Tracking ALMA System Temperature with Water Vapor Data at High Frequency

Hao He1, William R. F. Dent2, and Christine Wilson1
1 McMaster University, 1280 Main St W, Hamilton, ON L8S 4L8, Canada; heh15@mcmaster.ca
2 Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile
Received 2022 May 11; accepted 2022 November 28; published 2022 December 19

Abstract
As the world-leading submillimeter telescope, the Atacama Large Millimeter/submillimeter Array observatory is now putting more focus on high-frequency observations at Band 7–10 (frequencies from 275 to 950 GHz). However, high-frequency observations often suffer from rapid variations in atmospheric opacity that directly affect the system temperature $T_{\text{sys}}$. Current observations perform discrete atmospheric calibrations (Atm-cals) every few minutes, with typically 10–20 occurring per hour for high frequency observation and each taking 30–40 s. In order to obtain more accurate flux measurements and reduce the number of atmospheric calibrations (Atm-cals), a new method to monitor $T_{\text{sys}}$ 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 $T_{\text{sys}}$ continuously. We find a tight linear correlation between $T_{\text{sys}}$ measured using the traditional method and $T_{\text{sys}}$ 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 $T_{\text{sys}}$ measurements to fit the linear relation and use this heuristic relationship to derive $T_{\text{sys}}$ every 10 s. Furthermore, we successfully reproduce the observed correlation using atmospheric transmission at microwave modeling and demonstrate the viability of a more general method to directly derive the $T_{\text{sys}}$ from the modeling. We apply the semi-continuous $T_{\text{sys}}$ 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 $T_{\text{sys}}$ methods give us consistent flux measurements with differences up to 5%. Furthermore, this method has significantly reduced the flux uncertainty due to $T_{\text{sys}}$ variability for one data set, which has large precipitable water vapor fluctuation, from 10% to 0.7%.

Unified Astronomy Thesaurus concepts: Astronomical instrumentation (799); Astronomical techniques (1684); Flux calibration (544); Radio telescopes (1360); Interferometry (808)

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 (Remjjan 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 10,000 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 & Guilloteau (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 method 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 timescales 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), ALMA 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, ALMA generally perform 10–20 ATM calibrations per hour. So the time overheads just due to ATM calibration can become quite significant—up to 15%–20% at the highest bands (9 and 10, at 602–950 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)

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

where

1. \( T_{rx} \) is receiver temperature
2. \( T_{sky} \) is sky temperature
3. \( T_{amb} \) is ambient temperature where spillover is assumed to be terminated
4. \( \eta_f \) is the forward efficiency. This is equal to the fraction of the antenna power pattern that is contained within the forward hemisphere and is currently assumed to be 0.95
5. \( e^{-\tau_{sky}} \) is the fractional transmission of the atmosphere, where \( \tau_{sky} \) 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, Equation (6)). \( T_{sky} \) and \( T_{amb} \) can be further expressed as

\[
\begin{align*}
T_{sky} &= T_{atm}(1 - e^{-\tau_{sky}}) \\
T_{amb} &\approx T_{atm}
\end{align*}
\]

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

\[
T_{sys} = \frac{1}{\eta_f e^{-\tau_{sky}}}[T_{rx} + T_{amb}(1 - \eta_f e^{-\tau_{sky}})] \\
\approx \frac{1}{e^{-\tau_{sky}}}(T_{rx} + T_{sky}) \quad \text{(where } \eta_f \sim 1).
\]

This equation suggests that the key parameters to measure \( T_{sys} \) are \( T_{rx} \) and \( T_{sky} \) (as \( \tau_{sky} \) can be derived from \( T_{sky} \) using Equation (2)).
In ALMA flux calibration, the intensity of the observed source is directly proportional to $T_{\text{sys}}$ by the following equation (e.g., Brogan 2018)

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

(4)

where $S_0$ and $S_{\text{final}}$ are the fluxes measured before and after the flux calibration. $\Gamma$ 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, $T_{\text{sys}}$ is associated with each antenna and $S_{\text{final}}$ is associated with each baseline.

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

$$\text{rms} \approx \frac{cT_{\text{sys}}}{\sqrt{\Delta \nu t_{\text{int}}}}$$

(5)

where $\Delta \nu$ is the frequency bandwidth, $t_{\text{int}}$ 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 $T_{\text{sys}}$ 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 $T_{\text{sys}}$ as

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

(6)

1.3. Traditional Method to Measure $T_{\text{sys}}$

As noted above, $T_{\text{sys}}$ is highly dependent on the sky opacity (Equation (2)). ALMA antennas use a two-load system for $T_{\text{sys}}$ 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 $T_{\text{sys}}$ 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, $T_{\text{sys}}$ 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 (Casalta et al. 2008). Typically this process takes 30–40 s, including antenna slew time and overheads. At frequencies below about 400 GHz, where the system temperatures are more stable (except in the 183 and 325 GHz water lines), an ATM-cal scan is made every 10–20 minutes. However, at higher frequencies, and wherever the opacity is large and more variable, every scan on the astronomical target will have an associated ATM-cal measurement (Remijan 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 $T_{\text{rx}}$. In this case, the measured power can be expressed as (Mangum 2002)

$$P_{\text{hot}} = K (T_{\text{rx}} + T_{\text{hot}})$$

$$P_{\text{amb}} = K (T_{\text{rx}} + T_{\text{amb}})$$

$$P_{\text{sky}} = K (T_{\text{rx}} + T_{\text{sky}})$$

(7)

where $K$ is the gain to convert the temperature to the measured power. $T_{\text{hot}}$ and $T_{\text{amb}}$ are generally about 350 and 290 K. Based on the equation above, we can express $T_{\text{rx}}$ and $T_{\text{sky}}$ as

$$T_{\text{rx}} = \frac{T_{\text{hot}}P_{\text{amb}} - T_{\text{amb}}P_{\text{hot}}}{P_{\text{hot}} - P_{\text{amb}}}, \quad Y_1 = \frac{P_{\text{hot}}}{P_{\text{amb}}}$$

$$T_{\text{sky}} = \frac{P_{\text{sky}}T_{\text{amb}} - (P_{\text{amb}} - P_{\text{sky}})T_{\text{rx}}}{P_{\text{amb}}} = \frac{Y_2 T_{\text{amb}}}{1 - Y_2}$$

