or gain variations 3. For AC, no linearity correction in FDM mode.

Columns: 1. Water vapor radiometer data. 2. Auto-correlation and square law detector data.

References: a. [Hills et al. \(2001\)](#). b. [Payne et al. \(2001\)](#).

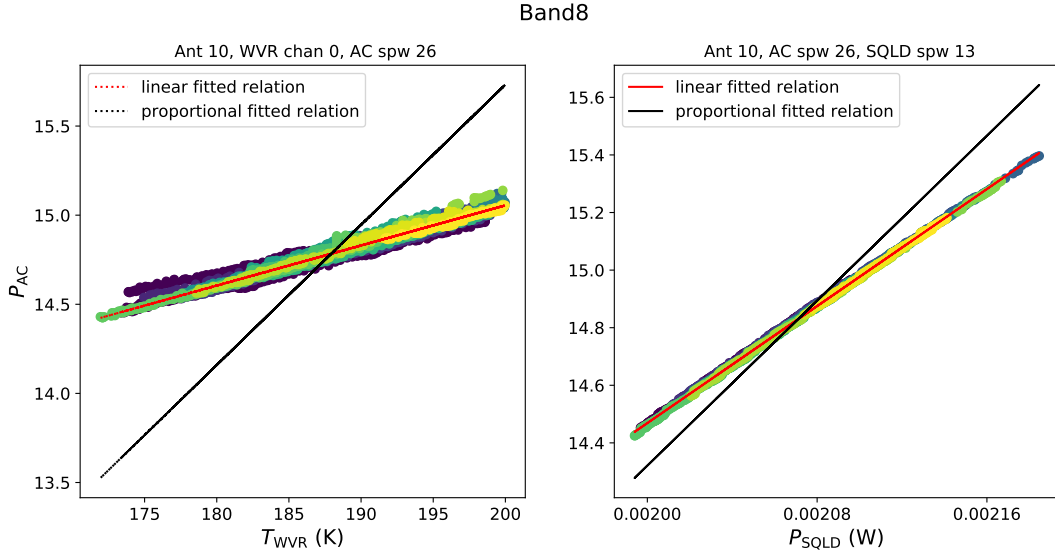


Figure 1. The correlation between water vapor radiometer (WVR) data, auto-correlation (AC) data and square law detector (SQLD) data for antenna 10 for dataset Band8 data. The scatter plots are color coded by the scan number. (Left) AC data versus WVR data. The red line is the best linear fit and black line is the proportional fit through 0. We have excluded WVR data greater than 200 as those are from samples taken on the hot load and ambient load. (Right) AC data versus SQLD data.

data do not follow the same proportional relation. This can be caused by various reasons summarized in Table 1. In particular, the distribution in the AC data at similar WVR levels on the left panel is indicative of slightly different system gains in different scans during the observation, or the differing wet and dry opacity contributions at different airmasses. In this case, we need to compare the two types of data to explore which one is better in tracking T_{sys} .

1.5. Outline of This Paper

In the following sections, we will explore how well different data track T_{sys} measurements. In Section 2, we explore the viability of using T_{WVR} to track T_{sys} . In Section 3, we use the Atmospheric Transmission at Microwave (ATM) modeling to test the theory behind the tight T_{sys} vs T_{WVR} correlation. In Section 4, we explore the viability to use AC or SQLD data to track T_{sys} .

What we find is that those two types of data do not work well in tracking T_{sys} . In Section 5, we describe our new calibration method to use alternative T_{sys} derived from T_{WVR} and how it compares to the original discrete calibration method.

For our analysis, we use measurements sets from several projects in Bands 7, 8, 9 and 10. We also include two projects with multiple measurements sets from Band 7 and 9. The summary of the data we use is given in Table 2.

2. WVR DATA TO TRACK T_{sys}

In this section, we examine how well T_{WVR} tracks T_{sys} and explore how the correlation is affected by various parameter choices. We mainly use the dataset Band8 (uid://A002/Xdb7ab7/X1880b) for illustration purposes. Examples from additional datasets are given in Appendix A.

Table 2. Summary of Data

Dataset Label	Project	Band	Target	Data uid	PWV (mm)	Elev. (deg)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Band10	2015.1.00271.S	10	Arp 220	uid://A002/Xbe0d4d/X12f5	0.28	43
Band9a	2016.1.00744.S	9	IRAS16293-B	uid://A002/Xbf792a/X14cc	0.37	63 – 76
Band8	2018.1.01778.S	8	SPT0311-58	uid://A002/Xdb7ab7/X1880b	0.85	53 – 55
Band7a	E2E8.1.00003.S	7	HT-Lup	uid://A002/Xec4ed2/X912	0.59	52 – 66
Band7b1	2018.1.01210.S	7	AS205A	uid://A002/Xda1250/X2387	0.69	50 – 64
Band7b2				uid://A002/Xda1250/X32df	0.53	80 – 85
Band7b3				uid://A002/Xda845c/X35d1	0.5	55 – 70
Band7b4				uid://A002/Xda1250/X3e39	0.51	60 – 75
Band7b5				uid://A002/Xda1250/X4db3	0.49	35 – 50
Band7b6				uid://A002/Xd99ff3/X15d7b	0.42	65 – 80
Band7b7				uid://A002/Xd99ff3/X1702c	0.51	63 – 77
Band7b8				uid://A002/Xd99ff3/X17da2	0.53	40 – 55
Band9b1	2019.1.00013.S	9	Circinus	uid://A002/Xed9025/X769c	0.43	34 – 42
Band9b2				uid://A002/Xed8123/X7b1	0.37	37 – 43
Band9b3				uid://A002/Xed4607/X1208a	0.34	31 – 39

Columns: (1) Label for each dataset (2) ALMA project code (3) Observed Band. (4) The target name to be observed (5) The uid for each measurement set (6) The precipitable water vapor (PWV). (7) Elevation range of the science target

2.1. T_{sys} versus T_{WVR}

To check whether T_{sys} is tracked by the WVR data, we first need to match the WVR data taken at the same time as the T_{sys} measurements. We then average the WVR values that are within 10 s around the time when T_{sys} is measured and compare the averaged T_{WVR} with its corresponding T_{sys} . 10 s is a typical time for one Atm-cal scan and hence is the shortest timescale we expect T_{sys} to change. We also note that T_{sys} recorded in the measurement set is a spectrum with two polarizations. In our analyses to compare T_{sys} with T_{WVR} , we average T_{sys} from both polarizations and also along the spectral axis within one spectral window (spw).

We first plot T_{sys} versus T_{WVR} from all antennas for each data set. One example of T_{sys} versus T_{WVR} is shown in the left panel of Fig. 2. For this case, we select T_{WVR} from WVR channel 1. We will discuss in Section 2.2 how the selection of different WVR channels affects the relation between T_{WVR} and T_{sys} . As we can see, there is a significant correlation between T_{sys} and T_{WVR} for each spw. However, the scatter is large along the direction perpendicular to the trend, as expected due to the differences in the receiver (T_{rx} and sideband gains) and WVR between antennas. Furthermore, since bandpass, phase-cal and science observations are observing targets at different elevations, it is possible that the scatter is also caused by data from different types of observations. Therefore, to see if the WVR tracks the time variation of T_{sys} , we normalize T_{sys} and T_{WVR} by the first measurement for each observing target (band-

pass, phase-cal and science) of each antenna as

$$\begin{aligned}\hat{T}_{\text{sys,obs}}(t) &= \frac{T_{\text{sys,obs}}(t)}{T_{\text{sys,obs}}(1^{\text{st}})} \\ \hat{T}_{\text{WVR,obs}}(t) &= \frac{T_{\text{WVR,obs}}(t)}{T_{\text{WVR,obs}}(1^{\text{st}})}\end{aligned}\quad (12)$$

where \hat{T}_{sys} and \hat{T}_{WVR} is the normalized values of T_{sys} and T_{WVR} , the subscript 'obs' is the generalized term for each type of observing target (bandpass, phase and science) and i^{th} in the bracket means the value when i^{th} T_{sys} for each observing target is measured.

The correlation between the \hat{T}_{sys} and \hat{T}_{WVR} is shown in the right panel of Fig. 2. We can see that these two variables have a tight linear correlation, with scatter less than 1%. This tight linear correlation is also seen in other data sets, as illustrated in Appendix A. This indicates that T_{WVR} can be used to track the T_{sys} if the slope and intercept can be determined for each spw or frequency. As described by eq. 10, the relation is expected to be frequency dependent, and it further differs from 1-to-1 due to the other contributions to T_{sys} apart from T_{sky} (e.g. eq. 9). In Section 3, we explore the relationship using atmospheric opacity model, but here we take a heuristic approach to determine the linear relationship from the data itself.

We can then use the fitted linear relation to extrapolate the continuous T_{sys} based on the first T_{sys} value for each observing target and the stream of T_{WVR} values.

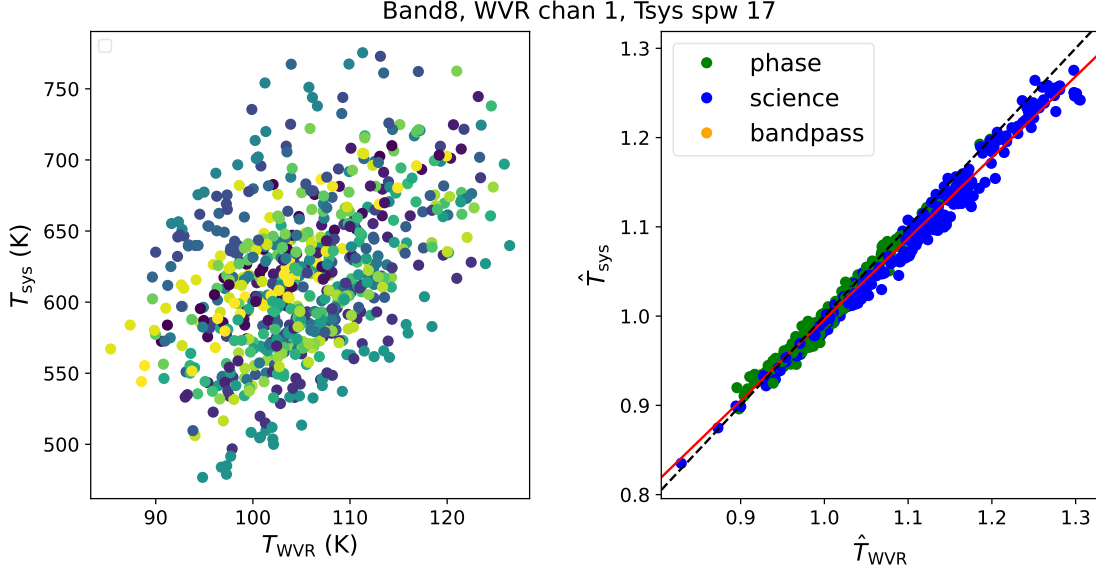


Figure 2. (Left) T_{sys} vs T_{WVR} for dataset Band8 color coded by different antennas. (Right) T_{sys} and T_{WVR} normalized to the value of first scan of each target (bandpass, phase, science) for each antenna. We can see the normalized \hat{T}_{sys} and \hat{T}_{WVR} follows a tight linear correlation.

The exact equation can be expressed as

$$\begin{aligned} T_{\text{sys}}(t) &= T_{\text{sys}}(1^{\text{st}}) \cdot \hat{T}_{\text{sys,obs}}(t) \\ &= T_{\text{sys}}(1^{\text{st}}) \cdot [m \hat{T}_{\text{WVR,obs}}(t) + b] \end{aligned} \quad (13)$$

where $T_{\text{sys}}(1^{\text{st}})$ and $T_{\text{WVR}}(1^{\text{st}})$ are T_{sys} and T_{WVR} values used to normalize each antenna and each type of observing targets. m and b are slope and intercept of the fitted linear function. For making the extrapolation, we also sample and average the WVR data every 10 seconds to be consistent with our fitting parameter choice. An example of extrapolated T_{sys} for one antenna is shown in Fig. 3. As we can see, the extrapolated continuous T_{sys} is consistent with the original discrete T_{sys} values for all 4 spectral windows. The trend is also quite continuous with no obvious glitches due to the measurement noise. The trends for all 4 spectral windows are similar.

Examples of fitting and T_{sys} extrapolation for other datasets are shown in Appendix A. We can see for all datasets, \hat{T}_{sys} and \hat{T}_{WVR} have a tight linear correlation but with different slopes and intercepts. The extrapolation also works well for most of the data sets.

2.2. Extrapolate T_{sys} with Other WVR Channels and PWV

The ALMA WVRs have 4 filter channels at frequencies (184.19, 185.25, 186.485 and 188.51 GHz respectively) close to the 183GHz H_2O line (Hills et al. 2001). These channels have different sensitivities to the line-of-sight water content (PWV) and hence T_{sys} , depending on the actual PWV at the observing time. In this section, we explore how the different parameter choices will

Table 3. RMS of the T_{sys} residual from the fitting

T_{sys} spws WVR chans	T_{sys} spws			
	17	19	21	23
0	1%	1.1%	0.9%	0.9%
1	0.9%	1.1%	0.8%	0.8%
2	1%	1.1%	0.9%	0.9%
3	1.4%	1.7%	1.3%	1.3%
PWV _{los}	0.8%	0.9%	0.7%	0.7%

The root mean square (RMS) of the residual from the linear fitting using T_{WVR} from different WVR channels and calculated PWV values along the line of sight (without elevation correction) for dataset Band8.

affect the correlation between T_{sys} and T_{WVR} . In Section 2.1, we selected T_{WVR} in channel 1 to track T_{sys} . Here we compare how well T_{WVR} from different WVR channels track the T_{sys} from different spectral windows by calculating the scatter of the data residual from the fitting (Table 3). As we can see, normalized \hat{T}_{WVR} from different WVR channels generally have the similar tight correlation with normalized \hat{T}_{sys} with scatter of $\sim 1\%$. For dataset Band8, \hat{T}_{WVR} from WVR channel 1 gives us the tightest linear correlation. We will discuss the reason later in this section.

