Rejection of Tanh Models for the Global 21-cm Signal: PART 4, χ^2 Sliding Window Approach *(ongoing work)*

Raul Monsalve
CASA, University of Colorado Boulder
SESE, Arizona State University

February 16, 2017

This document describes the continuation of the simulation analyses presented in Reports #80, #82, and #83 (`loco.lab.asu.edu/memos/`).

Here we present the approach of rejecting a set of Tanh models based on the χ^2 statistics. The χ^2 statistics is computed from the residuals after fitting the EDGES foreground polynomial model to a window of data after subtracting the Tanh 21-cm model. The Tanh expression is:

$$T_{21}(z; z_r, \Delta z) = \frac{28\ \mathrm{mK}}{2}\ \left[\tanh\left(\frac{z - z_r}{\Delta z}\right) + 1\right]\sqrt{\frac{1+z}{10}}. \tag{1}$$

The EDGES polynomial used to model the foregrounds, with N_{fg} terms, is:

$$T_{\mathrm{fg}}(\nu) = \sum_{i=0}^{N_{\mathrm{fg}}-1} a_i \nu^{-2.5+i}. \tag{2}$$

First, we compute a reference $\chi^2_\dagger$ from the residuals after fitting the foreground model of equation 2 to the N_i data points in the window:

$$\chi^2_\dagger = k \cdot \sum_{i=0}^{N_i-1} w_i [d_i - T^\dagger_{\mathrm{fg}}(\nu_i)]^2, \tag{3}$$

where k is a multiplicative constant, and v_i, d_i, and w_i represent the frequency, data, and weights within the window. In this expression, the model $T^\dagger_{\mathrm{fg}}(\nu)$ previously fitted the data directly, assuming that there is no 21-cm signal in the data.

We probe the Tanh models by computing the following χ^2_*:

$$\chi^2_*(\nu_r, \Delta z) = k \cdot \sum_{i=0}^{N_i-1} w_i [d_i - T_{21}(\nu_i; \nu_r, \Delta z) - T^*_{\mathrm{fg}}(\nu_i)]^2. \tag{4}$$

In this expression, the model $T^*_{\mathrm{fg}}(\nu)$ has previously fitted the data minus the Tanh expression we are probing.

We normalize χ^2_* using the value of k that results in $\chi^2_\dagger = \mathrm{DF}$ in equation 3, i.e.,

$$k = \frac{\mathrm{DF}}{\sum_{i=0}^{N_i-1} w_i [d_i - T^\dagger_{\mathrm{fg}}(\nu_i)]^2}, \tag{5}$$

where $\mathrm{DF} = N_i - N_{\mathrm{fg}} - 1$ are the degrees of freedom in the fit. This normalization is equivalent to stating that, on average, the uncertainty of the data points is equal to the residuals after fitting $T_{\mathrm{fg}}^{\dagger}$ to the data, assumed not to contain a Tanh model.

The rejection probability of a Tanh 21-cm model is given by:

$$\mathrm{prob}(\chi_*^2; \mathrm{DF}) = 1 - \int_{\chi_*^2}^{\infty} \mathrm{pdf}(\mathrm{DF}) \cdot d\chi^2, \tag{6}$$

where pdf is the Probability Distribution Function for the χ^2 distribution. As the χ_*^2 statistics increases for a fixed DF, the rejection probability increases. Here, we set a model rejection threshold of 95%.

The χ^2 sliding window algorithm works as follows:

(i) Sweep over window width

(ii) Sweep over number of foreground terms

(iii) Sweep window across frequency

(iv) Fit $T_{\mathrm{fg}}^{\dagger}$ to window data, and compute $\chi_{\dagger}^2$

(v) Sweep over reference frequency ν_r within window

(vi) Sweep over reionization duration Δz

(vii) Evaluate $T_{21}(\nu_r, \Delta z)$

(viii) Fit T_{fg}^{*} to window data minus $T_{21}(\nu_r, \Delta z)$, and compute χ_*^2

(ix) Normalize χ_*^2 by $\chi_{\dagger}^2$

(x) Compute rejection probability for $T_{21}(\nu_r, \Delta z)$

(xi) Find Δz with 95% probability at ν_r

(xii) Find maximum Δz with 95% probability at ν_r

(xiii) Find maximum Δz with 95% probability at ν_r

(xiv) Find maximum Δz with 95% probability at ν_r

The result of the process described above used on simulated data is shown in Figures 1-4.