(8)

where $Y_1 \equiv P_{\text{hot}}/P_{\text{amb}}$ and $Y_2 \equiv P_{\text{sky}}/P_{\text{amb}}$. Unlike the atmosphere, $T_{\text{rx}}$ is relatively constant throughout the observation. Measurements performed by ALMA show fluctuations of $T_{\text{rx}}$ are generally smaller than 1% during normal sidereal tracking. With the measurement of $T_{\text{sky}}$, we can further derive the optical depth based on Equation (2) and calculate the $T_{\text{sys}}$ based on Equation (3). In summary, the expressions for the key quantities to measure $T_{\text{sys}}$ are

$$T_{\text{rx}} = \frac{T_{\text{hot}} - Y_1 T_{\text{amb}}}{Y_1 - 1} \approx \text{const}$$

$$T_{\text{sky}} = \frac{Y_2 T_{\text{amb}} - (1 - Y_2)T_{\text{rx}}}{1 - Y_2}$$

$$T_{\text{sys}} \approx \frac{1}{e^{-\tau_{\text{sky}}}(T_{\text{rx}} + T_{\text{sky}})}.$$ 

(9)

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

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

As mentioned above, the current method takes extra time to obtain a spot measurement of $T_{\text{sys}}$ every few minutes. If we want to continuously track $T_{\text{sys}}$, 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 ($T_{\text{WVR}}$) at four frequencies (184.19, 185.25, 186.485 and 188.51 GHz respectively; Hills 2004) around the 183 GHz water line taken every 1.152 s (Remjian et al. 2019, Section A.6). By comparing $T_{\text{WVR}}$ with $T_{\text{amb}}$, we can calculate the precipitable water vapor (PWV), which is proportional to the atmospheric opacity caused by the water absorption. Since $T_{\text{sky}}$ at our observing frequency (Equation (9)) and $T_{\text{WVR}}$ are tracking sky temperatures at different frequencies, we would expect

$$T_{\text{sky}} = C \times T_{\text{WVR}} + \tau_{\text{sky,dry}}$$

(10)

where $\tau_{\text{sky}}$ is the sky opacity at the observed frequencies, $\tau_{\text{WVR}}$ is the optical depth at the WVR channel frequency, $\tau_{\text{dry}}$ and $T_{\text{sky,dry}}$ are the optical depth and sky temperature contribution for the dry component at the observing frequency, and C is a constant. A small dry contribution to $\tau_{\text{WVR}}$ and $T_{\text{WVR}}$ is not explicitly shown but does not change the form of the relationship. The overall sky opacity includes contributions from the wet component (H$_2$O lines from the troposphere which are relatively wide due to pressure broadening), and from the dry component (mostly due to lines of O$_2$ and 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 $T_{\text{sys}}$ is caused by the variation in $T_{\text{sky}}$, we would expect that $T_{\text{WVR}}$ is tracking $T_{\text{sys}}$. The major advantage of using the WVR data to trace $T_{\text{sys}}$ is that the radiometer is constantly monitoring the sky and internally calibrating itself. Therefore, we can extrapolate $T_{\text{sys}}$ throughout the entire observation based on the WVR data. Furthermore, since $T_{\text{WVR}}$ 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 $T_{\text{sys}}$ 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 $T_{\text{sys}}$ given the optical depth $\tau_{\text{sky}}$ using Equation (9) (see detailed discussion in Section 4). The total power signal received by each antenna is measured by a SQLD built into the ALMA signal path, whose data is also recorded in the data sets. Additionally, the autocorrelation data recorded in the measurement set should also give us the total signal received by each antenna. We would expect

$$T_{\text{rx}} + T_{\text{sky}} \propto P_{\text{AC}} \propto P_{\text{SQLD}}$$

(11)

where $P_{\text{AC}}$ and $P_{\text{SQLD}}$ are the power of auto-correlation data and SQLD data read from the measurement set, respectively. If $\tau_{\text{sky}}$ is small, we would expect direct proportionality between $T_{\text{sys}}$ and the total power received which could help us derive continuous $T_{\text{sys}}$. In addition, both AC and SQLD data cover the same frequency ranges 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 constant $C$ in Equation (10) and the explicit dry contributions).

In Figure 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 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 $T_{\text{sys}}$.  

### Table 1

| Advantage | Disadvantage |
|-----------|--------------|
| 1. The data is continuously calibrated to measure $T_{\text{WVR}}$ | 1. Has different frequency coverage as the science target |
| 2. Does not suffer from internal electronic gain drift | 2. The data is not calibrated. |

**Notes:**
- 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).

### References
- Hills et al. 2001
- Payne et al. 2001
- WVR (1)
- AC and SQLD (2)

| Advantage | Disadvantage |
|-----------|--------------|
| 1. Directly proportional to $T_{\text{sys}}$ when $\tau_{\text{sky}}$ is small | 1. The data is not calibrated. |
| 2. Has the same frequency coverage as the science target | 2. Suffers from the electronic gain drift (b) or gain variations |
| 3. For AC, no linearity correction in FDM mode. | |
1. 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 $T_{sys}$ versus $T_{WVR}$ 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 measurement sets from several projects in Bands 7, 8, 9 and 10 (Sekimoto et al. 2008;...
We also include two projects with multiple measurement sets from Band 7 and 9. The summary of the data we use is given in Table 2.

2. WVR Data to Track T\textsubscript{sys}

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

2.1. T\textsubscript{sys} versus T\textsubscript{WVR}

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

We first plot T\textsubscript{sys} versus T\textsubscript{WVR} from all antennas for each data set. One example of T\textsubscript{sys} versus T\textsubscript{WVR} is shown in the left panel of Figure 2. For this case, we select T\textsubscript{WVR} from WVR channel 1. We will discuss in Section 2.2 how the selection of different WVR channels affects the relation between T\textsubscript{WVR} and T\textsubscript{sys}. As we can see, there is a significant correlation between T\textsubscript{sys} and T\textsubscript{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\textsubscript{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\textsubscript{sys}, we normalize T\textsubscript{sys} and T\textsubscript{WVR} by the first measurement for each observing target (bandpass, phase-cal and science) of each antenna as

\[
\hat{T}_{\text{sys,source}}(t) = \frac{T_{\text{sys,obs}}(t)}{T_{\text{sys,obs}}(1^\circ)} \\
\hat{T}_{\text{WVR,obs}}(t) = \frac{T_{\text{WVR,obs}}(t)}{T_{\text{WVR,obs}}(1^\circ)}
\]

(12)

where \(\hat{T}_{\text{sys}}\) and \(\hat{T}_{\text{WVR}}\) are the normalized values of T\textsubscript{sys} and T\textsubscript{WVR}, the subscript “obs” is the generalized term for each type of observing target (bandpass, phase and science) and 1st in the bracket means the value when the first T\textsubscript{sys} for each observing target is measured.