For single dish telescopes such as APEX and JCMT, WVR data has been used to continuously track optical depth at the observed frequencies (e.g. Dempsey et al. 2013). The method converts T_{WVR} values from multiple WVR channels into a single PWV value and use it to track the optical depth at any given time, which reduces

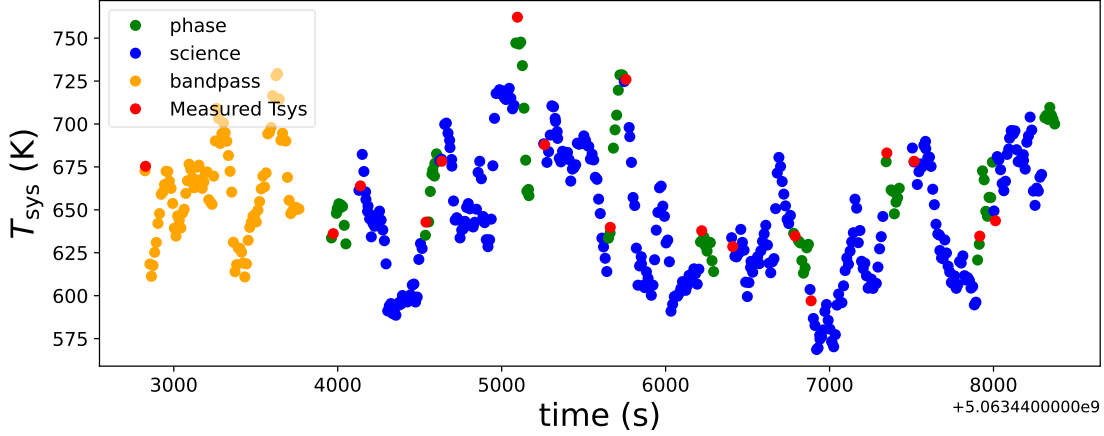


Figure 3. The extrapolated T_{sys} versus the original T_{sys} for antenna 10 of the dataset Band8 at spw 17. The orange, green and blue points are extrapolated continuous T_{sys} for each observing target based on Eq. 13 while the red points are original T_{sys} measurements.

the effect of measurement noise from a single channel. Given what we are doing is similar, as T_{sys} is mostly affected by the change in atmospheric optical depth, we can try to use PWV along the line of sight (PWV_{los}) instead of T_{WVR} from a specific channel to track T_{sys} . We calculate the PWV_{los} by fitting the Lorentz profile for the water line given the T_{WVR} from multiple WVR channels. We then normalize the PWV_{los} values the same way as we do for T_{WVR} (Eq. 12). The scatter of fit residual using PWV_{los} is also listed in Table 3. As we can see, PWV_{los} actually gives a tighter correlation, which is consistent with our expectation since it is less affected by the measurement noise from the single channel.

In our later analysis to apply the continuous T_{sys} in data calibration, we select WVR channels to maximize the following weighting function

$$w = \bar{T}_{\text{WVR}}(\bar{T}_{\text{WVR}} - 275) \quad (14)$$

where \bar{T}_{WVR} is the averaged T_{WVR} in one channel and 275 K is the approximate atmosphere temperature. The principle for this selection criterion is to make T_{WVR} neither too small to be robust against noise (in the case of low opacity) nor too large to be saturated (in the case of high opacity). Based on this criterion, we generally select WVR channel 0 or 1 for datasets in our analysis.

2.3. Fewer Atm-cal scans to Fit the Relation

As discussed in Section 2.1, the T_{sys} in different spectral windows have different linear relations with T_{WVR} . As mentioned, a goal is to reduce the number of T_{sys} measurement scans within each observation to increase observing efficiency. However, this leads to less data to fit the relations of T_{sys} to T_{WVR} or PWV. In Section 2.5 we investigate determining the relations from atmosphere opacity models, but here we test the reliability of

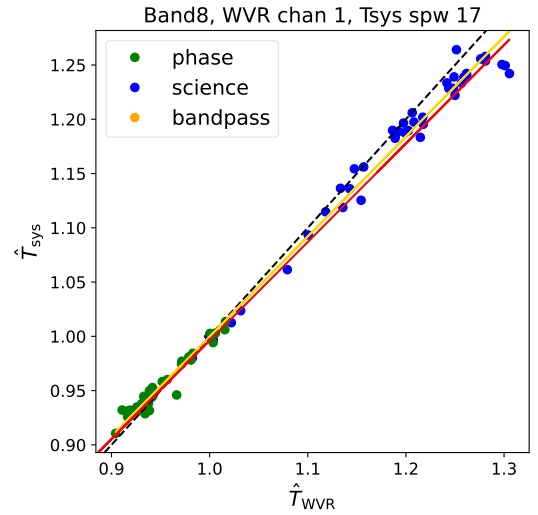


Figure 4. The correlation between \hat{T}_{sys} and \hat{T}_{WVR} for the 4 Atm-cal scans we selected to fit the linear relation. The red and gold lines are the fitting relation with all Atm-cal scans or just 4 Atm-cal scans. The black dashed line indicates the 1-to-1 relation. We can see the fitting using all Atm-cal scans are almost the same as just using 4 scans.

fitting the relations to a small number of T_{sys} measurement scans. Since we need to calculate the normalized T_{sys} , we need at least two Atm-cal scans to fit the linear correlation. To make the fitting more robust, we use 4 Atm-cal scans for the fitting with 2 from phase target and 2 from the science target. The 4 Atm-cal scans give us 2 independent \hat{T}_{sys} values if we normalize the T_{sys} from phase and science target independently. For the Atm-cal scan selection, we select 2 Atm-cal scans at the start and 2 Atm-cal scans in the middle. One example of the fits using 4 Atm-cal scans is shown in Fig. 4. As we can see, the fits based on data from all Atm-cal scans have almost no difference from the fits based on only 4

Atm-cal scans. The scatter of all the data points around the new relation has the almost the same scatter of 1%.

Fits using 4 Atm-cal scans for other data sets are also shown in Appendix A. We can see that the fits do not change much for almost all data sets except for that in dataset Band9b. We will discuss dataset Band9b1 in Section 2.5. In Fig. 5, we plot the relative scatter around the fit versus the maximal difference divided by mean value of the T_{sys} for every T_{sys} spw of all the data sets. As we can see, the scatter using both fitting methods is generally below 3%. Besides, fitting with just 4 scans only slightly increases the scatter compared with fitting with all scans. From this quantitative comparison, we can see it is viable to reduce the number of discrete T_{sys} measurements when using T_{WVR} to track T_{sys} .

2.4. Normalize Only to the Science target

For some ALMA data, T_{sys} for the phase-cal target is not measured. Instead, the calibration uses the nearest science T_{sys} values as the phase-cal T_{sys} . If we can normalize all T_{sys} values to the first T_{sys} of the science target instead of the first T_{sys} of each type of observing target itself, we can further reduce the number of T_{sys} measurements and thus no longer need to measure T_{sys} for phase-cal with our new method.

In this case, the normalized T_{sys} is calculated as

$$\begin{aligned}\hat{T}_{\text{sys,obs}}(t) &= \frac{T_{\text{sys,obs}}(t)}{T_{\text{sys,sci}}(1^{\text{st}})} \\ \hat{T}_{\text{WVR,obs}}(t) &= \frac{T_{\text{WVR,obs}}(t)}{T_{\text{WVR,sci}}(1^{\text{st}})}\end{aligned}\quad (15)$$

where $\hat{T}_{\text{sys,obs}}(t)$ is the normalized T_{sys} averaged along the spectral axis and $\hat{T}_{\text{WVR,obs}}(t)$ is the normalized T_{WVR} . Examples of \hat{T}_{sys} versus \hat{T}_{WVR} using the new normalization method are shown in Fig. 9 and 10. We can see that phase-cal and science target generally follows the same linear trend, which is consistent with our expectation since phase-cal and science targets are close in elevation. In contrast, we see offsets between bandpass targets and phase-cal/science trend, specifically for dataset Band8. We also expect this to happen since bandpass target usually has significant different elevations from the phase-cal/science targets. We will further discuss the cause of the offsets with the help of atmospheric modeling in Section 3.2. In general, these tests show that we can further reduce the phase-cal and bandpass T_{sys} measurements as we can derive it from T_{sys} measurements for science target.

Note that for early ALMA cycles, the T_{sys} measurements are purely done for the phase-cal target. The T_{sys} for the science target is then assumed to be the same as the T_{sys} for the closest phase-cal scan. As we can see

from this section, even though the phase-cal and science targets have different elevations, they generally follow the same T_{sys} vs T_{WVR} linear relation. Therefore, we can better extrapolate the science T_{sys} from the phase-cal T_{sys} using on the fitted linear relation.

2.5. Test with significant opacity and large T_{sys} variation

We would expect the linear relation between T_{sys} and T_{WVR} holds when τ_{sky} is small. In this case, we would have

$$T_{\text{sys}} \approx T_{\text{rx}} + T_{\text{sky}} = T_{\text{rx}} + C \times T_{\text{WVR}} + T_{\text{sky,dry}} \quad (16)$$

where C is a constant. However, at higher frequencies such as Band 9 and 10, τ_{sky} are quite high and we can no longer ignore the opacity term in T_{sys} equation (eq. 3. In this case, the increase in T_{sys} is dominated by the increase in τ_{sky} in the exponential form. Combining Eq. 2 and 3, T_{sys} at high opacity limit can be expressed as

$$T_{\text{sys}} = \frac{T_{\text{sky}}}{T_{\text{amb}} - T_{\text{sky}}} (T_{\text{sky}} + T_{\text{rx}}) \quad (17)$$

We can see when T_{sky} is large, the increase in T_{sys} is dominated by $T_{\text{sky}}^2/(T_{\text{amb}} - T_{\text{sky}})$ term, which has increasing slopes as T_{sys} increases.

We test if the T_{sys} vs T_{WVR} linear relation still holds on dataset Band9b, which has large τ_{sky} and T_{sys} range ($\sim 50\%$). In Fig. 6, we show T_{sys} fitting and extrapolation using all Atm-cal scans or just 4 Atm-cal scans for one measurement set in this project. We can clearly see there is a difference in the fitting functions derived from all Atm-cal scans or just 4 Atm-cal scans. It seems the slope becomes steeper due to data points with higher T_{sys} values, which are not included if we use just 4 scans. This is also reflected in the extrapolation plot at right side of Fig. 6, as the predicted T_{sys} is lower than the measured T_{sys} for higher T_{sys} values.

This is consistent with our expectation that the slope of T_{sys} vs T_{WVR} relation is increasing. We would expect the curving-up feature also happens to other datasets but we simply do not have large enough T_{sys} ranges.

3. ATMOSPHERIC TRANSMISSION AT MICROWAVE (ATM) MODELING

In the previous sections, we fit T_{sys} vs T_{WVR} heuristically and use the fitted relation to extrapolate the T_{sys} continuously. We find in most cases the correlation between T_{sys} and T_{WVR} is linear. However, the exact slopes and intercepts of the correlations vary across different frequencies. Although we can use less than 4 Atm-cal scans to fit the relation for each spw in each dataset, there might be cases when our selected Atm-cal scans

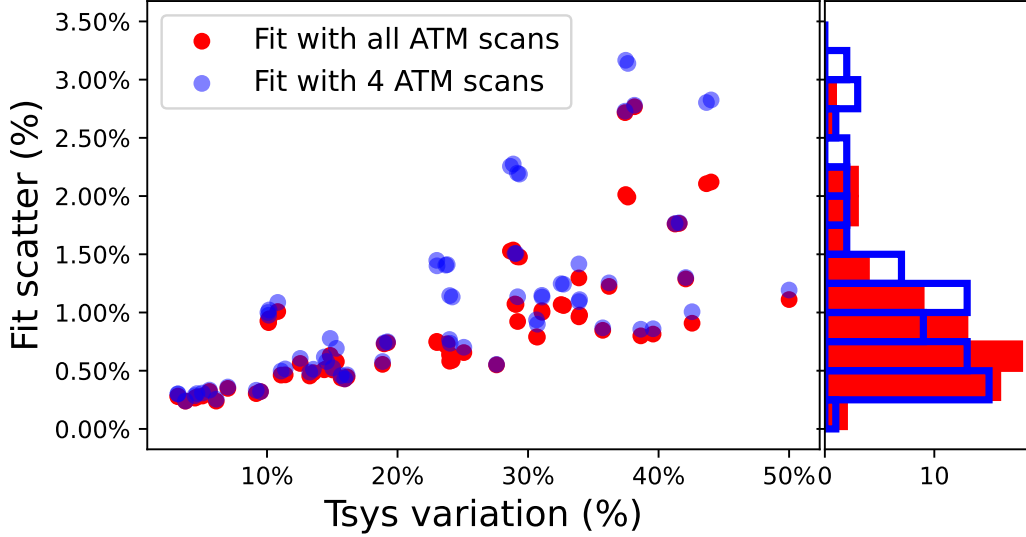


Figure 5. The scatter of data points around the T_{sys} vs T_{WVR} fits with all the antennas versus the maximal difference in \hat{T}_{sys} values for each T_{sys} spw of each dataset. The red and blue points are from fitting with all Atm-cal scans or just 4 Atm-cal scans respectively. The histogram at right side shows the distribution of the fit scatters using the two different methods. We can see the scatter of the fitting only increases slightly using just 4 Atm-cal scans.

have similar T_{sys} values, and hence might give us inaccurate fitting relation among small T_{sys} ranges. Furthermore, as described in Section 2.5, the linear approximation becomes insufficient when the opacity becomes significant or T_{sys} variations become large. A more robust method is to use an atmospheric opacity modeling code and an estimate of the T_{rx} and other static contributions to predict the T_{sys} vs T_{WVR} relation at the relevant frequencies, elevation and PWV ranges of the observation. In this section we use the Atmospheric Transmission at Microwave frequencies (ATM) model (Pardo et al. 2001) to predict the T_{sys} vs T_{WVR} relation for various datasets and compare the results with our heuristic method.

