



Universiteit
Leiden
The Netherlands

Time domain imaging of transient and variable radio sources

Cendes, Y.N.

Citation

Cendes, Y. N. (2020, May 12). *Time domain imaging of transient and variable radio sources*. Retrieved from <https://hdl.handle.net/1887/87646>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/87646>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/87646> holds various files of this Leiden University dissertation.

Author: Cendes, Y.N.

Title: Time domain imaging of transient and variable radio sources

Issue Date: 2020-05-12

FLAGGING OF RADIO FREQUENCY INTERFERENCE: IMPLICATIONS FOR SHORT-DURATION TRANSIENTS

Cendes, Y. et al.,

Astronomy and Computing, Volume 23, article id. 103., 2018

Abstract

With their wide fields of view and often relatively long coverage of any position in the sky in imaging survey mode, modern radio telescopes provide a data stream that is naturally suited to searching for rare transients. However, Radio Frequency Interference (RFI) can show up in the data stream in similar ways to such transients, and thus the normal pre-treatment of filtering RFI (flagging) may also remove astrophysical transients from the data stream before imaging. In this paper we investigate how standard flagging affects the detectability of such transients by examining the case of transient detection in an observing mode used for Low Frequency Array (LOFAR; [van Haarlem et al., 2013b]) surveys. We quantify the fluence range of transients that would be detected, and the reduction of their SNR due to partial flagging. We find that transients with a duration close to the integration sampling time, as well as bright transients with durations on the order of tens of seconds, are completely flagged. For longer transients on the order of several tens of seconds to minutes, the flagging effects are not as severe, although part of the signal is lost. For these transients, we present a modified flagging strategy which mitigates the effect of flagging on transient signals. We also present a script which uses the differences between the two strategies, and known differences between transient RFI and astrophysical transients, to notify the observer when a potential transient is in the data stream.

3.1 Introduction

Modern radio astronomy is particularly suited for the search of many classes of extreme transient object. This is due to a combination of astrophysical and instrumental effects. Astrophysically, extreme objects are often compact, non-thermal emitters, implying they can vary on short time scales and have significant luminosity at radio frequencies. At long radio wavelengths, the blackbody limit severely constrains the allowed luminosity, but often compact objects are seen to have coherent emission at those long wavelengths, which can exceed the (incoherent) blackbody limit by many orders of magnitude. Instrumentally, modern radio telescopes can have very wide fields of view and thus are much more likely than previous instruments to detect rare events. Further, modern computing techniques and hardware infrastructure enable the near real-time processing of the large data streams from these telescopes, and automated generation of alerts to transients. This has enabled rapid multi-wavelength follow-up of candidate transients, essential for understanding the transient source parameters. However, the large data volumes lead to a requirement of automated flagging, which can be pernicious to transient searches.

There are multiple ways of using a radio interferometer to search for transient events. One may use it to image the sky, and search for variations of brightness in the sources detected in the image [Stewart et al., 2016]. One may also add up the signals from all the array elements coherently or incoherently to form a beam on the sky and examine the time series of the signal from that beam [Coenen et al., 2014]. The latter method is preferred for looking at very brief (sub-second) events and periodic signals, and is the method of choice for detecting pulsars and Fast Radio Bursts (FRBs) [Rane & Lorimer, 2017; Lorimer et al., 2007]. When pushed to their limits, both methods require front-line data processing and supercomputing technology in order to cope with the data flow [van Haarlem et al., 2013b]. In response to this, other techniques are now also being developed, such as searching for suspected transient events in raw correlation data products and then imaging very short selected data segments with candidate transients [Bower et al., 2013; van Velzen et al., 2013].

In this paper we focus on the first method, searching for transients in a time series of images. The effectiveness of image domain transient searches has already been demonstrated by the discovery of a bright, low-frequency transient of several minutes duration by Stewart et al. [2016], as well as earlier discoveries by Hyman et al. [2009] and Jaeger et al. [2012]. Interferometers spend a large fraction of their time collecting data in imaging mode, and at the same time sample visibilities initially at order of seconds intervals for the purposes of avoiding time smearing, rejecting bad data, and calibration. This means that a large amount of data are collected in imaging mode that may be used to search for interesting transient sources, thus generating more science at no extra expense of observing time (but considerable extra processing time).

Typical image domain searches for transients of a given duration consist of tracking observations sampled at short timescales, followed by offline flagging and imaging at a much longer timescale in order to increase sensitivity. An example of this is the LOFAR survey mode, in which visibilities are integrated over 10 seconds in real-time, but imaging is carried out in 11 minute snapshots.

A key step in processing such radio data is the removal of Radio Frequency Interference (RFI)—terrestrial, generally human-generated, radio noise that may corrupt celestial signals. In

the large data volumes of e.g., LOFAR, automated algorithms are needed to manage this process in realistic time. Generally these algorithms use the fact that RFI is much stronger than cosmic signals in the data, and usually forms stripes in a dynamic spectrum that are either narrow in frequency or in time, on a smooth, much lower background. Simple removal of these data from consideration is adequate for most radio imaging, and even in a relatively RFI-rich environment like the Netherlands this results in only a few percent data loss [Offringa et al., 2012].

However, a bright and short astrophysical transient may resemble RFI in the data, and thus be filtered out before any image is created. In this paper, we investigate the effect this has on the likelihood of transient detection. We quantify the effect of various RFI flagging techniques on transients by injecting simulated transients into an actual LOFAR observation, then flagging and imaging the data.

The breakdown of the paper is as follows. In Section 2, we will present an overview of current standard methods of RFI flagging, and of potential problems faced while flagging data from transient searches. In Section 3, we will present our observations with the LOFAR Observatory and our transient simulations. In Section 4, we will discuss the effect of various flagging strategies on simulated transients. In Section 5, we will discuss the findings of our transient simulation results, along with a potential solution to identify astrophysical transients mistakenly flagged as RFI. In Section 6, we will present our conclusions from this research.

3.2 RFI Flagging Overview

3.2.1 Radio Frequency Interference and Celestial Radio Transients

Radio Frequency Interference (RFI) from man-made sources is emission that can severely contaminate celestial observations by swamping the comparatively weak sky signal received by a sensitive radio telescope. The signature of received RFI in an observation depends on several factors, including the type of radio telescope (single dish or interferometer), the type of observation (continuum, spectral line, beam-formed or time resolved), and the kind of RFI itself. The latter can range from impulse-like bursts, narrow-band, or wide-band RFI, and can be ground-, air-, or space-based (e.g. terrestrial emitters, airplanes, spacecraft).

Temporally restricted RFI bursts are especially pernicious to transient searches due to their resemblance to celestial transients. These can be both band limited (e.g. Jovian bursts) as well as wide band. Thus, extra care needs to be taken to distinguish such transients from RFI, and special RFI detection strategies have been developed based on the class of transient [Ryabov et al., 2004; Kocz et al., 2012].

Interferometers used for imaging observations possess several natural defences to RFI. A long enough spatial separation of the antennas making up a baseline results in a different RFI environment at each antenna, leading to a natural suppression of the RFI signal. Further, the delay tracking of a fixed location in the sky also leads to a delay decorrelation of the RFI signal. The phase tracking of the interferometer also leads to attenuation via fringe washing. Finally, interferometers usually have high enough spatial resolution to be able to isolate a coherent RFI source in the image plane.