The correlation between the \(\hat{T}_{\text{sys}}\) and \(\hat{T}_{\text{WVR}}\) is shown in the right panel of Figure 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\textsubscript{WVR} can be used to track the T\textsubscript{sys} if the slope and intercept can be determined for each spw or frequency. As described by Equation (10), the relation is expected to be frequency dependent, and it further

![Figure 2.](image-url)
Table 3

| 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% |

Note. 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 data set Band8. The PWV_{los} value for data set Band8 is \sim 1.05 mm.

differs from 1-to-1 due to the other contributions to $T_{sys}$ apart from $T_{sky}$ (e.g., Equations (1) and (3)). In Section 3, we explore the relationship using an 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. The exact equation can be expressed as

$$T_{sys}(t) = T_{sys}(1^m) \cdot \hat{T}_{sys, \text{obs}}(t) = T_{sys}(1^m) \cdot [m\hat{T}_{WVR, \text{obs}}(t) + b]$$

(13)

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

Examples of fitting and $T_{sys}$ extrapolation for other data sets are shown in Appendix A. We can see for all data sets that $\hat{T}_{sys}$ 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 four filter channels at frequencies (184.19, 185.25, 186.485 and 188.51 GHz respectively) close to the 183 GHz H$_2$O 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 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). We can see that for these observing conditions (PWV_{los} of 1.05 mm), the normalized $\hat{T}_{WVR}$ from different WVR channels have a similarly tight correlation with normalized $\hat{T}_{sys}$ with scatter of \sim 1%. For data set 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 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 $T_{WVR}$ (Equation (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

$$w = \frac{1}{T_{WVR} - 275}$$

(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 $\bar{T}_{WVR}$ 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 data sets 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 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 four Atm-
cal scans for the fitting with two from phase target and two from the science target. The four Atm-cal scans give us two independent $T_{\text{sys}}$ values if we normalize the $T_{\text{sys}}$ from phase and science target independently. For the Atm-cal scan selection, we select two Atm-cal scans at the start and two Atm-cal scans in the middle. One example of the fits using four Atm-cal scans is shown in Figure 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 four Atm-cal scans. The scatter of all the data points around the new relation has almost the same scatter of 1%.

Fits using four 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 the data sets except Band9b1, which we will in Section 2.5. In Figure 5, we plot the relative scatter around the fit versus the maximal difference divided by mean value of the $T_{\text{sys}}$ for every $T_{\text{sys}}$ spw of all the data sets. As we can see, the scatter using both fitting methods is generally below 3%. Fitting with just four 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 $T_{\text{sys}}$ measurements when using $T_{\text{WVR}}$ to track $T_{\text{sys}}$.

2.4. Normalize Only to the Science Target

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

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

$$\hat{T}_{\text{sys,obs}}(t) = \frac{T_{\text{sys,obs}}(t)}{T_{\text{sys,sci}}(\text{F1})}$$

$$\hat{T}_{\text{WVR,obs}}(t) = \frac{T_{\text{WVR,obs}}(t)}{T_{\text{WVR,sci}}(\text{F1})}$$

where $\hat{T}_{\text{sys,obs}}(t)$ is the normalized $T_{\text{sys}}$ averaged along the spectral axis and $\hat{T}_{\text{WVR,obs}}(t)$ is the normalized $T_{\text{WVR}}$. An example of $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ using the new normalization method are shown in Figure 6. 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 the trends of the bandpass target and phase-cal/science targets.
targets. 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 $T_{\text{sys}}$ measurements as we can derive it from $T_{\text{sys}}$ measurements for only the science target.

Note that for early ALMA cycles, the $T_{\text{sys}}$ measurements are purely done for the phase-cal target. The $T_{\text{sys}}$ for the science target is then assumed to be the same as the $T_{\text{sys}}$ 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 $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ linear relation.

Therefore, we can better extrapolate the science $T_{\text{sys}}$ from the phase-cal $T_{\text{sys}}$ using the fitted linear relation.

2.5. Test with Significant Opacity and Large $T_{\text{sys}}$ Variation

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

$$T_{\text{sys}} \approx T_{\text{rx}} + \tau_{\text{sky}} = T_{\text{rx}} + C \times T_{\text{WVR}} + T_{\text{sky, dry}}$$  \hspace{1cm} (16)

where $C$ is a constant. However, at higher frequencies such as Band 9 and 10, $\tau_{\text{sky}}$ are quite high and we can no longer ignore the opacity term in $T_{\text{sys}}$ (e.g., Equation (3)). In this case, the increase in $T_{\text{sys}}$ is dominated by the increase in $\tau_{\text{sky}}$ in the exponential form.

We test if the $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ linear relation still holds on data set Band9b, which has large $\tau_{\text{sky}}$ and $T_{\text{sys}}$ range (~50%). In Figure 7, we show $T_{\text{sys}}$ fitting and extrapolation using all Atm-cal scans or just four 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 four Atm-cal scans. It seems the slope becomes steeper due to data points with higher $T_{\text{sys}}$ values, which are not included if we use just four

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{The scatter of data points around the $T_{\text{sys}}$ vs. $T_{\text{WVR}}$ fits with all the antennas vs. the maximal difference in $\hat{T}_{\text{sys}}$ values for each $T_{\text{sys}}$ spw of each data set. The red and blue points are from fitting with all Atm-cal scans or just four Atm-cal scans respectively. The histogram at the 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 four Atm-cal scans.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{$T_{\text{sys}}$ vs. $T_{\text{WVR}}$ normalized to the first science Atm-cal scan for the data set Band8. The black dashed line is the 1-to-1 relation. The brown solid line is the linear fitting to the data excluding the bandpass data. The red solid line is the original fitting relation to the data normalized to each type of observing target. We can see the two fitting relation are almost the same. The bandpass data for spw 0 has a significant offset from the fitted relation, which is due to the elevation difference between bandpass target and phase-cal/science targets (see discussion in Section 3.2).}
\end{figure}
scans. This is also reflected in the extrapolation plot at the right side of Figure 7, as the predicted $T_{\text{sys}}$ is lower than the measured $T_{\text{sys}}$ for higher $T_{\text{sys}}$ values.

This is consistent with our expectation that the slope of $T_{\text{sys}}$ versus $T_{\text{WVR}}$ relation is increasing. We would expect the curving-up feature also happens to other data sets but we do not have large enough $T_{\text{sys}}$ ranges in the other data sets we analyzed.

### 3. Atmospheric Transmission at Microwave (ATM) Modeling

In the previous sections, we fit $T_{\text{sys}}$ versus $T_{\text{WVR}}$ heuristically and use the fitted relation to extrapolate the $T_{\text{sys}}$ continuously. We find in most cases the correlation between $T_{\text{sys}}$ and $T_{\text{WVR}}$ is linear. However, the exact slopes and intercepts of the correlations vary across different frequencies. Although we can use less than four Atm-cal scans to fit the relation for each spw in each data set, there might be cases when our selected Atm-cal scans have similar $T_{\text{sys}}$ values, and hence might give us inaccurate fitting relation among small $T_{\text{sys}}$ ranges. Furthermore, as described in Section 2.5, the linear approximation becomes insufficient when the opacity becomes significant or $T_{\text{sys}}$ variations become large. A more robust method is to use an atmospheric opacity modeling code and an estimate of the $T_{\text{a}}$ and other static contributions to predict the $T_{\text{sys}}$ versus $T_{\text{WVR}}$ relation at the relevant frequencies.