PRELIMINARY AND INCOMPLETE results of the process applied to sky measurements with EDGES High-Band are shown in Figure 5.

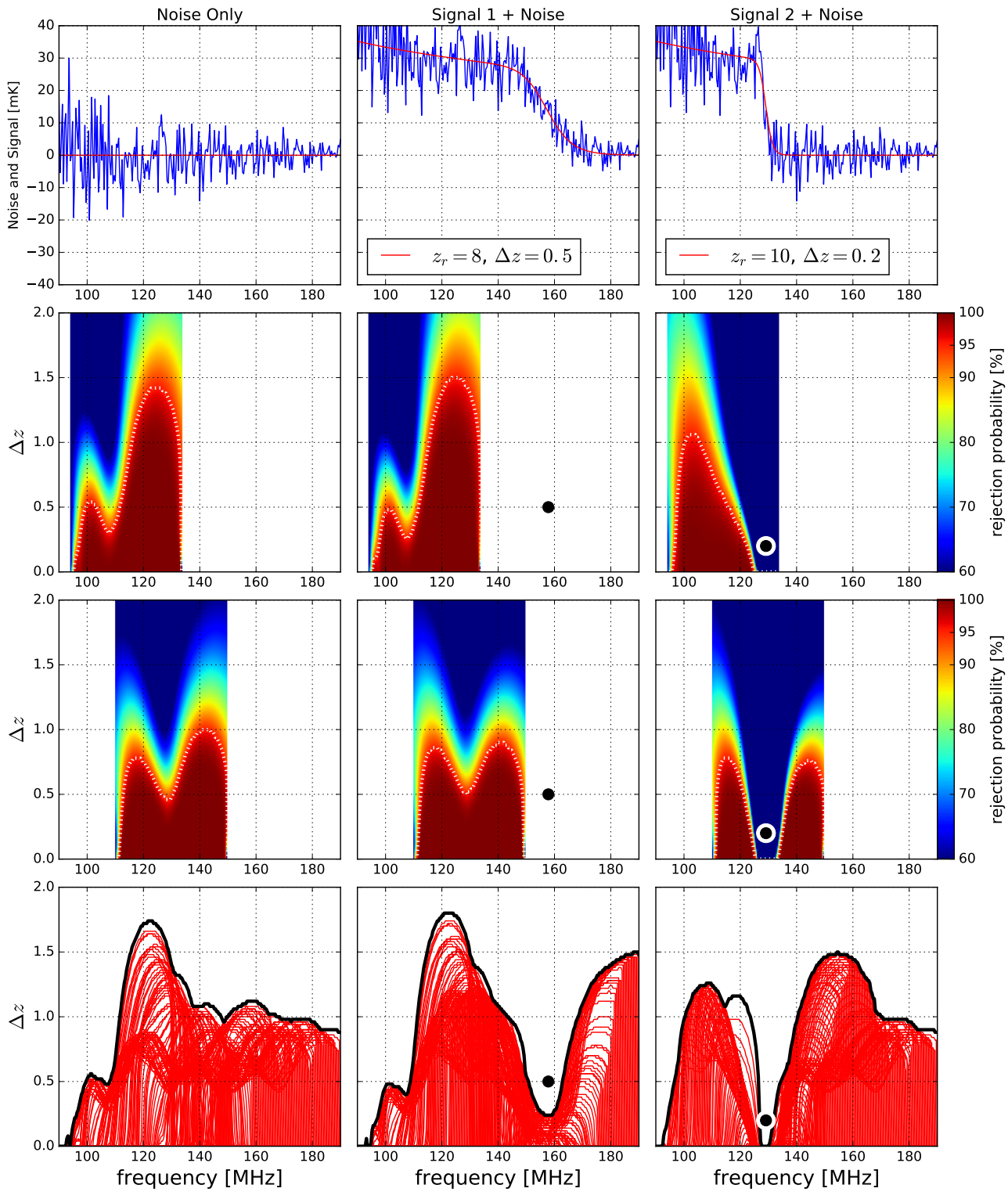

Figure 1: Summary of the χ^2 sliding window approach using simulated data. LEFT: Only noise added to the baseline power law spectrum. CENTER and RIGHT: Two Tanh models are added to the simulated data. TOP ROW: Signal. SECOND-THIRD ROWS: Rejection probabilities for two sample windows of 40-MHz width, modeled with 2 foreground terms. BOTTOM ROW: 95% rejection limits for all the 40-MHz windows modeled with 2 terms (IN RED), and maximum 95% rejection limits at each ν_r (IN BLACK). The black dots indicate the location of the input parameters. The rejected regions correctly fail to include these locations.