3.1. Modeling T_{sys} spectrum

We note that in previous sections when we explore the correlation between T_{sys} and T_{WVR} , we average T_{sys} for each spectral window along its spectral axis. Therefore, when we extrapolate the continuous T_{sys} , we are under the assumption that the T_{sys} spectrum does not vary significantly. In this subsection, we test this assumption with ATM modeling on the two representative datasets we have, Band8 and Band7a. We use the version of ATM included in CASA, accessed via a helper function `plotAtmosphere`¹ to generate T_{sys} and T_{WVR} spectra at frequency ranges of a given spw in the data. We set most of the parameters to the default for the ALMA

site (height 5000 m, pressure 557 mb and temperature 274 K). For each dataset, we set the PWV and elevation values to be the same as the value of first Atm-cal scan for the science target as our start point. Note that `aU.plotAtmosphere` only gives the T_{sky} and τ_{sky} . Therefore, we calculate the T_{sys} spectrum from the modeled T_{sky} and τ_{sky} using Eq. 1 by assuming T_{rx} to be 100 K.

We show our modeled results in Fig. 7. As we can see, spw 17 for dataset Band8 has significant dry opacity contribution as it sits at one of the O_2 lines. On the contrary, dataset Band7a is dominated by the wet component. We then increase the PWV and airmass ($1/\sin m_{\text{el}}$, m_{el} is the elevation) by a factor of 1.2 to see if they have different effects on increasing T_{sys} spectrum. For both dataset Band8 and Band7a, we can see increasing airmass is more effective in increasing the overall values of the T_{sys} spectrum. This is what we expect since increasing airmass will increase both wet and dry opacity while increasing PWV values only increase the wet opacity. If we compare the Band8 with Band7a, we can see the difference between increasing PWV and airmass is more significant for Band8 as the spectral window has significant dry opacity contribution.

To the first order, we see no significant change in the shape of T_{sys} spectrum by increasing either PWV or the airmass by 20%. But we do see some small change in T_{sys} spectrum shape while increasing airmass for dataset Band8. The lower frequency part has higher increase

¹ <https://safe.nrao.edu/wiki/bin/view/ALMA/PlotAtmosphere>

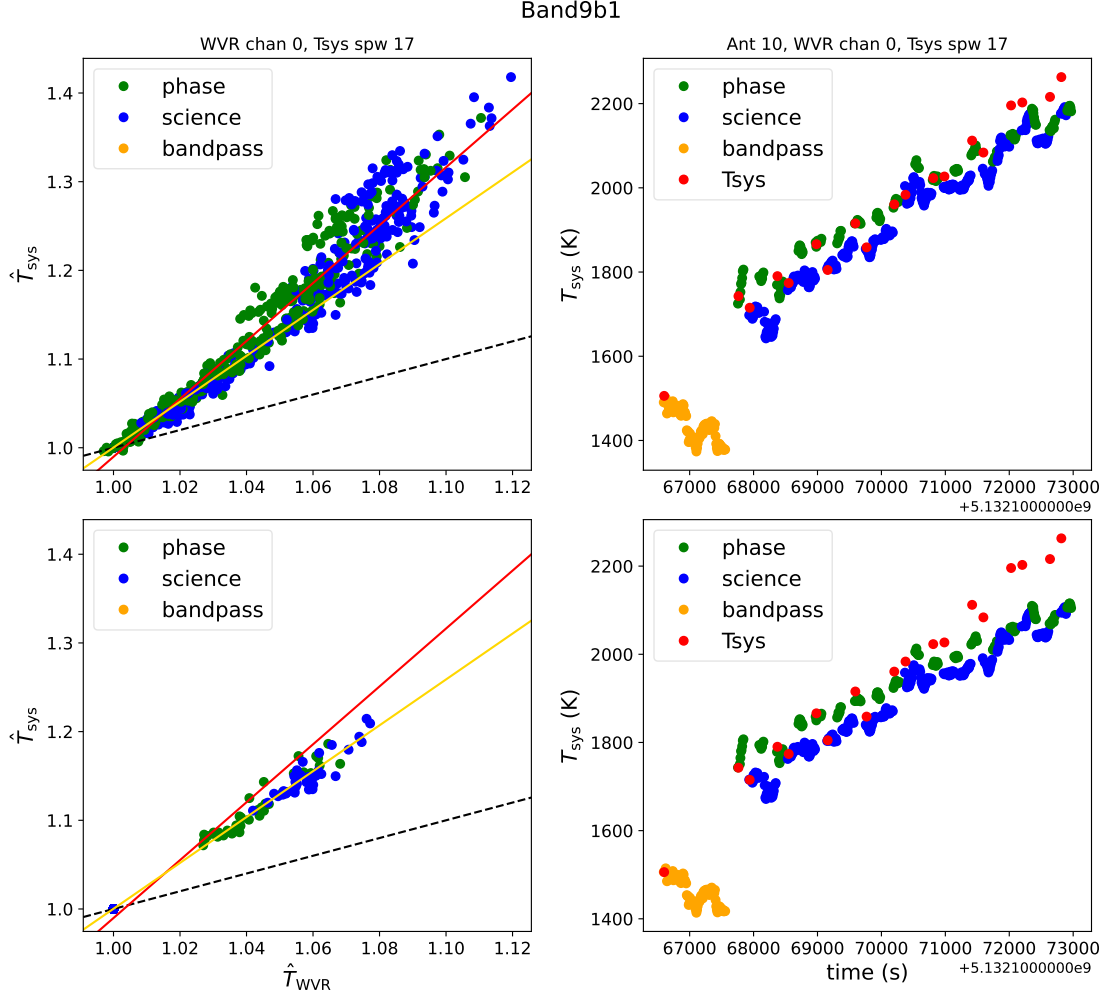


Figure 6. (Left) \hat{T}_{sys} (spw 17) versus \hat{T}_{WVR} (channel 0) for dataset Band9b1. The upper left panel shows \hat{T}_{sys} versus \hat{T}_{WVR} in spw 17 using all Atm-cal scans while the lower left panel shows the correlation with selected 4 Atm-cal scans. The red and yellow lines are the fits derived using all Atm-cal scans and just 4 Atm-cal scans. (Right) The measured and extrapolated T_{sys} for different observing targets as a function of time. Red points are measured T_{sys} values. T_{sys} in upper panel is extrapolated based on fits with all Atm-cal scans while T_{sys} in lower panel is extrapolated based on fits using only selected 4 Atm-cal scans. We can see in this case we will underestimate the T_{sys} value if we just use part of Atm-cal scans to fit the correlation.

than the higher frequency part. This is consistent with the dry opacity distribution in this spw as lower frequency has high dry opacity contribution. However, this is secondary effect over a narrow range of frequency coverage (2 GHz) and we do not expect it to affect our calibration result especially for the continuum. For spectral lines, the error in the extrapolation based on the T_{WVR} or PWV will be $\sim \pm 3\%$ (see Section 5.4 for more discussion). A future improvement might be correct the data spectrally rather than using a single channel-averaged value per timestamp.

3.2. Reproduce the Observed \hat{T}_{sys} vs \hat{T}_{WVR} correlation

In this section we will try to reproduce the observed \hat{T}_{sys} vs \hat{T}_{WVR} correlation in various cases. According to Eq. 3, T_{sys} are determined by T_{rx} and T_{sky} . T_{sky} can be

further determined by the measurement of PWV and elevation through ATM modeling. Therefore, we can generate modeled T_{sys} spectrum with given T_{rx} , PWV and elevation at a certain frequency range. To reproduce the T_{sys} measured in observation, we set the frequency range to be the same as the T_{sys} spectral window we want to model, with total bandwidth of ~ 2 GHz. The modeled T_{sys} spectrum are then averaged to a single T_{sys} value as what we did in observation. We also use the similar method to generate T_{WVR} at different WVR channels with given PWV and elevations. Note that T_{WVR} is essentially the sky temperature at frequency close to 183 GHz water line so we do not need to consider T_{rx} in this case.

For most of ALMA data, T_{rx} stays relatively constant throughout the observation. However, different anten-

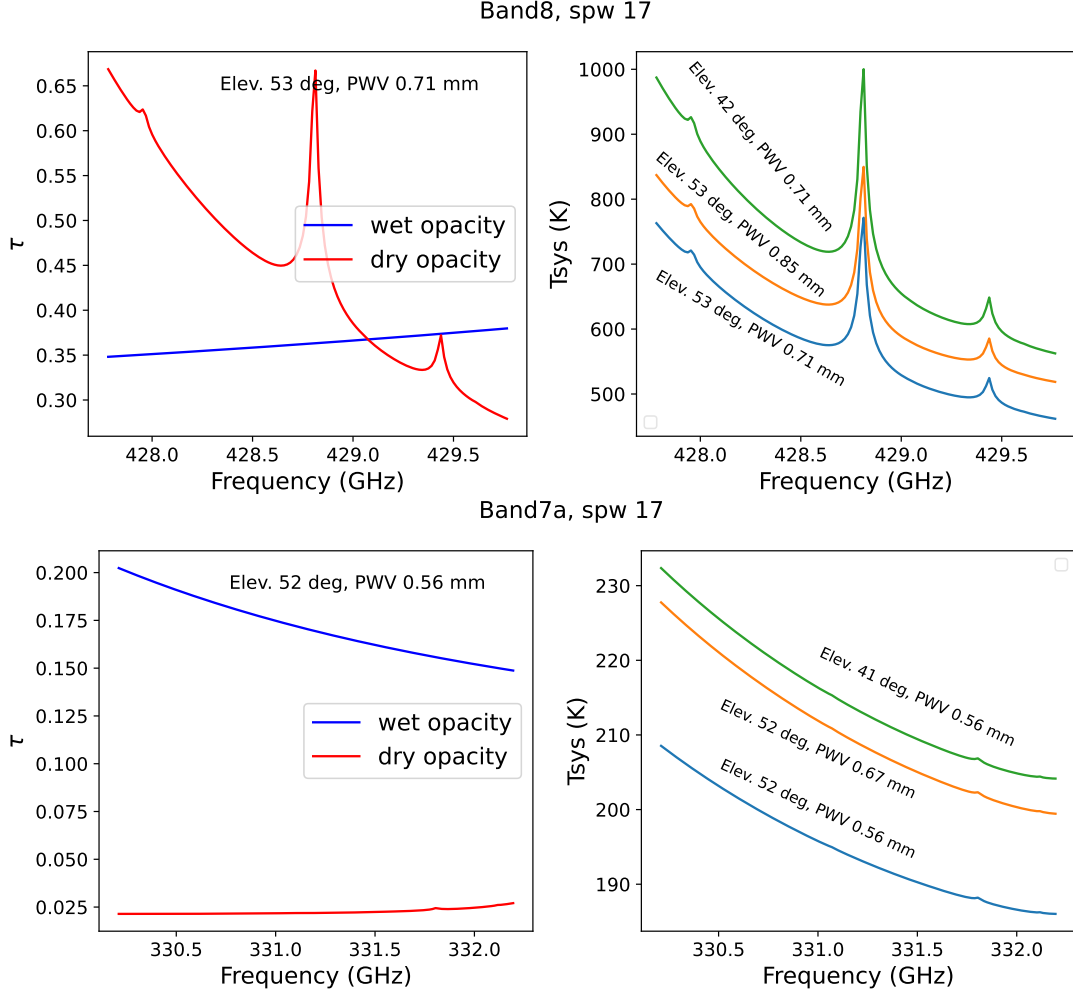


Figure 7. The ATM modeling opacity τ_{sky} and T_{sys} spectrum for dataset Band8 (upper) and Band7a (lower). For each row, the left panel shows the τ_{sky} spectrum for the wet and dry component. We can see that the dry component is significant for dataset Band8 but the water component is dominant for Band7a. The right panel shows the modeled T_{sys} spectrum for 3 set of different PWV and elevation values. The blue line is the baseline while we increase PWV or airmass value by 1.2 respectively for orange and green lines. We can see increasing airmass will increase T_{sys} faster since it increase T_{sky} from both wet and dry components.

nas generally have different T_{rx} values, which might give us slightly different shapes of correlation. Therefore, we first test how varying T_{rx} could affect the shape of \hat{T}_{sys} vs \hat{T}_{WVR} correlation for dataset Band8 spw 17 (Fig. 8). The T_{sys} and T_{WVR} are generated with varying T_{rx} and PWV values but with fixed elevation of 53 deg. We note that dataset Band8 has relatively constant elevations (see Table 2) for the science target throughout the entire observation, hence the T_{sys} variation across time are mostly due the the change in PWV values. After generating the modeled T_{sys} and T_{WVR} , we then normalize both quantities to the values when the PWV value is equal to that of the first science Atm-cal scan for each T_{rx} value. As we can see in Fig. 8, the ATM modeling shows a slight non-linear curvature for a high range

of T_{sys} , which depends slightly on the assumed T_{rx} and the elevation. However, the modeling curve is generally within the range of the data scatter. Varying T_{rx} also gives us similar correlation within data scatter of $\sim 1\%$. Therefore, T_{rx} values do not seem to affect the \hat{T}_{sys} vs \hat{T}_{WVR} correlation we get.