In spite of these effects, interferometric observations can still be sensitivity limited by residual RFI, due to the high sensitivity of the telescopes. Interferometers typically use offline post-correlation flagging methods, which can be applied over multiple passes on the data. Of the available post-correlation methods, flagging using thresholding is by far the most common approach taken due to its ease of implementation and its predictable effect on the image noise, where the thresholding occurs on the visibility amplitude per baseline [Offringa et al., 2010]. Another approach uses filtering techniques, where a parametric model of the RFI signal can be built up and subsequently subtracted from the data, but this is computationally expensive and model dependent [Offringa et al., 2012]. Adaptive interference cancellation using an independent estimate of the RFI signal obtained from reference channels also exist, but may not be as effective as the other methods [Barnbaum & Bradley, 1998]. Spatial filtering can be an effective method, where the RFI source is isolated and subtracted from the visibility data [Raza et al., 2002]. However, this is effective only if the RFI is significantly correlated at the antenna sites. Higher order statistics can also be used for RFI flagging [Fridman, 2001], and Spectral Kurtosis in particular has been tested for transient detection [Nita, 2016; Nita & Gary, 2010]. However, the statistical methods such as Spectral Kurtosis require much more data processing in order to calculate the mean and variance, which is less feasible than thresholding methods for large-scale transient surveys.

For telescopes with large data volumes, such as LOFAR or the Murchison Widefield Array (MWA; Tingay et al. [2013]), automated algorithms are needed to manage this process in realistic time. Generally these algorithms use the fact that RFI is much stronger than cosmic signals in the data, and usually forms stripes in a dynamic spectrum that are either narrow in frequency or in time, on a smooth, much lower background. Simple removal of these data from consideration is adequate for most radio imaging.

Visibility based flagging would eliminate only the transients bright enough to cross a high threshold on a single visibility, leaving weaker transients relatively unaffected, and available in the image domain. The latter could be many more in number, based on existing source counts [Cordon et al., 1984]. However, the current lack of a population of such transients and their unknown origins make the reliable and unambiguous detection of such events important, especially for triggering multi-wavelength observations. Detecting the brightest transients is thus essential in the discovery phase.

An illustration of the effects of thresholding based flagging on a bright transient can be seen in Figure 1, where the visibility amplitudes on a single baseline are visualized in the time-frequency domain. A simulated transient of 30 seconds duration has been injected into the visibilities. Through the sequence described in more detail below, the flagger will flag such a transient automatically as contaminated data. As such an example illustrates, it is crucial to understand whether such transient signals spanning a range of amplitudes could potentially be flagged before the data are even examined by the observer. This aim is also important in light of transients such as the one reported by Stewart et al. [2016], and surveys for FRBs using the MWA such as Tingay et al. [2015] and Rowlinson et al. [2016], which all relied on the RFI flagging algorithm, AOflogger [Offringa et al., 2012], for their analysis. We analyze the performance of this algorithm via simulations.

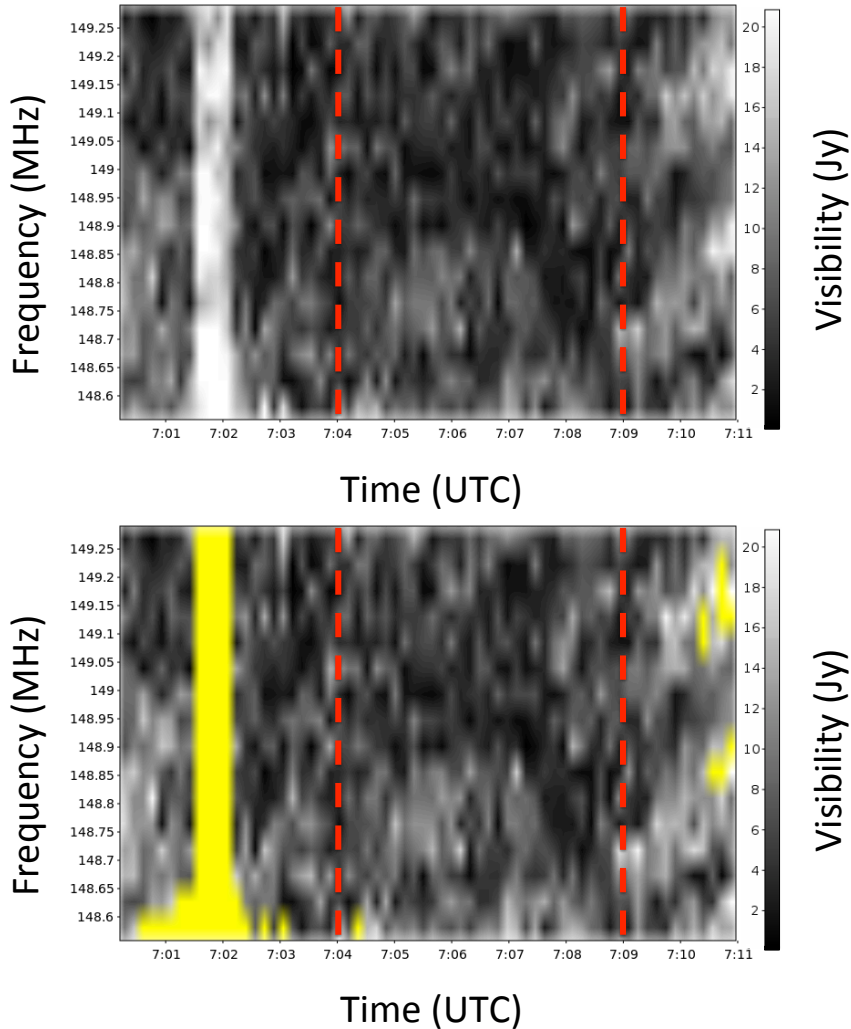


Figure 3.1 A visual representation of the observed visibility amplitude in the time-frequency domain on the baseline between two LOFAR stations in the visualization program “rfigui,” where the top is an unflagged data set and the lower left has been flagged by AOFlogger (yellow). Here we have an 11 minute LOFAR data set with a 30 second top hat simulated signal of amplitude 10 Jy injected at 07:01:30. This signal was flagged as RFI when the flagging algorithm is run with default parameters, showing the potential perils of using automatic flagging during transient searches. For this baseline, the 10 Jy signal is added on top of the median amplitude values. During the period where no transient is present, indicated as the space between the red dotted lines, when looking at the amplitudes of the visibility time series the data has a standard deviation of ~ 2.6 Jy and a median value of 5.5 Jy. Here the scale is 20 Jy due to bright amplitude values at the end of a few channels, which are later flagged.

3.2.2 RFI Flagging Algorithms

3.2.3 AOFlagger

The current default RFI flagging software for radio imaging at multiple observatories, which is used routinely during transient surveys such as MWA [Offringa et al., 2015] and LOFAR, is called the AOFlagger. In recent years, AOFlagger has become one of the standard RFI flagging algorithms, and is increasingly used on many radio telescopes. It has been successfully used with data from several interferometric arrays, such as Very Large Array [VLA; Williams et al., 2017], Westerbork Synthesis Radio Telescope [WRST; Adebahr et al., 2017], the Australia Telescope Compact Array [ATCA; Smolčić et al., 2016], the Boolardy Engineering Test Array of the Australian Square Kilometre Array Pathfinder [ASKAP BETA; Allison et al., 2017]. It has also been successfully implemented on single dish telescopes, such as Parkes [Offringa et al., 2012] and Arecibo.¹

In brief, the AOFlagger is an iterative algorithm which operates on a frequency resolved timeseries of visibilities from a single baseline, and can operate on either a single polarization or a chosen Stokes parameter. It fits a surface to the visibility amplitudes in the time-frequency plane in order to eliminate systematics, e.g., due to fringes from bright sources in sidelobes, and then clips out visibilities with amplitudes crossing a chosen threshold. The flagged visibilities are ignored in the next iteration. The algorithm detects low level RFI using a set of decreasing thresholds, implemented via the SUMTHRESHOLD algorithm. It has been empirically determined that two iterations of SUMTHRESHOLD are adequate for flagging [Offringa et al., 2012]. The algorithm consists of the following steps:

- **SUMTHRESHOLD:** This is a combinatorial thresholding algorithm where the same dataset is subjected to a set of decreasing thresholds. A threshold χ_M is determined by M , the number of samples in the data set surrounding a candidate data point (visibility amplitude for one time/frequency sample), where M is predefined, and a chosen false detection rate [Offringa et al., 2010]. If the sum of amplitudes of M contiguous visibilities exceeds the threshold χ_M , then all M visibilities are flagged.

