CAMBRIDGE
UNIVERSITY PRESS

Research Article

χ^2 from redundant calibration as a tool in the detection of faint radio-frequency interference

Theodora Kunicki ![ORCID] and Jonathan C. Pober

Department of Physics, Brown University, Providence, RI, USA

Abstract

Radio-frequency interference detection and flagging is one of the most difficult and urgent problems in 21 cm Epoch of Reionisation research. In this work, we present χ^2 from redundant calibration as a novel method for RFI detection and flagging, demonstrating it to be complementary to current state-of-the-art flagging algorithms. Beginning with a brief overview of redundant calibration and the meaning of the χ^2 metric, we demonstrate a two-step RFI flagging algorithm which uses the values of this metric to detect faint RFI. We find that roughly 27.4% of observations have RFI from digital television channel 7 detected by at least one algorithm of the three tested: 18.0% of observations are flagged by the novel χ^2 algorithm, 16.5% are flagged by SSINS, and 6.8% are flagged by AOFlagger (there is significant overlap in these percentages). Of the 27.4% of observations with detected DTV channel 7 RFI, 37.1% (10.2% of the total observations) are detected by χ^2 alone, and not by either SSINS or AOFlagger, demonstrating a significant population of as-yet undetected RFI. We find that χ^2 is able to detect RFI events which remain undetectable to SSINS and AOFlagger, especially in the domain of long-duration, weak RFI from digital television. We also discuss the shortcomings of this approach and discuss examples of RFI which seems undetectable using χ^2 while being successfully flagged by SSINS and/or AOFlagger.

Keywords: Radio interferometry; chi-squared statistic; outlier detection

1. Introduction

During the Universe's Epoch of Reionisation (EoR), which took place when $20 \gtrsim z \gtrsim 6$, the intergalactic medium (IGM) transitioned from being almost entirely neutral to an effectively completely ionised state. While the process of reionisation is generally expected to have been driven by ionising radiation from early stars, this period of the Universe's history has yet to be directly observed. The details of reionisation are sensitive to other important cosmological factors such as the process of early galaxy formation, so its observation is of high potential value to experimental cosmology. Two review articles which describe the EoR in depth are Choudhury (2022) and Wise (2019).

The most promising experimental probe into the EoR is 21 centimeter radiation emitted by neutral hydrogen gas in the IGM. Neutral hydrogen gas emits radiation at a wavelength of 21 centimeters owing to the spin-flip transition in the atom; notably, this transition requires both an electron and a proton, so 21 cm radiation is *not* emitted by ionised hydrogen. By mapping the EoR's 21 cm radiation, cosmologists hope to learn about the spatial distribution of neutral hydrogen in the IGM as the EoR progressed. For reviews of 21 cm cosmology, see, for example, Morales & Wyithe (2010), Furlanetto, Oh, & Briggs (2006), Liu & Shaw (2020) and Pritchard & Loeb (2012).

Efforts to observe the 21 cm power spectrum from the EoR are underway at LOFAR[a] (van Haarlem et al. 2013), HERA[b] (DeBoer et al. 2017), and the Murchison Widefield Array[c] (Tingay et al. 2013; Wayth et al. 2018).

Detecting the 21 cm EoR power spectrum is difficult for multiple reasons, and chief among them is the contamination introduced to the data by anthropogenic radio-frequency interference (RFI). Removing RFI-contaminated data from a power-spectrum analysis is essential if the measurement is to be sensitive to the 21 cm EoR signal (Wilensky et al. 2020). While LOFAR is situated in the Netherlands, where RFI is omnipresent, HERA and the MWA are located in more radio-quiet sites in South Africa and Western Australia, respectively. Even in these radio-quiet sites, there are many sources of RFI, including digital television (DTV) (which may even be reflected by passing aircraft Wilensky et al. 2019), ORBCOMM satellite transmissions (Sokolowski, Wayth, & Lewis 2015), and FM radio, which may reflect from satellites (Zhang et al. 2018). The MWA, in particular, is located far from any transmitters, and still observes DTV regularly, as well as narrow- and broad-band RFI from time to time (Offringa et al. 2015).

There are two software packages regularly used in RFI detection at the MWA, AOFlagger (Offringa et al. 2015), and SSINS

[a]https://www.astron.nl/telescopes/lofar/.
[b]https://reionisation.org/.
[c]https://www.mwatelescope.org/.

(Wilensky et al. 2019). While these software tools are both powerful, it has been demonstrated that they fail to excise all RFI, leaving behind so-called 'ultra-faint' RFI in the data (Wilensky et al. 2023). In Wilensky et al. (2023), statistical methods were used to demonstrate that this 'ultra-faint' RFI is present in existing datasets that have been used for analysis, and that even this very weak residual RFI is extremely deleterious to the effort to accurately measure the 21 cm EoR power spectrum.

In a radio interferometer, a baseline is defined as the displacement vector which separates two antennas whose signals are to be correlated by the instrument. Both HERA and the MWA have baseline redundancies, i.e., there are baseline vectors which are repeated multiple times throughout the array. This baseline redundancy is visible as translationally symmetric layouts in all (in the case of HERA) or a subset (in the case of the MWA) of antennas.[d] Baseline redundancy was built into these interferometers in order to increase sensitivity on these baselines (Parsons et al. 2012), and also to make use of redundant calibration (Wieringa 1992), which uses the fact that identical baseline vectors should measure identical visibilities. In a redundant calibration algorithm, a metric called χ^2, which measures a baseline's visibility's noise-normalised square deviation from a calculated 'consensus' visibility for that baseline group, is minimised for each baseline by adjusting antenna gains (see Section 2.3 for a formal definition of the χ^2 used in redundant calibration).

In this work we present an alternative use of the χ^2 metric from redundant calibration as an RFI detection tool. This approach has been proposed before in Zheng et al. (2014), but has not yet been investigated as a standalone technique. Li et al. (2019) observed elevated χ^2 in MWA data corresponding to RFI and used it to hand-flag some contaminated data, but did not include a systematic search using χ^2 to automatically flag data or a comparison to the RFI detected by SSINS and AOFlagger. In this work, we will demonstrate that χ^2 from redundant calibration is not only effective at detecting RFI in MWA data, but can also detect RFI which is not visible to either SSINS or AOFlagger.

The structure of this paper is as follows: Section 2 describes our dataset and preprocessing, provides an overview of redundant calibration, and introduces the χ^2 metric, which will be the focus of this work. Section 3 discusses χ^2's sensitivity to RFI, and describes our method used to generate flags based on RFI detected in χ^2. Section 4 briefly describes AOFlagger and SSINS, the performance of which is to be compared to χ^2 flagging. Section 5 presents the results of our analysis, including describing the specific RFI that is detected with this method, and presenting a comparison of χ^2 with AOFlagger and SSINS on their performance flagging the most common types of RFI observed in the dataset. Section 6 provides a discussion of the results and their implications for 21 cm cosmology, and concludes the work. This paper has three short appendices; in Appendix A, we discuss the effects of time-averaging on χ^2 generated by redundant calibration, in Appendix B, we discuss the specific parameters used to run redundant calibration, and Appendix C is a collection of interesting and exemplary RFI events, visualised using all three flagging algorithms.

---

[d]Antennas in the MWA are usually referred to as 'tiles' and consist of 16 crossed dipoles of less than a meter in any dimension arranged in a 4×4 grid over a ground screen. In this work, we will use the terms 'tile' and 'antenna' interchangeably to refer to MWA tiles.
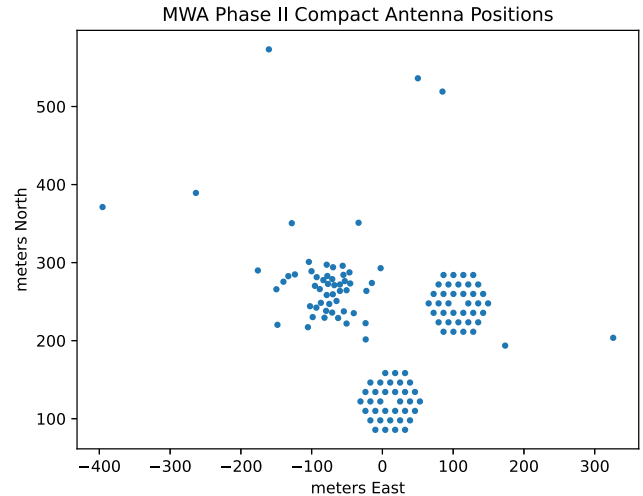

Figure 1. Antenna positions in the MWA Phase II compact configuration.

2. Data and redundant calibration

2.1 Data

Data were downloaded from the MWA All-Sky Virtual Observatory (ASVO)[e] service in MWA correlator FITS format. The data lie in the band between 167.055 MHz and 197.735 MHz and were collected by the MWA between the dates of October 15th and December 15th of 2016. This period coincides with the beginning of MWA Phase II operations, and these data were taken with the array in its 128-antenna 'compact' configuration (Wayth et al. 2018), which includes 71 antennas in two redundant, compact hexagons of 36 and 35 antennas (both hexagons were intended to comprise 36 antennas, but one antenna location was inaccessible on site Li et al. 2018).

Fig. 1 illustrates the antenna positions of the Phase II compact configuration of the MWA.

Within the dataset, there are observations of both the EoR0 (R.A. = 0.0°, Dec. = -27.0°) and EoR1 (R.A. = 60.0°, Dec = -30.0°) fields (Jacobs et al. 2016).

The EoR0 observations in this set are the same data analysed in Li et al. (2019), and (apart from three nights of data analysed in Zhang et al. 2020) the EoR1 data are heretofore unanalysed. The dataset comprises 51.8 h of observations, spanning 1 665 approximately two-minute observation files.

For both the EoR0 and EoR1 datasets, the data include several pointings of the MWA, which are accomplished instrumentally by introducing relative delays into the analog signals from each of the 16 individual dipoles which make up each MWA tile, changing the angle of the beam's phase center on the sky. In a single night of observing, the same field on the sky is tracked by allowing it to drift through the beam several times as the beam is phased in discrete increments to follow it. According to the grid numbering scheme in Beardsley et al. (2016), the pointings included in the dataset are -3 to $+2$ for the EoR0 observations, and -3 to $+3$ in the EoR1 observations (where pointing 0 is an observation with the beam center at zenith).

---

[e]https://asvo.mwatelescope.org/.

The MWA correlator used during Phase II of operations introduced a coarse-band structure to the data, which is discussed in detail in Section 3.4. Briefly, due to the design of the correlator, there are certain frequency bins (corresponding to the edges and centers of 24 coarse bands) which always contain corrupted data. Throughout this paper, we will mask out corrupted data in these bins and will always note when an algorithm has been modified to work around these masked sections of data.

2.2 Data preprocessing and redundant calibration

Before they are run through redundant calibration, the data are downsampled in time using the python package pyuvdata[f] version 2.4.0 (Hazelton et al. 2017). The files are time-averaged from a time cadence of one visibility every 2 s to one visibility every 18 s; this is to reduce noise in the observations, which at a 2-s cadence prove too noisy to allow the redundant calibration algorithm to arrive at reasonable gain solutions. The rationale for this time-averaging is discussed further and example images are provided in Appendix A.