We then explore how varying PWV or elevation can affect the \hat{T}_{sys} vs \hat{T}_{WVR} correlation. As we have mentioned in Section 3.1, varying PWV or elevation might have different effects on changing T_{sys} values depending on how significant the dry opacities are at given frequency. Dataset Band8 and Band7a represent two cases where one has significant dry opacity contribution while the other are dominated by the wet opacity. Therefore, it is natural for us to explore what drives the \hat{T}_{sys}

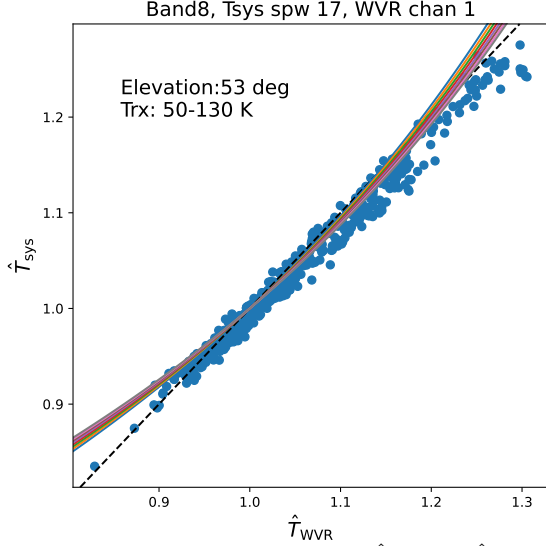


Figure 8. The correlation between \hat{T}_{sys} and \hat{T}_{WVR} (Fig. 2) overlaid by the predicted correlation curves from ATM modeling. Curves of different colors represent modeling using different T_{rx} values. The modeled T_{sys} and T_{WVR} are normalized to the value when PWV is 0.71 mm, which is the PWV value for the first Atm-cal scan for the science target. The T_{rx} range we use for ATM modeling is similar to the T_{rx} range of dataset Band8. We can see that varying T_{rx} values generally does not affect the correlation we get.

vs \hat{T}_{WVR} correlation in these two cases. Fig. 9 shows the comparison between modeling and observation for these two datasets where red and magenta line represents changing PWV and elevation respectively. In this comparison, the measured T_{sys} and T_{WVR} is normalized to the values in the first science Atm-cal scan instead of first Atm-cal scan for each target (see Section 2.4 for more description). The modeled T_{sys} and T_{WVR} is then normalized to the values when PWV and elevation is equal to those of the first science Atm-cal scan. We can see in both cases, the ATM modeling successfully reproduces the observed \hat{T}_{sys} vs \hat{T}_{WVR} correlation. For dataset Band8, the correlation is mainly driven by varying PWV values as the elevation for the science target stays relatively constant. By varying the elevation, we see a steeper slope of the correlation between \hat{T}_{sys} and \hat{T}_{WVR} . This is due to the fact that changing elevations (and therefore airmass) will be more effective to change T_{sys} values when the dry opacity contribution is significant (see discussion in Section 3.1). T_{sys} for bandpass target shows significant offsets from the main trend of phase/science target mainly due to the elevation difference, and hence sits at the red track of a different constant elevation value. Since the single bandpass T_{sys} measurement for each antennae has the same elevation but might point towards slightly different part of skys

with different PWV values, we see the bandpass data points still follow the track of constant elevation. On the other hand, we get similar \hat{T}_{sys} vs \hat{T}_{WVR} correlation by varying PWV or elevation for dataset Band7. This is probably due to the fact that wet component is dominant at this spectral window, hence changing PWV or elevation achieve the similar effect. We can also see in this case the T_{sys} for bandpass targets lie along the same trend as the phase-cal/science targets, which is what we expect since there is no specific parameter variation that could bring these data points out of the linear track.

As we have discussed in Section 2.5, dataset Band9b1 shows non-linear feature of \hat{T}_{sys} vs \hat{T}_{WVR} correlation as the T_{sys} variation becomes significantly large ($\sim 40\%$). Therefore, it is also worth testing if we can reproduce the curving feature for this dataset. We show the comparison in Fig. 10 using the same method described in the previous paragraph. As we can see, this dataset also has significant contribution from the dry opacity and hence shows different slopes when varying PWV or the elevation values. The observed T_{sys} generally agrees well with the ATM modeling relation of fixed PWV values. This suggests the T_{sys} variation shown for this dataset is mainly due to the elevation change. This is consistent with what we see in Fig. 6 as T_{sys} is smoothly increasing as the function of time without any short-time fluctuation. The bandpass data points also lie along the fixed PWV trend but with smaller value, which is probably due to larger elevation of the bandpass target. We also see some second-order scatter around the fixed PWV line for higher T_{sys} values, which might be due to the intrinsic scatter of PWV values during the observation. However, to the first order we can just measure T_{sys} once and predict the following T_{sys} based on the elevation change during the observation.

3.3. General Applicability of the current method and future direction for the improvement

In the previous section, we have successfully reproduced the \hat{T}_{sys} vs \hat{T}_{WVR} correlation for the 3 datasets with ATM modeling. However, we also find the correlation is not driven by a single parameter. For both dataset Band8 and Band9b1 where dry opacity is significant, the correlation is driven either by varying PWV or elevation. If both PWV and elevation have significant variation, we would not be able to get the tight correlation since the data points are driven up and down along different tracks. In contrast, for dataset Band7a where wet opacity is dominant, we can expect a tight \hat{T}_{sys} vs \hat{T}_{WVR} correlation even though both PWV and elevation have significant variation, as they are moving data up and down along the similar track. Therefore, it is safer

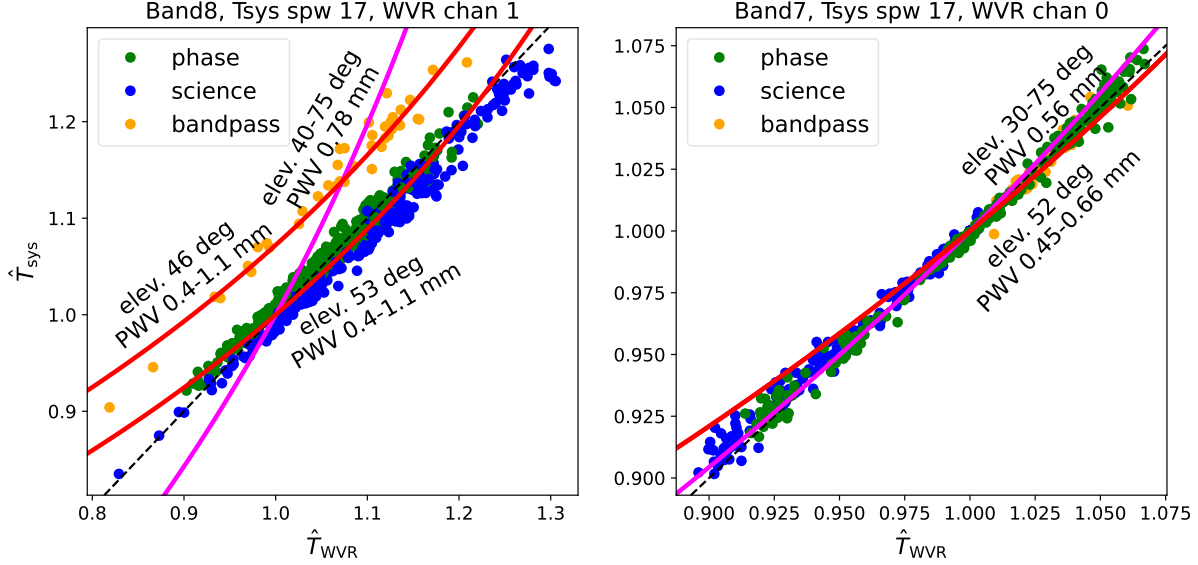


Figure 9. The correlation between \hat{T}_{sys} and \hat{T}_{WVR} overlaid by the predicted correlation curves from ATM modeling for dataset Band8 (left) and Band7a (right). Note that in this case both T_{sys} and T_{WVR} are all normalized to the first Atm-cal scan for the science targets (Eq. 15). The magenta line represent varying elevations while keeping the PWV value to be the same and the red line is vice versa. The T_{rx} are set to be equal to that of the first antenna of the data. The ATM modeled T_{sys} and T_{WVR} are normalized to the value when PWV and elevation is equal to those of first science Atm-cal scans in the data. We can see that for dataset Band8 where dry component is significant, varying PWV or elevation gives us correlations of different slopes. Since bandpass and science/phase targets have different elevations, we therefore see the offsets between T_{sys} for these two targets. In contrast, for dataset Band7a where wet component is dominant, varying PWV or elevation gives us similar \hat{T}_{sys} vs \hat{T}_{WVR} correlation. In this case, all 3 targets follow the same linear trend.

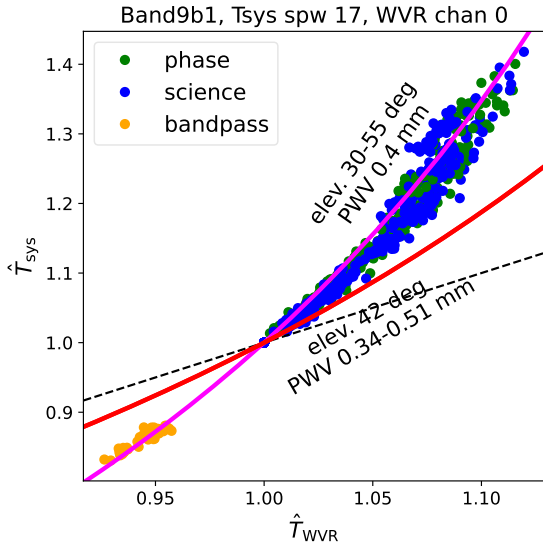


Figure 10. Similar plot as Fig. 9 but for dataset Band9b1. The dry component is also significant for this dataset. We can see that Band9b1 have a relatively constant PWV while the T_{sys} and T_{WVR} variation is mainly due to changing elevations.

to apply our current heuristic method to the dataset observed at frequency where wet opacity is dominant. In

Fig. 11, we model the wet and dry opacity spectrum covering the entire ALMA high frequency bands from Band 7 to Band 10. We assume the PWV of 0.5 mm, which is a typical value for the ALMA site, and elevation of 50 deg. For the wet opacity, we see several smooth line features, which indicate the presence of the H_2O line. For the dry opacity, we generally see a lot of Ozone lines featured as narrow spikes. These Ozone lines are generally much narrower than the typical bandwidth of a spw of 2GHz and hence have relatively small effect on the averaged T_{sys} values. On the other hand, we also see some broader spikes caused by O_2 lines. One of our dataset, Band8, sits right at the wings of one O_2 line and thus has significant dry opacity contribution. For all of our dataset used in this paper, only dataset Band8 and Band9b shows significant dry opacity contribution, which is consistent our expectation as all of the rest datasets in this paper exhibits a tight linear correlation (Fig. A1, A2 and A3).

However, we can expect that in lots of cases, the \hat{T}_{sys} vs \hat{T}_{WVR} might not follow a tight linear correlation due to various reasons mentioned above. In these cases, the best way is to directly derive T_{sys} from the ATM modeling. As shown in Section 3.2, with known T_{rx} , PWV and elevations for each Atm-cal scan, we can successfully

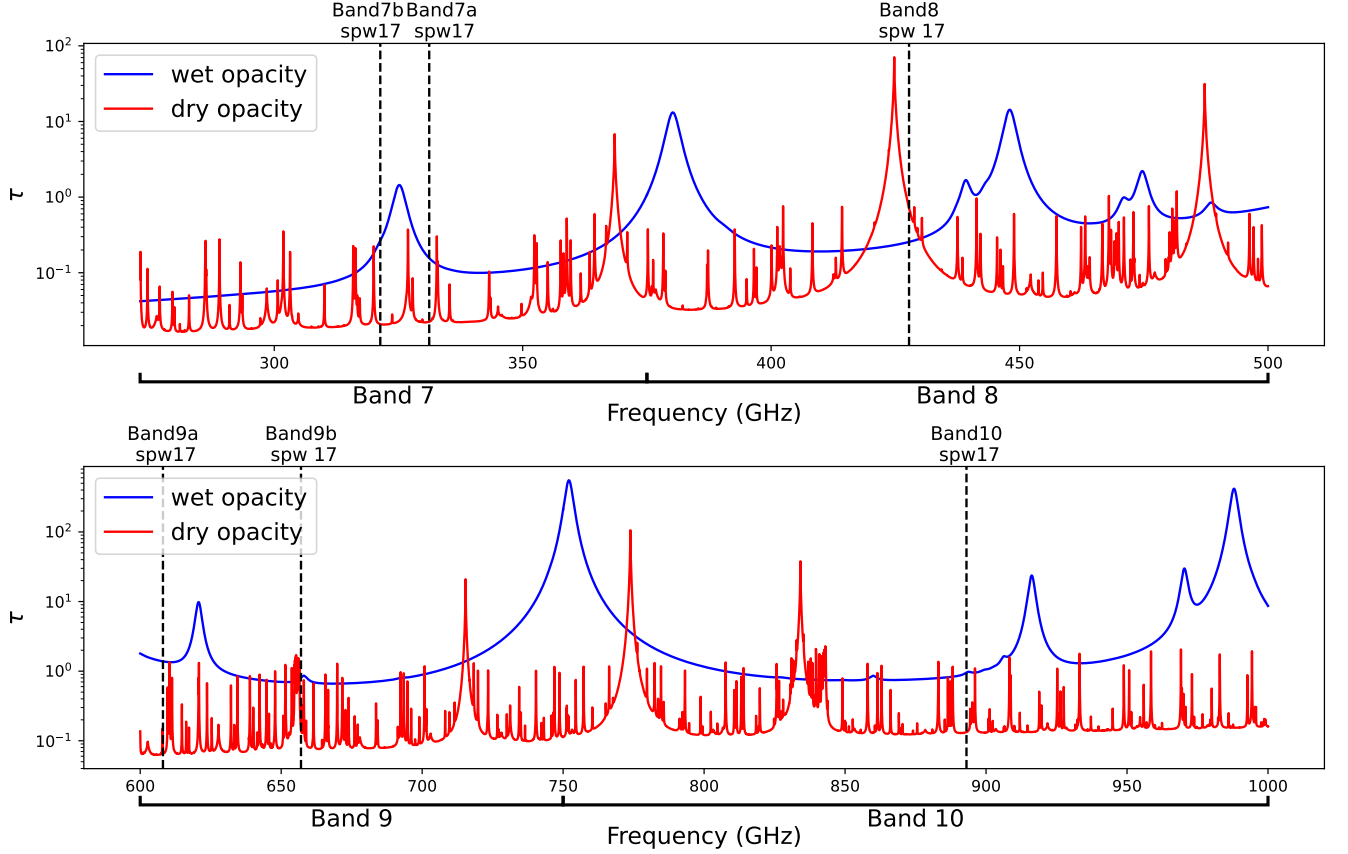


Figure 11. The wet and dry opacity from ATM modeling for the entire ALMA high frequency bands (from Band 7 to 10). We set the PWV value to be 0.5 mm and elevation to be 50 deg. The dashed line indicates the position of spectral window of ALMA data we have. We can see all of our dataset except for Band8 and Band9b are dominated by the wet component, which means changing PWV or elevations should give the similar \hat{T}_{sys} vs \hat{T}_{WVR} correlation for these datasets.

reproduce the T_{sys} measured in observation. *In the future, the best strategy for tracking T_{sys} is to derive PWV values from continuous T_{WVR} measurements at different WVR channels. By combining PWV, elevation and T_{rx} values, we can then reproduce the continuous T_{sys} throughout the observation.*

4. AC & SQLD DATA IN TRACKING T_{sys}

As mentioned in Section 1.4, we can also explore whether to use AC or SQLD data to track T_{sys} variation. Since AC and SQLD data are equivalent to one another (see Fig. 1 right panel), we only need to compare T_{sys} with one of the two quantities. We use AC data for comparison since the data size is much smaller. We use a similar method to normalize the AC data and compare it with the normalized T_{sys} . Fig. 12 shows the comparison between T_{sys} versus AC and T_{sys} versus WVR correlation for dataset Band8 and Band7a. We can see that the AC data also has a tight correlation with T_{sys} for dataset Band7. However, the correlation does not work well for dataset Band8 with a large scat-

ter. Therefore, we cannot just fit the relation to several Atm-cal scans to calculate the continuous T_{sys} with good precision.

