ALMA Memo (X Not Submitted X), 1–5; (2011) Printed 7 July 2011

Quality Control of WVR Phase Correction Based on Differences Between WVR Channels

B. Nikolic, R. C. Bolton & J. S. Richer
Astrophysics Group, Cavendish Laboratory, Cambridge CB3 0HE, UK
email: b.nikolic@mrao.cam.ac.uk
http://www.mrao.cam.ac.uk/~bn204/

r 425

ABSTRACT
Once phase correction coefficients have been estimated, each of the channels of ALMA water vapour radiometers can be used to independently estimate the path fluctuations to each of the antennas. By computing the *differences* between these estimates for each antenna, it is possible estimate the quality of the resulting phase correction, on an antenna-by-antenna basis, without need for specialised observing or data from the astronomical receivers. This computation has been implemented in the `wvrgcal` program and is expected to be useful for quality control of WVR-based phase correction.

1 INTRODUCTION

Water Vapour Radiometer (WVR) based phase correction appears to be working well in majority of ALMA observations. However, there is a fraction of observations in which it is not working well for a variety of reasons including atmospheric conditions (e.g., cloud), hardware issues, shadowing by other antennas, etc. These problems in WVR phase correction manifest themselves in a larger overall phase noise than expected (e.g., from the ALMA specifications), problems on longer baselines only, or problems on individuals antennas only. Each of these issues needs to be dealt with in a different way.

Interpreting the quality of WVR data by eye is however not easy because:

(i) Some problems are quite subtle and only identifiable through statistical calculation
(ii) The WVRs are specialised devices and not all scientists working at/with ALMA have sufficient knowledge to interpret their data
(iii) There will be eventually up to 54 WVRs working on ALMA, creating visualisation problems
(iv) The number of measurement sets produced by ALMA is large and manual inspection is time-consuming

Additionally, it is useful to have a quantitative estimate of how well the WVR phase correction is working which can be used in automated processing and for long-term tracking of observatory performance.

One way of doing quality assurance on the expected quality of phase correction derived from the WVR observations is to consider the four channels of the WVRs separately. Each of these channels can be used to compute a separate, independent, estimate of the path fluctuation to each antenna. In the ideal case in which our atmospheric model is exactly correct and there are no unexpected atmospheric, software or hardware effects, these path estimates will be very similar, with the difference only due to the intrinsic "thermal" noise within the WVR receiver itself. Many problems will however cause the estimates of path from the four channels to diverge. For example, cloud emission affects the outer WVR channels proportionally much more than the inner channels; and errors in the estimate of the phase correction coefficients are very unlikely to be the same for each of the channels and therefore again lead to a divergence of the path estimates between the channels.

A simple measure of the difference between path estimates from different channels is an indicator of how well we are likely able to convert the WVR temperature measurements into reliable path estimates, and is therefore a potential indicator of the quality of WVR phase correction, and can be computed without access to the astronomical data or any specialised observing. In this short memo we describe the performance of this estimator, explain its implementation in the `wvrgcal` program and show some examples.

2 EXPECTED DISCREPANCY

2.1 Intrinsic ("thermal") WVR noise

The intrinsic, thermal-like, noise in the mixer and LNA of the WVR is uncorrelated between the channels (note that this is not true for 'noise' due to gain fluctuations of the receiver). Therefore path estimates from different channels will differ by at least the combination of their individual path errors due to the thermal-like noise.

For 1 s integration times, the specified noise for the four channels of the WVRs is 0.09, 0.07, 0.08 and 0.1 K RMS (starting with the inner most channel, number 1). This internal noise is translated to effective path noise by dividing it by the path correction coefficients which are computed by the `wvrgcal` program. The variations of the path correction coefficients with water vapour column for a simple model of the atmosphere are plotted in Figure 1.