The algorithm processes data in order of decreasing thresholds $\chi_1, \chi_2, \dots, \chi_M$, where the thresholds are in units of the visibility amplitude standard deviation, σ . The single sample threshold, χ_1 is the highest, and eliminates really bright and sporadic visibilities. Visibilities flagged by a χ_i threshold are ignored when computing the statistics for applying the next level threshold χ_{i+1} . These lowered thresholds are used to eliminate low-lying RFI not filtered by χ_i . Visibilities comprising such low-level RFI are usually connected to their neighbors in time and frequency, i.e., the connected visibilities form islands of high amplitudes. In the SUMTHRESHOLD algorithm, the range of applied thresholds depends on the extent of connectivity in all four directions on the time/ frequency visibility amplitude plane.

For a given baseline, SUMTHRESHOLD is applied twice in an AOFlagger iteration, first prior to the surface fit in order to ignore RFI when fitting, and then again when the fitting has converged in order to establish the actual flags. The χ_i thresholds are decreased exponentially in every iteration of SUMTHRESHOLD to increase the sensitivity to low level RFI.

¹A full list of telescopes which have successfully used AOFlagger, along with the software itself, is available at: <https://sourceforge.net/projects/aoflagger/>

- **Channel and Time Selection:** Next there is a channel and time selection step which flags problematic channels and time steps which may be fully contaminated but have not yet been flagged. In the case of channel selection, the RMS value of visibilities across channels for a given time slice is computed. The time series of these RMS values are Gaussian smoothed, and the standard deviation, σ , of the difference between the RMS time series and the smoothed time series is computed. Any RMS value found to be $> 3.5\sigma$ results in the entire set of channels for that time slice to be flagged. The same occurs for the time axis. This step, akin to conventional thresholding, is implemented in order to facilitate faster convergence.
- **Surface fitting:** Surface fitting then occurs, which removes fringes caused by strong sources in the side lobes in order to increase accuracy. This is done with a Gaussian kernel sliding window in both time and frequency space, and is a time-consuming step when compared to the other steps.

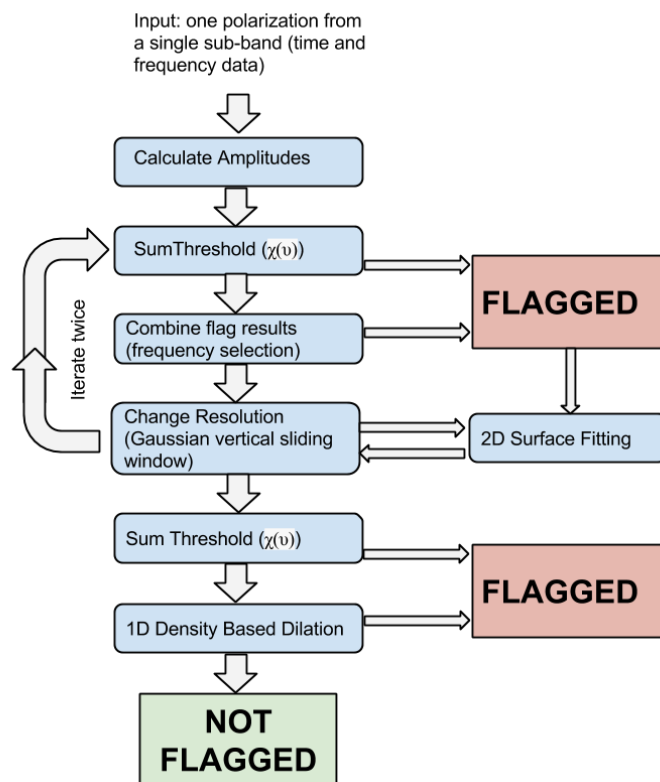
Overall, for routine data collection and subsequent imaging, the AOFlagger has proven to be a fast and accurate automatic flagger, particularly in the relatively high-RFI LOFAR radio environment [Offringa et al., 2013b], although it can, in some cases, flag data that is not RFI [Offringa et al., 2013b]. Normally, however, this is not an issue in routine observatory observations because mistakenly flagged data are a small percentage of the entire data set [Offringa et al., 2013b, 2015].

A flow chart of AOFlagger can be seen in Figure 3.2.

3.2.4 Concerns for Transient Searches

The aim of this paper is to investigate the effect of AOFlagger on transient observations. As such, in addition to investigating the impact of the default AOFlagger settings on transients, we also focused on modifying the flagger to see whether an automatic RFI flagging method could be adapted for the needs of transient searches. In particular we examined the following:

- One particular concern was that the thresholding algorithm, SUMTHRESHOLD, works in both the time and frequency domains. This means brief, bright transients that are broadband, and a focus for image-plane surveys, could be flagged due to their sudden increase of power in the time domain.
- Time selection steps that rely on sigma clipping may inaccurately flag data that are not RFI. For example, a sudden, brief transient in an otherwise quiet radio field would have its amplitude suddenly change by a high value compared to the data around it, and it could be mistaken for RFI and flagged at this stage.
- In addition to this, the lowered thresholds in later steps of SUMTHRESHOLD for data connected in either frequency or time are also a concern for transient searches. This is because such lower thresholds may flag fainter transients which are normally not affected by the high single sample threshold χ_1 .



Figuur 3.2 A flow chart of the steps in AOFlogger. The steps that we focused on in particular for our transient specific analysis for modification are the SUMTHRESHOLD and time selection steps. We should note the input can also be a Stokes parameter of interest.

3.3 Observations and Transient Simulations

3.3.1 Observations

As we were interested in examining the consequences of RFI flagging on real data, we wanted to examine a field observed with LOFAR. For this purpose, we chose data taken with the LOFAR telescope as an extra target observation during routine transient monitoring observations (Fender et al., in prep) using one sub-array pointing of the telescope, where the beam is formed using the innermost antennas from each station. We then focused on a low-RFI 11 minute measurement set, taken on February 10, 2013, using one beam at 149 MHz with a bandwidth of 781 kHz, along with calibrator source Cygnus A.

First, the data were pre-processed using the standard methods with LOFAR: they were flagged for RFI using AOflogger [Offringa et al., 2012], and the data calibration and imaging were carried out using the method described by Broderick et al. [2016]. We used a model of the source for the calibration of the calibrator sub-bands, and then the gain amplitudes and phases were transferred to the target field data as outlined by Heald et al. [2011]. We also refined the calibration further by performing phase-only self-calibration on the target field, as outlined by van Haarlem et al. [2013b].