The other thing we find is that AC and T_{sys} do not follow the proportional correlation as might be expected. The major reason is that AC data track the total signal received after the atmosphere attenuation while T_{sys} tracks the total signal before it comes through the atmosphere, as shown in the diagram in Fig. 14. The AC data is directly proportional to the total signal received by the antenna, which is mainly comprised of emission from the sky ($\eta_{\text{f}} T_{\text{sky}}$), the receiver itself (T_{rx}), and other fixed losses terminating at ground ($(1 - \eta_{\text{f}}) \times T_{\text{amb}}$). However, based on Eq. 1, T_{sys} is not directly proportional to these 3 components added together. Instead, T_{sys} can be thought of as brightness temperature of a fake source in space that generates a signal equal to the 3 components added together after atmosphere attenuation. In other words, for a single band setting,

$$P_{\text{AC}} \propto \eta_{\text{f}} e^{-\tau_0 \sec z} T_{\text{sys}} \approx T_{\text{rx}} + \eta_{\text{f}} T_{\text{sky}} + (1 - \eta_{\text{f}}) \times T_{\text{amb}} \quad (18)$$

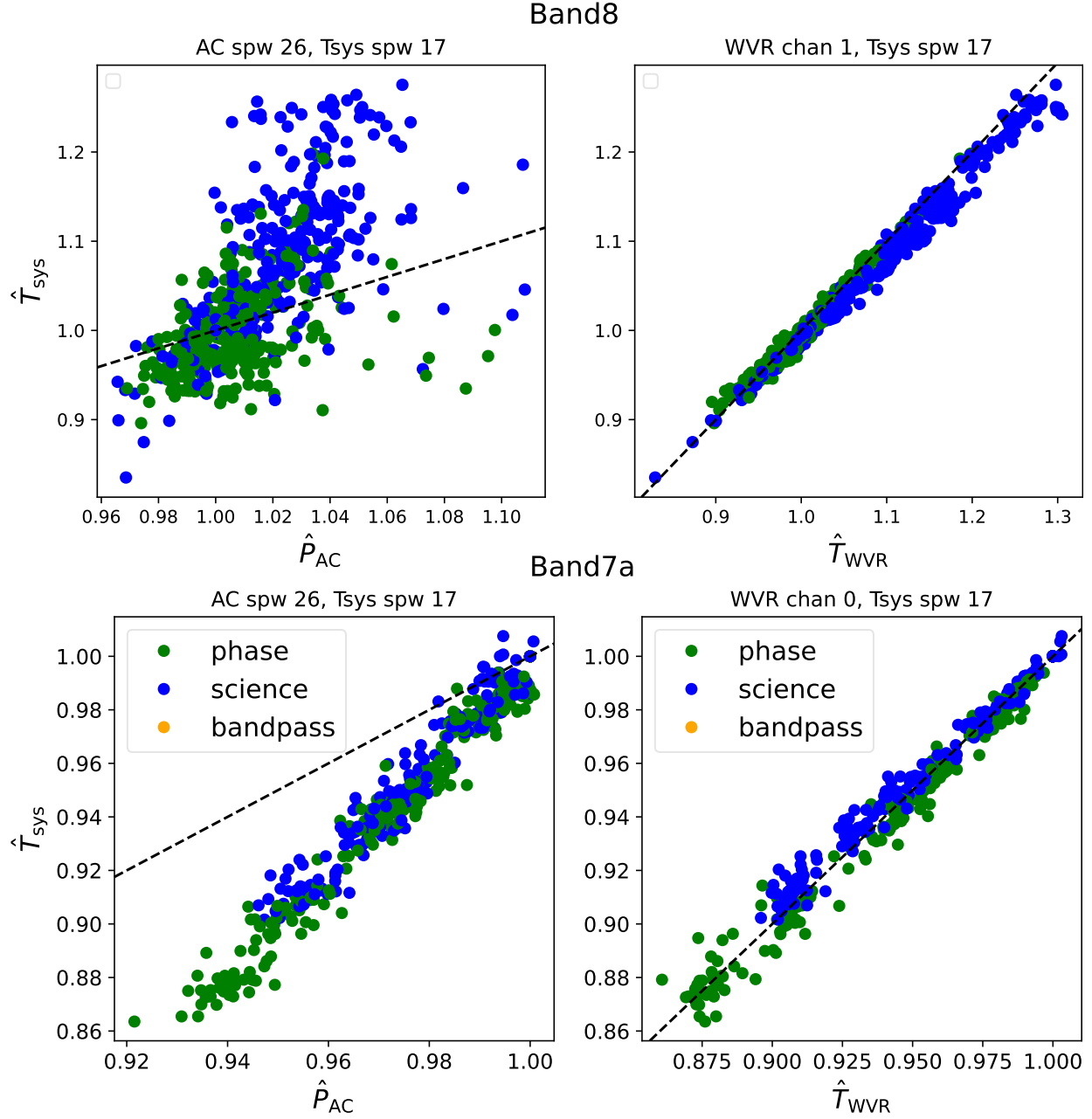


Figure 12. The correlation between the \hat{T}_{sys} and matched \hat{T}_{WVR} and normalized P_{AC} for all the antennas. Both WVR and auto-correlation data is averaged over 10 seconds.

We call the right side of the equation attenuated T_{sys} . We also normalize the attenuated T_{sys} the same way as we do for the original T_{sys} and compare it with normalized AC data. The comparison between T_{sys} and attenuated T_{sys} versus AC correlation is shown in Fig. 13. As we can see, the normalized attenuated T_{sys} follows the 1-to-1 relation with normalized AC as suggested by Eq. 18. If we rearrange Eq. 18, it becomes

$$T_{\text{sys}} \propto \frac{P_{\text{AC}}}{e^{-\tau_{\text{sky}}}} \approx \frac{P_{\text{AC}}}{1 - T_{\text{sky}}/T_{\text{amb}}} \quad (19)$$

Therefore, even though we have the AC data, we still need a method to continuously determine T_{sky} or τ_{sky} to obtain T_{sys} . A possible work-around is to use AC data to track T_{sky} first by combining Eq. 3, 2 and 19 as we can generally assume T_{amb} and T_{rx} to be constant. This technique has been applied to correct T_{sys} values in Agliozzo et al. (2017). However, in our case to extrapolate continuous T_{sys} , we need to note that the AC and SQLD data also suffer from gain drift and gain step changes between scans, as noted above and seen in Fig. 1.

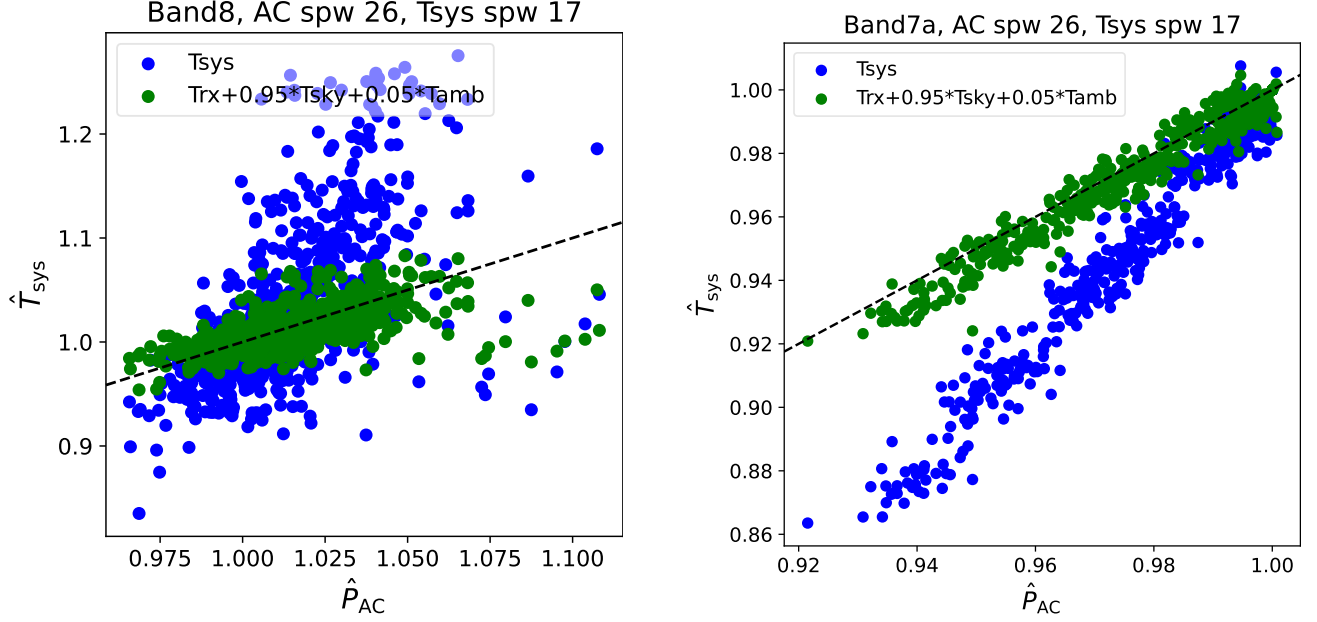


Figure 13. The correlation between the normalized T_{sys} (blue) and attenuated T_{sys} (green) and normalized AC data for the dataset Band8 and Band7a we have for all antennas of one spectral window. As we can see, the attenuated T_{sys} follows the 1-to-1 correlation with AC data, which proves equation 18 to be right.

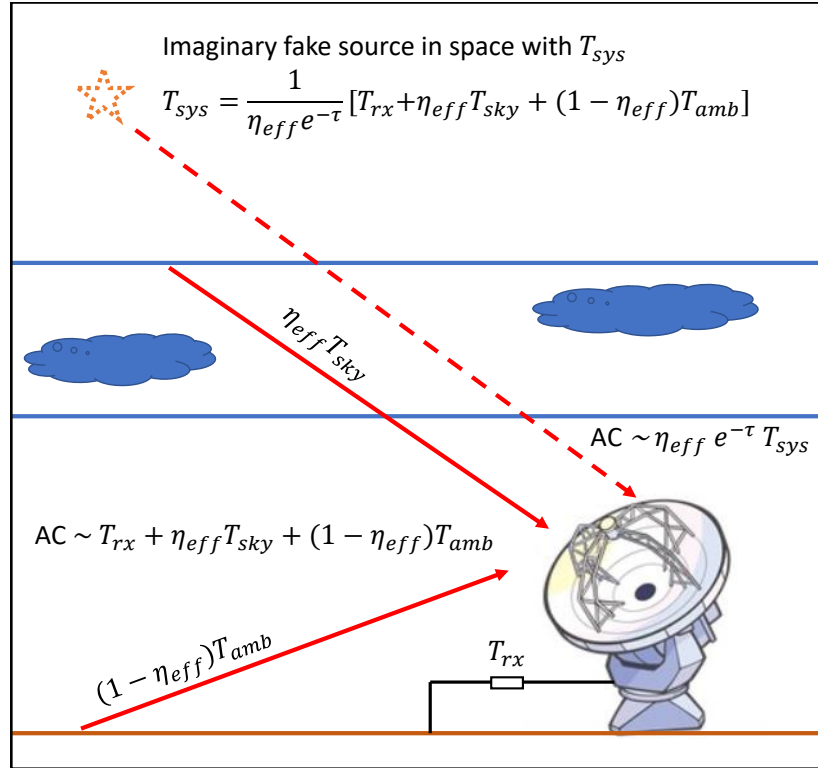


Figure 14. Illustration of the theoretical relationship between the autocorrelation signal AC and T_{sys} . AC is proportional to the total power signal received by an antenna, which includes receiver noise (T_{rx}), sky noise (T_{sky}), and thermal noise due to losses and spillover terminating around the ambient temperature $((1 - \eta_f) \times T_{\text{amb}})$. However, T_{sys} is corrected for a source outside the atmosphere, and includes an extra term due to the atmospheric attenuation.

5. APPLYING CONTINUOUS T_{sys} TO THE CALIBRATION

In Section 2, we demonstrated the viability to use WVR data to track T_{sys} continuously. In this section,

we apply the extrapolated continuous T_{sys} in calibration to test whether our new method for measuring T_{sys} works. We calibrate each dataset with the original T_{sys} table, the new continuous T_{sys} table extrapolated using all Atm-cal scans and that using just 4 Atm-cal scans with CASA (McMullin et al. 2007) package. We then make images from data calibrated using these 3 different methods and see if the measured fluxes for the same target are more consistent with each other using our new methods. The detailed description of the scripts we use for the data processing can be found at https://github.com/heh15/ALMA_intern_Tsys.git.