In Figure 2 we have combined the specification of noise of WVRs and the path-correction coefficients to show the discrepancy of the path estimate between channels 1 to 3 and channel 4. This estimate of path discrepancy is close to the minimum we would expect to see in practice; however, some of the WVRs are significantly better than the specification and the usual integration time is 1.152

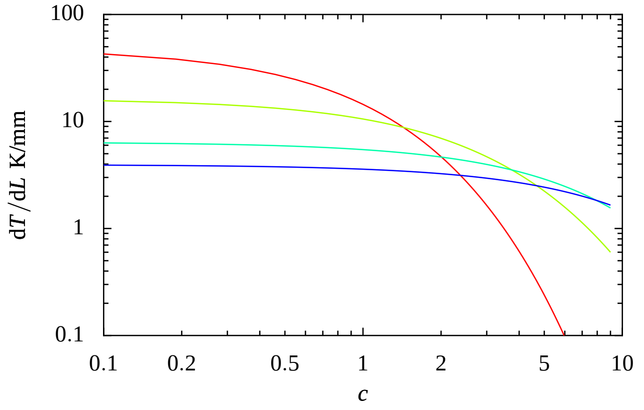

Figure 1. Model phase correction coefficients versus the total column of water vapour. Red to blue lines (following the spectrum) are channels 1 to 4, innermost to outermost channel.

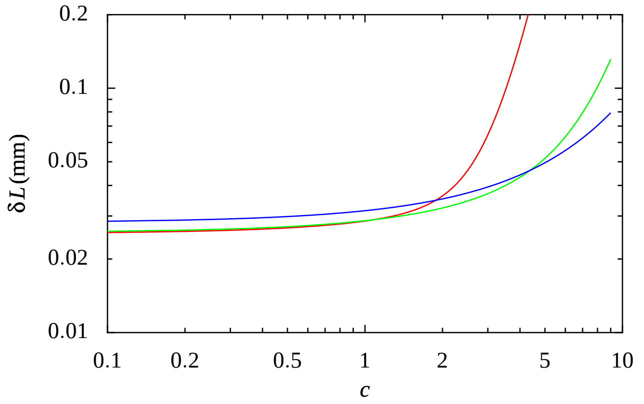

Figure 2. Discrepancy (due to intrinsic noise of the radiometers only) between path estimates in channels one to three versus channel 4 (red line: channel 1 vs channel 4; green line: channel 2 vs channel 4; blue line: channel 3 vs channel 4).

so sometimes smaller discrepancies can be seen in the very best conditions.

2.2 Filter response errors

The relationship between changes in the column of water vapour along the line of sight of an antenna and the signal recorded by the WVRs depends sensitively on the shape and centre frequencies of the WVR filters. This sensitivity is transferred to the path correction coefficients which likewise depend on the filter shapes and centre frequencies. The ALMA production WVRs however have filters with a significant manufacturing spread. Although the characteristics of the filters are measured after the manufacturing process, the conveyance of this information to the `wvrgcal` application has not yet been implemented. Therefore, although it is in principle possible to take into account the difference in filter characteristics during the phase correction application, this is not done at the moment, and this therefore creates a source of discrepancy between the path estimates from different channels.

To first order the effect of variation of filter characteristics can be approximated by the change of the effective centre frequency of the filter. In Figure 3 we have plotted the *relative* change in the path correction coefficients when each of the filter centre frequencies

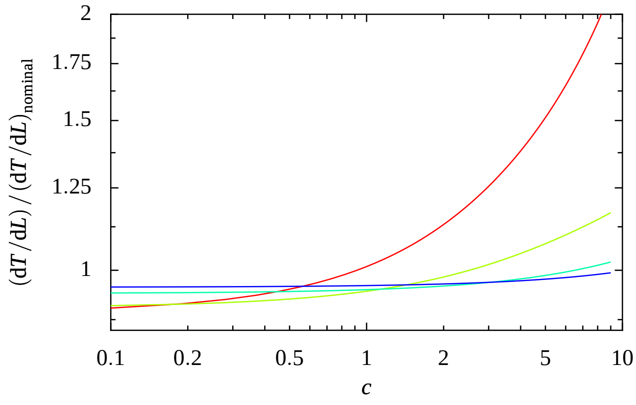

Figure 3. Fractional change in the model phase correction coefficients for a 200 MHz shift in the centre frequency of each filter. Red to blue lines (following the spectrum) are channels 1 to 4.

is increased by 200 MHz (the specification for centre frequency manufacturing accuracy is 5% which means maximum acceptable manufacturing difference from nominal is 60 to 360 MHz, depending on the filter).