Pyuvdata is likewise used to select down to only the 71 antennas in the redundant section of the array. This is because the software package used in this work for running redundant calibration, hera_cal, was developed for use with the HERA telescope, which is fully redundantly laid out (Dillon & Parsons 2016), as opposed to the MWA which has some redundant antennas alongside pseudorandomly positioned antennas. The software is therefore likely to calculate spuriously 'redundant' baselines when all of the pseudorandomly positioned antennas are included.

2.3 Redundant calibration

Once the data have been preprocessed, hera_cal[g] (Dillon et al. 2020) is used to run redundant calibration, using hera_cal's redcal_run() method.

Redundant calibration takes advantage of the fact that baselines which have the same vector separating their component antennas should in principle measure the same visibility from the sky. We assume that a measured visibility for a given time t, frequency ν, and polarisation p, between antennas i and j, v_{ij} can be expressed as the product of two antenna-dependent complex gains, g_i and g_j and a 'true' sky visibility y_{ij}, with an added noise term n_{ij}

$$v_{ij}(t, \nu, p) \approx g_i(t, \nu, p)g_j^*(t, \nu, p)y_{ij}(t, \nu, p) + n_{ij} \quad (1)$$

We drop the explicit time, frequency, and polarisation dependencies in much of the following discussion, but it should be emphasised that redundant calibration provides an independent solution for the gains and visibilities at every unique combination of time, frequency, and polarisation. This also means that we get a value for χ^2 at each time, frequency, and polarisation.

The goal of calibration is to solve for the gains g_i. We go about this by minimising χ^2, defined as

$$\chi^2 = \sum_{i<j} \frac{\left| v_{ij} - g_i g_j^* y_{ij} \right|^2}{\sigma_{ij}^2} \quad (2)$$

where y_{ij} has subtly changed in meaning to be a 'consensus' visibility calculated for the baseline group (a 'baseline group' being

a collection of baselines with identical baseline vectors), and σ_{ij}^2 is an estimation of the noise variance of that baseline type. The consensus visibility, y_{ij}, of a group of identical baselines is a free parameter to be solved for in our system of linear equations, alongside the gains and visibilities, and its value is such that the overall χ^2 values for that baseline group are minimised.

Minimising these χ^2 values is our objective as we adjust parameters in a large system of linear equations. In the redundant subsection of the MWA, there are 71 antennas, and if we only consider redundant baseline groups containing two or more baselines, that leaves 181 unique baseline types; 71 complex gains and 181 complex visibilities make for 252 complex free parameters to solve for, but the number of equations is much larger at $(71 \times 70)/2 = 2\,485$ (the total number of baselines in the redundant subsection of the array), making this system overdetermined.

χ^2 and its sensitivity to radio-frequency interference will be the focus of this work. In order to obtain χ^2 for analysis, we must run redundant calibration using hera_cal. However, it is vital to note that the results of this redundant calibration process – the gains and visibilities generated by the algorithm – will not be used in our analysis. Only the χ^2 values are of interest in finding and flagging RFI.

There are three different algorithms, run in series on the uncalibrated data, used to solve this system of equations in hera_cal. The first is aptly named firstcal, the second is known as logcal, and the third is called omnical.[h] For an in-depth description of how these algorithms work, please consult Li et al. (2018) and Dillon et al. (2020).

3. RFI identification and flagging with χ^2

3.1 χ^2 and degrees of freedom

As hera_cal implements the above algorithms to estimate gains g_i and visibilities y_{ij}, it does so with the goal of minimising the metric χ^2, defined in equation (2). This is the overall χ^2 for the entire array (at each time, frequency, and polarisation), found by summing over each baseline ij, comparing a baseline's measured visibility with the visibility calculated by multiplying the antenna gains g_i and g_j^* by the calculated 'consensus' visibility y_{ij}. The difference of these quantities is squared, and then divided by an estimation of the noise variance σ_{ij}^2. If the measured visibility and calculated visibility only differ by a factor of noise, this quotient should be of order 1. When summed over the entire system, we would expect χ^2 of a well-calibrated array to be of order nDoF, or the number of complex degrees of freedom in the system, calculated by Dillon et al. (2020) to be

$$\text{nDoF} = N_{bl} - N_{ubl} - N_{\text{ants}} + 2.5 \quad (3)$$

where N_{bl} is the total number of baselines, N_{ubl} is the number of unique baseline types, or alternatively the number of y_{ij}

---

[f] Available at https://github.com/RadioAstronomySoftwareGroup/pyuvdata.

[g] Available at https://github.com/HERA-Team/hera_cal, version 3.4.1.dev2+g102e14f8 was used in this work.

[h] Further complicating the history of these algorithms, the software package described in Zheng et al. (2014) is called OMNICAL, and contains three algorithms: logcal, lincal, and omnical (where we use lowercase letters to distinguish the algorithm omnical from the software package OMNICAL). lincal and omnical perform the same task with different solvers. Previous work applying redundant calibration to the MWA (Li et al. 2018; Li et al. 2019, and Zhang et al. 2020) use the OMNICAL package to run logcal and lincal but not omnical. To be explicit, in this work we use the hera_cal software package (not OMNICAL) and the omnical algorithm (not lincal).

that need to be solved for, and N_{ants} is the total number of antennas, understood as the number of individual gains g_i to be solved for.

The additive factor of 2.5 accounts for the degeneracies of redundant calibration in the array, which reduce the total number of parameters which can be solved for in our system of linear equations, increasing the total number of degrees of freedom. In Dillon et al. (2020), this constant factor is calculated to be 2 in a classical redundant array. In such an array, these 2 complex numbers or 4 real parameters correspond to overall amplitude, overall phase, tip/tilt of the array in the N/S direction, and tip/tilt of the array in the E/W direction. However, due to having two spatially separated redundant hexagons, the MWA introduces an additional real degenerate parameter – the phase separation between the two redundant hexagons – which increases the number of real degeneracies in the system from 4 to 5, and the number of complex degenerate parameters in turn from 2 to 2.5 (Dillon 2024).

Taking χ^2/nDoF, we expect a number close to 1.0 for each time and frequency if redundant calibration has been successfully run and the resulting calibration is accurate.

χ^2 is intrinsically a measure of how redundant the array actually is. Said another way, χ^2 measures how well the calibrated baseline groups agree with each other on a measured visibility; in a perfectly calibrated, perfectly constructed array, each baseline should measure identical visibilities. Higher χ^2 values for a given time or frequency indicate that nominally 'redundant' baseline groups do not agree on their measured visibilities.

We have observed that χ^2/nDoF is elevated by radio-frequency interference (c.f. the lower plot in Fig. A1). There are two main ways this could occur: either the presence of RFI increases χ^2 by increasing non-redundancy (the numerator of χ^2), or the presence of RFI leads to an underestimation of noise variance σ_{ij}^2 (the denominator of χ^2).

In order to determine which of these mechanisms was responsible for heightened χ^2 in the presence of RFI, we separated the numerator and denominator of equation (2). Within `hera_cal`, the antenna autocorrelations, which measure the total power incident on each antenna element, are used as a proxy for σ^2. We do find that some powerful RFI events slightly elevate the total power seen in the autocorrelations, and hence elevate `hera_cal`'s estimation of σ^2. However, this has the effect of suppressing χ^2, opposite to the overall effect in χ^2 we report here. It follows that RFI elevates χ^2 by magnifying or exacerbating non-redundancy in the array, i.e., it increases the numerator of equation (2).

3.2 Modified z-score as a tool to flag outlying χ^2 data

The simplest approach to flagging RFI in χ^2 data would be to set an absolute threshold for χ^2/nDoF, and to flag data which exceed that threshold. However, both the EoR0 and EoR1 fields have an overall Local Sidereal Time (LST) dependence in the χ^2 values. This makes it difficult to define an absolute χ^2/nDoF cutoff for flagging data, since what may be an elevated χ^2 value for one observation can be the χ^2 value of perfectly uncontaminated data for another. Fig. 2 illustrates this LST-dependence with the North-North polarisations of EoR1 observations. Y-axis locations represent the mean of χ^2/nDoF for each observation, and the error bars signify their standard deviations. Points are color-coded by integer modified Julian date, such that data taken during the same 24-h period

are of the same color. Different marker shapes represent different pointings of the MWA beam.

In the EoR1 data especially, there is a clear upward trend in each individual pointing's χ^2/nDoF means. This sub-pointing trend is likely due to bright celestial objects moving through the sidelobes of the individual tile beams, which are much less redundant between the tiles of the MWA than the beam's main lobe (Line et al. 2018; Chokshi et al. 2021); see Choudhuri, Bull, & Garsden (2021) for a similar effect seen in simulations of HERA. Having the same bright sources cross our beam multiple times as the beam pointing is adjusted is natural in 'drift and shift' observations such as these, where the primary beam pointing is adjusted to different angles at regular intervals throughout the observing period.

Taking the modified z-scores of the χ^2/nDoF data greatly reduces their LST-dependence, as shown in Fig. 3, which represents the same data as Fig. 2. Modified z-score for the ith element of a dataset x is calculated according to the following formula:

$$z_i = 0.6745 \times \frac{(x_i - \tilde{x})}{(\text{MAD})} \tag{4}$$

where $\tilde{x}$ represents the dataset's median, and MAD is the median absolute deviation of the data from the median.

$$\text{MAD} = \widetilde{|x_i - \tilde{x}|} \tag{5}$$

Here, the tilde (~) operator has been used to signify 'take the median of the dataset.' The factor of 0.6745 is included to scale modified z-scores to have the same step size as standard z-score, i.e., if the standard z-score and modified z-score are both taken of the same perfectly Gaussian dataset, each integer step in modified z-score is the same size as a step in standard z-score, and corresponds to 1 standard deviation of the dataset.

As opposed to the standard z-score, modified z-score uses medians instead of means. As a result, the measure is less affected by outlier values, and more appropriate for datasets that have a mostly Gaussian distribution, aside from some high z-score outliers (in this case caused by RFI).

Modified z-score assumes that the data have a distribution that is close to Gaussian; this is a valid assumption for these data because, even though a χ^2 distribution is not Gaussian, due to the central limit theorem, as nDoF increases, so does the distribution's resemblance to a Gaussian. For the redundant sub-array of the MWA, as calculated with equation (3), nDoF = 2 228, which is more than enough to warrant making this simplification.

Because it almost completely eliminates LST-dependence of χ^2, we base our RFI-flagging algorithm on the modified z-scores of χ^2/nDoF generated by redundant calibration.

As shown in Fig. 3, most files have a mean z-score which is near 0.0 and a standard deviation of z-scores which is close to 1.0. There are also files with noticeably higher mean z-scores, and wider standard deviations, which are almost always caused by strong RFI events. However, rather than further analyse the distribution of χ^2 values within a single observation, we turn to the time, frequency, and polarisation information of the χ^2's, which help us localise and characterise the source of the interference.