5.1. Creation of T_{sys} Table

In this subsection we discuss how we construct the new T_{sys} table used for the calibration. We note that the original T_{sys} table used for calibration is a spectrum with two polarizations. Recording all the extrapolated T_{sys} spectra in one big table would take a lot of disk space. Based on our check of the T_{sys} spectrum plots generated using the original calibration script, the shape of the T_{sys} spectrum of each spectral window does not vary much as a function of time. Therefore, we can just record the initial T_{sys} for each observing target in the T_{sys} table and record the ratio of the extrapolated T_{sys} relative to the initial T_{sys} into an amplitude gain table. In this case, the two tables we provide for T_{sys} calibration are

$$T_{\text{sys}}(t, \nu) = T_{\text{sys,obs}}(1^{\text{st}}, \nu)$$

$$G(t) = \sqrt{1 / \left[\frac{T_{\text{sys}}(t)}{T_{\text{sys}}(1^{\text{st}})} \right]_{\text{fit}}} \quad (20)$$

where $T_{\text{sys}}(t, \nu)$ is the recorded T_{sys} spectrum as a function of time t and frequency ν , and G is the derived gain as a function of t , $T_{\text{sys,obs}}(1^{\text{st}}, \nu)$ is the first T_{sys} spectrum measured for each type of observation of given antenna and spectral window and $\left[\frac{T_{\text{sys}}}{T_{\text{sys}}(1^{\text{st}})} \right]_{\text{fit}}$ is the extrapolated normalized T_{sys} from the fitting. We note that G is not directly equal to the \hat{T}_{sys} values. This is due to the different methods that CASA uses to handle T_{sys} and gain table. For each baseline, the correlated amplitude is

$$S(i, j) \propto \sqrt{T_{\text{sys}}(i)T_{\text{sys}}(j)} \propto \frac{1}{G(i)G(j)} \quad (21)$$

Therefore, the G is written so that it can be properly translated to the variation in T_{sys} .

In Section 2.3, we tested using only 4 T_{sys} measurements to fit the linear relation between \hat{T}_{sys} and \hat{T}_{WVR} . We saw that the difference between this method and using all T_{sys} measurements is small. However, we still need to quantify if the small difference in the linear fits

makes much difference in the measured flux of the image product. In this case, we also apply Eq. 20 to create the alternative T_{sys} table with the fitting relation derived from 4 Atm-cal scans. Therefore, we also extrapolate T_{sys} based on the fitting relation from 4 Atm-cal scans.

5.2. Calibrating and Imaging the Data

After we create the T_{sys} table, we then apply the continuous T_{sys} in calibration. The calibration script we use is generated from the command `es.generateReducScript()`. We then modify the script to use the alternative T_{sys} and amplitude gain tables created (see details in https://github.com/heh15/ALMA_intern_Tsys.git). We also run the original calibration script to calibrate the data with the original discrete T_{sys} method for comparison.

After calibration, we then proceed with making continuum images. We generally adopt the default settings using the command `tclean`. We set the `robust` parameter to be 2.0 instead of the default 0.5 to maximize the sensitivity and hence flux accuracy. We also set the number of iterations to be 0 to only make the dirty image. This process will reduce any effects from `tclean` itself when making comparisons between fluxes from different calibration datasets. For projects with multiple datasets, we directly compare the measured fluxes from different datasets to see if they are consistent with each other. For projects with just one measurement set, we further make images using just the first half or the second half of the science scans and then compare fluxes among these 3 images. The top panel in Fig. 15 shows an example of images made with the 3 different methods for dataset Band8. We can see that the structure of these images looks almost the same. This is what we expect since T_{sys} should only affect the intensity scale of the final image. The bottom panel of Fig. 15 shows the images of the same target made with all scans or just half of the scans using the new method to extrapolate T_{sys} from all Atm-cal scans. We can see the structures of these images are also almost the same, which is also what we expect as images from the same dataset should have the similar uv coverage. On the other hand, if we compare images made from different datasets, we can see there is a larger difference in image structure. This is most obvious in the Band 9 project 2019.1.00013.S shown in Fig. 16. The beam sizes for the 3 different images are also significantly different from each other. For dataset Band7b, all images have similar beam shapes so differences among the image structures are not that significant.

Once the images are made, we draw an aperture around the central point source to measure the flux.

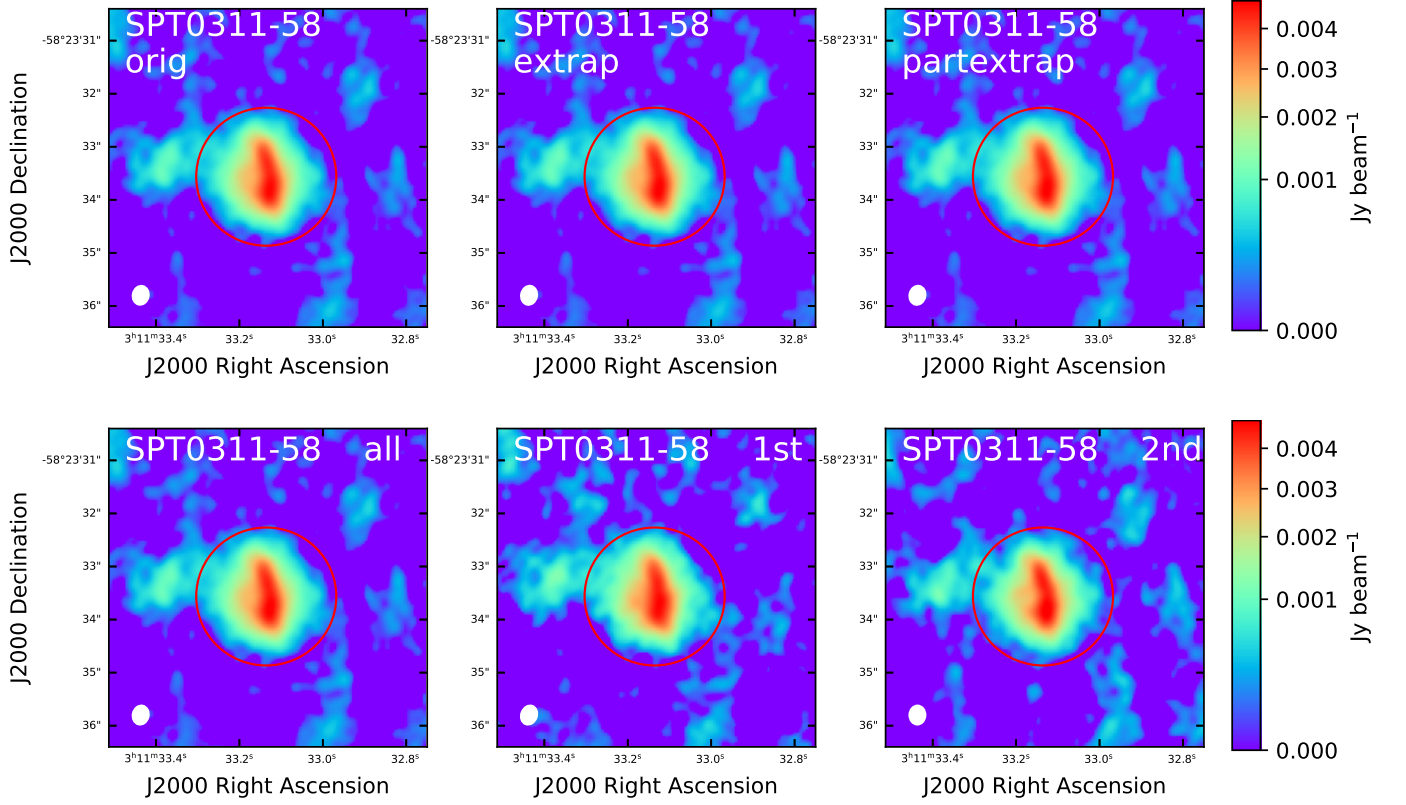


Figure 15. (Top) dirty images made using originally measured T_{sys} (orig), extrapolated continuous T_{sys} with all Atm-cal scans (extrap) and that with just 4 Atm-cal scans (partextrap) for dataset Band8 of SPT0311-58. The red circle is the aperture used for flux measurements. The fluxes for these 3 images are 0.0425, 0.0429 and 0.0407 Jy. (Bottom) dirty images made using alternative continuous T_{sys} table derived from fitting all Atm-cal scans. The 3 columns are images made using all, 1st half and 2nd half of the science scans. The fluxes for these 3 images are 0.0429, 0.0432 and 0.0428 Jy.

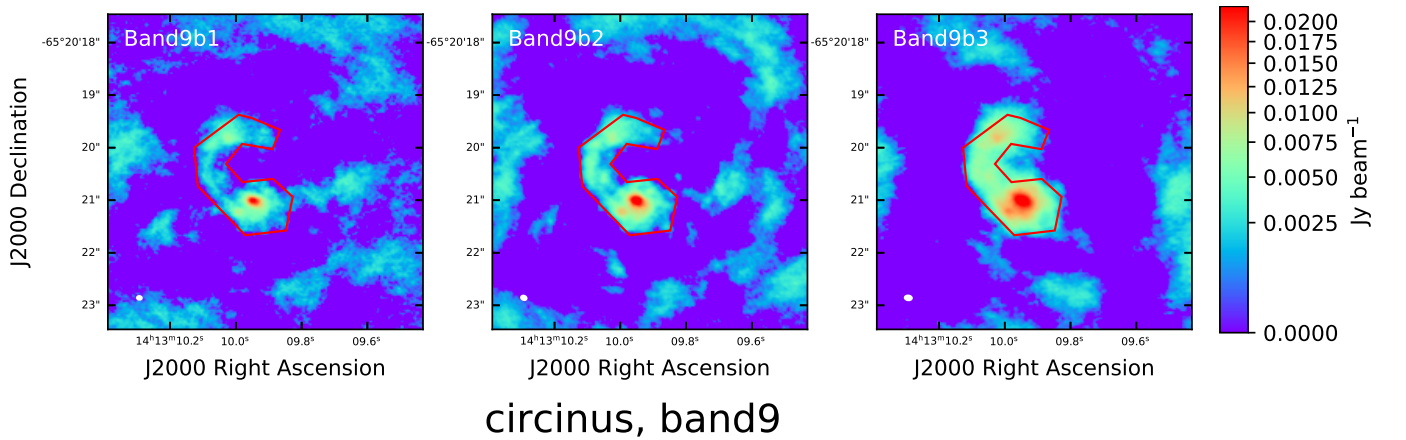


Figure 16. Dirty images made using alternative continuous T_{sys} table using all Atm-cal scans for the 3 data sets in the Band 9 project 2019.1.00013.S. The red polygons are apertures used to measure the flux. The fluxes for these 3 images are 0.875, 1.0 and 1.35 Jy.

Since the change in T_{sys} does not change the structure of the continuum image, we can compare the fluxes measured from the same aperture for the same target (as long as the aperture is not missing any flux, or there is no decorrelation – as decorrelation will reduce the total flux potentially). The apertures we used to measure the fluxes are shown in Fig. 15, 16 and B1. Note that as well as changes in T_{sys} , changes in phase decorrelation during a single observation and between observations may also affect the measured fluxes, and account for some of the scatter in fluxes in Table 4, 5 and 6. However, we assume this effect is the same in all reductions, independent of the T_{sys} calibration method.

5.3. Flux Comparison

The measured fluxes are recorded in Table 4, 5 and 6. The flux uncertainty can be separated into two parts, the measurement error and the calibration error. The measurement error is calculated as

$$\text{Meas.Err} = \text{rms} \sqrt{N_{\text{beam}}} \quad (22)$$

where the rms is the measured noise of the image and N_{beam} is the number of beams across the aperture used to measure the flux. On the other hand, our alternative method to measure T_{sys} should mainly work on reducing the calibration error. To test if our new method improves the flux calibration accuracy, we need to quantify the calibration error for each method we use. For projects with single dataset, we compare fluxes of images made with 1st half, 2nd half and all scans and calculate the maximal difference between the 3 flux values as the calibration error. For projects with multiple datasets, the calibration error is calculated as the standard deviation of fluxes of the different datasets. Both measurement and calibration errors are recorded in Table 4, 5 and 6.

To further compare how our methods work for datasets from different frequency bands, we normalize the flux values to the flux value using the original discrete T_{sys} calibration method with all Atm-cal scans. The relative uncertainties are calculated as the calibration error divided by the flux value for each method using all Atm-cal scans. The comparison is shown in Fig. 17. We can see that fluxes using the original T_{sys} table do not differ significantly from the fluxes using alternative T_{sys} and gain tables, with maximal differences smaller than 5 %. We also demonstrate that the extrapolated T_{sys} using 4 Atm-cal scans gives us fluxes that are almost the same as when using all Atm-cal scans, which proves it is viable to significantly reduce the number of T_{sys} measurements by using WVR-tracked T_{sys} . Furthermore, it seems for most of the datasets,

our new methods give better flux consistency, especially for dataset Band8 which brings down the flux calibration uncertainty from $\sim 10\%$ to 0.7% . As we have shown in Section 3.2, the T_{sys} variation for this dataset is mainly driven by the PWV variation at short time-scale. In this case, the discrete T_{sys} measurements probably did not catch the fluctuations of the real-time T_{sys} . Our new method instead catches the variation in T_{sys} between the discrete ATM calibrations, and thus keeps the flux consistent.

This method also works for dataset Band7b with multiple data sets with relative flux uncertainties reduce from 2.5% to 1.9%. The only project that gives us larger flux uncertainties using our new methods is the dataset Band9b. For this project, the uncertainties using all 3 methods are quite large ($\sim 15\%$). The large uncertainties are probably due to different uv-coverages, the complex target structure and relatively high phase noise in these datasets. This is also a tricky dataset for which the linear fitting does not work as well as for other data sets. In our future work, we will explore if the larger uncertainty is caused by imperfect fitting of T_{sys} versus T_{WVR} relation.