The actual observed discrepancy due to filter centre frequency shifts depends on the *total* path fluctuations seen by the antenna and the *difference* between path coefficient errors for the two filters being compared. For example, the total path fluctuation on time scales of tens of minutes at the AOS is 0.2 mm RMS and the maximum fractional difference between channels is about 0.05, and therefore the maximum contribution to discrepancy due to the filter error is 10 μmRMS.

2.3 Clouds

When conditions at the AOS are overcast it is expected that clouds will have a dominant contribution to the discrepancy and this is one of the primary motivations for implementing this discrepancy calculation.

The reason why they create a large discrepancy is that clouds produce a similar signal in all four WVR channels (because their intrinsic spectrum, proportional to frequency squared, combined with the double side-band nature of WVR mixers give a close-to flat response) while the signal from water vapour is very different in the four channels. Therefore converting the observed signal on the assumption that is due to water vapour only creates very different estimates of the path from each channel.

The magnitude of the discrepancy can be estimated by assuming that the clouds have a radiation temperature of 250 K and that the opacity has a standard deviation of 2% (expressed as opacity itself, i.e., $\sigma_\tau = 0.02$). The corresponding brightness fluctuations can be converted to path discrepancy using the coefficients plotted in Figure 1. The resulting RMS path error is shown in Figure 4, where the negative values simply indicate that channel 4 path estimates are greater than the other channel estimates. The implication of Figure 4 is that, for example, the discrepancy between channels 2 and 4 is almost 1 mm RMS under good conditions (pwv < 2 mm) and if the cloud opacity is varying by 2%.

The impact of errors on the actual phase fluctuations in the data after phase correction depends on several factors:

- Overall precipitable water vapour, which determines the phase correction coefficients. As channels most affected by cloud are the

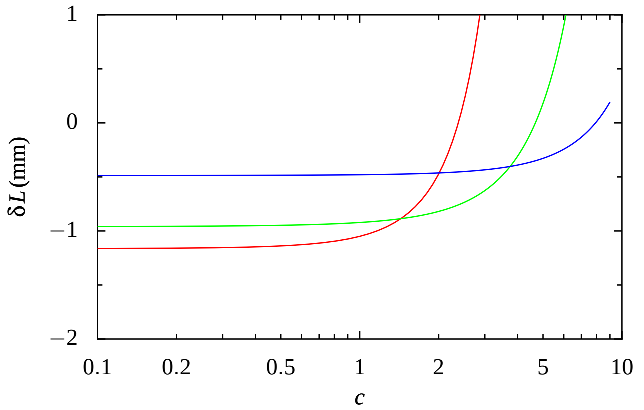

Figure 4. Expected discrepancy of path estimates between channels 1, 2 and 3 and channel 4 due to effect of clouds. Cloud opacity is assumed to have standard deviation of 2% and cloud temperature is assumed to be 250 K.

outermost channel and they are used proportionally more when conditions are humid, the impact of clouds *increases* when the PWV is high

- Baseline length and structure of the clouds, because the errors due to cloud will largely cancel out when antennas have substantially significant lines of sight

The discrepancy calculation is however to an extent an estimate of the maximum path error that could be expected. Anecdotal evidence suggests that the actual errors in fairly compact configurations are about 10% to 30% of the discrepancy but this needs substantial further investigation.

2.4 Modelling errors

Another possible source of discrepancy between path estimates in the different channels is modelling error, leading to incorrect estimates of the path correction coefficients. There are two sources of errors which should be considered: the internal model error due to the limited accuracy with which the retrieval of atmospheric model parameters can be made, and the external modelling error due simply to a wrong model being used. The internal model error is computed by `wvrgcal` and for currently the used models it is quite small and likely to be dominated by the external modelling error.