For imaging, we used the AWImager [Tasse et al., 2013]; a maximum projected baseline of 6 km was chosen to ensure a reliable image from the given u,v coverage for the 11 min snapshot, leading to an angular resolution of approximately 34 arcsec. This image has a total field of view of 11.35 deg^2 in a circular region. We measure a noise standard deviation of 25 mJy by generating a noise map of the image using the sourcefinding script PySE (Carbone et al., submitted)

3.3.2 Observational Detection Bounds on Transient Fluence

We quantified the effect of a threshold based flagging strategy on a bright transient by simulating the presence of such a transient on a dataset with a realistic noise and RFI instantiation. Assuming no sources in the field, the measured LOFAR System Equivalent Flux Density (SEFD) for a core station in HBA mode is 3 kJy at 150 MHz [van Haarlem et al., 2013b]. To calculate the expected noise on a single visibility amplitude of a single polarization, ΔS , under the assumption of a high SNR, we use the equation

$$\Delta S = \frac{1}{\eta} \frac{SEFD}{\sqrt{2} \sqrt{\delta\nu t_{int}}} \quad (3.1)$$

where $\delta\nu$ is the channel width, and t_{int} is the integration time. For our data, the system efficiency η is assumed to be 1, and the two stations making up the baseline are assumed to have the same SEFD. The subband of 781 kHz was spectrally resolved into 16 channels, with visibility integration times of 10 seconds. This corresponds to a noise ΔS of 3.04 Jy per visibility of a single polarization. Thus, if the time selection threshold is set at its hard-coded default as 3.5σ , we can calculate that any transient signal exceeding $S_{flag} = 10.6 \text{ Jy}$ will be flagged for these observing parameters.

For determining the theoretical flagging bounds, we used the single sample threshold due to its effect on short duration (1 sample) transients. At the single sample level, two thresholds are

applied to the data by AOFlogger: χ_1 , and the time selection threshold of 3.5σ . We calculate the flagging bound based on the time selection threshold due to it being the more conservative (lower) of the two.

We estimated the standard deviation of the visibility amplitude over the 11 minute observation for a typical baseline, and this was found to be ~ 2.5 Jy. We chose to carry out our analysis with the per visibility noise derived from the empirical SEFD due to the expectation of LOFAR sensitivity to be of that order, and the fact that there can be systematics in the visibility time series such as fringes due to bright sources which can bias the measured statistics.

When flagging is applied on high time and frequency resolution visibilities and then transient detection is carried out on the integrated image, transients can be lost due to two reasons:

- The transient flux per visibility exceeds the used flagging threshold, or
- The transient fluence (flux \times duration) is lower than the image detection threshold.

Thus, in order to be detected via imaging, transients need to lie in the region where they are bright enough to be visible in the integrated image, but not bright enough to be flagged.

3.3.2.1 Effect of transient duration on its detection

We can compute the duration a transient should have to be visible in an integrated snapshot by limiting the transient flux to be < 10.6 Jy to prevent its visibility flagging. Due to computational limitations, typical searches for these brief signals in the image domain will be in integrated images, such as the 11 minute observations of LOFAR. As outlined by Trott et al. [2013], if an image is longer than the width of the transient signal duration, then we can estimate the minimum transient signal flux density, $S_{min,w}$, that LOFAR is sensitive to using:

$$S_{min,w} = S_{min,11min} \frac{\Delta t}{w} \quad (3.2)$$

where $S_{min,11min}$ is the sensitivity of one snapshot image multiplied by the detection threshold, $\Delta t = 660$ s is the snapshot integration time, and w is the duration of the transient. The background noise of the 11 minute integrated image is 25 mJy beam $^{-1}$, as measured by our observations, and we set a 5σ detection limit for a source in the image plane, corresponding to a source flux of $S_{min,11min} = 0.125$ Jy in an image with $\Delta t = 660$ s. From Equation 3.1, we showed sources with a flux greater than 10.6 Jy will be flagged. Therefore, using Equation 3.2, we can show that transients with a duration $w \lesssim 8$ s would not be detected due to the channel selection step, although transients of longer duration could also be flagged due to other parts of AOFlogger such as SUMTHRESHOLD or surface fitting.

3.3.2.2 Fraction of transients lost

We can quantify the fraction of transients lost due to flagging from time selection and fluence by assuming that the number density of transient sources is constant in a Euclidean universe. Therefore, knowing the cumulative flux distribution, $\log(N) - \log(S)$, of the sources, and assuming a simple power law, $N(> S) \propto S^{-\alpha}$, we can determine the fraction of sources expected to be lost due to flagging via:

$$\frac{N(S > S_{flag})}{N(S > S_{min,11min})} = \frac{S_{flag}^{-\alpha}}{(S_{min,11min} \frac{\Delta t}{w})^{-\alpha}} \quad (3.3)$$

For $w < 8$ s, all transients will be lost. For longer durations, let us take the concrete example of a Euclidean universe, where transient sources are evenly distributed and radio source count for transient sources brighter than a given flux would follow a cumulative distribution for $\log N - \log S$ $\alpha = 1.5$. For a transient which lasts 5% of the measurement set, or 33 seconds, $\sim 11\%$ of transients detectable in the image will be flagged based on Equation 3.3. These transients are significant because they give the highest S/N, and are the best candidates for multi-wavelength follow up and understanding the physics behind the source population.

The fraction of transients expected to be rejected via a flagging operation is dependent upon the transient duration and the extent of the imaging dwell time. It increases as the transient duration reduces to approach the hard limit of 8 seconds, after which all transients are undetectable. The fraction flagged then decreases as the transient duration increases to that of the snapshot integration. With theoretical boundaries set up, we populated the range of fluences expected to be affected by a default flagging strategy in order to quantify the effect of flagging on transient detection.

3.3.3 Transient Simulations

In order to quantify the effect of a flagging algorithm based on applying a set of thresholds on the visibility amplitude, we added simulated transients onto an actual LOFAR observation which would have a realistic noise and RFI instantiation. We focused our efforts on an observation at 149 MHz consisting of a single snapshot of 11 minutes and with a total bandwidth of 781 kHz, where the visibilities were already calibrated. These data were available at a spectral resolution of 48.8 kHz and a temporal resolution of 10 seconds. As described earlier, data were already flagged once with AOflogger, and the visibilities calibrated. The tests we describe here involve a secondary RFI flagging step after the transient was simulated and added to the calibrated visibilities, but before imaging. This was done in order to study the effect of the flagger on this one injected transient signal in particular, as this data was very low in the amount of RFI it contained.

For these simulations a transient source was injected in the center of the image by taking a model point source, transforming it to visibility space, and adding this to the recorded visibilities. This was done using BlackBoard Selfcal (BBS) software [Loose et al., 2008] where the temporal profile of the point source was a top hat signal of a given strength and duration in the image's sky model. After this injection, the data were processed as usual through the imaging pipeline described in Section 3.3.1, and the resulting image was then analyzed.

This transient's amplitude was varied from 0.5-10 Janskys in half Jansky intervals, and the duration was varied from one second to 2 minutes on a logarithmic scale. This range, based on estimates laid out in Section 3.3.2, was selected to probe the range of χ thresholds which depend on both the amplitude and connectivity of visibilities. The resulting images for these data were then run through the sourcefinding script PySE where the integrated flux at the location of the injected transient was measured via a forced fit of a point source. We should note that the data we injected the transient into had already been run through RFI processing during pre-processing

when it was first obtained for analysis— that is, the typical RFI processing for LOFAR imaging. The tests we describe here involve a secondary RFI flagging step after the transient was simulated and injected into the data but before imaging in order to study the effect of the flagger on this one injected transient signal in particular, especially as these data contained a very low amount of RFI.

The results of this test before any secondary flagging steps were introduced can be seen in Figure 3.3. Here, we see the brightest detected transients correspond with the brightest simulated transients with the longest duration, and the measured flux of the transient (indicated by the color bar) is consistent with expectations from Equation 3.2. This measured flux then decreases for transients of shorter durations and brightnesses, until they are undetectable because they are either too faint, too brief, or a combination of the two. After this, the same tests were conducted but with an automatic flagging step added after the transient was injected into the data in order to measure the effects of RFI flagging on the transient signal.

