The Effects of Weather on MUSTANG Data Quality

Brian S. Mason & Dan Perera

Draft 1: 01sep10

Abstract

We examined 1623 scans of MUSTANG data from the past two seasons of GBT 90 GHz science observations to quantify the quality of the data and how it was affected by weather. We then correlated these data with Ron Maddalena's NWS-forecasted weather products to search for good predictors of periods of weather that are useful to MUSTANG observers. From this analysis we conclude that: a) while the observed trends are similar to those seen in earlier continuum stability analyses, the 90 GHz T_{sys} (from the atmosphere + hydrosols) at the target elevation is a better predictor of good observing conditions than long-IR wavelength downard irradiance, precipitable water vapor, zenith opacity, or the opacity in hydrosols; b) a 90 GHz T_{sys} of 35 K provides excellent data 59% of the time, and usable data 95% of the time; and c) a more lax criterion of $T_{sys} < 50$ K provides useful data 97% and excellent data 52% of the time, with substantially more total time available. These cutoffs correspond to 90 GHz opacities of $\tau \sim 0.13$ and $\tau \sim 0.19$, respectively. We were not able to identify a criterion among the existing weather forecast products that more efficiently selected the *most* stable conditions; this is an open area for future research into the forecasts which could deliver a substantial benefit to the GBT's 90 GHz science effectiveness. Finally, we applied these criteria, along with the current PTCS-recommended solar constraint and our best-estimate 3mm mapping wind limit, to determine the average amount of time available per year, its variation between years, and its average availability by month. **In all there are on average 354 hours per year of weather available by the strict (35 K) condition, and 539 hours per year available by the more lax (50 K) criterion. If daytime observing becomes feasible this increases to 985 hours (35 K) and 1607 hours (50 K) for the given wind cut of 11 mph**. We recommend implementation of a variable 90 GHz atmospheric T_{sys} criterion (including hydrosols) to be evaluated at the target elevation, along with variable forecasted wind and solar criteria. The default criteria for these for MUSTANG science projects should be: $T_{sys} < 35$ K; expected ground wind speed < 9 mph; and sunset+3h to sunrise. For projects targeting compact sources ($< 1'$ in diameter) the $T_{sys} < 50$ K criterion can be used.

1 Introduction

The sensitivity of continuum observations with made with single-dish telescopes are often limited by fluctuations in the atmospheric emission along the telescope sight-lines rather than intrinsic photon statistics (the radiometer equation, or at higher frequencies and lower loading levels, shot noise). Focal plane arrays have a powerful handle on the contaminating signal from atmospheric fluctuations since the fluctuations tend to be strongly correlated between different detectors in the focal plane, allowing their efficient removal by means of some form of "common mode" subtraction in which a moving average of all detectors is subtracted from each detectors' output as a function of time. The penalty for subtracting a common mode from the data is the systematic loss of sky signal on angular scales comparable to or larger than the camera instantaneous field of view. When the atmosphere is particularly stable, much less aggressive algorithms can be used for the common mode subtraction, with the result that the net impact of sky emmission varitions is to limit the ability to reliably reconstruct large scale sky structure. When the sky stability degrades too far the "common mode" approximation itself breaks down, preventing an efficient removal of atmospheric emission and resulting in low quality maps. In order to assess the effect of weather on MUSTANG observations we have examined the past two years of MUSTANG science observations (1623 scans in all). Our experience with MUSTANG on the GBT indicates that the common mode is typically dominated by atmospheric effects; our analysis therefore focuses on quantifying the stability of the common mode on a 5 second timescale. The approach is conceptually similar to that adopted by D. Balser in the 9 GHz continuum stability analysis presented in DSS SPN 14.2.

The organization of this memo is as follows. § 2 describes the MUSTANG dataset and data reduction procedures. § 3 the results from this analysis. § 4 summarizes the weather statistics over the period May 1, 2004 to July 30, 2010. All weather information, unless otherwise specified, is obtained from the observing weather forecast tool in CLEO (Maddalena 2010) retrieved in August 2010.

2 MUSTANG Dataset & Analysis

We analyzed all scans from a representative set of 29 MUSTANG science telescope periods (observing sessions). There were 1623 scans having a median scan duration of 110 seconds; in all, 51 hours of GBT data went into the analysis.The telescope periods were (omitting "AGBT" prefixes) : 08A_056_12, 08A_056_13, 08C_026_02, 08C_026_03, 08C_026_04, 08C_026_05, 08C_026_06, 08C_026_07, 08C_026_08, 09A_044_01, 09A_044_02, 09A_044_03, 09A_052_01, 09A_052_03, 09A_052_04, 09A_052_05, 09A_052_06, 09C_020_00, 09C_031_01, 09C_031_02, 09C_031_03, 09C_059_01, 09C_059_02, 09C_059_04, 09C_059_05, 10A_056_01, 10A_056_02, 10A_056_03,& 10A_056_04.