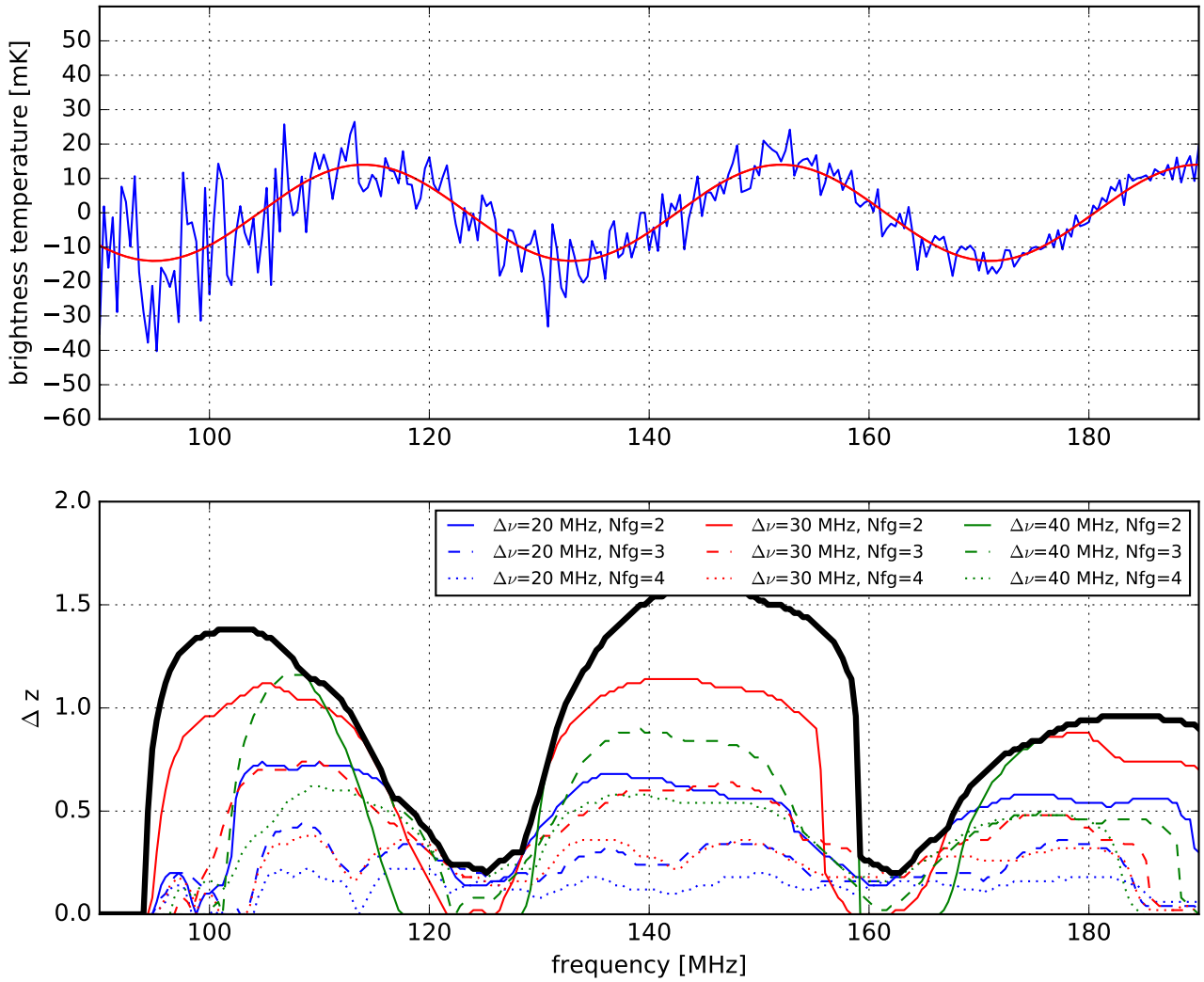

Figure 2: 95% rejection limits for the Tanh 21-cm model, for a simulated measurement that only includes noise and systematics in the form of a cosine function. Rejection limits are computed for three window widths (20, 30, and 40 MHz) and three number of foreground terms (2, 3, and 4). The black line represents the maximum rejection limits at each ν_r from all the cases considered.