Figure 7. (Left) $T_{\text{sys}}$ (spw 17) vs. $T_{\text{WVR}}$ (channel 0) for data set Band9b1. The upper left panel shows $T_{\text{sys}}$ vs. $T_{\text{WVR}}$ 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 four Atm-cal scans. (Right) The measured and extrapolated $T_{\text{sys}}$ for different observing targets as a function of time. Red points are measured $T_{\text{sys}}$ values. $T_{\text{sys}}$ in upper panel is extrapolated based on fits with all Atm-cal scans while $T_{\text{sys}}$ in lower panel is extrapolated based on fits using only selected four Atm-cal scans. We can see in this case we will underestimate the $T_{\text{sys}}$ value if we just use part of Atm-cal scans to fit the correlation.
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 $T_{\text{sys}}$ versus $T_{\text{WVR}}$ relation for various data sets and compare the results with our heuristic method.

### 3.1. Modeling $T_{\text{sys}}$ Spectrum

We note that in previous sections when we explore the correlation between $T_{\text{sys}}$ and $T_{\text{WVR}}$, we average $T_{\text{sys}}$ for each spectral window along its spectral axis. Therefore, when we extrapolate the continuous $T_{\text{sys}}$, we are assuming that the $T_{\text{sys}}$ spectrum does not vary significantly. In this subsection, we test this assumption with ATM modeling on the two representative data sets we have, Band8 and Band7a. We use the version of ATM included in CASA (McMullin et al. 2007; Emonts et al. 2020, The CASA Team et al. 2022), accessed via a helper function plotAtmosphere to generate $T_{\text{sys}}$ and $T_{\text{WVR}}$ spectra for the 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 data set, we set the PWV and elevation values to be the

---

Figure 8. The ATM modeling opacity $\tau_{\text{sky}}$ and $T_{\text{sys}}$ spectrum for data set Band8 (upper) and Band7a (lower). For each row, the left panel shows the $\tau_{\text{sky}}$ spectrum for the wet and dry component. We can see that the dry component is significant for data set Band8 but the water component is dominant for Band7a. The right panel shows the modeled $T_{\text{sys}}$ spectrum for 3 sets 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 $T_{\text{sys}}$ faster since it increase $\tau_{\text{sky}}$ from both wet and dry components.
same as the value of the first Atm-cal scan for the science target as our start point. Note that \texttt{aU.plotAtmosphere} only gives \( T_{\text{sky}} \) and \( \tau_{\text{sky}} \). Therefore, we calculate the \( T_{\text{sys}} \) spectrum from the modeled \( T_{\text{sky}} \) and \( \tau_{\text{sky}} \) using Equation (1) by assuming \( T_{\text{rx}} \) to be 100 K.

We show our modeled results in Figure 8. As we can see, spw 17 for data set Band8 has a significant dry opacity contribution as it sits at one of the O2 lines. On the contrary, data set Band7a is dominated by the wet component. We then increase the PWV and airmass (1/ \( \sin m_{\text{cl}}, m_{\text{cl}} \) is the elevation) by a factor of 1.2 to see if they have different effects on increasing the \( T_{\text{sys}} \) spectrum. For both data set Band8 and Band7a, we can see increasing airmass is more effective in increasing the overall values of the \( T_{\text{sys}} \) spectrum. This is what we expect since increasing airmass will increase both wet and dry opacity while increasing PWV only increases the wet opacity. If we compare data set 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.

The change of \( T_{\text{sys}} \) spectrum shape is generally small by increasing the PWV or airmass by 20%. However, a small change is noticeable when increasing airmass for spw 17 in data set Band8, due primarily to the significant \( \tau_{\text{sky}} \) from the dry component and its large variation across the spw. Not tracking such small \( T_{\text{sys}} \) spectrum shape changes will have negligible impact on continuum observations, and for spectral lines the error in the extrapolation based on \( T_{\text{WVR}} \) or PWV will be within \( \sim \pm 3\% \) (see Section 5.4 for more discussion). A future improvement might be to correct the data spectrally rather than using a single channel-averaged value per timestamp.

### 3.2. Reproduce the Observed \( \hat{T}_{\text{sys}} \) versus \( \hat{T}_{\text{WVR}} \) Correlation

In this section we will try to reproduce the observed \( \hat{T}_{\text{sys}} \) versus \( \hat{T}_{\text{WVR}} \) correlation in various cases. According to Equation (3), \( T_{\text{sys}} \) are determined by \( T_{\text{rx}} \) and \( T_{\text{sky}} \). \( T_{\text{sky}} \) can be further determined by the measurement of PWV and elevation through ATM modeling. Therefore, we can generate modeled \( T_{\text{sys}} \) spectrum with given \( T_{\text{rx}} \), PWV and elevation at a certain frequency range. To reproduce the \( T_{\text{sys}} \) measured in observation, we set the frequency range to be the same as the \( T_{\text{sys}} \) spectral window we want to model, with total bandwidth of \( \sim 2 \) GHz. The modeled \( T_{\text{sys}} \) spectrum is then averaged to a single \( T_{\text{sys}} \) value as we did with the observations. We also use a similar method to generate \( T_{\text{WVR}} \) at different WVR channels with given PWV and elevations (\( T_{\text{WVR}} \) is just \( T_{\text{sky}} \) at the WVR channel frequencies).

For most of the ALMA data, \( T_{\text{rx}} \) stays relatively constant throughout the observations. However, different antennas generally have different \( T_{\text{rx}} \) values, which might give us slightly different shapes of correlation. Therefore, we first test how varying \( T_{\text{rx}} \) could affect the shape of the \( T_{\text{sys}} \) versus \( \hat{T}_{\text{WVR}} \) correlation for data set Band8 spw 17 (Figure 9). The \( T_{\text{sys}} \) and \( T_{\text{WVR}} \) are generated with varying \( T_{\text{rx}} \) and PWV values but with fixed elevation of 53 deg. We note that data set Band8 has relatively constant elevations (see Table 2) for the science target throughout the entire observation, hence the \( T_{\text{sys}} \) variations across time are mostly due the the change in PWV values. After generating the modeled \( T_{\text{sys}} \) and \( T_{\text{WVR}} \), 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 \( T_{\text{rx}} \) value. As we can see in Figure 9, the ATM modeling shows a slight nonlinear curvature for a high range of \( T_{\text{sys}} \), which depends slightly on the assumed \( T_{\text{rx}} \) and the elevation. However, the modeling curve is generally within the range of the data scatter. Varying \( T_{\text{rx}} \) also gives a similar correlation within data scatter of \( \sim 1\% \). Therefore, \( T_{\text{rx}} \) values do not seem to affect the \( T_{\text{sys}} \) versus \( \hat{T}_{\text{WVR}} \) correlation we get.