Estimating the external modelling error is complex and so in this section we only consider some plausible errors in inferred basic atmospheric parameters and how they would affect the discrepancy of path estimates between channels. The parameters considered are total water vapour column (Figure 5), temperature of the water vapour layer (Figure 6) and pressure of the water vapour layer (Figure 7).

3 IMPLEMENTATION IN `WVRGCAL`

Normally, `wvrgcal` computes the path correction from a weighted average of estimates from each of the channels, with the weighting selected to minimise the error from the intrinsic noise in each of the filters.

Computing the discrepancy involves computing additional estimates of the path correction which are based on only one of the WVR channels. This computation is, at top level, implemented in the `computePathDisc` function in the `wvrgcal.cpp` file.

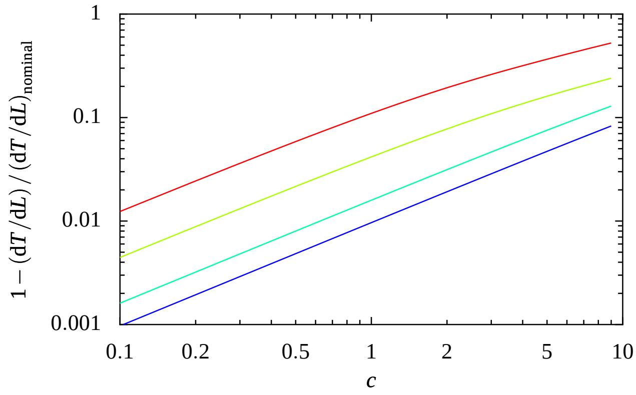

Figure 5. Fractional change in the model phase correction coefficients for a 10% change in the water vapour column. Red to blue lines (following the spectrum) are channels 1 to 4.

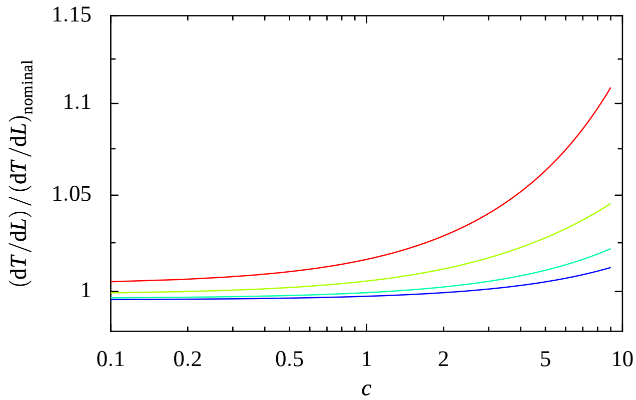

Figure 6. Fractional change in the model phase correction coefficients for a 2 K change in the water vapour temperature. Red to blue lines (following the spectrum) are channels 1 to 4.

The computation is made easier by the feature to *mask* channels in the `dTdLCoeffsBase` class using the field `chmask`. This field represents an additional weighting which is applied to path estimates from each of the channels. For computation of the discrepancy, the field is set to all zeros except for one channel and an estimate of path is computed (by constructing an appropriate object and calling the