Because we are using data from scheduled 90 GHz telescope periods, they are biased (by the existing, sometimes nebulous and subjective) scheduling decision criteria towards periods of good weather. There is, however, a tail of poor weather periods — due to inaccuracies in the forecasts, human error, etc. — which allow us to get a handle on the impact of poor weather on our data. The distribution of total cloud cover (low+medium+high, from the forecasts) is shown in Figure 1, and the distribution of forecast 90 GHz zenith T_{sys} is shown in Figure 2.

Standard MUSTANG data reduction procedures were used to calibrate and compute the common-mode signal, which is typically a good estimator of fluc-

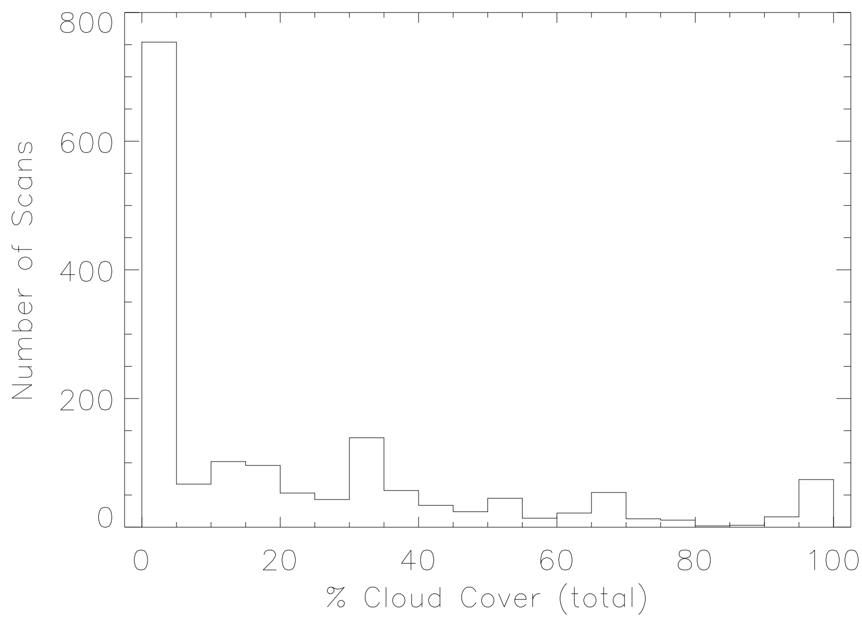

Figure 1: The distribution of forecast cloud cover for the scans in our analysis.46% of scans have forecast total cloud cover of 5% or less.

tuations in atmospheric emission. First, a $\sim$ 30 sec scan with the internal calibration lamp flashing is used to flag bad (optically non-responsive) detectors, to flat-field the detectors (*i.e.*, apply relative gain corrections), and to put the data in units of "nominal Janskys/beam" based on a canonical, astronomically determined value of 28.5 Janskys equivalent flux density for the calibration lamp. This calibration does not take into account variations in the antenna gain due to different active surfaces in use (although the 28.5 Jansky number is appropriate for those in use for most of the data in question), or variations due to the quality of OOF solution in place, for instance. The factors, however, do not significantly affect the antenna temperature which will be seen due to atmospheric emission, so the "nominal Jansky/beam" calibration can be taken as a stable relative indicator of the antenna temperature due to the atmosphere—exactly analogous to calibrating a GBT heterodyne receiver with the archival T_{cal} values, without reference to celestial calibration for a given telescope period.

Atmospheric emission, and a number of other systematic signals, contribute nearly identical signals to all detectors at a given point in time. We estimate this "common mode" signal as a function of time as follows. Consider that we have a series of integrations at times t_i from a detector indexed by j, d_{ij}. The