3.3 Examples of χ^2 and its sensitivity to RFI

To demonstrate the sensitivity of χ^2 to RFI, we have included here a few χ^2 time/frequency waterfall plots and histograms of the corresponding data in Fig. 4.
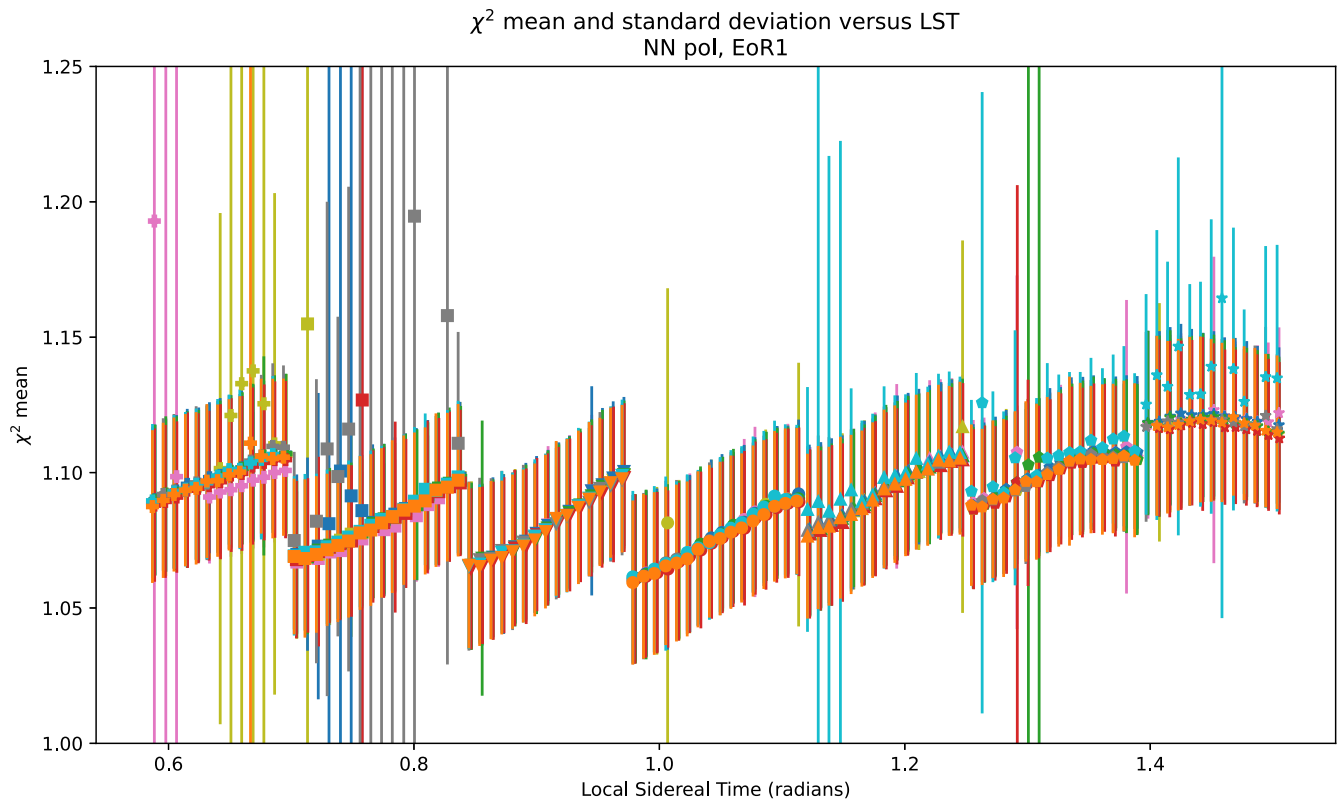
Figure 2. The mean value of χ^2/nDoF versus local sidereal time; each point represents one two-minute observation, with the standard deviation of χ^2/nDoF within the observation represented as an error bar. Each color represents data from a different night, while each shape represents data from a different pointing. The χ^2 means have a pronounced LST dependence. Within individual instrument pointings, there is also an upward trend. Because of the dependence on LST, a single χ^2/nDoF cutoff value applied across all observations in sub-optimal for flagging RFI.

While the RFI-free file produces a distribution of χ^2 that is close to Gaussian, the presence of RFI leads to outliers. Included are four examples of DTV RFI of differing strengths. Common to all four of them is a bimodal distribution of χ^2, which generally remains below $z = 40$, but, even in the example of the weakest RFI, have bins above $z = 4$.

On the other hand, the RFI in the final waterfall plot of Fig. 4 may be difficult to see, because it only spans one frequency-bin at 196.175 MHz. It is worth noting, however, that the χ^2 z-score values are exceptionally high for these six bins, as demonstrated in the histogram.

3.4 RFI flagging

RFI appearing in the χ^2 data is flagged with a two-step process consisting of iterative modified z-score flagging followed by watershed flagging.

First, the modified z-scores of the χ^2 values are calculated for the dataset. Each point with a modified z-score greater than 4.0 is first masked out of the dataset. After high-z-score datapoints have been flagged, z-scores are recalculated without the outlier points, and the process is repeated until it produces no update in flags from its most recent iteration.

Starting from this preliminary set of flags, we follow the watershed RFI flagging algorithm described in Appendix A of Kerrigan et al. (2019). Briefly, this flagging algorithm works on

the assumption that RFI-contaminated bins tend to be contiguous, and flags every bin with a modified z-score greater than 2.0 which lies orthogonally adjacent to an already-flagged bin. This process is iteratively repeated until the flags remain unchanged from the last iteration, thereby 'flooding' the lower-z-score data which are directly connected to higher outliers.

This algorithm has to be slightly modified in order to handle the coarse-band flagging which is necessary for Phase II MWA data. In general, when working with MWA Phase II data, within each 1.28 MHz, 32-channel coarse band, two channels on either edge of the band as well as the central channel are flagged and excluded from analysis due to data corruption arising from aliasing from the poly-phase filter bank (for the edges) and DC offsets (for the center) (Offringa et al. 2015). These flags are apparent in the masked values in the waterfall plots of Fig. 4. As a result, the function for finding 'adjacent' bins has to be modified to ignore the coarse band edges. When a bin being processed by the watershed algorithm is adjacent to a coarse band flag, the algorithm skips the flagged coarse feature and considers the next bin beyond it as adjacent to the flagged bins on the other side for 'flooding.'

This two-step flagging algorithm is far from perfect. Even after watershed flagging, there are datapoints which are clearly contiguously part of an RFI event which remain unflagged due to their low z-scores. One such event is shown in Fig. 5. While the two-step algorithm has nearly completely flagged contiguous points with high modified z-score, we can identify this RFI event as a DTV
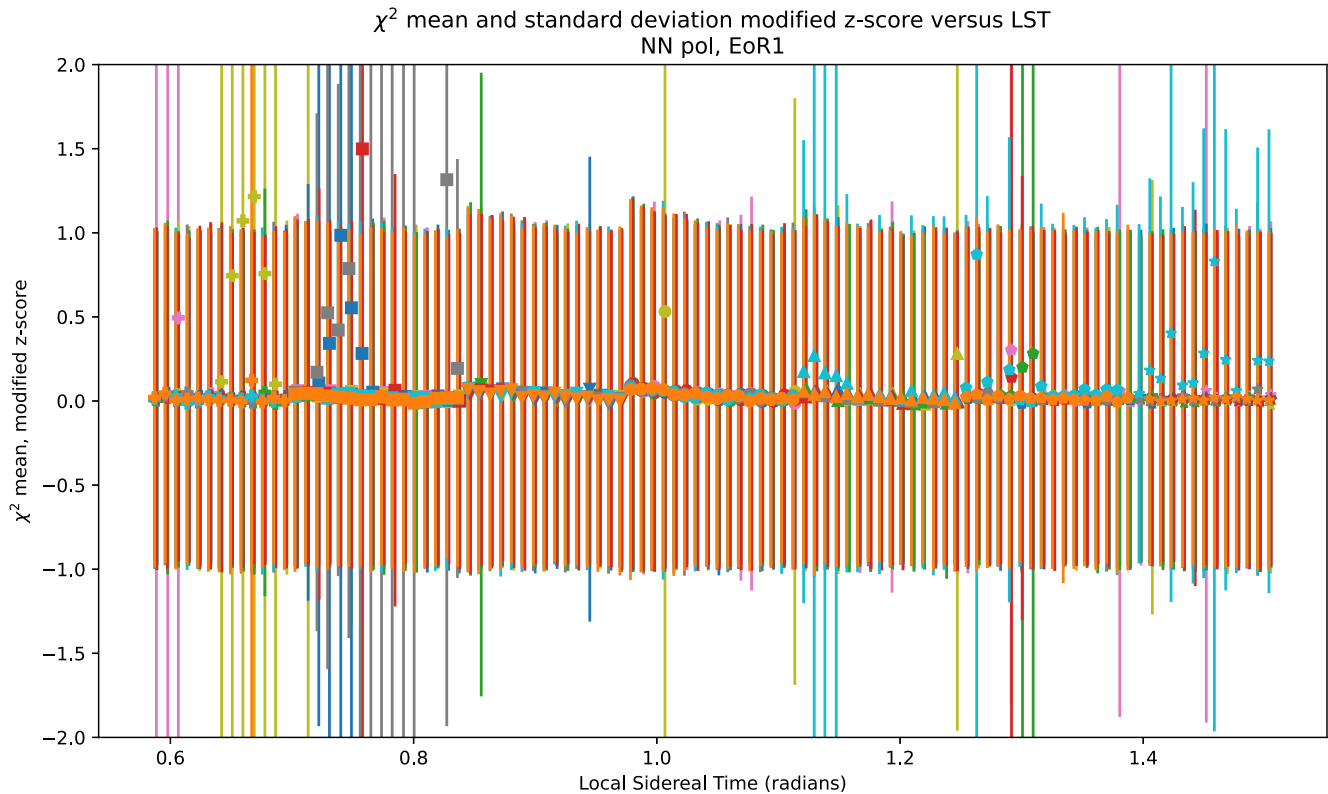
Figure 3. This plot represents the same data as Fig. 2, except now modified z-scores have been taken of the χ^2/nDoF data. The mean of the modified z-scores of χ^2/nDoF is plotted against local sidereal time, with the standard deviation of the modified z-scores of the data represented as error bars. The LST dependence of χ^2 present in Fig. 2 is effectively eliminated by using modified z-score. As with Fig. 2, color denotes the day data were taken, and different point shapes represent different antenna pointings of the instrument.

channel 7 broadcast due to its frequency band. Using this deduction, we can assume *a priori* that the RFI should occupy the entire channel 7 band, which is not fully flagged by this code.

In this work, we focus on demonstrating the potential for redundant calibration χ^2 to detect RFI events. Future work is likely to improve on the exact algorithm for calculating flags from the χ^2 values themselves. One might, for instance, explore averaging similar data together to reduce noise in χ^2, since we can expect noise to reduce with averaging whereas consistent RFI will not. Furthermore, the prospect of using a machine-learning approach to flag RFI events is especially interesting, based on the heuristic that they are easy for the human eye to identify, while still being difficult to algorithmically define.

4. Other RFI-flagging algorithms: AOFlagger and SSINS

In order to compare χ^2 flagging with existing RFI flagging algorithms, we must obtain these algorithms' flags for the same dataset.

For the purposes of this work, we have compared the performance of χ^2 flagging with AOFlagger as part of the MWA `cotter` pipeline (Offringa et al. 2015) and SSINS (Wilensky et al. 2019), run locally.