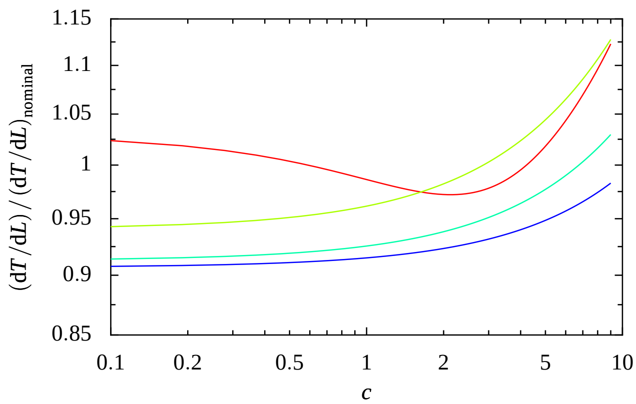

Figure 7. Fractional change in the model phase correction coefficients for a 50 mBar change in the water vapour pressure. Red to blue lines (following the spectrum) are channels 1 to 4.

```
     Retrieved parameters
--------------------------------------------------------
Evidence        PWV      PWV Error        dTdL    dTdL
1.95645e-15     0.531694 0.00877767       23.7037 12.7433

     Antenna/WVR information:
--------------------------------------------------------
#       Name    WVR?    Flag?   RMS (um)        Disc (um)
0       DV01    Yes     No      99.8            26.8
1       DV02    Yes     No      98.3            55.1
2       DV04    Yes     No      99.5            25.3
3       DV05    Yes     No       101            19.1
4       DV06    Yes     No      97.3              20
5       DV07    Yes     No      96.5            18.7
6       DV08    Yes     No      98.9            28.3
7       DV09    Yes     No      94.6            19.4
8       DV10    Yes     No      96.9            21.4
9       DV11    Yes     No      95.5            19.5
10      DV12    Yes     No      97.7            21.9
11      DV13    Yes     No      94.6            20.6
12      PM01    Yes     No      95.7            20.7
13      PM03    Yes     No      97.3            18.9
```

Figure 8. Sample output of `wvrgcal` showing the discrepancy computation (last column). The retrieved parameter section has been truncated so it does not show the 3rd and 4th path correction coefficient. The dataset used was uid___A002_X219601_X5c7.

`ArrayGains::calc` function). Estimates of path for different channels are then computed using the `ArrayGains::pathDiscAnt` function.

An important feature for the discrepancy calculation is the `--statfield` option which allows the user to restrict the computation of WVR statistics for only one of the fields in the input measurement set. This feature, which is also applicable to other "statistics" outputs from `wvrgcal`, i.e., the total path RMS and the greatest baseline path fluctuation, is useful because observations with multiple fields often involve large changes in airmass which can dominate the actual atmospheric path fluctuation and make interpretation of statistics difficult.

In the current implementation the discrepancy is computed **between channels 2 and 4**. The reason for using channel 2 is that it is less optically thick and therefore more sensitive in the relatively poor conditions which are also associated with cloud and other atmospheric effects which make WVR phase correction difficult. An example output of the discrepancy computation is shown in Figure 8.

4 EXAMPLES

4.1 Noisy WVR

The example in Figure 8 shows that all WVRs have discrepancies of 27 μm or less except the WVR attached DV02 which is showing 55 μm of discrepancy. The total PWV for this observation was very low (0.5 mm) and therefore the thermal contribution to the discrepancy should be only about 20 μm. In general, in cases like this of a high discrepancy on one antenna the WVR data should be carefully examined for antenna-specific problems. In this case it was found that the unit attached to this antenna is showing excess noise by a factor of two. Note that the this excess noise is *not* reflected in the absolute path RMS column as this is dominated by the atmospheric fluctuations.

The impact on the science data of higher noise in one of the WVRs is to slightly increase the de-correlation on all baselines involving this antenna. This can be quite a subtle effect and should be carefully considered in absolute flux calibration of observations.

```
--------------------------------------------------------
Evidence        PWV      PWV Error        dTdL    dTdL
1.11918e-09     3.74126 0.145589          0.64603 3.20351

     Antenna/WVR information:
--------------------------------------------------------
#       Name    WVR?    Flag?   RMS (um)        Disc (um)
0       DV01    Yes     No       385            37.4
1       DV04    Yes     No       350            40.4
2       DV06    Yes     No       379            36.3
3       DV07    Yes     No       345            40.7
4       DV08    Yes     No       339            49.6
5       DV10    Yes     No       347            37.2
6       PM01    Yes     No       332            42.9
7       PM02    Yes     No       365            40.9
8       PM03    Yes     No      8.94e+07        2.29e+08
```