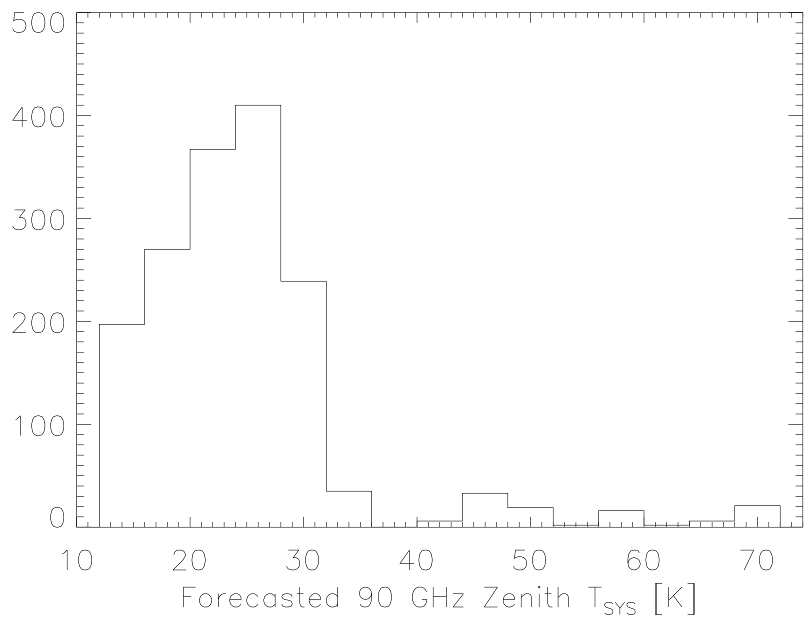

Figure 2: The distribution of forecast 90 GHz zenith system temperature due to the atmosphere and aerosols for the scans in our analysis. 6.5% of scans have $T_{sys} > 35\,\mathrm{K}$.

common mode signal is computed as

$$c_i = Median(d_{ij}) \tag{1}$$

where the median is computed for all good (optically responsive) detectors j, for a fixed integration i.

This common mode template (or time-stream) for each scan is then processed as follows:

- The known 1.411 Hz signal due to the pulse tube (PT) refrigerator is fit out with a sine and a cosine. This signal has a slowly varying amplitude so in fact several nearby frequencies are also added to the sine+cosine to allow for the finite width of this feature in fourier space. An overall gradient is also removed in this fit.
- The PT-cleaned common mode signal c_i is broken into as many 5-second chunks as will fit into the scan in question; let these chunks be indexed by k.
- A robust linear fit of c_i vs time is performed for each chunk, resulting in a set of slopes m_k.

- The mean absolute value of the slopes is computed resulting in an overall 5-second stability measure for this scan, $M = mean(|m_k|)$, where the mean is taken over the chunks indexed by k.

This number M represents the mean *fluctuation* (not gradient) in the common mode for the scan in question and has whatever units the data were calibrated to (in this case, nominal Janskys). It is more representative of factors affecting image quality than the overall gradient since gradients will be removed in any case.

This set of 1623 M values, measuring the 5 second stability of the atmosphere, are our basic quantifier of observed 90 GHz atmosphere stability. By examining the data we identified two thresholds in M which separate qualitatively "excellent" data, from "useful" data (for bright and/or compact source projects), from clearly "useless" or poor data. These thresholds were defined by examining the maps, common modes, and M values for a representative sampling of the scans in our analysis. We designate scans with $M < 0.01$ Jy/sec as "Excellent"; scans with $0.01 < M < 0.1$ Jy/sec as "Usable"; and data with $M > 0.1$ Jy/sec as "Useless".

Occasional glitches can substantially affect the common mode and our stability quantifier. For instance, there was a period of several weeks during which the radome air compressor was periodically introducing substantial vibrations into the MUSTANG data. To avoid biasing our results we took no special measures to exclude such events. They should not be correlated with the weather, so should only add noise to our analysis. They invariably *reduce* stability, so will tend to make the weather appear slightly worse than it is.

2.1 Forecast Data Products

CLEO weather forecasts were retrived with 1 hour time resolution covering the period of MUSTANG observations analyzed, and associated with each GBT scan. The primary data products used in this analysis are 90 GHz sky brightness at zenith (including aerosols) and the wind speed.

3 Results

A scatter plot of the mean 5 second fluctuations M for all 1623 scans is shown in Figure 3. Figure 4 shows the fraction of "Excellent" data in each Tsys bin, and the fraction of "Useless" data in each bin. Both are shown as functions of the predicted 90 GHz sky brightness along the line of sight for the scan in question. For $T_{sys} < 35$ K, 51% of the data are Excellent; only 3% are Useless. For data collected with forecast $35\,\mathrm{K} < T_{sys} < 50\,\mathrm{K}$, 20% are Useless and only ~ 5% are Excellent. For data with forecast $T_{sys} > 80$ K, fully 50% of the data are useless. From these data we conclude that selecting periods with forecast 90 GHz atmospheric system temperatures along the expected sightline of 35 K provides good enough weather for 90 GHz continuum projects on the GBT, and many projects will be able to obtain useful data when scheduled with a 50 K