4.1 AOFlagger

AOFlagger operates in three broad steps. First, a high-pass filter is applied to the visibility amplitudes. The filtered visibility

amplitudes are then run through a straight-line detecting algorithm called SumThreshold, introduced in Offringa et al. (2010), based on the observation that RFI often has sharp edges in the frequency-time domain. Finally, the scale-invariant rank (SIR) operator, introduced in Offringa, Van De Gronde, & Roerdink (2012), is applied to the SumThreshold-flagged visibilities, which acts to propagate those flags into adjacent contaminated bins.

AOFlagger flags are generated by `cotter` (the processing pipeline in use by the MWA during Phase II operations), and included when downloading MWA data from the data download service ASVO. The AOFlagger algorithm generates flags by baseline, so in order to visualise the flagging patterns for an entire individual observation, we average over the baseline axis and view the average occupation fraction for each frequency/time bin for a given polarisation.

4.2 SSINS

SSINS is short for Sky-Subtracted Incoherent Noise Spectra, and is an RFI-flagging algorithm first introduced in Wilensky et al. (2019).[i] SSINS begins with 'sky-subtraction,' whereby every visibility is subtracted from the next visibility adjacent in time, removing features that vary much more slowly than the two-second cadence of most MWA data, including the sky. Once sky-subtraction has been applied to every baseline's visibilities, the

---

[i]https://github.com/mwilensky768/SSINS.
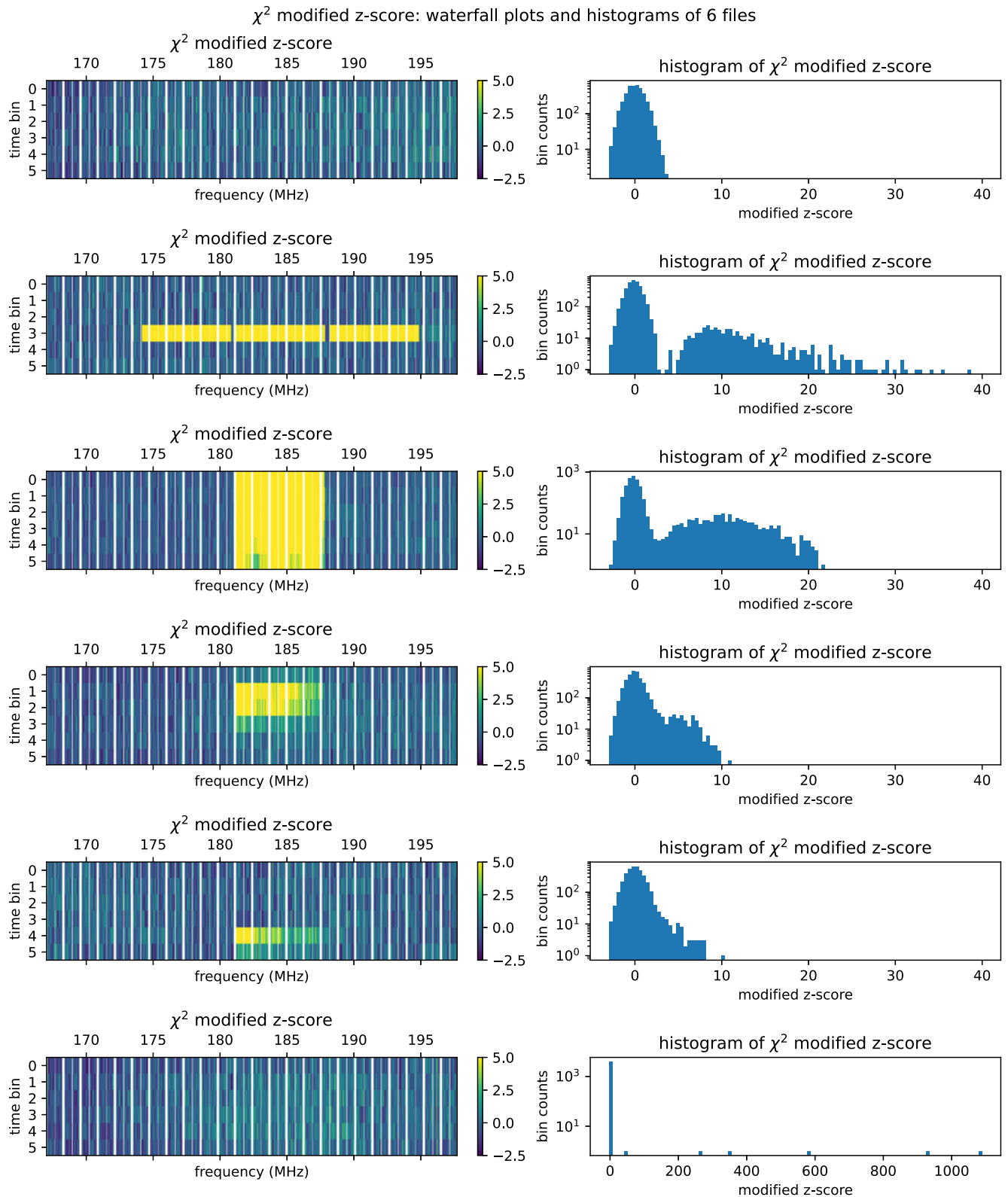
Figure 4. Waterfalls and histograms of the modified z-scores of χ^2/nDoF of 6 two-minute observations. The waterfall plots, left, show the modified z-score of χ^2/nDoF as a function of time and frequency. Elevated values are indicative of RFI. To the right of each waterfall plot is the corresponding histogram, showing the distribution of the modified z-scores of the data. The first observation is RFI-free and shows a nearly Gaussian normal distribution, whereas the subsequent examples all contain RFI and have high-z-score outliers of varying degree. The second through fifth images all represent DTV events, while the sixth is a narrow-band event at 196.175 MHz.
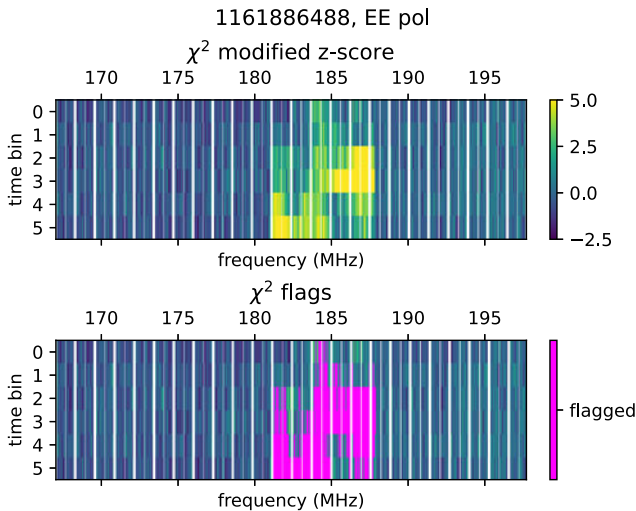
Figure 5. An example χ^2 waterfall plot of an incompletely-flagged channel 7 DTV event. The top panel shows the calculated z-scores, while the bottom panel shows the derived flags in magenta. This event was flagged using the algorithm described in Section 3.4, but the channel 7 DTV band is only partially flagged for any given time.

Figure 6. The fraction of data flagged using χ^2 across all observations as a function of frequency. The most commonly flagged RFI (in approximately 5 to 6% of the data) is channel 7 DTV, as indicated by the large feature between 181 and 188 MHz. Also visible are slight amounts of flagging in channels 6 (174–181 MHz) and 9 (195–202 MHz). Narrow-band RFI is apparent at 196.175 and 194.455 MHz. Most other narrow-band peaks correspond to MWA coarse band edges, indicated by the shaded regions and are unlikely to be actual RFI. The slight increase in flagging at the lowest frequencies is due to spurious flagging of elevated χ^2 stemming from the presence of the Galactic plane in the sidelobes of the primary beam (see Fig. 7).

data should consist of only noise and RFI. The sensitivity to RFI is increased by forming an 'incoherent noise spectrum,' or INS, of the sky-subtracted visibilities. This is done by averaging the visibility amplitudes (discarding phase) over all baselines in the array. One INS per polarisation is produced.

SSINS uses a frequency-matched flagging algorithm to find features in the sky-subtracted incoherent noise spectrum. This algorithm is aware of DTV-bands and will flag out the entire channel band if enough DTV RFI is detected in that channel. The flagging algorithm is also sensitive to narrow-band RFI and broad-band streaks.

SSINS version 1.4.7 was run locally on the downloaded data. In order to reduce the number of spurious flags, the incoherent noise spectrum (INS) object was calculated with a 2nd order polynomial fit for each frequency channel during mean subtraction.

5. Results

5.1 Average flagging occupation fractions

When the χ^2 flags are averaged across all of the files in the dataset, as visualised in Fig. 6, it becomes clear that RFI falling in the band allocated to Australia's digital television channel 7 (181–188 MHz) comprises the bulk of the RFI which is flagged. Looking closely, we see that DTV channels 6 (174–181 MHz) and 9 (195–202 MHz) are also visible as slight elevations in flagging fraction.

In addition to the strong peak at channel 7 frequencies, there are some noticeable narrow-band peaks. Most of these narrow-band peaks in bins adjacent to the masked-off bins corresponding to the MWA correlator's coarse bands, which are widely known to contaminate data, as described in Section 3.4. Coarse band flags are represented by the shaded zones in Fig. 6, and data from these frequency bins have been excluded from this plot. Most of the observed narrow-band peaks in flagging fraction are adjacent to these flagged channels, which perhaps indicates that the data corruption extends beyond the commonly-flagged channels, although

the flagging fraction is only 1 − 2%, so it follows that their χ^2 values are not consistently elevated. The cause of the elevated χ^2 numbers that occasionally appear around the coarse band edges remains a topic for further investigation.

Of the narrow-band flagging peaks, two prominent examples are *not* adjacent to coarse band edges. These represent real narrow-band RFI, each contained within a single 40 kHz frequency-bin at central frequencies of 196.175 and 194.455 MHz for the higher and lower peaks, respectively.

The last obvious feature of Fig. 6 is a small increase in flagging at low frequencies, which is due to the spurious flagging of low-frequency bins whose χ^2 values are elevated – not by RFI, but by the presence of the Galactic plane in the far sidelobes of the primary beam. Affected observations, like the one visualised in Fig. 7, are uniformly from gridpoint number −3 of the grid pointing scheme used in Beardsley et al. (2016), with az = 90°, el = 69.1655° – one of the more extreme pointings of the instrument used to observe the EoR0 field which appear in this dataset. The affected data were taken when the Galactic anti-center was still above the horizon, where it could be picked up by sensitivity in the sidelobes of the beam. Although using modified z-score for flagging decreases the number of spurious flags in this pointing, it does not reduce them to zero. This results in the uptick in χ^2 flagging at the lowest frequencies observable by the MWA. Future work looking to apply RFI flags generated from χ^2 values to MWA data may need to handle the analysis of this pointing separately.

5.2 Comparing χ^2 flagging with SSINS and AOFlagger

The most common type of RFI detected in this dataset was channel 7 DTV, which is broadcast between 181 and 188 MHz. This section focuses on flagging RFI when only this band is considered. Other types of RFI are discussed in Section 6.