We then explore how varying PWV or elevation can affect the \( T_{\text{sys}} \) versus \( \hat{T}_{\text{WVR}} \) correlation. As we have mentioned in Section 3.1, varying PWV or elevation might have different effects on changing \( T_{\text{sys}} \) values depending on how significant the dry opacities are at given frequency. Data Set Band8 and Band7a represent two cases where one has significant dry opacity contribution while the other is dominated by the wet opacity. Therefore, it is natural for us to explore what drives the \( T_{\text{sys}} \) versus \( \hat{T}_{\text{WVR}} \) correlation in these two cases. Figure 10 shows the comparison between modeling and observation for

![Figure 9](image-url)
these two data sets where red and magenta lines represent changing PWV and elevation respectively. In this comparison, the measured $T_{\text{sys}}$ and $\hat{T}_{\text{WVR}}$ are normalized to the values in the first science Atm-cal scan instead of the first Atm-cal scan for each target (see Section 2.4 for more description). The modeled $T_{\text{sys}}$ and $T_{\text{WVR}}$ are then normalized to the values when PWV and elevation are equal to those of the first science Atm-cal scan. We can see in both cases the ATM modeling successfully reproduces the observed $\tilde{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ correlation. For data set 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 $\tilde{T}_{\text{sys}}$ and $\hat{T}_{\text{WVR}}$. This is due to the fact that changing elevations (and therefore airmass) will be more effective to change $T_{\text{sys}}$ values when the dry opacity contribution is significant (see discussion in Section 3.1). $T_{\text{sys}}$ for the 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 $T_{\text{sys}}$ measurement for each antennae has the same elevation but might point toward 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 a similar $\tilde{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ correlation by varying PWV or elevation for data set Band7. This is probably due to the fact that the wet component is dominant at this spectral window, hence changing PWV or elevation achieves a similar effect. We can also see in this case the $\tilde{T}_{\text{sys}}$ for the bandpass target 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, data set Band9b1 shows a nonlinear $\tilde{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ correlation as the $T_{\text{sys}}$ variation becomes significantly large ($\sim 40\%$). Therefore, it is also worth testing if we can reproduce the curving feature for this data set. We show the comparison in Figure 11 using the same method described in the previous paragraph. As we can see, this data set also has significant contribution from the dry opacity and hence shows different slopes when varying PWV or the elevation values. The observed $T_{\text{sys}}$ generally agrees well with the ATM modeling relation with fixed PWV values. This suggests the $T_{\text{sys}}$ variation shown for this data set is mainly due to the elevation change. This is consistent with what we see in Figure 7 as $T_{\text{sys}}$ 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 values, which is probably due to the larger elevation of the bandpass target. We also see some second-order scatter around the fixed PWV line for higher $T_{\text{sys}}$ values, which might be due to the intrinsic scatter of PWV values during the observation. However, to first order we can just measure $T_{\text{sys}}$ once and
the H$_2$O line. For the dry opacity, we generally see a lot of smooth line features, which indicate the presence of the H$_2$O line. For the dry opacity, we generally see a lot of smooth line features, which indicate the presence of

**Figure 11.** Similar plot as Figure 10 but for data set Band9b1. The dry component is also significant for this data set. We can see that Band9b1 has a relatively constant PWV while the $T_{sys}$ and $T_{WVR}$ variation is mainly due to changing elevations.

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 Improvement

In the previous section, we have successfully reproduced the $T_{sys}$ versus $T_{WVR}$ correlation for three data sets with ATM modeling. However, we also find that the correlation is not driven by a single parameter. For both data set 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 data set Band7a where wet opacity is dominant, we can expect a tight $T_{sys}$ versus $T_{WVR}$ 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 to apply our current heuristic method to data sets observed at frequencies where wet opacity is dominant. In Figure 12, we model the wet and dry opacity spectrum covering the entire ALMA high frequency bands from Band 7 to Band 10. We assume PWV of 0.5 mm, which is a typical value for high-frequency ALMA observations, and elevation of 50 deg. For the wet opacity, we see several smooth line features, which indicate the presence of the H$_2$O line. For the dry opacity, we generally see a lot of Ozone lines as narrow spikes. These Ozone lines are generally much narrower than the typical bandwidth of a spw of 2 GHz and hence have a 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 data sets, Band8, sits right at the wings of one O$_2$ line and thus has significant dry opacity contribution. For all of our data sets used in this paper, only data sets Band8 and Band9b show significant dry opacity contribution, which is consistent with our expectation as all of the other data sets in this paper exhibit a tight linear correlation (Figures A1–A3).

However, we can expect that in lots of cases, the $T_{sys}$ versus $T_{WVR}$ 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 reproduce the $T_{sys}$ measured in observations. 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 $T_{sys}$ values, we can then reproduce the continuous $T_{sys}$ throughout the observation.

4. AC and 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. Instead, AC is proportional to $e^{-\tau} T_{sys}$ (Figure 14). Since AC and SQLD data are equivalent to one another (see Figure 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 $T_{sys}$. Figure 13 shows the comparison between $T_{sys}$ versus AC and $T_{sys}$ versus WVR correlation for data set Band8 and Band7a. We can see that the AC data also has a tighter correlation with $T_{sys}$ for data set Band7. However, the correlation does not work well for data set Band8 with a large scatter. 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 Figure 15. The AC data is directly proportional to the total signal received by the antenna, which is mainly comprised of emission from the sky ($\eta_1 T_{sky}$), the receiver itself ($T_{rx}$), and other fixed losses terminating at ground ($\left(1-\eta_1\right) T_{amb}$). However, based on Equation (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,
\[ P_{AC} \propto \eta_I e^{-\tau_{sec}} T_{sys} \]
\[ \approx T_{rx} + \eta_I T_{sky} + (1 - \eta_I) \times T_{amb}. \]  
(17)

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 Figure 14. As we can see, the normalized attenuated \( T_{sys} \) follows the 1-to-1 relation with normalized AC as suggested by Equation (17). If we rearrange Equation (17), it becomes

\[ T_{sys} \propto \frac{P_{AC}}{e^{-\tau_{sky}}} \approx \frac{P_{AC}}{1 - T_{sky}/T_{amb}}. \]  
(18)

Therefore, even though we have the AC data, we still need a method to continuously determine \( T_{sky} \) or \( \tau_{sky} \) to obtain \( T_{sys} \). A possible work-around is to use AC data to track \( T_{sky} \) first by combining Equations (3), (2) and (18) 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 Figure 1.