Figure 9. Output for a WVR which is producing spurious data. (Dataset uid___A002_X1cf8d4_X11).

```
-------------------------------------------------
Evidence        PWV      PWV Error        dTdL
2.10641e-10     5.82428 0.378964          0.0316761 1.01058

     Antenna/WVR information:
-------------------------------------------------
#       Name    WVR?    Flag?   RMS (um)        Disc (um)
0       DV01    Yes     No      1.64e+03        701
1       DV04    Yes     No      1.23e+03        470
2       DV07    Yes     No      1.23e+03        401
3       DV08    Yes     No      1.2e+03         406
4       DV10    Yes     No      1.22e+03        409
5       PM01    Yes     No         0              0
6       PM02    Yes     No      1.26e+03         440
7       PM03    Yes     No      4.13e+08       8.02e+08
```

Figure 10. Dataset showing evidence of cloudy conditions, and problematic WVRs.(Dataset uid___A002_X1ce2d9_X1e).

4.2 Faulty WVR

Figure 9 shows output of `wvrgcal` when one of the WVRs in the observation is faulty and producing spurious data. It can be seen that both the absolute path RMS and the discrepancy are very large, and in cases such as this the WVR must be completely flagged (use the `--wvrflag` option).

4.3 Cloudy conditions

Figure 10 shows output of `wvrgcal` when the conditions at the AOS were cloudy and very wet (6 mm of water vapour). Several effects can be see in this output:

(i) WVR attached to PM01 is showing 0 for both the RMS and discrepancy. This is an example of a software bug affecting few observations around the time that these data were taken in which the WVR is not properly initialised and continuously returns zero for the observed sky temperatures in all channels. For further analysis this radiometer would need to be flagged

(ii) WVR attached to PM03 is faulty as described above hence the large RMS and discrepancy values

(iii) WVRs attached to antennas DV04-DV10 are showing discrepancy values around 400 μm. As can be seen from Figure 2 we would expect a higher discrepancy than usual due to the wet conditions but the measured value is almost ten times higher than then expected. This indicates that the data are affected by significant, time variable, cloud cover and that WVR-based phase correction is likely to work very poorly for these observations

5 ON-LINE QUALITY CONTROL

The discrepancy calculation is useful for identifying situations in which the WVR phase correction is not expected to work very well. However, in most cases, there is little that can be done to improve the quality of phase calibration once the data have been taken. Therefore, when applied off-line this technique is only useful in estimating the overall phase stability and impact on the science data.

However, the computation itself is very simple and can also be easily done on-line, or for example in the quick-look pipeline. The only requirement is that antennas are observing a single field for a sufficient period of time (the expected minimum is about five minutes). If any problems are identified on-line, the observing strategy and observing project can be adjusted to take into account the expected performance of the phase correction; for example, the cycle time for observing the phase calibration can be shortened or a new project at a lower frequency can be selected.

The simplest way of implementing this would be to run the `wvrgcal` program on Measurement Sets generated by the real-time filler. This would result in a short but tolerable delay as weather conditions of course change quite slowly.

6 SUMMARY

After computing the phase correction coefficients (which requires all four WVR channels to work reasonably well), it is possible to produce four independent estimates of the path fluctuations by using data from each channel of the WVRs separately. These four estimates are affected very differently by various problems that can affect the WVR data – probably most importantly the effect of cloud cover. For this reason, it is possible to compute a simple indicator of quality of the phase correction by computing RMS of the difference of path estimates from two channels (we call this simply the discrepancy).

In Section 2 we have estimated the discrepancy that is expected in normal operating conditions and the additional discrepancy that is expected to arise due to errors and problems in the data. The expected discrepancy is a function of the total precipitable water vapour column and of the magnitude of fluctuations of the water vapour along the line of sight of the antennas. Therefore these factors should be taken into account when interpreting the discrepancy computation.

The discrepancy calculation, based on comparison between channels 2 and 4, has been implemented in the `wvrgcal` program starting with version 0.22. The details of this implementation are given in Section 3. Also implemented in this version is the `--statfield` option which allows the observing field for which this and other statistics are computed.

We have shown in Section 4 some examples of use of the discrepancy calculation output to identify problems in the data. We expect that with experience we will be able to increase our ability to interpret this information and make a judgement on the quality of phase calibration that can be anticipated. We also suggest that it would be useful to take this information into account when deciding on observing strategy or the project that is suitable for observing.