1161173776, NN pol
χ^2 modified z-score
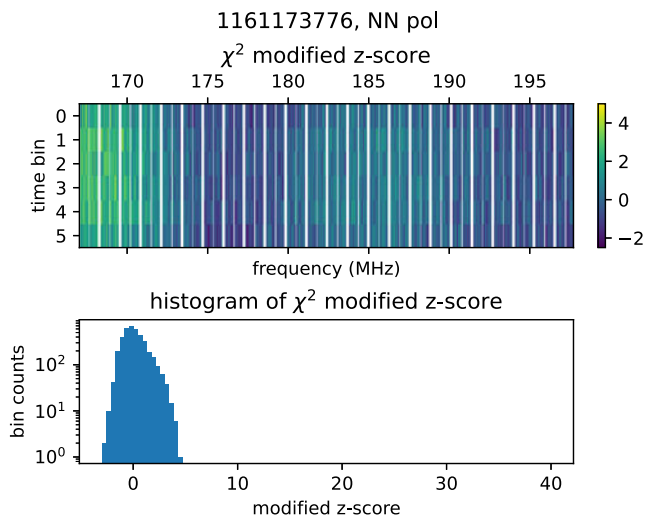
Figure 7. An example of an observation with elevated low-frequency NN χ^2 values due to the presence of the Galactic plane in the sidelobes of the MWA primary beam. In the corresponding histogram, we can see that the distribution deviates from a Gaussian normal. Observations like this (all corresponding to the -3 pointing with the instrument phased to EoR0) cause the uptick in flagging at very low frequencies in Fig. 6.

In Fig. 8, a filled dot indicates that channel 7 DTV RFI was detected in a particular observation, whereas an empty dot indicates that channel 7 DTV RFI was not detected in that observation. In order to qualify as 'detected' by χ^2 or AOFlagger, a file must clear a threshold value of 5% of channel 7 bins being flagged by the algorithm in question, as well as verifying that the DTV channel 7 band is more densely flagged than the file as a whole (which distinguishes channel-7-only RFI from broad-band events). To qualify as 'detected' by SSINS, a file must be tagged as containing

channel 7 RFI by SSINS's match filter. Although Fig. 6 indicates that between 5% and 6% of the data are flagged for channel 7 RFI by χ^2, as indicated in Table 1, roughly 27.4% of the two-minute observation *files* are flagged for this type of RFI. That is to say, 27.4% of files have 5% or more of the DTV channel 7 band flagged (if the detection was with χ^2 or AOFlagger) or was tagged by SSINS's match filter as containing DTV channel 7, but if all of the data is considered in bulk, between 5% and 6% of the total data are contaminated by channel 7 DTV (according to χ^2-based flagging).

When the flagging fraction of the channel 7 band is compared across algorithms and plotted versus time, as in Fig. 8, we observe that there are certain DTV events which are flagged by χ^2 but not the other algorithms, and that the same could be said for SSINS. The number of files flagged in the channel 7 band for each algorithm is summarised in Table 1 and in the Venn diagram in Fig. 9.

While every event caught by AOFlagger was also found by either SSINS, χ^2, or both, there is less overlap between the events flagged by SSINS and χ^2. Of all events flagged by χ^2, about 56.5% are unflagged by other algorithms, and of all the events flagged by SSINS, roughly 52.2% are unflagged by other algorithms. This indicates that the SSINS and χ^2 algorithms are both suited to spotting exclusive categories of events, whereas there is a significant overlap between AOFlagger and either SSINS or χ^2.

Using only these three algorithms, it is impossible to determine whether there are events which are missed by *all three*; it is of course possible that a novel algorithm will be able to detect RFI events invisible to χ^2, SSINS, and AOFlagger. However, for the data and methods we have, it seems that using a combination of χ^2 and SSINS will flag all detectable RFI, and the addition of AOFlagger does not measurably add to our sensitivity.
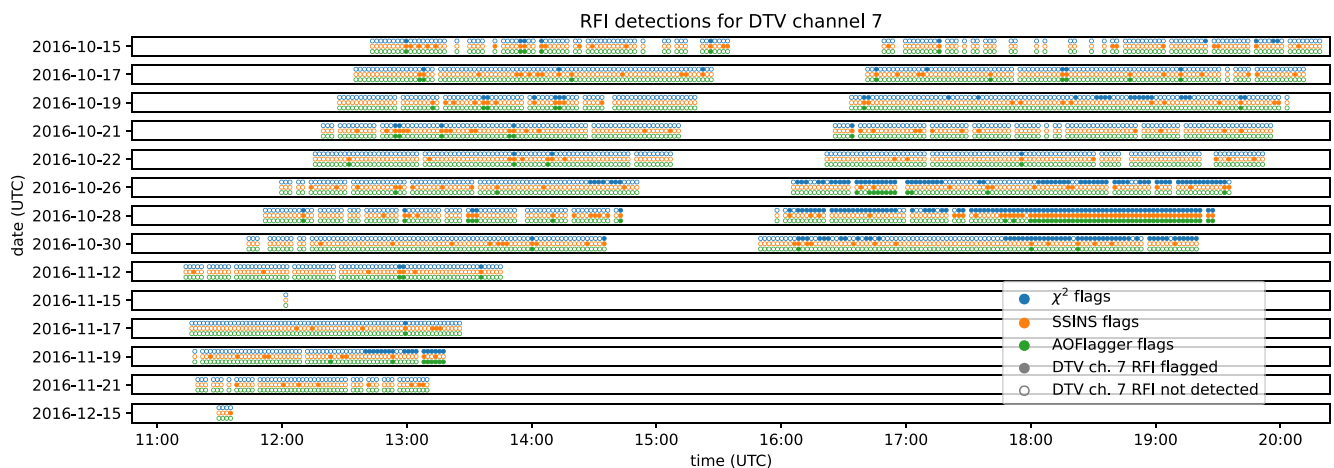
Figure 8. A comparison of χ^2, SSINS, and AOFlagger performance on flagging DTV channel 7, the most commonly-detected source of RFI in our dataset. The x-axis represents the local time of the observation in UTC, while each vertical panel represents observations from a different date. Within a panel, each observation is represented by three circles: the top row (blue circles) is our χ^2 method, the middle row (orange circles) is SSINS, and the bottom row (green circles) is AOFlagger. Filled circles indicate that RFI was detected by an algorithm in the channel 7 band for that file, whereas empty circles indicate that the algorithm did not detect any RFI in that observation file. "Detected" meaning, in the case of χ^2 and AOFlagger, that > 5% of possibly contaminated bins were flagged, and in the case of SSINS, that the match filter identified DTV channel 7 in this file. χ^2 seems well-suited to detecting long-duration DTV events which are picked up only sporadically by SSINS and AOFlagger. Meanwhile, SSINS detects isolated events which are not detected by χ^2 and, often, AOFlagger.

Table 1. Number of files flagged in the DTV channel 7 band for χ^2, SSINS, and AOFlagger.

Flagged by	Flagged by none	Flagged by χ^2 only	Flagged by SSINS only	Flagged by AOFlagger only	Flagged by χ^2 and SSINS only	Flagged by χ^2 and AOFlagger only	Flagged by SSINS and AOFlagger only	Flagged by all
Number of observations	1 209	169	143	0	30	13	14	87
% of total	72.61%	10.15%	8.59%	0.00%	1.80%	0.78%	0.84%	5.23%
% of flagged	n/a	37.06%	31.36%	0.00%	6.58%	2.85%	3.07%	19.08%
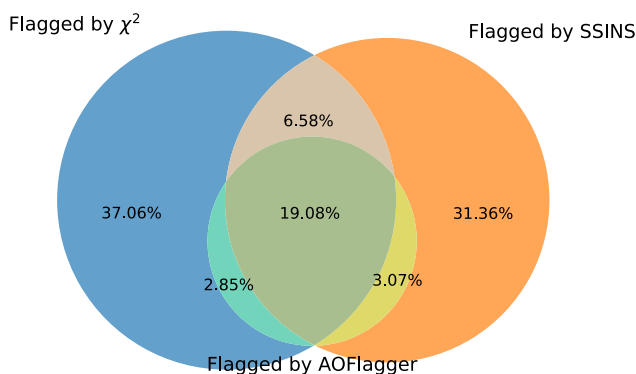
Figure 9. A Venn diagram comparing the percentage of DTV events flagged by the χ^2, SSINS, and AOFlagger algorithms. Most of the events flagged by AOFlagger are detected by either χ^2 or SSINS, but there are a significant number which are flagged by χ^2 and missed by SSINS, and vice-versa. The circles in this diagram are to scale, although overlaps are not.

6. Discussion and conclusion

6.1 RFI classification

Several distinct types of RFI can be seen in χ^2 data. The most common type of RFI is digital television (DTV) signals, usually in Australian channel 7 (181–188 MHz), although channel 9 (195–202 MHz) is sometimes also present. Channels 6 (174–181 MHz) and 8 (188–195 MHz) are the rarest to see, usually only visible in very strong DTV RFI events which contain 7 and 9 as well. Channel 7 is flagged in roughly 5-6% of the data, channels 6, 8, and 9 are all flagged in less than 1% of the data.

We also observe broad-band, short-duration RFI which is visible through most of the frequencies observed by the MWA. In general, SSINS seems much more sensitive to these events, which seem difficult to flag using χ^2.

Finally, we can see long-duration, narrow-band RFI which fits entirely into a single 40 kHz frequency-bin at central frequencies of 196.175, and 194.455 MHz, as discussed in Section 5.1. An example event is visualised in Figure C7 in Appendix C.

When flagging occupancy for the 196.175 MHz narrow-band RFI is plotted against time, as in Fig. 10, it seems to represent a repeating signal with a period of around 105 min. The source of this radiation is unclear, although the 105 min period does fall in a range of periods typical for low Earth orbit satellites. Further investigation is needed to track down this signal's provenance.

6.2 Comparing events flagged by different algorithms

Referring to Fig. 8, we observe a few patterns:

- Observations flagged by only χ^2 tend to occur clustered in time, and are more common in the data taken later at night (which correspond to EoR1 observations).
  - an example of an observation that was flagged by χ^2 alone, and was a part of one of these late-night clusters is Fig. C1 (this and other examples in this section may be found in Appendix C). This observation is from the second half of the night on 2016-10-26, around 18:15 UTC.
- SSINS flags also cluster in time somewhat, but less consistently than χ^2 flags, and the clusters seem shorter in duration on average (consistent with the time clustering properties found in Wilensky et al. 2023).
  - Figs. C2 and C3 are part of a short cluster of SSINS flags, at around 14:30 on 2016-10-30.
  - While Fig. C3 is a bright signal in SSINS, it is brief enough to be reduced in strength by time-averaging. Fig. C2, on the other hand, is much weaker when visualised in SSINS, even though it has a longer duration. It is unclear why the event in Fig. C2 was visible to SSINS but not to χ^2.
- There are also many observations flagged by SSINS alone which are isolated from other flags in time.
  - An example of an event which is isolated from other detections may be found at Fig. C4. Like Fig. C3, it is bright and brief. It occurred on 2016-11-12 shortly after 11:45.
- AOFlagger flags for this dataset overlap entirely with SSINS flags, χ^2 flags, or both.
  - It seems as though AOFlagger is not uniquely sensitive to any of the RFI in this dataset.
  - An event which was flagged by AOFlagger and χ^2 but not seen by SSINS can be seen at Fig. C5. This event occurred just after 13:15 on 2016-11-19.