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 data set with the original \( T_{sys} \) table, the new continuous \( T_{sys} \) table extrapolated using all Atm-cal scans and that using just four Atm-cal scans with CASA 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\textsubscript{sys} Table

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

$$
\begin{bmatrix}
T_{\text{sys}},1, & 1, & 1, & 19 \\
T_{\text{sys}},, & st, & st, & \text{fit} \end{bmatrix}
$$

Figure 13. The correlation between the $T_{\text{sys}}$ and matched $\hat{T}_{\text{WVR}}$ and normalized $P_{\text{AC}}$ for all the antennas. Both WVR and auto-correlation data is averaged over 10 s.

5.1. Creation of $T_{\text{sys}}$ Table

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

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

$$
G(t) = \sqrt{1 - \left(\frac{T_{\text{sys}}(t)}{T_{\text{sys}}(1^\text{st})}\right)_{\text{fit}}} 
$$

(19)
where $T_{\text{sys}}(t, \nu)$ is the recorded $T_{\text{sys}}$ spectrum as a function of time $t$ and frequency $\nu$, and $G$ is the derived gain as a function of $t$, $T_{\text{sys,obs}}(1^\text{st}, \nu)$ is the first $T_{\text{sys}}$ spectrum measured for each type of observation of given antenna and spectral window and $\left[\frac{T_{\text{sys}}}{T_{\text{sys}}(P)}\right]_{\text{fit}}$ is the extrapolated normalized $T_{\text{sys}}$ from the fitting.

We note that $G$ is not directly equal to the $\hat{T}_{\text{sys}}$ values. This is due to the different methods that CASA uses to handle $T_{\text{sys}}$ and gain table. For each baseline, the correlated amplitude is

$$S(i, j) \propto \sqrt{T_{\text{sys}}(i)T_{\text{sys}}(j)} \propto \frac{1}{G(i)G(j)}.$$ \hfill (20)

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

In Section 2.3, we tested using only four $T_{\text{sys}}$ measurements to fit the linear relation between $T_{\text{sys}}$ and $T_{\text{WVR}}$. We saw that the difference between this method and using all $T_{\text{sys}}$ 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 Equation (19) to create the alternative $T_{\text{sys}}$ table with the fitting relation derived from four Atm-cal scans.

### 5.2. Calibrating and Imaging the Data

After we create the $T_{\text{sys}}$ table, we then apply the continuous $T_{\text{sys}}$ in calibration. The calibration script we use is generated from the command `es.generateReducScript()` (Petry et al. 2014). We then modify the script to use the alternative $T_{\text{sys}}$ and amplitude gain tables created (see details in [https://github.com/heh15/ALMA_intern_Tsys.git](https://github.com/heh15/ALMA_intern_Tsys.git)). We also run the original calibration script to calibrate the data with the original discrete $T_{\text{sys}}$ 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 data sets. For projects with multiple data sets, we directly compare the measured fluxes from different data sets 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 three images. The top panel in Figure 16 shows an example of images made with three different methods for data set Band8 and Band7a we have for all antennas of one spectral window. As we can see, the attenuated $T_{\text{sys}}$ follows the 1-to-1 correlation with AC data, which proves Equation (17) to be right.

---

![Figure 14](image-url)  
**Figure 14.** The correlation between the normalized $T_{\text{sys}}$ (blue) and attenuated $T_{\text{sys}}$ (green) and normalized AC data for the data set Band8 and Band7a we have for all antennas of one spectral window. As we can see, the attenuated $T_{\text{sys}}$ follows the 1-to-1 correlation with AC data, which proves Equation (17) to be right.

---
The beam sizes for the three different images are also significantly different from each other. For data set 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. Since the change in $T_{\text{sys}}$ 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 Figures 16, 17 and B1. Note that as well as changes in $T_{\text{sys}}$, 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 Tables 4–6. However, we assume this effect is the same in all reductions, independent of the $T_{\text{sys}}$ calibration method.

5.3. Flux Comparison

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

\[
\text{Meas. Err} = \text{rms} \sqrt{N_{\text{beam}}}
\]

where the rms is the measured noise of the image and $N_{\text{beam}}$ is the number of beams across the aperture used to measure the flux. On the other hand, our alternative method to measure $T_{\text{sys}}$ 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 a single data set, 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 data sets, the calibration error...
is calculated as the standard deviation of fluxes of the different data sets. Both measurement and calibration errors are recorded in Tables 4–6.

To further compare how our methods work for data sets from different frequency bands, we normalize the flux values to the flux value using the original discrete $T_{\text{sys}}$ calibration method.
| Data Label | Target     | Scans Used | Flux (Jy) | Meas. Err. (Jy) | Beam (") |
|------------|------------|------------|-----------|-----------------|-----------|
|            |            |            | Tsys_orig | Tsys_extrap     | Tsys_partextrap |
| Band7b     | Arp 220    | all        | 7.169     | 7.146           | 7.187     | 0.39 | 0.5 × 0.47 |
|            |            | 1st half   | 7.1638    | 7.135           | 7.171     | 0.42 | 0.52 × 0.45 |
|            |            | 2nd half   | 7.1077    | 7.09            | 7.133     | 0.4  | 0.48 × 0.48 |
|            |            | MAX. DIFF. | 0.06      | 0.056           | 0.054     |      |            |
| Band9a     | IRAS16293-B| all        | 10.43     | 10.36           | 10.324    | 0.34 | 0.34 × 0.27 |
|            |            | 1st half   | 10.4      | 10.32           | 10.32     | 0.36 | 0.35 × 0.27 |
|            |            | 2nd half   | 10.47     | 10.4            | 10.34     | 0.38 | 0.34 × 0.26 |
|            |            | MAX. DIFF. | 0.07      | 0.04            | 0.02      |      |            |
| Band8      | SPT0311-58 | all        | 0.0425    | 0.0429          | 0.0429    | 0.0015 | 0.35 × 0.3 |
|            |            | 1st half   | 0.0453    | 0.0432          | 0.0433    | 0.0019 | 0.36 × 0.29 |
|            |            | 2nd half   | 0.0407    | 0.0428          | 0.0429    | 0.0017 | 0.35 × 0.3 |
|            |            | MAX. DIFF. | 0.0046    | 0.0003          | 0.0003    |      |            |
| Band7a     | HT-Lup     | all        | 0.173     | 0.178           | 0.178     | 0.0038 | 0.22 × 0.12 |
|            |            | 1st half   | 0.169     | 0.175           | 0.175     | 0.0044 | 0.22 × 0.12 |
|            |            | 2nd half   | 0.175     | 0.18            | 0.18      | 0.0045 | 0.22 × 0.12 |
|            |            | MAX. DIFF. | 0.006     | 0.005           | 0.005     |      |            |