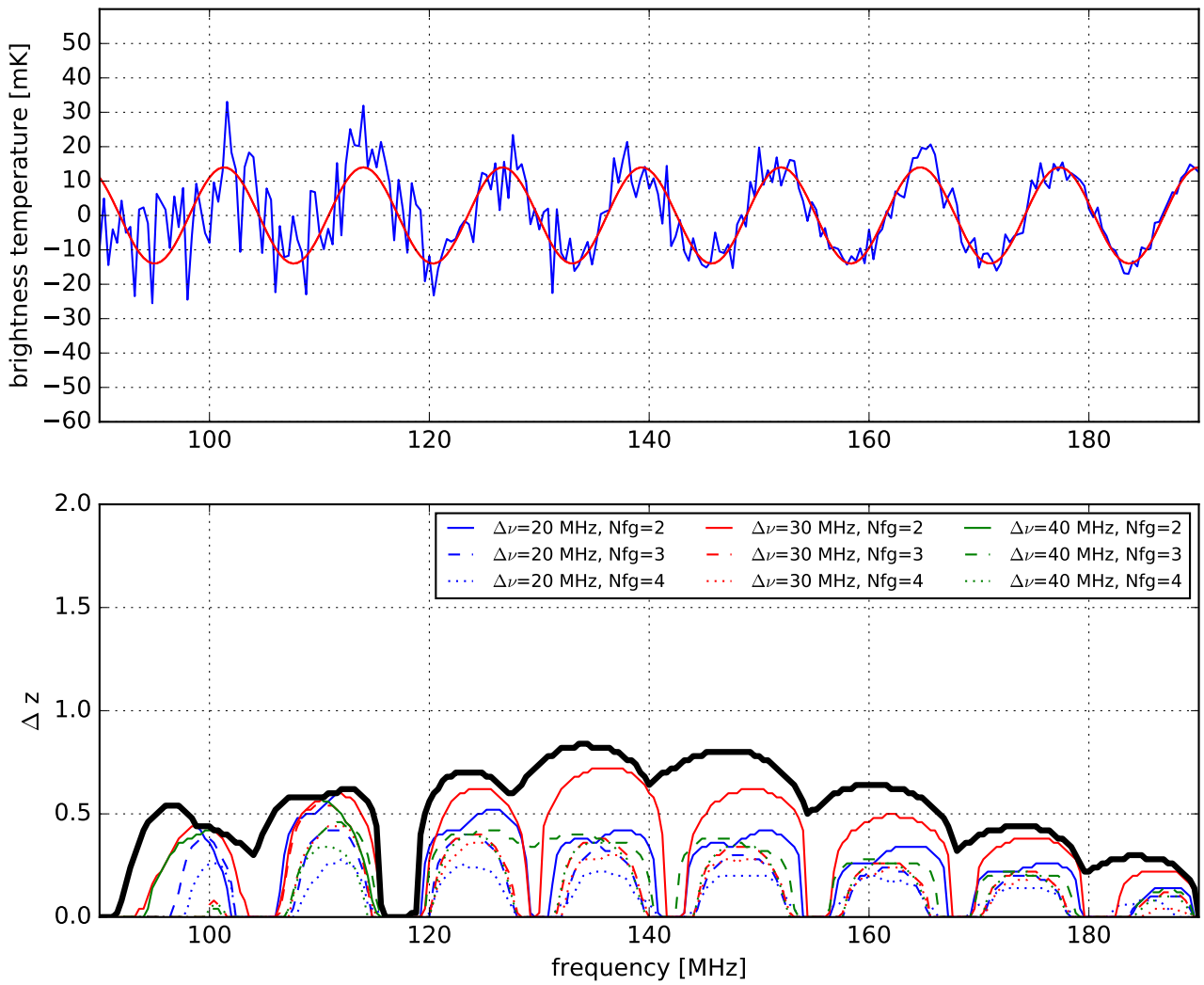

Figure 3: Same as Figure 2 but with systematics having a shorter period.