We also performed a similar simulation for radio transients of varying brightness but with 2 minute long snapshots, which was the same duration as the Rowlinson et al. [2016] MWA survey for brief transients, flagged and calibrated using 2 minute measurement sets but imaged in 30 second increments. This was done in order to understand how flagging for transients would affect shorter measurement sets, because the number of data samples is much smaller for such observations. For these measurement sets, we focused on transients that were up to 60 seconds duration in the data, or half the length of the observation. Further, we note that because AOFlagger has a recommendation of at least 1,000 time steps to be passed through the flagger, and this is impossible for MWA-length measurement sets, we edited the time window in which the measurement sets are divided for flagging into smaller increments than is recommended.

3.4 Effect of Flagging Strategies on Simulated Transients

3.4.1 Default AOFlagger Settings

We first considered only the effects of the standard AOFlagger settings on the flagging of transients, in order to characterize the typical effects of such flagging on bright transients. In an 11 minute measurement set, we found that if the simulated transient was of a longer duration than two minutes there was no statistically significant difference in the observed flux of the transient—that is, the transient was unaffected by flagging algorithms because its long duration would not trigger the thresholding algorithm. However, when considering brief transients of durations less than two minutes, as seen in Figure 3.3, differences between the two become apparent. If automated flagging is used, very brief transients ($\sim \Delta t < 60$ s) of a detectable brightness will be flagged out altogether by the automated flagging software because of the single sample threshold, χ_1 , designed to eliminate very bright and sporadic RFI.

It should be emphasized that the transient signal measured at the end of this procedure would be weaker than the original signal injected, as described by Equation 3.2. Thus, in Figure 3.3, for example, the transient signal flux measured is less than the amount injected.

Longer, brighter transients ($60 \text{ s} < \Delta t < 120 \text{ s}$), however, are no longer flagged completely by AOFlagger, but do appear to have a lower flux when compared to the original strength of the

3.4 EFFECT OF FLAGGING STRATEGIES ON SIMULATED TRANSIENTS

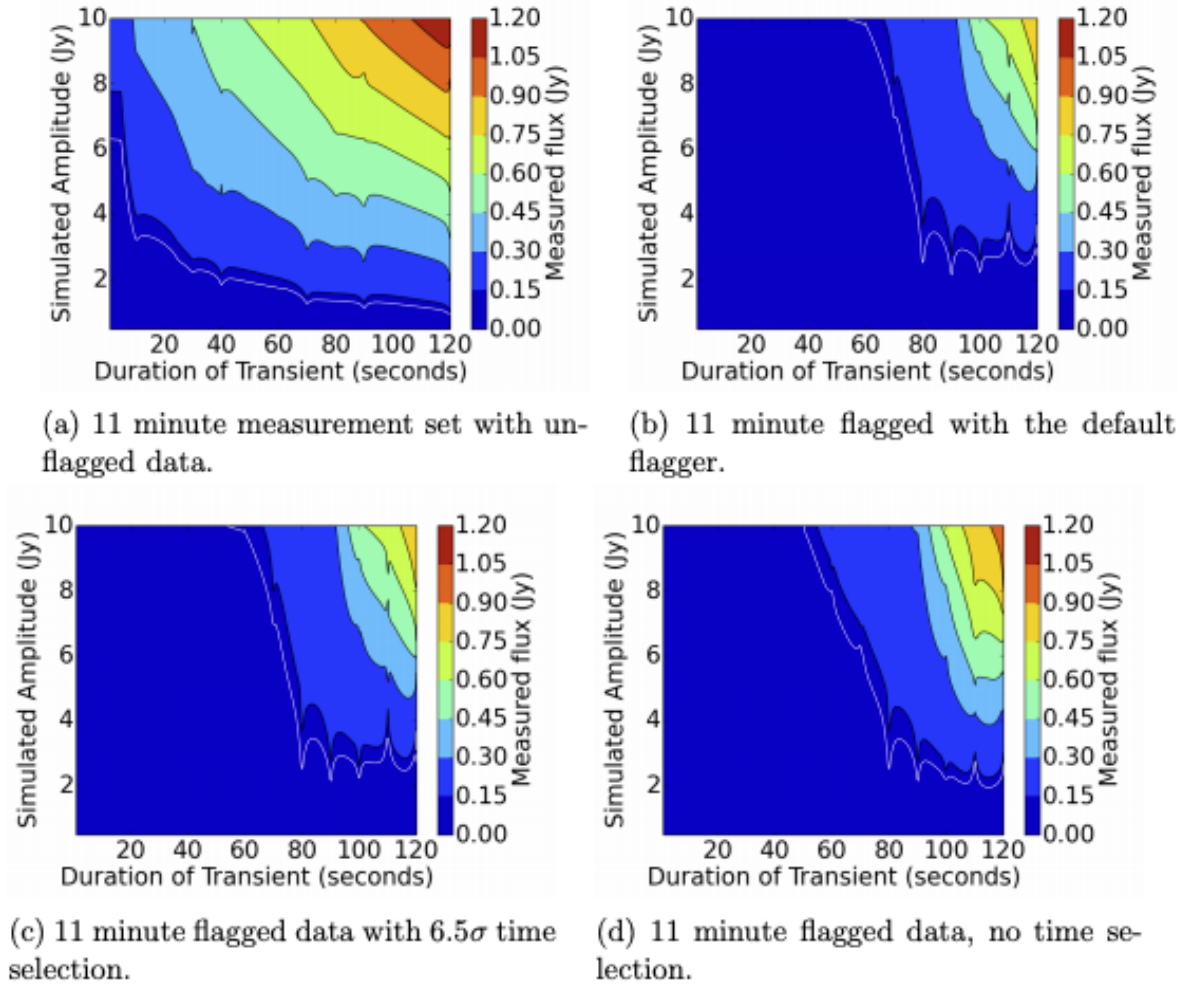
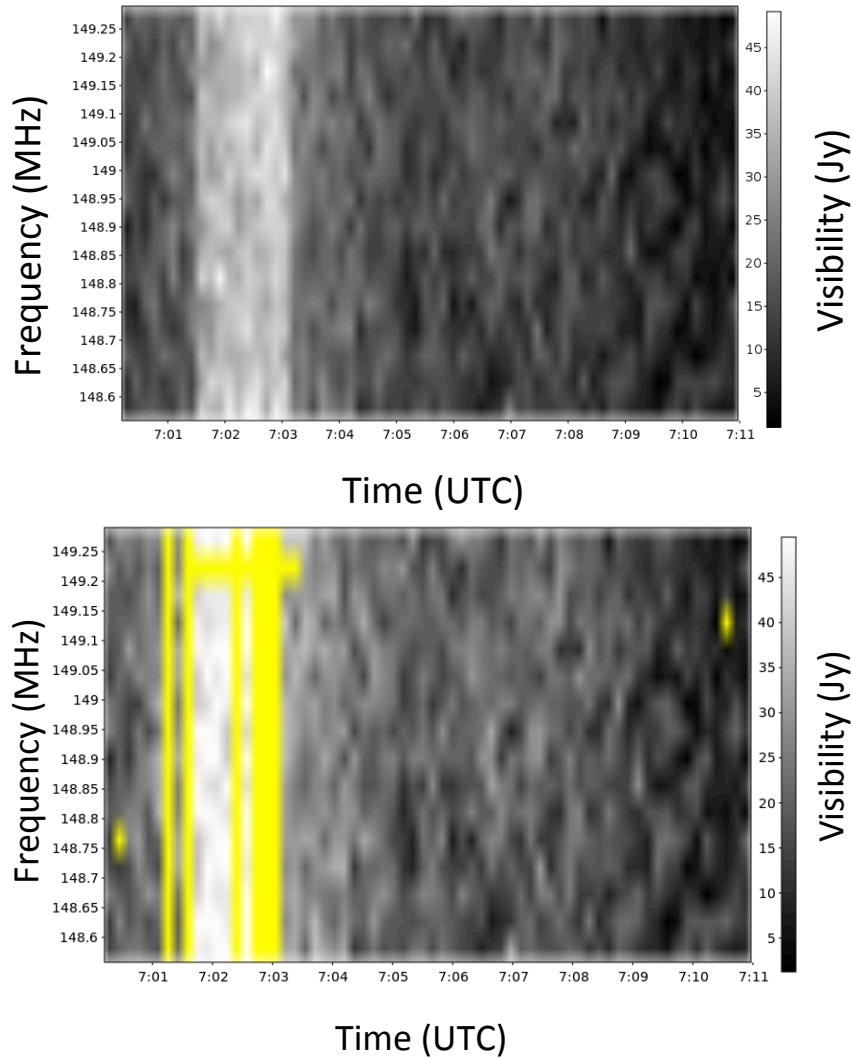


