A&A, 709, A253 (2026)
https://doi.org/10.1051/0004-6361/202659339

Astronomy & Astrophysics

A citizen-science catalogue of interplanetary Type III solar radio bursts and first statistical results

Aikaterini Pesini[1,2,★], Antonio Vecchio[2,1], Xavier Bonnin[1], Milan Maksimovic[1], Sandor Kruk[3], Nicole Vilmer[1], Corentin K. Louis[1], Sophie Musset[4], Marc Klein-Wolt[1], Vratislav Krupar[5,6], and Heino Falcke[2]

[1] LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CY Cergy Paris Université, CNRS, 92190 Meudon, France
[2] Radboud Radio Lab, Department of Astrophysics, Radboud University Nijmegen, Nijmegen, The Netherlands
[3] European Space Astronomy Centre, Madrid, Spain
[4] Johns Hopkins University Applied Physics Laboratory (JHU/APL), North Laurel, Maryland, USA
[5] NASA Goddard Space Flight Center (GSFC), Greenbelt, Maryland, USA
[6] University of Maryland, Baltimore County (UMBC), Baltimore, Maryland, USA

Received 5 February 2026 / Accepted 1 April 2026

ABSTRACT

Aims. We present the first extensive catalogue of interplanetary Type III solar radio bursts compiled through human participation, and provide a statistical analysis of burst occurrence, peak flux, and frequency drift.

Methods. The catalogue is based on Radio and Plasma Waves (RPW) observations from Solar Orbiter, collected via the citizen science campaign Solar Radio Burst Tracker on Zooniverse.org with contributions from 1952 volunteers between February 2020 and March 2025. A dedicated post-processing pipeline extracted each burst's time–frequency extent, peak flux, and quantitative uncertainties, enabling statistical analyses.

Results. The catalogue comprises 15 934 bursts. Burst occurrence strongly correlates with sunspot number and shows modulation with Solar Orbiter's heliocentric distance. Peak flux distributions across 23 frequencies follow power laws with an average slope of -1.56 ± 0.004, steepening at higher frequencies. This is consistent with earlier studies of smaller samples. At still higher frequencies, slopes plateau around -1.7 to -1.8, resembling self-organized criticality behaviour. Burst occurrence peaks at 1–2 MHz, reflecting efficient electron-beam transport and beam–plasma interaction. Frequency drift rates for 13 074 bursts scale as $\mathrm{d}f/\mathrm{d}t = -0.001\, f^{1.37}$, differing from prior relations over other frequency ranges.

Conclusions. This human-validated, large-scale catalogue provides a robust statistical resource for Type III bursts, capturing solar cycle and observational effects, extending analyses to faint and complex events, and constraining frequency-dependent flux and drift rates. It also offers a benchmark for future automated and machine-learning detection methods.

Key words. catalogs – Sun: activity – Sun: radio radiation

1. Introduction

Solar activity follows an approximately 11-year cycle and manifests through a wide variety of phenomena (e.g. Hathaway 2015). Among the most prominent are sunspots, which form as a result of strong magnetic fields emerging through the solar surface and appear darker than the surrounding photosphere due to their lower temperature (e.g. Solanki 2003). Other key manifestations of solar activity include solar flares, coronal mass ejections (CMEs), and solar jets, which release vast amounts of energy into the solar corona and the interplanetary medium (e.g. Benz 2017; Chen 2011; Raouafi et al. 2016). Solar flares are intense bursts of electromagnetic radiation produced by the rapid release of magnetic energy. They are observed across a broad range of wavelengths, including X-rays, extreme UV, Hα, and radio, and typically last from minutes to several hours (Fletcher et al. 2011; Shibata & Magara 2011; Benz 2017). A fraction of the released energy is transferred to the acceleration of charged particles up to semi-relativistic energies, resulting in enhanced X-ray and radio emission (Wild 1950; Ginzburg & Zhelezniakov 1958; Dulk 1985; Klein 2021).

Type III solar radio bursts are generated by beams of energetic electrons, typically in the 1–100 keV range, accelerated during episodes of magnetic energy release (e.g. Priest & Forbes 2000). These energy releases can originate from solar flares, jets, or CME-related phenomena, which accelerate electron beams that propagate along open magnetic field lines into the interplanetary medium. As these beams travel outwards, they produce the characteristic rapidly drifting radio emission observed as Type III bursts (Wild 1950; Ginzburg & Zhelezniakov 1958; Suzuki & Dulk 1985; Henri et al. 2009; Mann et al. 2018; Paipa-Leon et al. 2025). Due to these accelerated electrons, a bump in the local velocity distribution function is created, triggering the so-called 'bump-on-tail instability', which leads to the generation of Langmuir waves (e.g. Sarkar et al. 2015). These Langmuir waves ultimately produce radio emissions through wave–wave interactions, generating the fundamental component at the plasma frequency (f_p) and its harmonic at $2f_p$ (Ginzburg & Zhelezniakov 1958; Zheleznyakov & Zaitsev 1970; Melrose 1980). This mechanism leads to the characteristic Type III burst structure, which starts at higher frequencies and drifts rapidly towards lower frequencies in dynamic spectra (DSs) that map radio power over time and frequency. This signature reflects the motion of energetic electron beams through

---

[★] Corresponding author: `katerina.pesini@ru.nl` regions of decreasing plasma density in the corona and interplanetary space, as the emission frequency is proportional to the local electron density ($f_p \sim \sqrt{n_e}$) and thus decreases with increasing heliocentric distance, with higher-frequency components originating near the solar surface and lower-frequency ones corresponding to regions farther from the Sun.