In general, it seems that χ^2 is more consistent in finding longer-timescale RFI than SSINS, and conversely, SSINS seems better at detecting short-timescale RFI. An example of the former can be seen in Fig. C1, and an example of the latter in Fig. C3. Given that the first step to processing the χ^2 values of MWA data is to time-average the files, and this would have the tendency to 'wash out' shorter-timescale events, this pattern makes sense. Conversely, since SSINS relies on time-differencing to remove slowly varying signals, it is perhaps not surprising that it misses events with long durations.
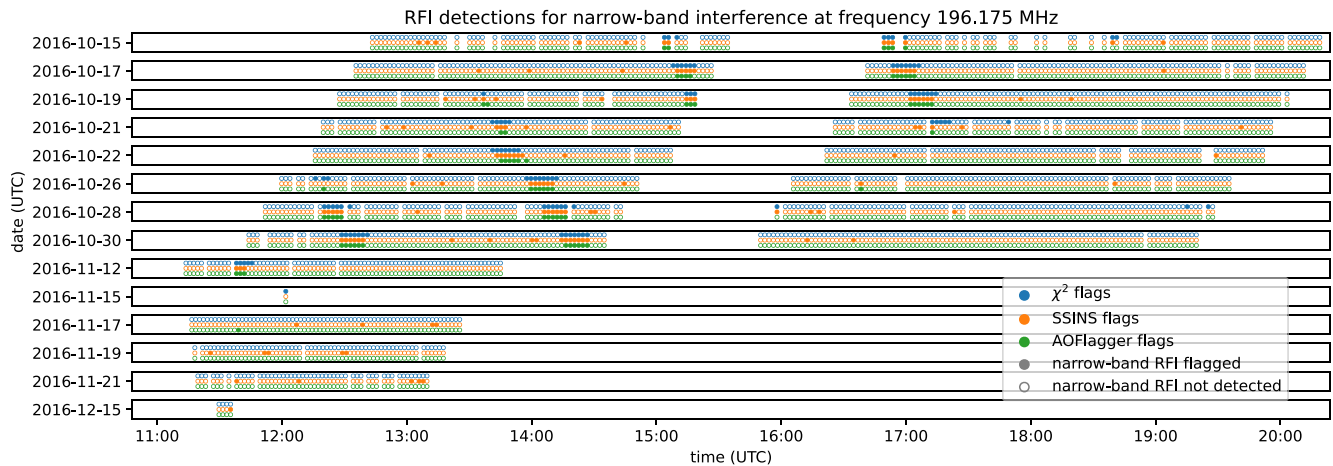
Figure 10. The same as Fig. 8, except for the 196.175 MHz narrow-band RFI signal (as opposed to channel 7 DTV). This signal appears to be periodic with a period of around 105 min.

SSINS also sees certain wide-band, short duration RFI, termed 'streaks', which are not as visible to χ^2. An example is Fig. C8.

6.3 χ^2 and RFI

In Section 3, we discussed how χ^2 is calculated and how it could be influenced by RFI. We came to the conclusion that the heightened χ^2 seen with RFI are not due to an underestimation of the noise variance σ^2, but instead due to genuinely less redundant visibilities between nominally redundant baselines.

RFI detected by χ^2 has in the past been hypothesised to be locally produced, which would lead to certain antennas detecting the signal more strongly than others, thus increasing the χ^2 metric. In fact, in our analysis, it became clear that the RFI picked up by χ^2 does not seem to be associated with local transmitters, as it doesn't peak in particular areas of the array for individual observations – when χ^2 is elevated for a certain observation, it is increased for every antenna in the array.

We offer the following theory for the elevated χ^2 values we observe. One source of non-redundancy in the MWA is the antenna-to-antenna variation of the the primary beam. Observationally, the MWA beams are generally consistent in the main lobe, but the sidelobes exhibit more variability (Line et al. 2018; Chokshi et al. 2021). RFI from sources in the sidelobes (e.g. signals propagated over the horizon via tropospheric ducting or reflections from airplanes far away from zenith) would appear with different strengths in the signals from different antennas, leading to more observed disagreement in nominally redundant visibilities and enhanced χ^2 values.

As a corollary to this theory, we might expect this χ^2 flagger to have a lower success-rate when RFI is located near the zenith, i.e., within the main lobe, since we expect significantly more agreement between redundant baselines. The RFI events detected by SSINS but not χ^2 could potentially be reflections off airplanes or other objects transiting near zenith, a detection-case which SSINS has been shown to handle well (Wilensky et al. 2019). In order to verify this hypothesis and its implications, direct imaging of RFI caught by a χ^2-based flagger is likely necessary. This is an avenue for future work on this topic.

6.4 Conclusion

RFI detection and flagging remain a formidable challenge in the realm of 21 cm EoR cosmology. Without an algorithm (or suite of algorithms) which can detect and excise RFI at a sensitivity heretofore unachieved, instruments like the MWA and HERA may be unable to deliver an accurate power spectrum of the EoR's 21 cm radiation.

In this work, we have presented χ^2 from redundant calibration as a new approach to flagging RFI in data from the MWA. We have demonstrated the effectiveness of this method and compared its efficacy at flagging different kinds of RFI than AOFlagger and SSINS.

χ^2 flagging as implemented here seems effective at flagging the DTV RFI events which are all too common at the MWA site. In the realm of long-duration, weak DTV events, χ^2 finds RFI that's missed by the current state-of-the-art algorithm SSINS. However, χ^2 should not be taken as a replacement for or improvement on SSINS; in fact, the algorithms seem quite complementary, each being able to detect RFI that remains invisible to the other.

There are many potential avenues for future work using this tool. The flagging algorithm based on the modified z-scores of $\chi^2/nDoF$ could be improved with features like those found in the SSINS flagging suite, such as applying a threshold to fully flag a DTV channel when sufficient RFI has been detected within it. Furthermore, a machine-learning approach to flag generation could potentially be sensitive to even weaker and more tenuous events, which are visible to the human eye but difficult to pick out algorithmically.

In the course of this work, a repeating narrow-band RFI source was identified at 196.175 MHz with a period of around 105 min. Identifying the source of this RFI is another opportunity for further work.

Producing a cosmological 21 cm power spectrum using data screened with χ^2 RFI flagging alongside SSINS is another near-term goal. Furthermore, it would be instructive to investigate whether this hypothetical power spectrum still carries the statistical footprint of ultra-faint RFI as described in Wilensky et al. (2023).

That the χ^2 values produced by redundant calibration are sensitive to weak RFI is a happy accident, which will hopefully become useful in improving the quality of the data that goes into serious analyses in the future. This work represents the first step in bringing this promising new method into the pipelines of new 21 cm data-processing efforts.

References

Beardsley, A., et al. 2016, ApJ, 833, 102
Chokshi, A., et al. 2021, MNRAS, 502, 1990
Choudhuri, S., Bull, P., & Garsden, H. 2021, MMNRAS, 506, 2066
Choudhury, T. R. 2022, GRG, 54, doi: 10.1007/s10714-022-02987-4
DeBoer, D. R., et al. 2017, PASP, 129, 045001
Dillon, J. S. 2024, personal communication
Dillon, J. S., & Parsons, A. R. 2016, ApJ, 826, 181
Dillon, J. S., et al. 2020, MNRAS, 499, 5840
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
Hazelton, B. J., Jacobs, D. C., Pober, J. C., & Beardsley, A. P. 2017, JOSS, 2, 140
Jacobs, D. C., et al. 2016, ApJ, 825, 114
Kerrigan, J., et al. 2019, MNRAS, 488, 2605
Li, W., et al. 2018, ApJ, 863, 170
Li, W., et al. 2019, ApJ, 887, 141
Line, J. L. B., et al. 2018, PASA, 35, doi: 10.1017/pasa.2018.30
Liu, A., & Shaw, J. R. 2020, PASP, 132, 062001
Morales, M. F., & Wyithe, J. S. B. 2010, ARA&A, 48, 127
Offringa, A., et al. 2010, MNRAS, 405, 155
Offringa, A., Van De Gronde, J., & Roerdink, J. 2012, A&A, 539, A95
Offringa, A., et al. 2015, PASA, 32, e008

Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012, ApJ, 753, 81
Pritchard, J. R., & Loeb, A. 2012, RPPh, 75, 086901
Sokolowski, M., Wayth, R. B., & Lewis, M. 2015, in 2015 IEEE Global Electromagnetic Compatibility Conference (GEMCCON) (IEEE)
Tingay, S. J., et al. 2013, PASA, 30, e007
van Haarlem, M. P., et al. 2013, A&A, 556, A2
Wayth, R. B., et al. 2018, PASA, 35, doi: 10.1017/pasa.2018.37
Wieringa, M. H. 1992, ExAs, 2, 203
Wilensky, M. J., Barry, N., Morales, M. F., Hazelton, B. J., & Byrne, R. 2020, MNRAS, 498, 265
Wilensky, M. J., et al. 2019, PASP, 131, 114507
Wilensky, M. J., et al. 2023, ApJ, 957, 78
Wise, J. H. 2019, CoPh, 60, 145
Zhang, X., et al. 2018, MNRAS, 477, 5167
Zhang, Z., et al. 2020, PASA, 37, e045
Zheng, H., et al. 2014, MNRAS, 445, 1084

Appendix A. The effects of time-averaging on χ^2/nDoF

In an initial analysis, the data were processed with hera_cal without the additional nine-fold time averaging step described in Section 2.2. The result was that the χ^2 values were noisy and elevated in regions where there is no RFI. A

The upper plot in Fig. A1 is a single time-slice of the χ^2/nDoF generated by hera_cal from a file that was not first downsampled in time. This file is known to have DTV channel 7 RFI present, visible between 181 and 188 MHz. Notably, the presence of RFI increases the signal-to-noise level in those frequencies, and so there is much less noise propagated through to χ^2 in that band.

The lower plot of Fig. A1, on the other hand, is also a single time-slice of the χ^2/nDoF generated by hera_cal from the same file at approximately the same time, when the data had been time-averaged before calibrating. Files are time-averaged from a cadence of one visibility every 2 s to one visibility every 18 s, which we would expect to lower the noise amplitude by a factor of $\sqrt{9} = 3$. This plot showcases the lower noise levels afforded by
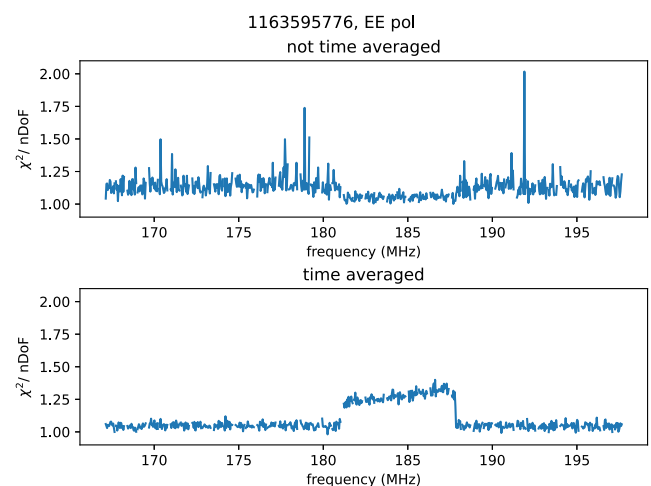
Figure A1. A single time-slice of χ^2/nDoF, generated by hera_cal without downsampling in time first (top) and after downsampling in time (bottom). The result without downsampling is noisy and difficult to process to identify RFI. After downsampling, noise in the χ^2 metric is significantly reduced, and the RFI is visibly elevated over the noise floor.

time-averaging, and the stark contrast between frequencies with and without RFI.