**Table 4**

Flux Measured for Project with only One Data Set

| Data Label | Target     | Scans Used | Flux (Jy) | Meas. Err. (Jy) | Beam (") |
|------------|------------|------------|-----------|-----------------|-----------|
|            |            |            | Tsys_orig | Tsys_extrap     | Tsys_partextrap |
| Band7b     | Arp 220    | all        | 0.7699    | 0.759           | 0.7594     | 0.031  | 1.13 × 0.79 |
|            |            | 1st half   | 0.713     | 0.7146          | 0.7129     | 0.029  | 0.96 × 0.77 |
|            |            | 2nd half   | 1.001     | 1.005           | 1.003      | 0.043  | 0.68 × 0.53 |
|            |            | MAX. DIFF. | 0.747     | 0.7387          | 0.7359     | 0.028  | 1.03 × 0.78 |
| Band7b3*   |            | all        | 0.556     | 0.58            | 0.58       | 0.029  | 1.31 × 0.72 |
| Band7b4    |            | all        | 0.749     | 0.7534          | 0.7533     | 0.03   | 0.91 × 0.08 |
| Band7b5*   |            | all        | 0.7355    | 0.7347          | 0.7348     | 0.029  | 1.01 × 0.81 |
| Band7b6    |            | all        | 0.768     | 0.737           | 0.7372     | 0.029  | 1.24 × 0.79 |
| AVG. (a)   |            | all        | 0.7471    | 0.7394          | 0.7389     |      |            |
| STD. (b)   |            | all        | 0.0194    | 0.0143          | 0.0149     |      |            |

**Table 5**

Flux Measured for Data for Band7b Project (A5205A)

**Note.** Columns: (1) The label for each data set (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 Tsys table. The Tsys is extrapolated from all Atm-cal scans. (6) The images made with modified script but Tsys is extrapolated from 4 Atm-cal scans. (7) The measured flux errors for the images made with original Tsys table (8) The beam size of images using original Tsys table. Rows: (a) The average value for each column. (b) The maximal differences for fluxes at each column.

---

...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 Figure 18. We can see that fluxes using the original Tsys table do not differ significantly from the fluxes using alternative Tsys and gain tables, with maximal differences smaller than 5%. We also demonstrate that the extrapolated Tsys using four 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 $T_{\text{sys}}$ measurements by using WVR-tracked $T_{\text{sys}}$. Furthermore, it seems for most of the data sets, our new methods give better flux consistency, especially for data set Band8 which brings down the flux calibration uncertainty contribution due to $T_{\text{sys}}$ variability from $\sim$10% to 0.7%. As we have shown in Section 3.2, the $T_{\text{sys}}$ variation for this data set is mainly driven by the PWV variation at short timescale. In this case, the discrete $T_{\text{sys}}$ measurements poorly sampled the fluctuations of the real $T_{\text{sys}}$ (Figure 3). Our new method instead catches the variation in $T_{\text{sys}}$ between the discrete ATM calibrations, and thus keeps the flux consistent.

This method also works for data set Band7b with multiple data sets for which 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 data set 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, the relatively high phase noise and maybe the time variability of the flux calibrator in these data sets. This is also a tricky data set 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 the $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ relation.

---

**Table 6**

| Data Set Label | Flux (Jy) | Meas. Err. (Jy) | Beam (") |
|----------------|----------|----------------|-----------|
|                | $T_{\text{sys}}$, orig | $T_{\text{sys}}$, extrap | $T_{\text{sys}}$, partextrap | (5) | (6) |
| Band9b1        | 0.85  | 0.875 | 0.849 | 0.018 | 0.086 $\times$ 0.065 |
| Band9b2        | 0.96  | 1    | 0.983 | 0.021 | 0.104 $\times$ 0.077 |
| Band9b3        | 1.27  | 1.35 | 1.34 | 0.016 | 0.131 $\times$ 0.085 |
| AVG.           | 1.027 | 1.075 | 1.057 |      |          |
| STD.           | 0.178 | 0.201 | 0.207 |      |          |

Note: Columns: (1) The label for each data set (see Table 2). (2) Flux of the data set using the original $T_{\text{sys}}$ table. (3) Flux of the data set using alternative continuous $T_{\text{sys}}$ table with the linear relation fitted using all Atm-cal scans. (4) Flux of the data set using continuous $T_{\text{sys}}$ 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.
5.4. Additional Considerations for the Continuous $T_{\text{sys}}$
Method

For some targets, the source brightness temperature in single-dish measurements can be significant compared with $T_{\text{sys}}$. 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 (2 GHz) used to measure $T_{\text{sys}}$, and the continuous $T_{\text{sys}}$ method uses a spectrally averaged broad-band $T_{\text{sys}}$, then the effect of bright lines in such cases will be negligible. But for very bright continuum sources such as planets, the spectrally averaged $T_{\text{sys}}$ 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 $T_{\text{sys}}$ using off-source data along with the normalization of the visibilities using the autocorrelation, introduces a calibration error for bright sources.\(^4\) The planned change is to measure $T_{\text{sys}}$ on-source. Again, this should not significantly affect the continuous $T_{\text{sys}}$ method, for reasons given above. However, we need to note that this technique cannot be applied to solar observing because $T_{\text{WVR}}$ from all WVR channels will be heavily saturated.

For spectral lines, an assumption is made that $T_{\text{sys}}$ is mainly affected by PWV, and the correlation of $T_{\text{sys}}$ with PWV uses $T_{\text{sys}}$ 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 Figure 8), 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 428.8 GHz in Figure 8, the error in the correction of $T_{\text{sys}}$ based on the PWV would be $\sim\pm 3\%$. A future improvement might be to 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 $T_{\text{sys}}$ method could be to correct for the increase in $T_{\text{sys}}$ 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 $T_{\text{sys}}$ 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 data streams available from WVR monitoring to track the atmospheric opacity and hence $T_{\text{sys}}$ 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.

1. There is a tight linear correlation between normalized $T_{\text{sys}}$ and $T_{\text{WVR}}$, with typical scatter of $\sim 1\%$. For the worst case of Band 9 data with large $T_{\text{sys}}$ variations (50%), the simple linear fit would give us scatter of $\sim 4\%$, which is due to the nonlinearity of the relation at high opacities with large $T_{\text{sys}}$ variations. Although the exact form of the linear relation varies among different spectral windows and different data sets, we can use as few as four 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 $T_{\text{sys}}$ 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 $T_{\text{sys}}$ method is able to track differences in $T_{\text{sys}}$ between the two.