Figure 3.3 Observed fluxes for transient signals of a given amplitude and duration in an 11 minute measurement set. Here, the top left plot is with unflagged data, the top right plot is for the default flagger, the bottom left is with the modified flagger with SUMTHRESHOLD only in the time direction but a time selection $\sigma > 6.5$, and the bottom right is with SUMTHRESHOLD only in the time direction and with no time selection steps. The unusual contours in the flagged data between levels is due to a sampling effect caused by the sampling window size of the RFI algorithm and brightness of the transient. The white contour corresponds with $S_{min,11min} = 0.125$ Jy, which is the 5σ limit derived from the RMS map of the image.



Figuur 3.4 An illustration of "partial flagging" in the time-frequency domain for two antennas from the program "rfigui," where the top is an unflagged data set and the bottom has been flagged by AOFlagger (yellow). Here we have an 11 minute data set with a top hat transient from an injected signal with 10 Jy amplitude and 90 seconds duration. A sizeable percentage of the transient is flagged as RFI but not the entire signal. As such, the transient would be detected, but not at its full brightness or duration. We note that some data around the transient which does not contain the injected signal are also flagged.

3.4 EFFECT OF FLAGGING STRATEGIES ON SIMULATED TRANSIENTS

injected transient. A check of what is being flagged in the visibilities reveals that the transient signals are becoming "partially flagged," where some, but not all, of the signal is being flagged as RFI. An example of such partial flagging can be seen in Figure 3.4. Exactly how much of the signal is being flagged is dependent on the intensity and duration of the signal. We also note that sections of the non-transient data are also being flagged by the flagger in order to eliminate any potential fringes of RFI within the data due to dilation, which can further affect the measured fluxes in the data stream. For example, for the unflagged data, the 10 Jy transient with 2 minutes duration had a measured flux of 1.24 Jy (Figure 3.3). For the flagged data, the same transient was measured as 1.44 Jy in brightness. This was because while $\sim 30\%$ of the transient signal was flagged, a large fraction of data surrounding the signal which did not contain a transient was also flagged by AOflogger, making the effective Δt of the measurement set shorter than for the unflagged data. This can have a significant effect: for example, a 4 Jy transient of 100 seconds duration is easily detectable in the unflagged data, but does not exceed the 0.125 Jy detection threshold in the image.

For fainter transients of $\lesssim 3$ Jy, we also see flagging occurring. This is because the connectivity of the data points making up the transient signal trigger the lower thresholds of the SUMTHRESHOLD algorithm. Excluding the parts of Figure 3.3 where transients are too diluted to be detected, we find that $\sim 10\%$ of transients are flagged due to this underlying connectivity. This percentage was arrived at by taking the ratio of the number of surviving transients to the total number of injected transients.

Similar effects are seen in the measurement sets of 2 minutes duration. These results can be seen in Figure 3.5, where again the upper left plot shows the fluxes measured in unflagged data, and the upper right plot shows results from the flagged data set. We see a similar plot to what was seen before with the longer measurement set in that the briefest transients ($\simeq \Delta t < 25$ s) are all entirely flagged. After this, transients of longer duration ($25 \text{ s} < \Delta t < 60 \text{ s}$) are visible, but not at their full strengths because they are partially flagged in visibility space. This can have consequences: for example, a 10 Jy, 20 second transient would be detectable in the unflagged data, but would not exceed the detection threshold in the flagged data.

Thus, flagging both template datasets results in loss of the brightest transients, as well as dilution of the recovered flux of unflagged transients.

3.4.2 Modified Flagger Results

In light of the results with default flagging, we carried out modifications of the flagger parameters, and tested the modified flagger on the same data. Specifically, this test focused on two points: adjusting SUMTHRESHOLD to work "in time direction" only, and to increase the σ cutoff value for the time selection steps. These tests were carried out for both the 2 minutes duration and 11 minutes duration measurement sets.

First, we note that only adjusting one parameter but not the other did not yield a significant difference in flagging results compared with the default AOflogger settings— that is, if SUMTHRESHOLD was adjusted the transient would still be flagged as RFI by the $\sigma > 3.5$ time selection step, and vice versa. When adjusting both SUMTHRESHOLD to only flag "in time direction" and increasing the sigma cutoff for time selection, however, does yield results that are noticeably different