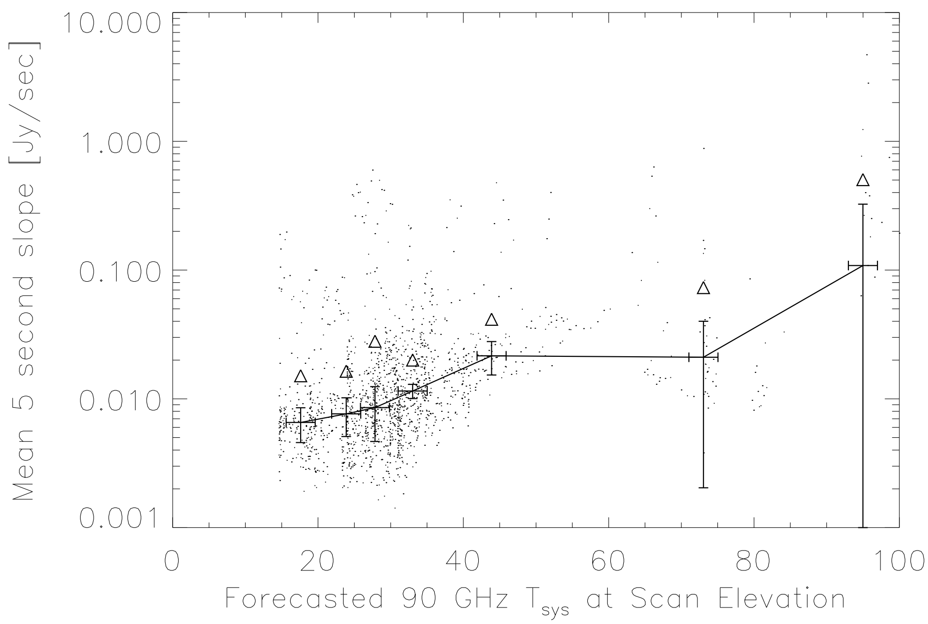

Figure 3: Observed 90 GHz stability on a 5 second timescale vs forecast 90 GHz atmospheric system temperature at the elevation of the scan in question. Plus signs show the medians within the chosen set of 6 bins; triangles show the mean in each bin. The vertical axis is our stability measure M.

threshold. We note that accounting for the observed elevation of the source provided notably clearer trends and correlations than simply using the forecast *zenith* brightness. The steep gradient in the fraction of “Excellent” weather vs sky brightness for $< 35\,\mathrm{K}$ is indicative that further weather forecast research is needed to optimally schedule projects needing those conditions, rather than merely usable conditions.

We also examined trends of M vs aerosol opacity, total zenith opacity τ_{90}, precipitable water vapor content, and downard IR irradiance. With the exception of aerosol opacity, all showed qualitatively similar trends, but 90 GHz system temperature showed the clearest trends and correlations. There were too few periods with nonzero forecasted aerosol opacities to form any conclusions. The trends for forecast precipitable water vapor and downward irradiance are shown in Appendix A.

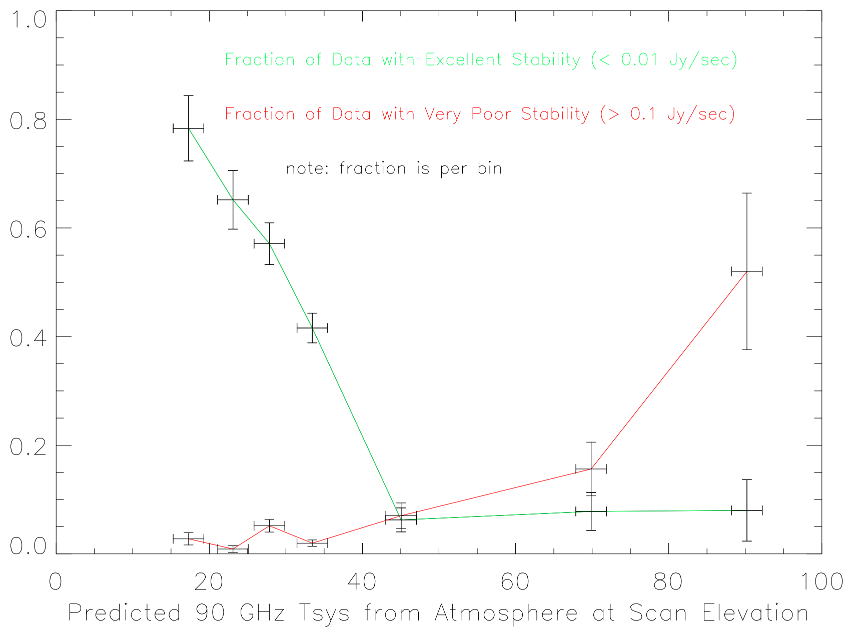

Figure 4: The fraction of "Excellent" (green) and "Useless" (red) data in each of 6 bins in forecast 90 GHz system temperature from the atmosphere at the elevation of the scan in question.