5.4. Additional considerations for the continuous T_{sys} method

For some targets, the source brightness temperature in single-dish measurements can be significant compared with T_{sys} . For example, this may occur for bright galactic targets in 12CO, bright masers, or for some Solar System objects in continuum. As the widths of galactic spectral lines are generally negligible ($< 1\%$) compared with the normal bandwidth (2GHz) used to measure T_{sys} , and the continuous T_{sys} method uses a spectrally-averaged broad-band T_{sys} , then the effect of bright lines in such cases will be negligible. But for very bright continuum sources such as planets, the spectrally-averaged T_{sys} will potentially be affected by the target brightness. However, the beam of the WVR unit on each antenna is offset from the optical axes of the receiver beams by several arc minutes (depending on the receiver band in use - see ALMA Technical handbook); this means that the WVRs are not pointing to the science target, and in general will not be affected by its strong continuum. Additionally, it has recently become apparent that the method used by ALMA to measure T_{sys} , using off-source data along with the normalisation of the visibilities using the autocorrelation, introduces a calibration error

Table 4. Flux measured for project with only one dataset

Data Label	Target	Scans used	Flux (Jy)			Meas. Err. (Jy)	Beam (")
			Tsys_orig	Tsys_extrap	Tsys_partextrap		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Band10	Arp 220	all	7.169	7.146	7.187	0.39	0.5 x 0.47
		1st half	7.1638	7.135	7.171	0.42	0.52 x 0.45
		2nd half	7.1077	7.09	7.133	0.4	0.48 x 0.48
		MAX. DIFF ^a	0.06	0.056	0.054		
Band9a	IRAS16293-B	all	10.43	10.36	10.324	0.34	0.34 x 0.27
		1st half	10.4	10.32	10.32	0.36	0.35 x 0.27
		2nd half	10.47	10.4	10.34	0.38	0.34 x 0.26
		MAX. DIFF.	0.07	0.04	0.02		
Band8	SPT0311-58	all	0.0425	0.0429	0.0429	0.0015	0.35 x 0.3
		1st half	0.0453	0.0432	0.0433	0.0019	0.36 x 0.29
		2nd half	0.0407	0.0428	0.0429	0.0017	0.35 x 0.3
		MAX. DIFF	0.0046	0.0003	0.0003		
Band7a	HT-Lup	all	0.173	0.178	0.178	0.0038	0.22 x 0.12
		1st half	0.169	0.175	0.175	0.0044	0.22 x 0.12
		2nd half	0.175	0.18	0.18	0.0045	0.22 x 0.12
		MAX. DIFF.	0.006	0.005	0.005		

Columns: (1) The label of each dataset (see Table 2) (2) The target name. (3) The science scans used to make images. (4) Fluxes of the source with images made using original calibration script. (5) Fluxes of the source with images made using modified script with alternative T_{sys} table. The T_{sys} is extrapolated based on all Atm-cal scans. (6) The images made with modified script but T_{sys} is extrapolated from 4 Atm-cal scans. (7) The measured flux errors for the images made with original T_{sys} table (8) The beam size of images using original T_{sys} table.

Rows: a. The maximal differences for fluxes at each column.

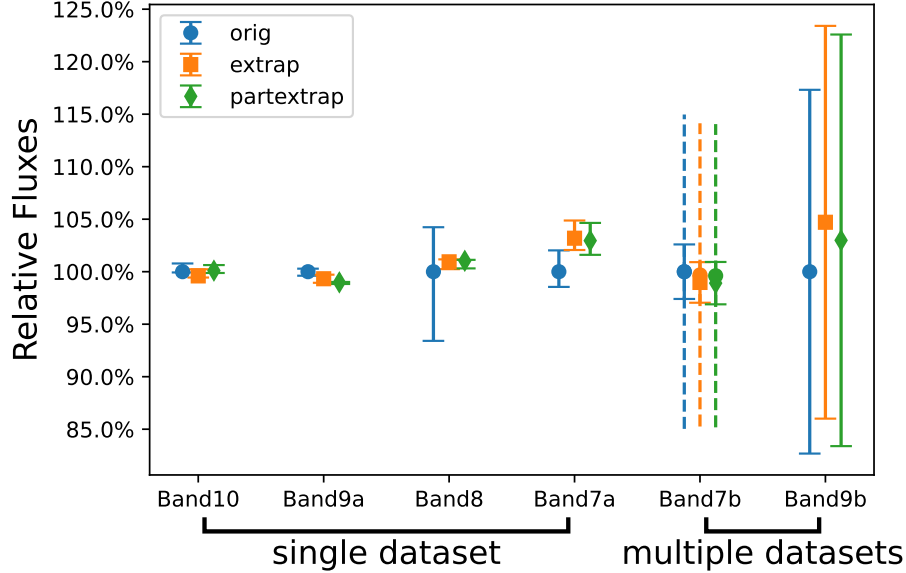


Figure 17. Comparison of fluxes of images made using the original T_{sys} (orig), continuous T_{sys} extrapolated from all Atm-cal scans (extrap) and that extrapolated from 4 Atm-cal scans (partextrap). The vertical axis are the ratio of fluxes to the fluxes using original method. The first 4 projects contain single dataset while the last two projects contain multiple datasets. The dashed line for Band7b project shows the error including the two datasets that has abnormal flux values (see Table .5). We can see the flux consistency is generally better for most of the projects with smaller uncertainties. See description in Section 5.3.

Table 5. Flux measured for data for Band7b project (AS205A)

Dataset Label	Flux (Jy)			Meas. Err. (Jy)	Beam (")
	Tsys_orig	Tsys_extrap	Tsys_partextrap		
(1)	(2)	(3)	(4)	(5)	(6)
Band7b1	0.7699	0.759	0.7594	0.031	1.13 x 0.79
Band7b2	0.713	0.7146	0.7129	0.029	0.96 x 0.77
Band7b3*	1.001	1.005	1.003	0.043	0.68 x 0.53
Band7b4	0.747	0.7378	0.7359	0.028	1.03 x 0.78
Band7b5*	0.556	0.58	0.58	0.029	1.31 x 0.72
Band7b6	0.749	0.7534	0.7533	0.03	0.91 x 0.08
Band7b7	0.7355	0.7347	0.7348	0.029	1.01 x 0.81
Band7b8	0.768	0.737	0.7372	0.029	1.24 x 0.79
AVG. ^a	0.7471	0.7394	0.7389		
STD. ^b	0.0194	0.0143	0.0149		

Columns: (1) The label for each dataset (see Table 2). (2) Flux of the data set using the original T_{sys} table. (3) Flux of the data set using alternative continuous T_{sys} table with the linear relation fitted using all Atm-cal scans. (4) Flux of the data set using continuous T_{sys} table with the linear relation fitted using part of Atm-cal scans. (5) Measured flux errors using original T_{sys} table. (6) Beams of the image using original T_{sys} table.

Rows: a. The average value for each column. b. The standard deviation for each column.

Notes: * denotes data with unusual fluxes. Fluxes from these data are not included in the calculation of the average and standard deviation value.

Table 6. Flux measured for data for dataset Band9b (Circinus)

Dataset Label	Flux (Jy)			Meas. Err. (Jy)	Beam (")
	Tsys_orig	Tsys_extrap	Tsys_partextrap		
(1)	(2)	(3)	(4)	(5)	(6)
Band9b1	0.85	0.875	0.849	0.018	0.086 x 0.065
Band9b2	0.96	1	0.983	0.021	0.104 x 0.077
Band9b3	1.27	1.35	1.34	0.016	0.131 x 0.085
AVG.	1.027	1.075	1.057		
STD.	0.178	0.201	0.207		

Columns: (1) The label for each dataset (see Table 2). (2) Flux of the data set using the original T_{sys} table. (3) Flux of the data set using alternative continuous T_{sys} table with the linear relation fitted using all Atm-cal scans. (4) Flux of the data set using continuous T_{sys} table with the linear relation fitted using part of Atm-cal scans.

Rows: a. The average value for each column. b. The standard deviation for each column.

for bright sources²³. The planned change is to measure T_{sys} on-source. Again, this should not significantly affect the continuous T_{sys} method, for reasons given above.

For spectral lines, an assumption is made that T_{sys} is mainly affected by PWV, and the correlation of T_{sys} with PWV uses T_{sys} averaged over the spectral window. This is considered reasonable for continuum and most spectral lines, but for calibration of spectral lines coincident with deep Ozone absorption (e.g. see Fig. 7), the correlation will have a slightly different slope and intercept. In general this is considered a second-order effect; for example, a line exactly coincident with the strong O_3 peak at 428.8GHz in Fig. 7, the error in the correction of T_{sys} based on the PWV would be $\sim \pm 3\%$. A future improvement might be correct the data spectrally rather than using a single channel-averaged value per timestamp. However, this would make the correction table significantly larger (see Section 5.1).

An additional use of the continuous T_{sys} method could be to correct for the increase in T_{sys} due to shadowing of the antennas. On ALMA, the default is that data taken with any slight blocking of the beam from an antenna, for example by a nearby antenna or building, is flagged and removed during data reduction. In general this cannot be corrected for using the gain calibrator amplitude solution, as this is not observed at the same sky location as the target. However, if the corresponding increase in T_{sys} due to shadowing is measured continuously, it may be possible to calibrate out some degree of shadowing. Further investigation of this technique should be done.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we explore a new method to use continuous datastreams available from WVR monitoring to track the atmospheric opacity and hence T_{sys} in mm and submm data. The aim is to improve flux calibration in conditions where the sky opacity is rapidly varying, and to reduce overheads needed for frequent discrete calibration using internal loads. Here we summarize our main conclusions regarding initial tests of this method.

- There is a tight linear correlation between normalized T_{sys} and T_{WVR} , with typical scatter of $\sim 1\%$. For the worst case of Band 9 data with large T_{sys} variations (50%), the simple linear fit would give us scatter of $\sim 4\%$, which is due to the

non-linearity of the relation at high opacities with large T_{sys} variations. Although the exact form of the linear relation varies among different spectral windows and different data sets, we can use as few as 4 Atm-cal scans to determine the slope and intercept of the linear relation, which suggests it is possible to significantly reduce the number of discrete T_{sys} measurements during observations, particularly at high frequencies. Furthermore it is not necessary to perform separate calibrations on the phase calibrator and science target, as the continuous T_{sys} method is able to track differences in T_{sys} between the two.

- We have successfully reproduced the observed tight \hat{T}_{sys} vs \hat{T}_{WVR} correlation using ATM modeling for several datasets. The ATM modeling suggests that changing elevation or PWV will give us \hat{T}_{sys} vs \hat{T}_{WVR} relation of different slopes when the dry opacity is significant at the observing frequencies. This suggests that we might not get the tight \hat{T}_{sys} vs \hat{T}_{WVR} correlation at these frequencies if both PWV and elevation varies significantly. A better strategy to track T_{sys} , especially in the cases mentioned above, would be to calculate the PWV using the continuous T_{WVR} measurements from different WVR channel and combine the measured PWV, elevation and T_{rx} for the dataset to derive the continuous T_{sys} based on the ATM modeling.
- We apply the continuous T_{sys} in calibration and find that it generally gives us more consistent fluxes for the same target. For the dataset Band8 which has the largest PWV variation, the flux accuracy improved from 10% to 0.7%. The only exception is the dataset Band9b as our new methods give higher flux uncertainties. We suspect part of reasons are due to the imperfect linear fitting of the T_{sys} vs T_{WVR} relation. Since the uv-coverages for the data sets in this project are significantly different, it is hard to confirm this scenario for this data set.
- If this method is used for sub-mm observatories such as ALMA, it can reduce the number of T_{sys} measurements required at high frequencies from 10 \sim 20 down to 5 (4 T_{sys} measurements for the fitting and 1 bandpass T_{sys}) or fewer. Assuming each observation block takes ~ 60 mins and each T_{sys} measurement takes about 30 – 40 seconds, it has the potential to save $\sim 10\%$ of observing time for high frequency observing, which is made more valuable as the amount of time in such good conditions is limited.

² <https://help.almascience.org/kb/articles/what-are-the-amplitude-calibration-issues-caused-by-alma-s-normalization-strategy>

³ <https://help.almascience.org/kb/articles/what-errors-could-originate-from-the-correlator-spectral-normalization-and-tsys-calibration>

We thank the referee for thoughtful comments and constructive suggestions, particularly regarding ATM modeling. This paper makes use of the following ALMA data:

ADS/JAO.ALMA #2015.1.00271.S
 ADS/JAO.ALMA #2016.1.00744.S
 ADS/JAO.ALMA #2018.1.01778.S
 ADS/JAO.ALMA #E2E8.1.00003.S
 ADS/JAO.ALMA #2018.1.01210.S
 ADS/JAO.ALMA #2019.1.00013.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of

Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This research made use of Astropy,⁴ a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). HH acknowledges the support of NSERC-CREATE NTCO training program. The research of C.D.W. is supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs program.