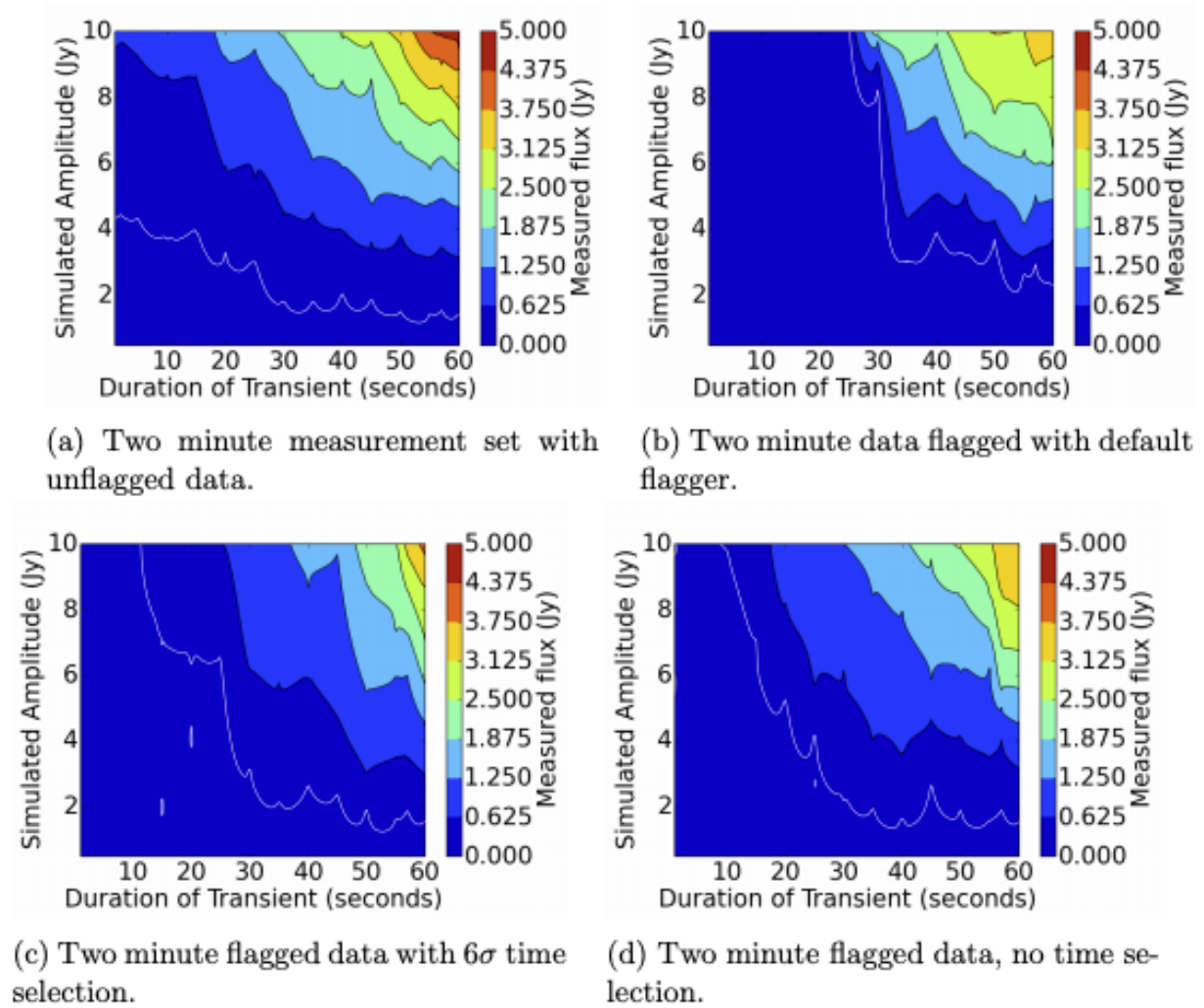


Figure 3.5 Observed fluxes for transient signals of a given amplitude and duration in a 2 minute measurement set. Here, the top left plot is with unflagged data, the top right plot is for the default flagger, the bottom left is with the modified flagger with SUMTHRESHOLD only in the time direction but a time selection $\sigma > 6.5$, and the bottom right is with SUMTHRESHOLD only in the time direction and with no time selection steps. The white contour corresponds with $S_{min,11min} = 0.55$ Jy, as measured from the RMS map as calculated by the sourcefinding script PYSE.

3.4 EFFECT OF FLAGGING STRATEGIES ON SIMULATED TRANSIENTS

from the default AOFlagger settings.

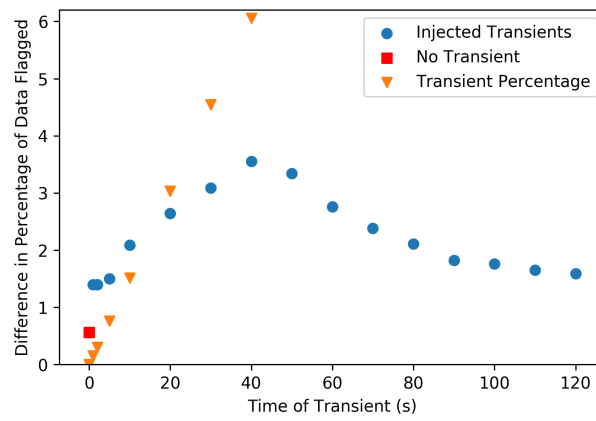
For the 11 minute measurement set, two such examples can be seen in the lower plots of Figure 3.3. In both plots SUMTHRESHOLD is set to “in time direction” only, but in the lower left panel the time selection step is set to $\sigma > 6.5$, which we chose as an arbitrary cutoff to demonstrate the effect of a higher threshold for time selection. In the lower right plot, the time selection is deleted from the algorithm altogether. We see that although all the transients less than 60 seconds in duration are still flagged, there is some improvement in the number of transients that are visible because they are only partially flagged in both cases. The scenario where the time selection steps are eliminated also shows better results than the panel with a $\sigma > 6.5$ time selection step included. For example, while in the lower left the $\lesssim 3$ Jy signals are all flagged by the lower threshold, in the lower right only the $\lesssim 2.5$ Jy signals are now affected. This can be attributed to the decoupling of the time dimension in the SUMTHRESHOLD connectivity based thresholding. Further, the 4 Jy, 100 s transient described earlier which was flagged out would now exceed the detection threshold with the modified flagger that had no time selection present in the default AOFlagger, although would still be flagged if 6.5σ time selection was applied.

We see similar results for the 2 minute measurement set, where the modified flagger results are shown in the lower plots of Figure 3.5, with the $\sigma > 6.5$ cutoff on the left hand side plot, and eliminated altogether in the lower right plot. Here, while all transients of a duration less than 20 seconds are still flagged, we have partial flagging occurring on a larger fraction of transients that are of ($25 \text{ s} < \Delta t < 60 \text{ s}$) duration. Once again, the lower right panel, which contains no time selection step, flags fewer transients than the lower left panel with a time selection step of $\sigma > 6.5$. Here, for example, we see that the 10 Jy, 20 second transient which did not exceed the detection limit with the default flagger now exceeds the detection limit in both modified flaggers.

3.4.3 Observer Alert Script

We have developed a script which can alert the observer when a potential transient is in the data stream. This is based on a comparison on the percentages of data flagged between the two RFI flaggers, and various criteria which are expected to be different between transient RFI and an astrophysical signal. In this script, first the RFI flagger checks the percentage of data flagged for the entire data set in each channel- a standard output given in AOFlagger- for both the default and modified strategies. Although we expect the modified flagger to flag slightly less data than the default flagger, due to less stringent criteria, a transient in the data stream would also create a difference in the percentage flagged between the two methods. As such, the difference between the percentage flagged can then be used to then determine whether a signal is in the data stream which may be worth further examination through further selection criteria.

To illustrate this phenomenon, we show the difference in percentages flagged for a 10 Jy signal of various durations injected into an 11 minute observation in Figure 3.6, illustrated with blue points. We include the difference in percentage between the two strategies when no transient is injected in the data stream (red), and the percentage that would occur if the default flagger fully flagged the transient but the modified flagger did not (yellow), for reference. What we find is even for the shortest duration transient we tested, of 1 second length, there is already a $\sim 1.5\%$ difference between the percentage of the measurement set flagged between the default and modified



Figuur 3.6 The difference in percentage of data flagged between the default and modified strategies for an injected 10 Jy signal in an 11 minute measurement set, for different durations of time (blue). We provide the difference for the same data set with no injected transient (red) for reference. For the injected flagged transients, the difference between the two strategies steadily increases until it peaks at 40 seconds, and then begins to decrease (there is 1-4% more data flagged by the default flagger, depending on the transient duration). We attribute this decrease to the emergence of partial flagging, described in Section 3.4.1, in transients of this duration and longer. For reference, we also provide the percentage of the transient if it was completely flagged by the default but not at all by the modified flagger (yellow) until 40 seconds, when partial flagging begins to occur. We will note that error bars in the average values in this plot were included, but are too small to be visible.

flaggers, which exceeds the amount seen with no transient (where there is a 0.53% difference). Visual inspection of the data flags in rfigui confirmed this is because while both flaggers will flag these shortest transients in our case (Figure 3.3), the default flagger will additionally flag extra data before and after the transient which is not part of the signal. The percentage flagged will then steadily increase as the percentage of time the signal is in the measurement set increases, until it peaks at 40 seconds, and begins to decrease. This is the duration where partial flagging becomes a factor, where a fraction of the transient signal is flagged instead of the entire signal. This results in the difference in percentages flagged beginning to decrease, although this difference still remains consistently higher than when there is no transient in the data.

Based on these tests, we decided that any difference in percentage of data flagged between the two strategies that exceeds 1% is of interest for further inspection for these transients. For this, we used PYRAP to sort the data by time series and to search for the following:

- First, we searched the data set for baselines where the columns FLAG_ROW were set, which occurs when all the data in a row of visibilities are flagged. We chose this parameter because a broadband transient signal is expected to be visible in all frequency channels in our data set, but a flag in only a few channels is expected to be RFI. We should note that because LOFAR data sets often include a flagged row in the final time step of data, we set our script to disregard the last ten seconds of the measurement set.
- Next, we searched the flagged row by baseline. Specifically, we were interested in occurrences when FLAG_ROW was true for a majority of baselines at a given time. We chose this parameter because a flagged row only present in a small number of baselines is likely caused by local interference, whereas an astronomical signal would be visible in most baselines (although not all, depending on the noise within that specific baseline).
- Finally, if all these criteria are met, the script returns a note to the observer, informing them a potential transient is present in their data.