Criteria	Hours/year
Night only; wind < 11 mph; $T_{sys} < 35\,\mathrm{K}$	354 h
Night only; wind < 11 mph; $T_{sys} < 50\,\mathrm{K}$	539 h
Day+Night; wind < 11 mph; $T_{sys} < 35\,\mathrm{K}$	985 h
Day+Night; wind < 11 mph; $T_{sys} < 50\,\mathrm{K}$	1607 h

Table 1: Average time available per year for $T_{sys} < 35\,\mathrm{K}$ and $T_{sys} < 50\,\mathrm{K}$ cuts, also applying our current wind and thermal (sunset + 3h) thresholds.

Year	Excellent Hours	Usable Hours
2005	314 (871)	466 (1422)
2006	316 (928)	522 (1623)
2007	237 (760)	343 (1211)
2008	375 (1028)	674 (1954)
2009	472 (1233)	647 (1803)

Table 2: Excellent and usable hours by year. All cuts are applied; totals for day+night are shown in parentheses.

4 Site Weather Statistics

We would like to evaluate the amount of time currently available for MUSTANG observations. We have established one criterion for scheduling MUSTANG projects on the GBT ($T_{sys} < 35\,\mathrm{K}$). Two other criteria must be considered: wind, and solar illumination. We adopt a wind limit of CLEO-forecasted wind limit of 11 mph, corresponding roughly to an expected ground wind speed of 9 mph. We also assume that MUSTANG projects are only scheduled from 3 hours after sunset to sunrise. With these criteria in hand we seek to assess the number of useful MUSTANG observing hours available per year and per month, on average, and its annual fluctuation.

To do so we use the CLEO weather forecast archival data for the period 01 May 2004 to 30 July 2010. For this analysis we have used the raw CLEO-predicted wind speeds, uncorrected for the known forecast bias.

The results of this analysis are shown in Tables 1 and 2 and Figure 5.

Note: All T_{sky} values in this section are at zenith.

4.1 Comparison of Forecast Data with 86 GHz Tipper Measurements

Since the weather information in our analysis come from the National Weather Service models at locations tens of miles from the Green Bank site, it is important to check these products against independent measurements of the weather conditions at the site. The most direct comparison is provided by atmospheric brightness temperature measurements provided by the 86 GHz tipper operated

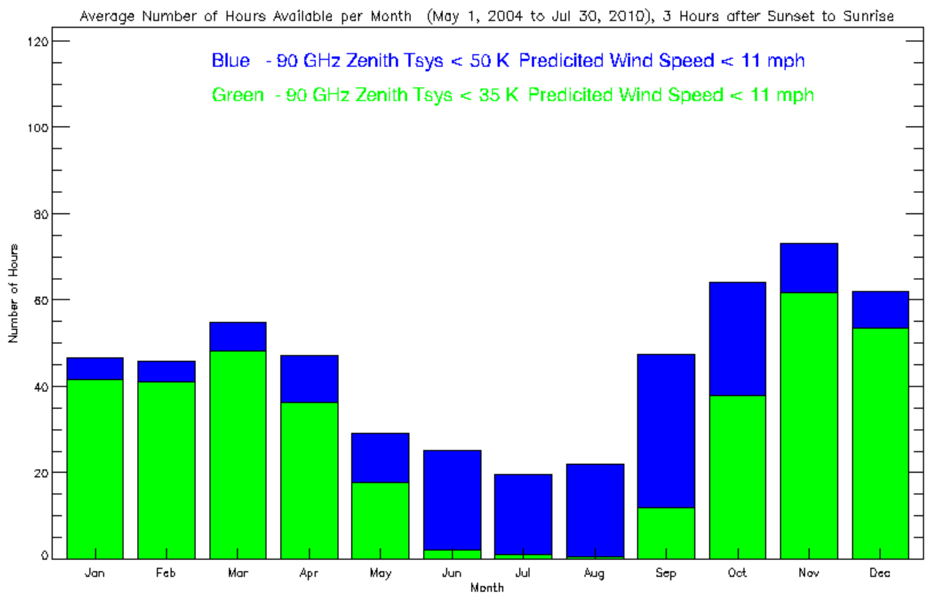

Figure 5: The number of hours on average per month, with "Excellent" (green; $T_{sys} < 35\,\mathrm{K}$) and "Usable" (blue; $T_{sys} < 50\,\mathrm{K}$) conditions. An 11 mph wind speed cut (uncorrected for the known bias; corresponding to ~ 9 mph expected ground winds) has also been applied, and only data between sunset+3h and sunrise are included.

for several years around the year 2000. The distribution of tipper measurements is compared with the forecast 86 GHz zenith sky brightness in Figure 6.