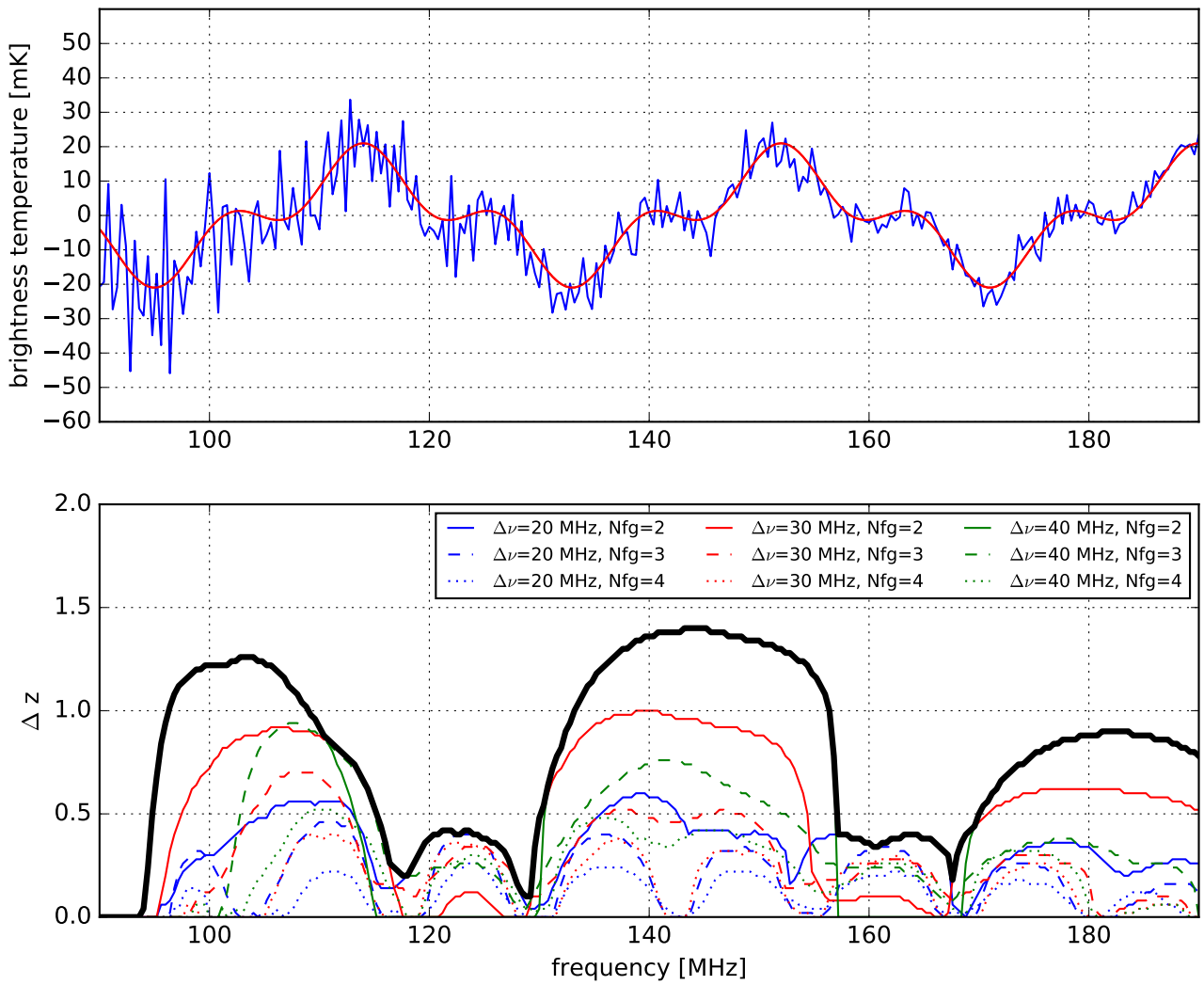

Figure 4: Rejections for a combination of short- and long-period systematics.