Appendix A.1. Weighting χ^2 by N Samples to correct noise discrepancy

When pyuvdata downsamples in time, it does not include flagged data in the calculated average. Furthermore, when working with MWA data, it is good practice to flag a few times at the beginning and end of the file, which often contain bad data. Noise is reduced by time-averaging, so we observe that χ^2 generated from MWA data which have been flagged at the beginning and end of the file in this manner have an elevated noise-floor.

We can correct for this discrepancy by multiplying each χ^2 value by the corresponding element in pyuvdata's NSample_array, which tracks the fraction of unflagged data which were averaged into a given time/frequency bin.

When averaged, noise is decreased by a factor of $\sqrt{N}$, where N is the total number of data-points averaged together. Because σ^2 scales with the square of the noise level, and forms the denominator of equation (2), we expect χ^2 to scale inversely with N. It is sufficient, then, to normalise the noise level by multiplying χ^2 by the corresponding element in the UVData object's NSample_array.

Appendix B. redcal_run() parameters for use on MWA data

A few custom parameters are necessary to make `hera_cal` work with MWA data. The parameters we used were

```
max_dims=3, oc_conv_crit=1e-10, oc_maxiter=2000,
    check_every=10, check_after=500,
    ant_z_thresh=100.0, gain=0.3
```

A few of these warrant further explanation – `max_dims=3` is absolutely necessary, because the spatial separation between the two hexagons introduces an extra dimension to the degenerate subspace inherent to redundant calibration solutions, corresponding to an arbitrary phase separation between the hexagons, in addition to the three other degenerate parameters, which correspond to the tip/tilt of the array and an overall phase. Setting `ant_z_thresh=100` eliminates the software's throwing out of high-z-score antennas when calculating solutions, because we are *not* planning to use the gain solutions for calibration. The remaining parameters are the result of some by-hand fine-tuning, but could probably be improved with a more systematic parameter-space search.

Appendix C. A selection of interesting RFI events

In the course of this research, we encountered many interesting individual RFI events. In this appendix, we have included a selection of particularly illustrative or interesting events.

Figs. C1 through C9 show time-frequency waterfall plots of χ^2/nDoF modified z-scores (top row), SSINS's Incoherent Noise Spectrum waterfall plot (INS, middle row), and AOFlagger flags averaged over baseline as described in Section 4.1 (bottom row). Each column shows one of the two linear polarisations measured by the MWA.
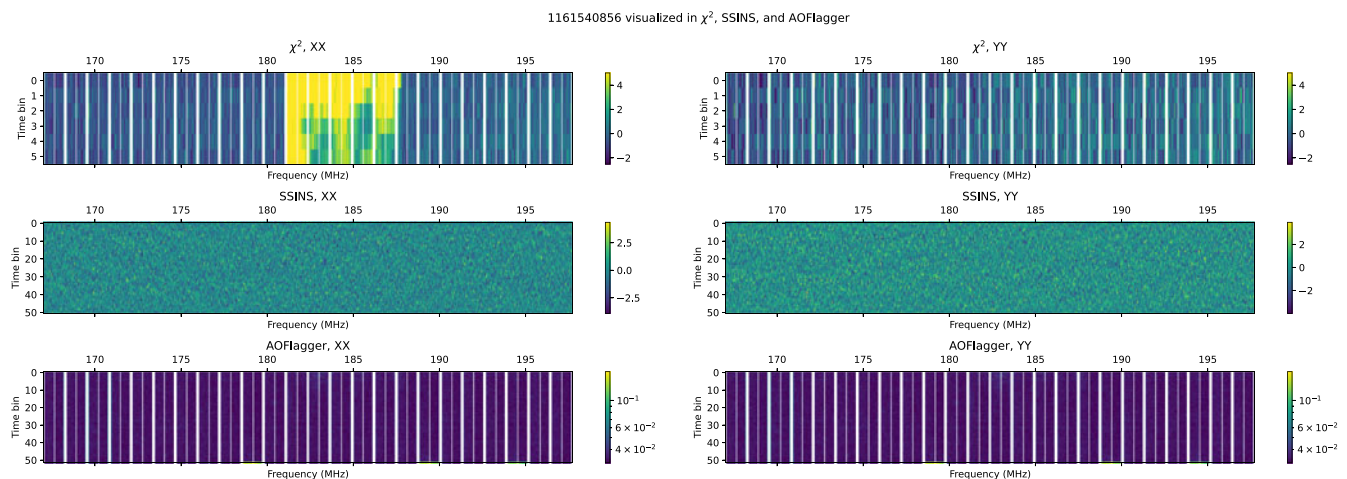

Figure C1. A typical example of an observation with RFI in DTV channel 7 which is detectable with χ^2, and not with other flagging algorithms. This observation is a part of a long string of DTV detections seen by χ^2 alone.
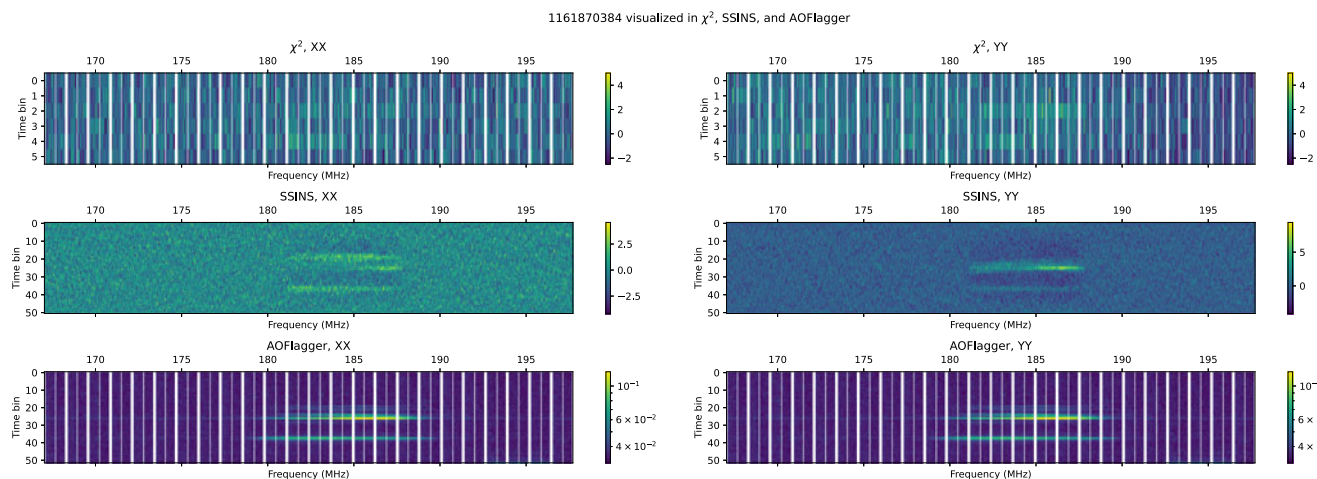
Figure C2. This SSINS DTV detection occurs as part of a small cluster of files flagged by SSINS alone. The event is also visible in AOFlagger, and a hint of it is visible in χ^2, but it did not reach the threshold for being listed as being 'detected' by any algorithm besides SSINS.
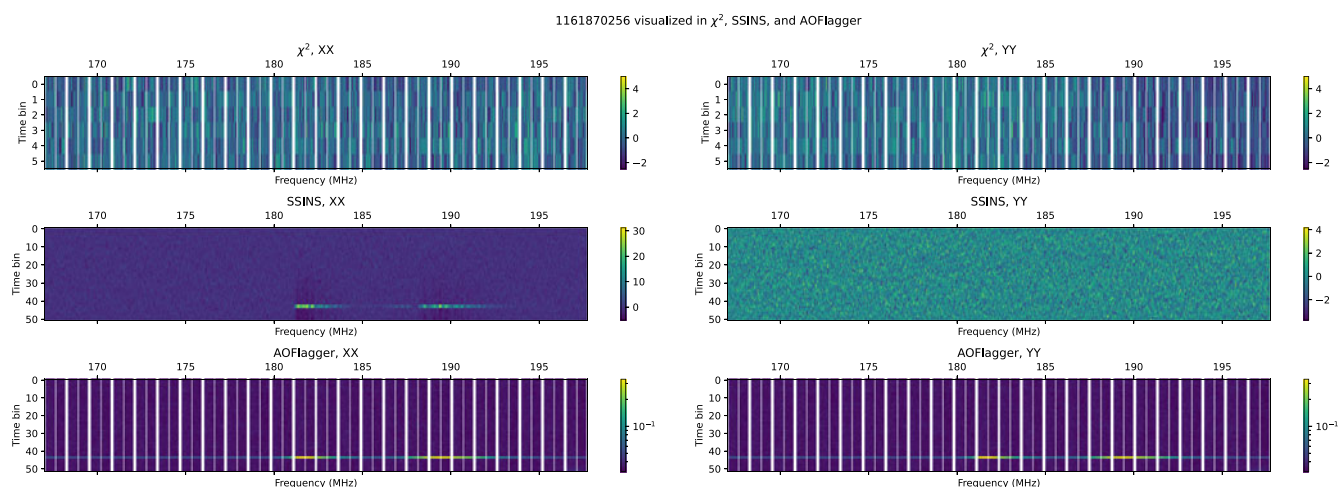

Figure C3. This SSINS DTV detection occurs as part of the same small cluster of SSINS detections as Fig. C2. The event is also visible in AOFlagger, but it did not reach the threshold for being listed as being 'detected' by AOFlagger or χ^2.
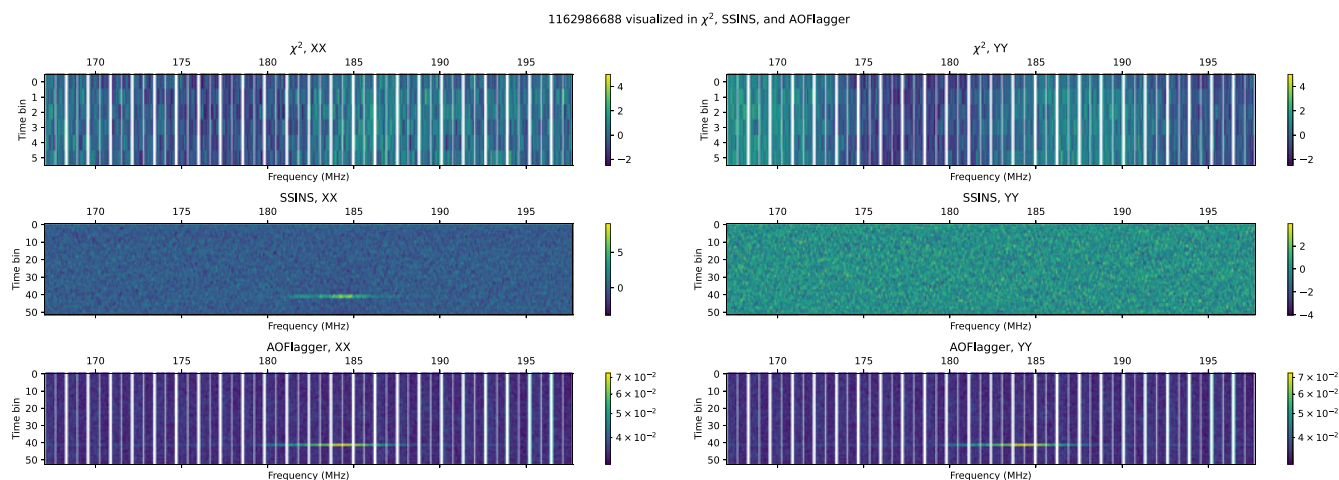