2. We have successfully reproduced the observed tight $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ correlation using ATM modeling for several data sets. The ATM modeling suggests that changing elevation or PWV will give us $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ relation of different slopes when the dry opacity is significant at the observing frequencies. This suggests that we might not get the tight $\hat{T}_{\text{sys}}$ versus $\hat{T}_{\text{WVR}}$ relation at these frequencies if both PWV and elevation varies significantly. A better strategy to track $T_{\text{sys}}$, especially in the cases mentioned above, would be to calculate the PWV using the continuous $T_{\text{WVR}}$ measurements from different WVR channel and combine the measured PWV, elevation and $T_{\text{sys}}$ for the data set to derive the continuous $T_{\text{sys}}$ based on the ATM modeling.

3. We apply the continuous $T_{\text{sys}}$ in calibration and find that it generally gives us more consistent fluxes for the same target. For the data set Band8 which has the largest PWV variation, the flux calibration uncertainty contribution due to $T_{\text{sys}}$ variability is reduced from 10% to 0.7%. The only exception is the data set Band9b as our new methods give

\(^4\) https://help.almascience.org/kb/articles/what-are-the-amplitude-calibration-issues-caused-by-alma-s-normalization-strategy
\(^5\) https://help.almascience.org/kb/articles/what-errors-could-originate-from-the-correlator-spectral-normalization-and-tsyst-sys-calibration
higher flux uncertainties. We suspect part of reasons are due to the imperfect linear fitting of the \( T_{\text{sys}} \) versus \( T_{\text{WVR}} \) 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.

4. If this method is used for sub-mm observatories such as ALMA, it can reduce the number of \( T_{\text{sys}} \) measurements required for high-frequency observations from 10~20 down to five (four \( T_{\text{sys}} \) measurements for the fitting and 1 bandpass \( T_{\text{sys}} \)) or fewer. Assuming each observation block takes \( \sim \)60 minutes and each \( T_{\text{sys}} \) measurement takes about 30–40 s, 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.

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, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). H.H. 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, The CASA Team et al. 2022).

Appendix A

\( T_{\text{sys}} \) versus \( T_{\text{WVR}} \) Relation

In this section we show the \( T_{\text{WVR}} \) versus \( T_{\text{WVR}} \) relation and the extrapolated \( T_{\text{sys}} \) for the rest of data we have in Figures A1–A3.

---

6 http://www.astropy.org
Figure A1. The summary of linear relation between $T_{\text{sys}}$ and $T_{\text{WVR}}$ and comparison between measured and extrapolated $T_{\text{sys}}$ for measurement sets in project 2018.1.01210.S. For each measurement set, the left subplot shows the linear correlation between normalized $T_{\text{sys}}$ and $T_{\text{WVR}}$. 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 four Atm-cal scans. The right plot shows the extrapolated $T_{\text{sys}}$ based on the fitting relation using all Atm-cal scans. The orange, green and blue points are extrapolated $T_{\text{sys}}$ for bandpass, phase and science targets. The red points are the original measured $T_{\text{sys}}$. 

Project: 2018.1.01210.S; Band 7
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 Figure 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 four projects used in this paper with just one measurement set. See details in Figure A1.
Appendix B
Dirty Images for the Rest of Targets

In this section we show the dirty image of the rest of data sets made with extrapolated continuous $T_{\text{sys}}$ using all Atm-cal scans and the aperture we used to measure the flux in Figure B1.

![Dirty Images](https://example.com/diagram.png)

**Figure B1.** Dirty images for other sources made using alternative continuous $T_{\text{sys}}$ 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.

**ORCID iDs**

Hao He @ https://orcid.org/0000-0001-9020-1858

**References**

Aglizzo, C., Trigilio, C., Pignata, G., et al. 2017, *ApJ*, 841, 130

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33

Bachiller, R., Carilli, C., Cox, P., et al. 2003, ALMA Science Advisory Committee (ASAC), https://nrao.edu/archives/items/show/34475

Baryshev, A. M., Hesper, R., Mena, F. P., et al. 2015, *A&A*, 577, A129

Brogan, C. 2018, Advanced Calibration Topics—I, https://science.nrao.edu/science/meetings/2018/16th-synthesis-imaging-workshop/talks/Brogan_Adv_Cal_1.pdf

Casalta, J. M., Molins, A., Bassas, M., et al. 2008, *Proc. SPIE*, 7018, 701838

Condon, J. J., & Ransom, S. M. 2016, Essential Radio Astronomy (Charlottesville, VA: NRAO)

Dempsey, J. T., Friberg, P., Jenness, T., et al. 2013, *MNRAS*, 430, 2534

Emonts, B., Raba, R., Moellenbrock, G., et al. 2020, in ASP Conf. Ser. 527, Astronomical Data Analysis Software and Systems XXIX, ed. R. Pizzo et al. (San Francisco, CA: ASP), 267

Gonzalez, A., Fujii, Y., Kameko, K., et al. 2014, *Proc. SPIE*, 9153, 91530N

Hills, R. 2004, ALMA Memo No. 495, https://library.nrao.edu/public/memos/alma/main/memo495.pdf
Hills, R., Gibson, J., Richer, J., et al. 2001, ALMA Memo 352, https://library.nrao.edu/public/memos/alma/main/memo352.pdf
Mahieu, S., Maier, D., Lazareff, B., et al. 2012, ITTST, 2, 29
Mangum, J. 2002, ALMA Memo No. 434, https://library.nrao.edu/public/memos/alma/main/memo434.pdf
Mangum, J. 2017, ALMA Memo No. 602, https://library.nrao.edu/public/memos/alma/main/memo602.pdf
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Ser. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco, CA: ASP), 127
Moreno, R., & Guilloteau, S. 2002, ALMA Memo 372, http://legacy.nrao.edu/alma/memos/html-memos/alma372/memo372.pdf
Pardo, J. R., Cernicharo, J., & Serabyn, E. 2001, ITAP, 49, 1683
Payne, J., Vaccari, A., Emerson, D., & Mangum, J. 2001, ALMA Construction Project Book, Chapter 3 Section 2 (Charlottesville, VA: NRAO)
Petry, D., Vila-Vilaro, B., Villard, E., Komugi, S., & Schnee, S. 2014, Proc. SPIE, 9152, 91520J
Remjian, A., Biggs, A., Cortes, P. A., et al. 2019, ALMA Technical Handbook, ALMA Doc. 7.3, ver. 1.1, 2019, Zenodo, doi:10.5281/zenodo.4511522
Sekimoto, Y., Iizuko, Y., Satou, N., et al. 2008, in Nineteenth Int. Symp. on Space Terahertz Technology, ed. W. Wild (Leiden, SH: ISSTT), 253
The CASA Team, Bean, B., Bhatnagar, S., et al. 2022, PASP, 134, 114501
Yun, M., Bastian, T., Holdaway, M., Mangum, J., & Welch, J. 1998, ALMA Memo No. 211, https://library.nrao.edu/public/memos/alma/main/memo211.pdf