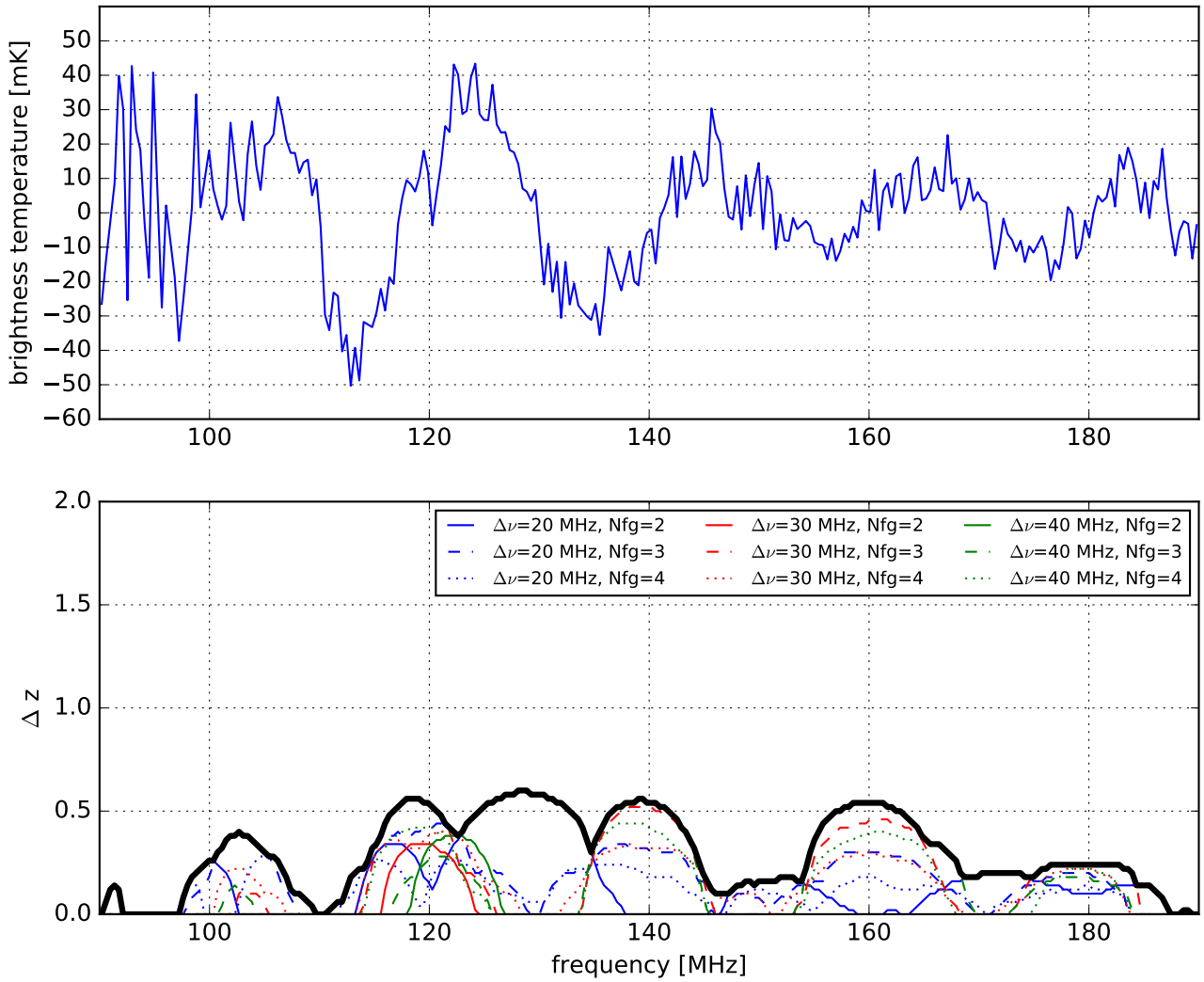

Figure 5: PRELIMINARY AND INCOMPLETE rejections for an average of EDGES High-Band data between days 2015-250 and 2015-299, over LSTs 0.26-6.26. We use the same, limited number of window sizes and foreground terms used as in Figures 2, 3, and 4. TOP PANEL: Residuals after a 5-term foreground fit to the test average spectrum.