Figure C4. This SSINS DTV detection occurs isolated in time from other RFI events.
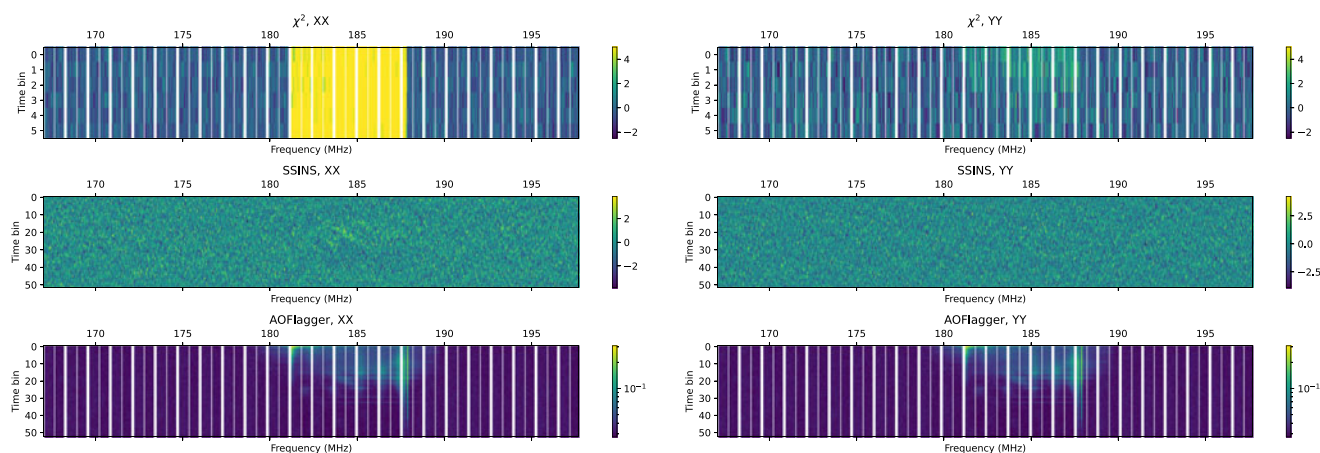
Figure C5. This DTV channel 7 event was detected by χ^2 and AOFlagger, but missed by SSINS.
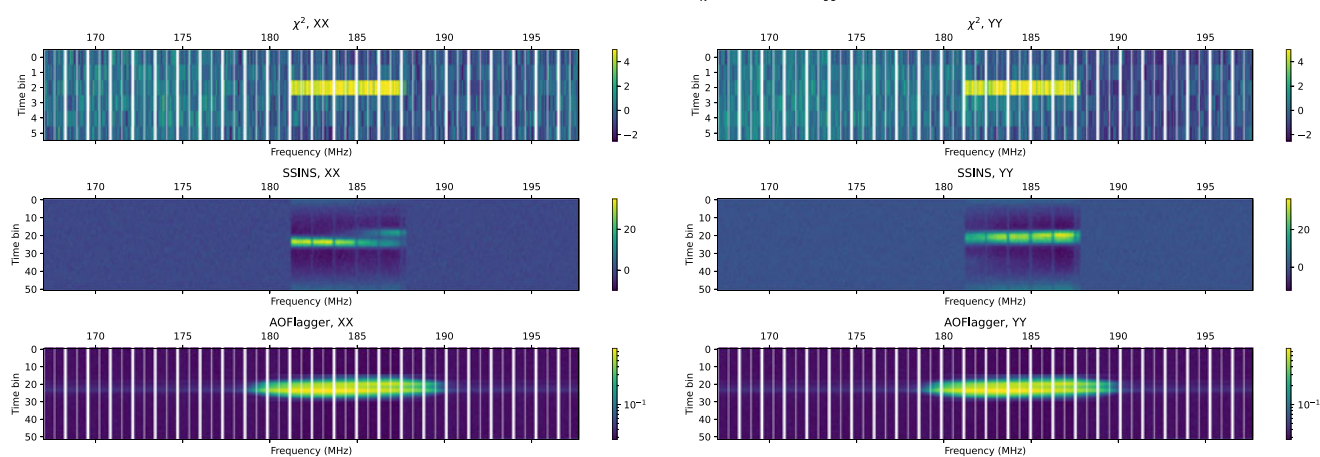
Figure C6. An example of DTV channel 7, detected by all three algorithms.
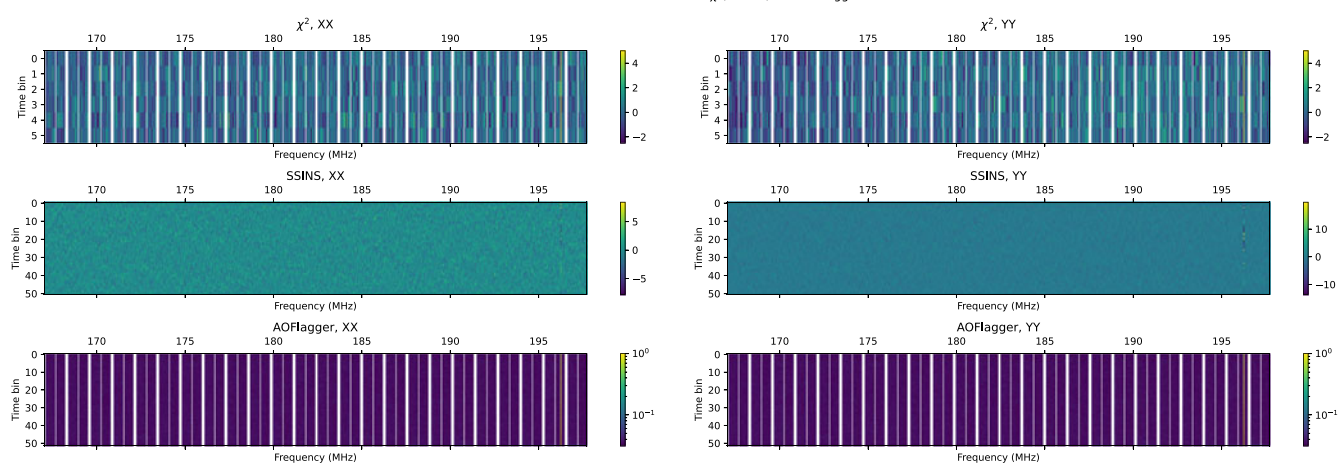
Figure C7. An example of narrow-band RFI at 196.175 MHz, detected by all three algorithms.
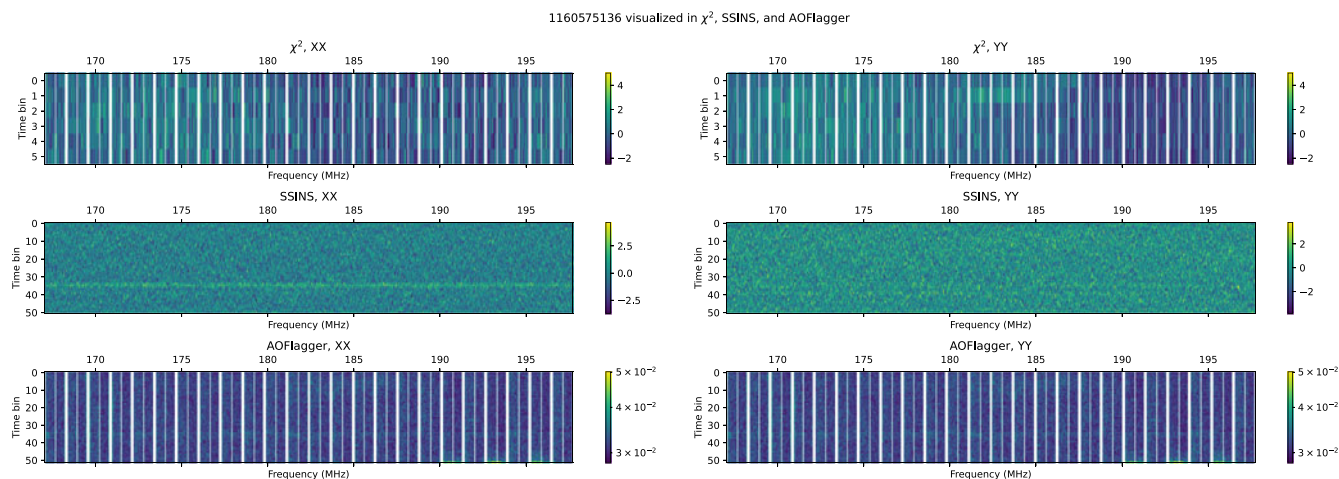
Figure C8. An example of a 'streak' found in SSINS, but invisible to χ^2.
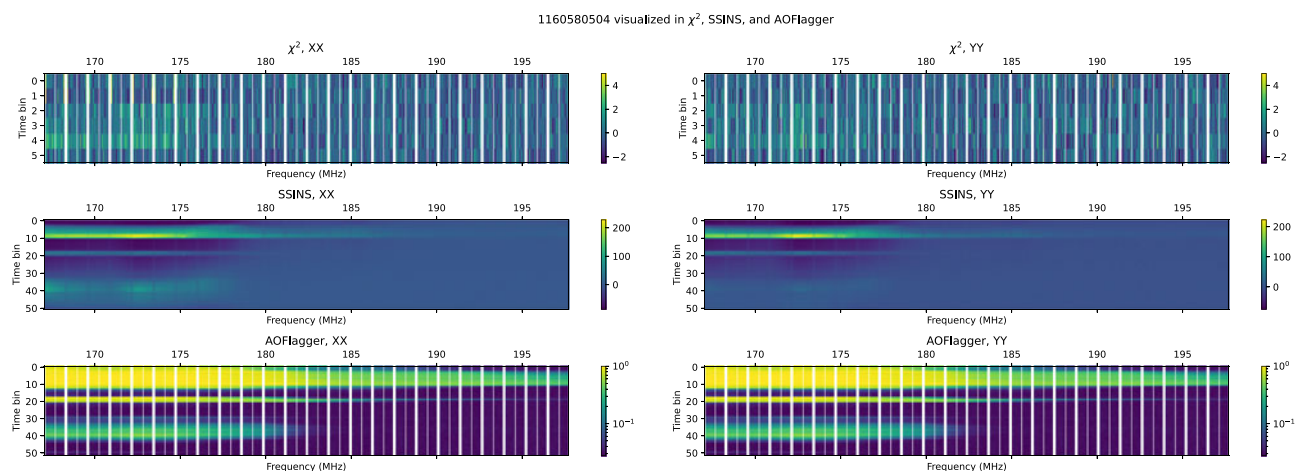
Figure C9. An unusual event, which might have been due to weather, is also nearly invisible to χ^2.