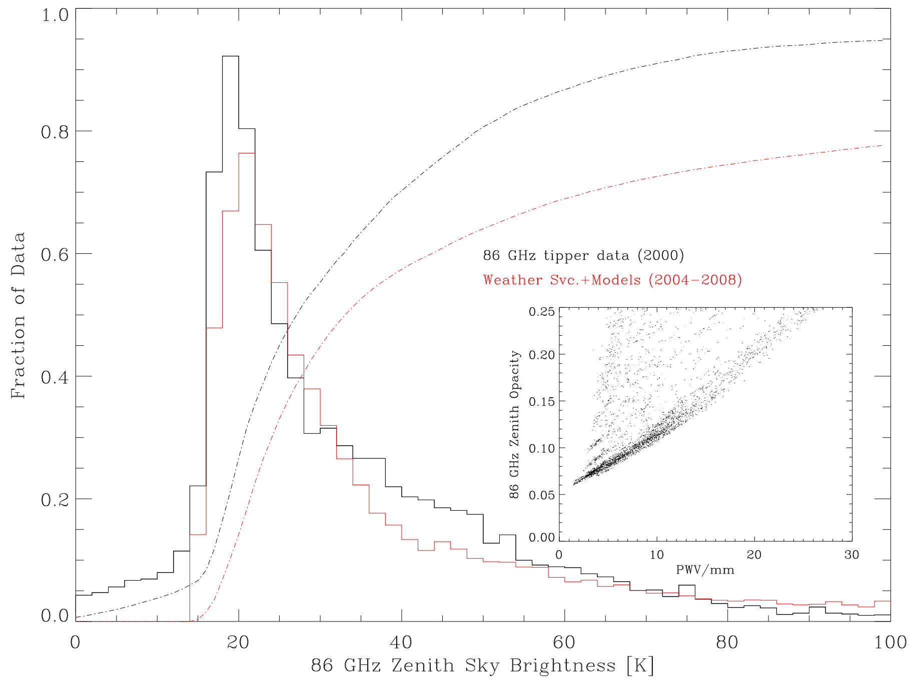

Figure 6: Distribution of 86 GHz zenith opacity measurements from the GB tipper in the year 2000, with NWS model-based predictions of the 86 GHz zenith opacity over the period 2004-2008. (Data courtesy of Ron Maddalena)

5 Conclusions & Scheduling Recommendations

Our main conclusions, and some notes and items for future work, are as follows.

- We have evaluated the stability of MUSTANG data and established that the 90 GHz sky brightness along the expected sight-line is a good criterion for selecting useful observing conditions.
- For projects which require the best observing conditions (faint sources with structure) we recommend a cutoff of $T_{sys} < 35\,\mathrm{K}$. For projects concerned with bright or compact sources, a looser criterion of $T_{sys} < 50\,\mathrm{K}$ can be used.
- We have applied a (CLEO forecasted) wind cutoff of 11 mph, corresponding to $\sim$ 9 mph expected ground wind, qualitatively based on our expe-

rience with MUSTANG observing. This wind limit needs to be better quantified and could have a significant effect on our bottom line.

- With wind and PTCS thermal (sunset+3h) thresholds, we have on average 354 hours per year of excellent weather, and 537 hours of usable weather.

- The time estimates in this memo are *schedulable* time. This analysis predicts that the amount of *useful* time will be a few percent lower for projects needing only "Usable" conditions. They will be lower by a factor of ~ 1.7 for projects needing the best conditions, but this factor is already accounted for in the overhead observers have been told to allow. Further research into forecasting atmospheric stability per se could improve the efficiency with which these projects can be executed.

- We have not tried to estimate operational losses (due to snow in the dish, maintenance, or pulsar monitoring, for instance). At some level these are also already implicit in the overheads observers are told to allow for.

- Raising the wind limits and opening up daytime 3mm mapping operation should be very high priorities for future GBT development. Opening up daytime observing, for instance, allows 985 hours per year of excellent conditions and 1607 hours of usable conditions (both with the same wind cut).

For scheduling, we recommend implementing a user-tunable sight-line 90 GHz atmospheric system temperature threshold (default values of 35 K or 50K, set by the project friend), together with a tunable wind threshold (default value 9 mph), and an "hours after sunset" thermal threshold (default value 3h, although allowing for project change-over and startup times, perhaps 2h would do). For the sight-line elevation, the elevation of the source at the mid-point of the telescope period should be used. Actual sunset rise and set over the mountains should be used. Care must be taken to avoid scheduling too close to the zenith, considering GBT performance limits.