We tested our script on our transient signals seen in Figure 3.6, which were 10 Jy signals in a 11 minute measurement set. We found that when we used the above criteria, our script returned a potential transient alert for all our signals. However, raising the difference in percentage flagged to 3% only returned a potential transient alert for the transients of 30-50 second duration, and no alert was returned when the difference was set in excess of 4%.

3.5 Discussion

3.5.1 Observed AOflogger Effects

From our simulation work, it appears that AOflogger, which is relied on by many radio surveys searching for brief transient signals, could be flagging out existing bright, brief transients in the data stream. In particular, it is possible all bright transients of just a few seconds' duration sufficiently bright to be detected in the integrated images have been flagged in previous surveys such as those of Stewart et al. [2016], if they were present. Re-examination of the data's flagged sections with the observer alert script could aid transient detection. Using the observer alert

script does not require the more computationally intensive imaging steps to identify previously flagged data which merits further inspection, but any astrophysical transient would originate from one point in the sky and converge to a point source upon imaging. RFI, on the other hand, is much more likely to originate from the ground, meaning it would result in a noise-filled image if only affecting a few baselines.

Even when a transient is detected using default flagging methods, however, consideration should be given that the signal can be subject to partial flagging by RFI detection algorithms. Further, some non-transient data is also flagged by the flagger, which can further affect measured fluxes. Checking any candidate transient signals for partial flagging in the data stream (if possible) should be carried out in any transient detection follow up to determine the true flux density.

3.5.2 Effect of the Duration of an Observation

Further, it is important to consider the merits behind searching for very brief transients on the order of a few seconds in an 11 minute observation. Although using such an observation length is not ideal for transients of this duration, often such observations are “piggybacked” for transient searches from observations such as sky surveys. For such a survey, even if imaging is done using shorter timescales than 11 minutes, the data are typically flagged over the full length of the measurement set before it is sliced into smaller increments for imaging in order to obtain a sufficiently large number of data points for flagging to be conducted. This means that even if imaging is done in smaller increments that are more ideal for transient searches, potential transient signals on the order of a few seconds duration would be flagged.

In addition to this, our tests using a 2 minute observation show transients on the order of a few seconds duration are flagged even when flagging is conducted over a shorter time span when using LOFAR data. There are surveys being conducted at these time scales, such as Rowlinson et al. [2016] which sought to constrain the transient rate by searching for signals on the order of a few seconds by taking 2 minute measurement sets from MWA and slicing them into 30 second images. However, due to the differences between the LOFAR and MWA telescopes in SEFD and channel width, while the brightest transient signals would have been flagged, if they existed, it seems likely that fainter transient signals would have survived the flagging process. Repeating the simulations in this work using an MWA measurement set would answer where precisely these cutoffs lie. Further, if the brightest transient signals were flagged, the subsequent transient rate estimates from this and other surveys currently do not incorporate the effects of RFI flagging, and this should be taken into account.

3.5.3 Effects of AOFlagger Modifications and Detection Script

There are some modifications to AOFlagger that can be implemented in order to detect more transients than one would with the default settings (though, it should be emphasized, not all transients). In particular, for AOFlagger, setting SUMTHRESHOLD to only operate in the time direction, and eliminating all time selection steps (i.e., only use frequency selection), would be advisable when seeking transients of \sim minute long durations within snapshots. For transients on the order of tens of seconds long, flagging on a shorter time scale is advised, but consideration

must be given to the number of time steps and resolution in the length of the data stream to ensure flagger accuracy.

We note that while these modifications to the RFI flagger will still result in automated flagging of narrow-band RFI, and not have significant effects on low-level RFI in general, implementing such a modified flagger opens up the possibility that bright, broadband RFI signals will no longer be flagged.

We have also presented in Section 3.4.3 a potential solution in which the differences between the default AOFlagger and modified version, along with the known distinguishing properties between RFI and transient signals, are used to alert the observer when a transient may be present in the data stream. This script could be further modified depending on the class of transients in which the observer is interested- for example, a search for minutes-long transients such as that described in Stewart et al. [2016] could use a higher difference in percentage flagged than 1%, which would decrease the number of false positives. Further, the data could be examined for the sequential occurrence of the script criteria in several time stamps, which would indicate a potential transient of longer duration.

We would recommend that an observer using this method choose their own parameters for the difference in percentages flagged, and the percentage of baselines with completely flagged visibilities at a given time, as these are dependent on the location of the telescope, and the resolution of the data acquired by the observer. In particular, because some transient-like RFI will still meet the criteria outlined above, it is inevitable that a percentage of false positives will be present in the data stream. We will investigate the effects of a modified flagger with alert script on a high-RFI observation with injected transients and different data resolutions in a future publication.

Another possibility could be to also implement the dictionary approach outlined by Czech et al. [2017], which focused on filtering so-called transient RFI (defined in Czech et al. [2017] as non-constant RFI). This work demonstrated the use of Markov models to identify transient RFI as a sequence of sub-events, and could reliably extract transient RFI events from a larger observation. Given the similarities between transient RFI and astronomical sources, investigating the differences between the two categories with this method could provide another useful tool in distinguishing between the two.

3.6 Conclusions

Using the typical LOFAR imaging survey mode as an example, we have investigated the effects of automated RFI flagging on transient signals of astronomical origin. We have demonstrated that RFI flagging with AOFlagger can flag out brief transients < 60 s in duration. Further, we have calculated that all bright transients > 10.6 Jy would be flagged when the integration is 10 s and the frequency integration is 48 kHz. Fainter signals of a few Jy brightness are also affected, depending on the extent of connectivity of the signal in the visibility plane.

For transient signals of > 60 s length in a measurement set of 11 minutes duration, the majority of the expected bright transient signals will survive the flagging process- albeit with diminished flux due to “partial flagging”- and we have a good chance of observing these transients. However,

the original data would need to be reexamined in order to determine whether part of the signal was mistakenly flagged. Reexamining the data may also reveal shorter transients already flagged in the data stream, which face a significant loss. For a shorter LOFAR measurement set of just 2 minutes duration, it is equally likely that 100% of radio transients of up to 20 seconds duration would be flagged. Approximately 60% of transients on the order of tens of seconds or longer will be detected, but partially flagged.

We have presented a script which can be used to alert an observer that the flagged data contains a potential astrophysical transient, which identified our simulated transients which had been flagged. This could be used for any additional investigations into whether there are any transients in already flagged RFI data, in both previous survey data and in any future data preflagged to default settings.

For future transient surveys, in order to minimize the flagging of bright transients in the image plane we recommend modifying AOflogger by only applying thresholding in the frequency direction, and eliminating the time selection steps in the algorithm that rely on sigma clipping. With these steps in place, a larger fraction of bright transient signals will be seen, though likely partially flagged. We also recommend the implementation of an observer alert system based on the differences between transient RFI and astronomical sources, to allow the correct identification of these signals.

Acknowledgements

We acknowledge support from the European Research Council via the Advanced Investigator Grant no. 24729. This work is also supported in part by European Research Council Advanced Grant 267697. LOFAR, the Low Frequency Array designed and constructed by ASTRON, has facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the International LOFAR Telescope (ILT) foundation under a joint scientific policy. We would like to thank the LOFAR Observatory staff for their assistance in obtaining and the handling of this large data set. We would also like to thank Andre Offringa for answering some initial questions. S.C. acknowledges funding support from the UnivEarthS Labex program of Sorbonne Paris Cité (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02).