Software: astropy (Astropy Collaboration et al. 2013, 2018), CASA (McMullin et al. 2007)

REFERENCES

- Aglozzzo, C., Trigilio, C., Pignata, G., et al. 2017, ApJ, 841, 130, doi: [10.3847/1538-4357/aa72a1](https://doi.org/10.3847/1538-4357/aa72a1)
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- Bachiller, R., Carilli, C., Cox, P., et al. 2003, ALMA Science Advisory Committee (ASAC), 2003, "<https://www.nrao.edu/archives/items/show/34475>"
- Brogan, C. 2018, Advanced Calibration Topics - I, "https://science.nrao.edu/science/meetings/2018/16th-synthesis-imaging-workshop/talks/Brogan_Adv_Cal.1.pdf"
- Condon, J. J., & Ransom, S. M. 2016, Essential Radio Astronomy
- Dempsey, J. T., Friberg, P., Jenness, T., et al. 2013, MNRAS, 430, 2534, doi: [10.1093/mnras/stt090](https://doi.org/10.1093/mnras/stt090)
- Hills, R. 2004, ALMA Memo No. 495
- Hills, R., Gibson, J., Richer, J., et al. 2001, ALMA memo 352
- Mangum, J. 2002, ALMA Memo No. 434
- . 2017, ALMA Memo No. 602
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127
- Pardo, J. R., Cernicharo, J., & Serabyn, E. 2001, IEEE Transactions on Antennas and Propagation, 49, 1683, doi: [10.1109/8.982447](https://doi.org/10.1109/8.982447)
- Payne, J., Vaccari, A., Emerson, D., & Mangum, J. 2001, ALMA Construction Project Book, Chapter 3 Section 2.
- Remjian, A., Biggs, A., Cortes, P. A., et al. 2019, ALMA Technical Handbook, ALMA Doc. 7.3, ver. 1.1, 2019, doi: [10.5281/zenodo.4511522](https://doi.org/10.5281/zenodo.4511522)
- Yun, M., Bastian, T., Holdaway, M., Mangum, J., & Welch, J. 1998, ALMA Memo No. 211

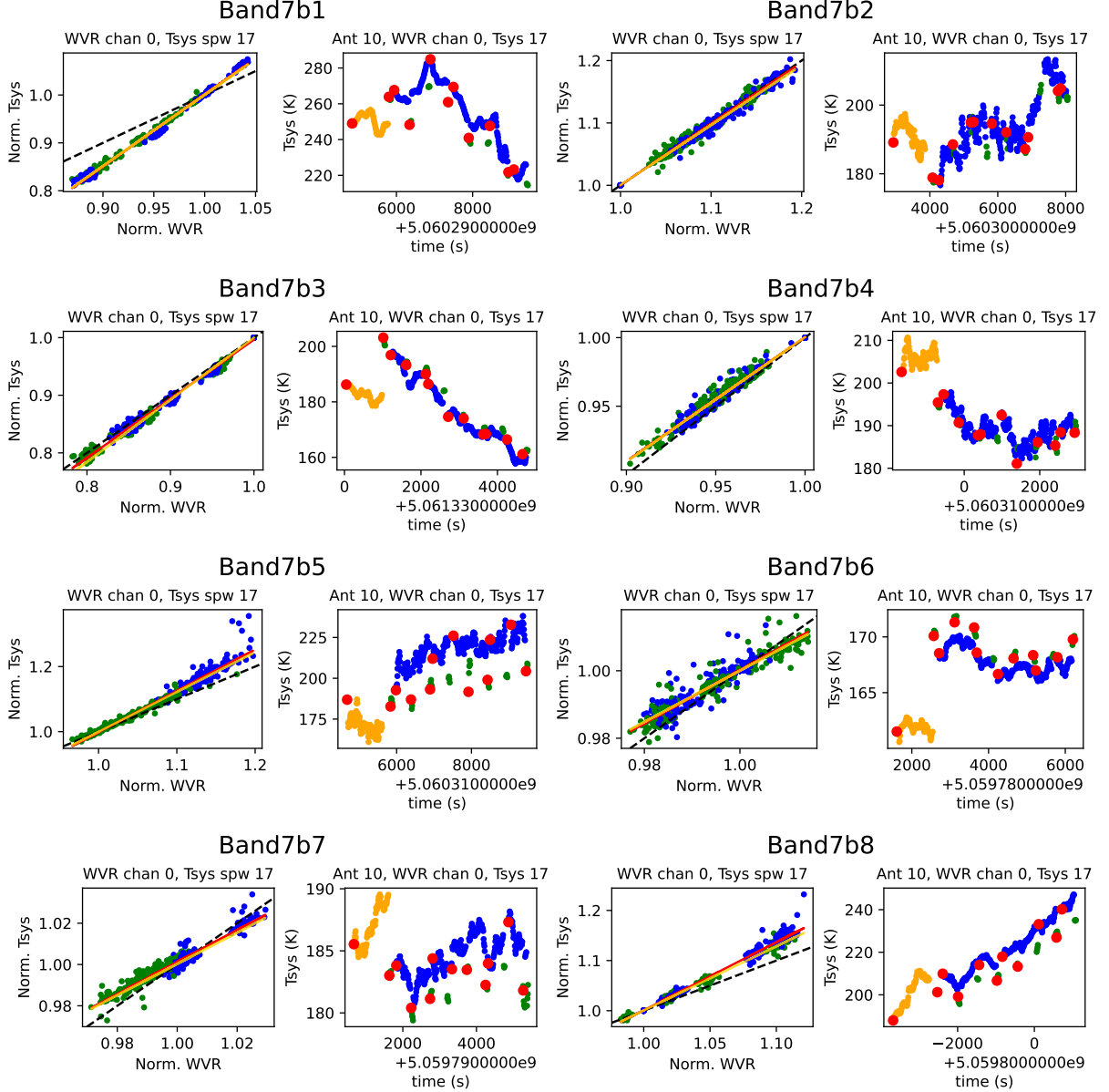
Moreno, R., & Guilloteau, S. 2002, ALMA memo 372

⁴ <http://www.astropy.org>

APPENDIX

A. T_{sys} VS T_{WVR} RELATION

In this section we show the T_{WVR} vs T_{WVR} relation and the extrapolated T_{sys} for the rest of data we have.



Project: 2018.1.01210.S; Band 7

Figure A1. The summary of linear relation between T_{sys} and T_{WVR} and comparison between measured and extrapolated T_{sys} for measurement sets in project 2018.1.01210.S. For each measurement set, the left subplot shows the linear correlation between normalized T_{sys} and T_{WVR} . The blue and green points are from science and phase scan. The dashed line shows the 1-to-1 relation. The red and golden solid line is the fitted linear relation using data from all Atm-cal scans or just 4 Atm-cal scans. The right plot shows the extrapolated T_{sys} based on the fitting relation using all Atm-cal scans. The orange, green and blue points are extrapolated T_{sys} for bandpass, phase and science targets. The red points are the original measured T_{sys} .

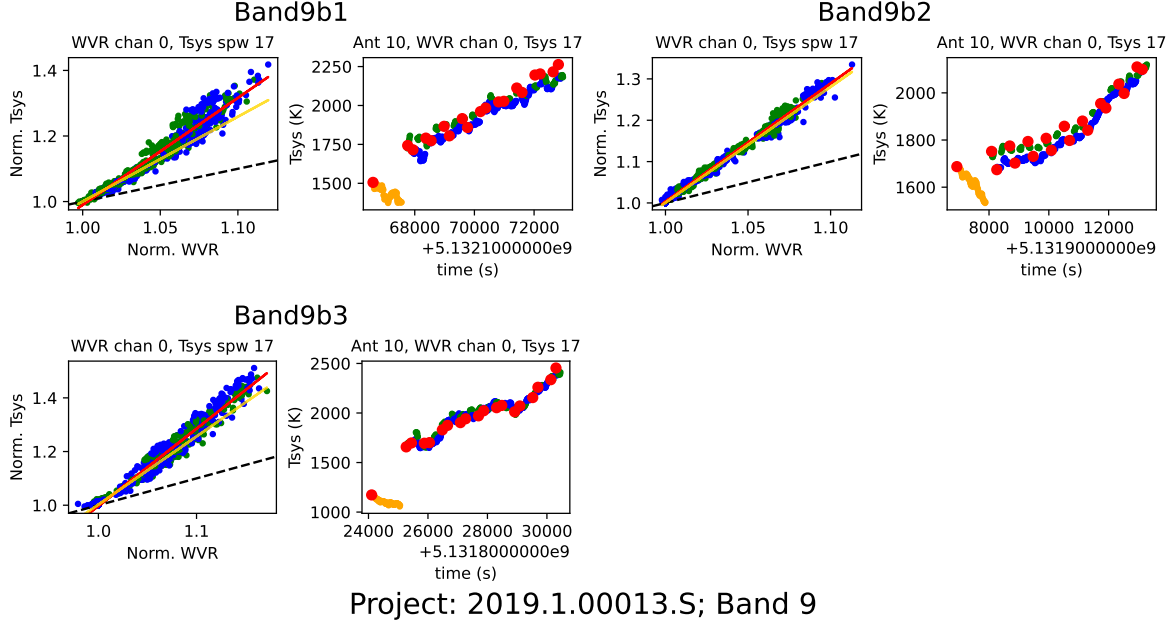


Figure A2. The summary of linear relation between T_{sys} and T_{WVR} and comparison between measured and extrapolated T_{sys} for measurement sets in project 2019.1.01210.S. See details in Fig. A1.

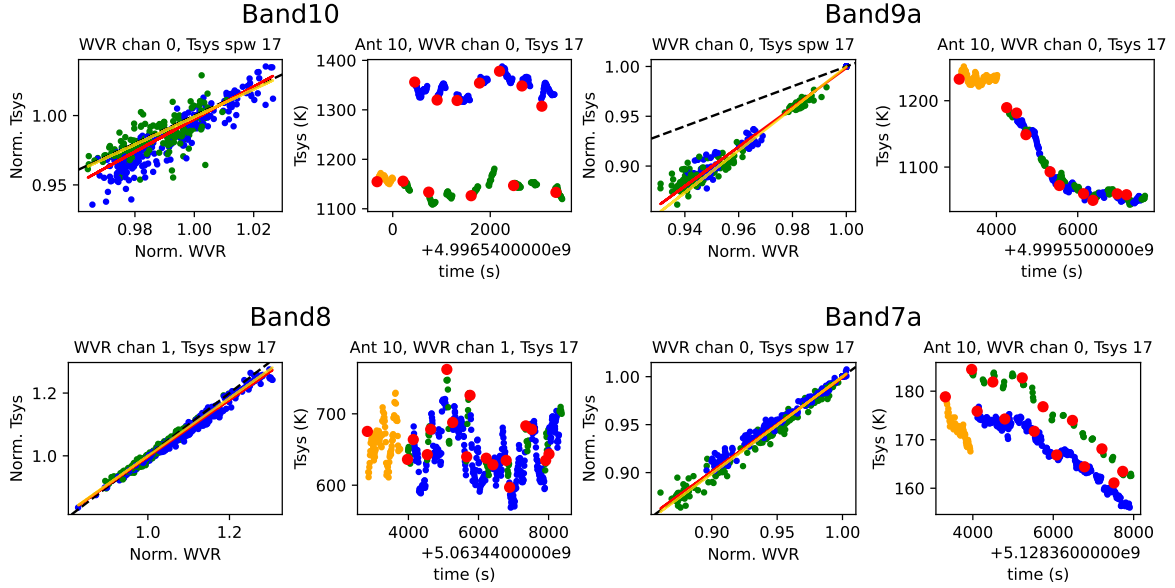


Figure A3. The summary of linear relation between T_{sys} and T_{WVR} and comparison between measured and extrapolated T_{sys} for the other 4 projects used in this paper with just one measurement set. See details in Fig. A1.

B. DIRTY IMAGES FOR THE REST OF TARGETS

In this section we show the dirty image of the rest of datasets made with extrapolated continuous T_{sys} using all Atm-cal scans and the aperture we used to measure the flux.

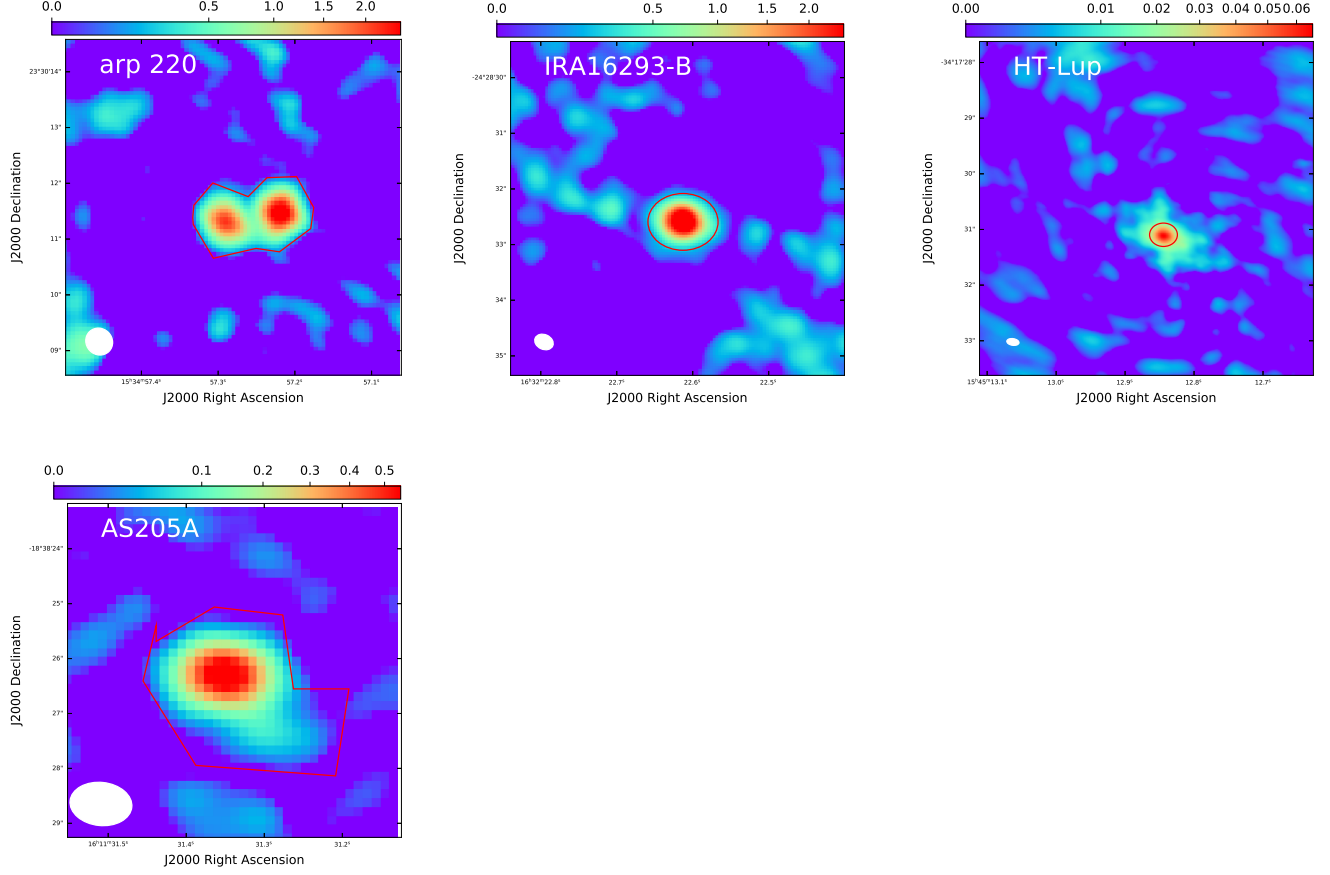


Figure B1. Dirty images for other sources made using alternative continuous T_{sys} table using all Atm-cal scans. For Band 7 project 2018.1.01210.S (AS205A), we show the image made from data uid://A002/Xda1250/X2387 (Band7b1). The red polygons are apertures used to measure the flux.