References & Further Reading
D.S. Balser, Dynamic Scheduling System Project Note 14.2 (April 2010), "Atmospheric Stability from GBT Continuum Observations"
J.J. Condon, PTCS System Note 3 "Scientific Requirements for High Frequency Observations with the GBT" (March 2003)
J.J. Condon, "Refined Scientific Requirements for the PTCS" (December 2003)
J.J. Condon & D.S. Balser, Dynamic Scheduling System Project Note 4.2 (May 2007), "Ranking Observing Sessions for Dynamic Scheduling"
R. Maddalena, `http://www.gb.nrao.edu/~ rmaddale/Weather/index.html` (Retrieved 25aug10)
R. Maddalena, "Accurate Weather Forecasting for Radio Astronomy" in the Bulletin of the American Astronomical Society, Vol. 42, p.406 (January 2010)
P. Ries, T. Hunter, & F. Ghigo, PTCS Project Note 64.4 "Quadrant Detector Practical Application" (12 Feb 2009)
P. Ries, PTCS Project Note 68.1 "Winds and their Effect on the GBT" (12 Feb 2009)

A Appendix: Other Metrics

Figures 7 through 10 show the stability and fractions of good and bad data as functions of forecast PWV and downward irradiance. These present a qualtiatively similar picture to that seen through forecast 90 GHz zenith system temperature, and to what was found at X-band in DSS SPN 14.2. Differences in the quality of these metrics presumably reflect a combination of their intrinsic appropriateness and the accuracy with which they can be predicted.

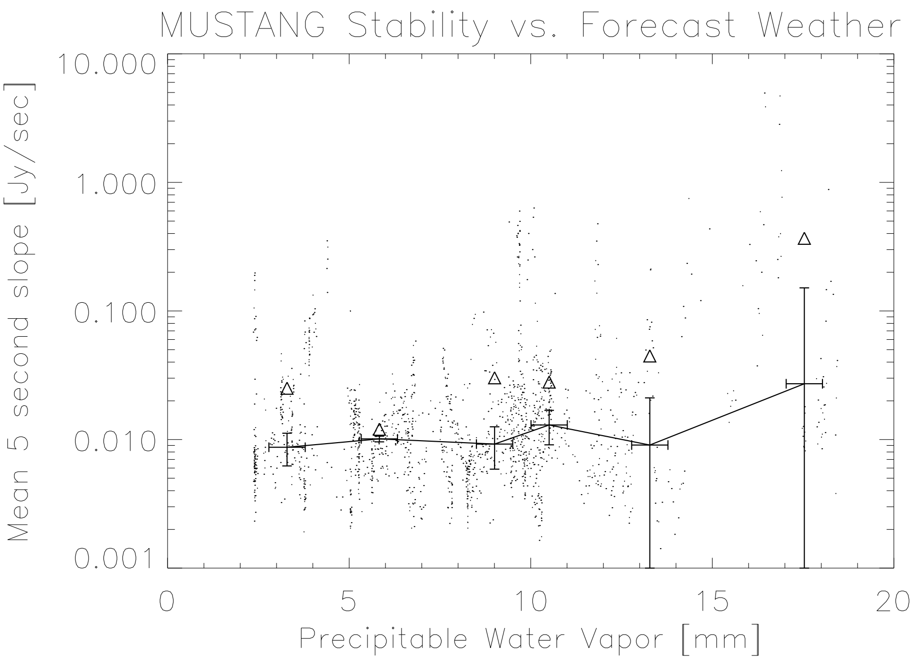

Figure 7: Observed 90 GHz stability on a 5 second timescale vs forecast PWV.

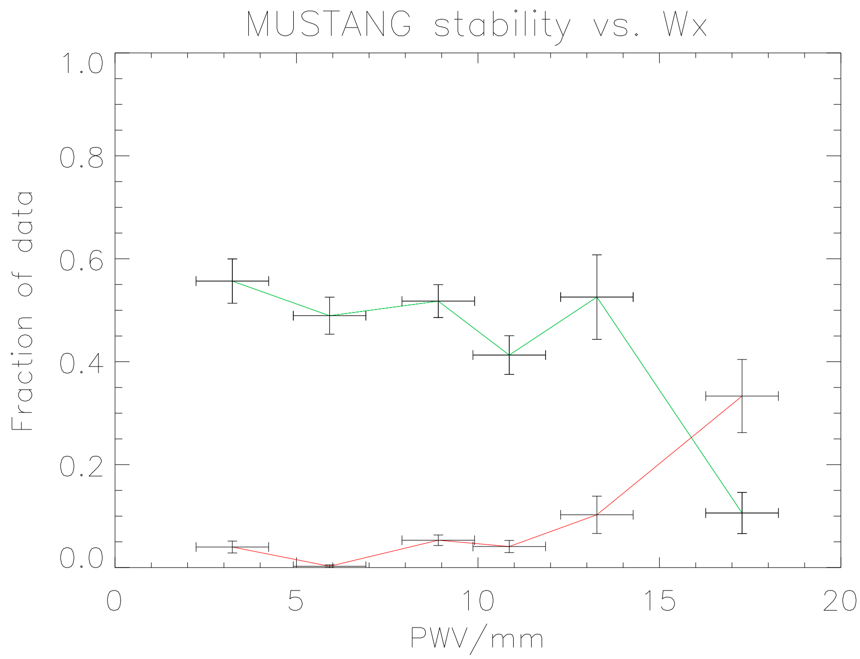

Figure 8: The fraction of “Excellent” (green) and “Useless” (red) data in each of 6 bins in forecast PWV.

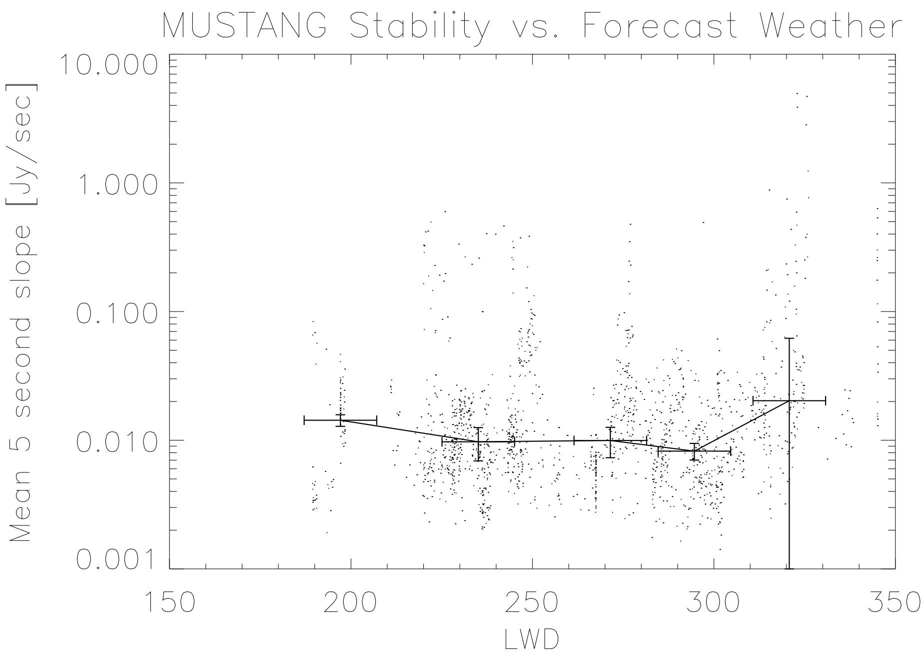

Figure 9: Observed 90 GHz stability on a 5 second timescale vs forecast long-wavelength downward irradiance (LWD).

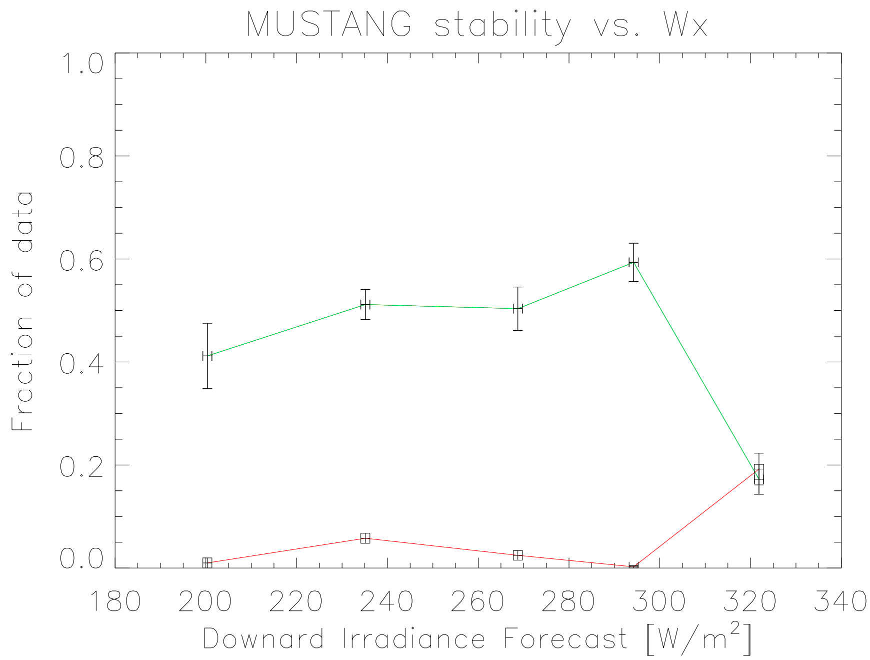

Figure 10: The fraction of "Excellent" (green) and "Useless" (red) data in each bin in forecast LWD.