Type III radio bursts exhibit significant variability in both the frequency range over which they occur and their duration, and are characterised by exceptionally high frequency drift rates compared to other classes of solar radio emission (Reid & Kontar 2015). Type III bursts typically originate at frequencies below ~100 MHz and can extend down to tens of kilohertz; such events are commonly referred to as interplanetary Type III bursts. The low-frequency component of the emission generally persists for longer durations than the high-frequency component, which appears earlier in time. Within the framework of the plasma emission mechanism, the observed frequency drift rate is controlled by both the radial velocity of the exciter electron beams and the gradient of the background electron density as a function of heliocentric distance (Alvarez & Haddock 1973).

Solar Orbiter, a joint ESA-NASA mission launched in 2020, carries the Radio and Plasma Waves (RPW) instrument, which is designed to measure in situ magnetic and electric fields as well as solar and heliospheric radio emissions up to 16 MHz (Müller et al. 2020; Maksimovic et al. 2020). Two of the RPW receivers are the Thermal Noise Receiver (TNR; 4 kHz–1 MHz) and the High Frequency Receiver (HFR; 500 kHz–16 MHz) (Maksimovic et al. 2020; Vecchio et al. 2021); together they enable continuous monitoring of Type III bursts across a broad frequency range, including bands inaccessible from the ground due to the ionospheric cutoff, which prevents the propagation of radio waves below 10 MHz.

Several efforts have aimed to automatically detect Type III radio bursts in solar radio data, employing a variety of techniques. These approaches range from threshold-based methods (e.g. Lobzin et al. 2009; Zhang et al. 2018; Salmane et al. 2018) to algorithms designed to identify characteristic signatures of Type III bursts, such as their high frequency drift rates (e.g. Lobzin et al. 2014). More recently, machine-learning and deep learning methods have been employed (e.g. Scully et al. 2021; Bussons Gordo et al. 2023; le Roux et al. 2025). In this project, we present a citizen-science approach for collecting Type III radio bursts, engaging numerous volunteers, and leveraging collective human pattern-recognition capabilities.

Through Solar Radio Burst Tracker[1] we aim to construct a highly detailed and inclusive catalogue of interplanetary Type III bursts at low frequencies, without imposing restrictive assumptions on signal intensity or morphology. In particular, we aimed to avoid the limitations inherent to automatic detection methods based on fixed intensity thresholds or predefined burst shapes, which can bias catalogues against faint events or complex emission patterns. By presenting zoomed-in views of DSs to multiple independent volunteers, the campaign enabled the identification of a wide range of Type III bursts, including weak events and complex cases such as Type III storms or bursts with multiple electron injections. This human-based approach maximises completeness at low frequencies and provides reliable uncertainty estimates through consensus statistics. The resulting catalogue not only offers improved statistical coverage of faint and complex Type III bursts, but also constitutes a valuable benchmark dataset for the development, validation, and performance improvement of future machine-learning and automated detection algorithms.

In Section 2, we describe the methodology of the citizen-science campaign, including its preparation, implementation, and post-processing analysis. Section 3 presents the Type III burst catalogue, detailing its contents and features, and provides an initial statistical analysis of the dataset. Finally, Section 4 summarises the conclusions drawn from these results and discusses potential future applications of the catalogue.

2. Method

In this section, we describe the dataset and the data processing steps designed to optimise the volunteers' experience. We also outline the tasks assigned to the volunteers. Finally, we present the post-processing analysis, which includes data aggregation methods and the extraction of key features from the identifications, ensuring that the resulting catalogue is optimised and suitable for scientific analysis.

2.1. Data preparation

For the citizen-science campaign, we used preliminary Level 3 RPW data products derived from calibrated Level 2 data (European Space Agency 2024), covering 12 February 2020 to 5 March 2025. These data provide radio flux density in physical units ($\mathrm{W\,m^{-2}\,Hz^{-1}}$), processed into regularly sampled time–frequency matrices with 1% background noise subtraction and flux interpolation. Data from HFR and TNR receivers were combined to achieve continuous frequency coverage, limited to 4 kHz–7 MHz due to spacecraft-generated noise at higher frequencies.

A dedicated processing pipeline enhanced the visibility of Type III bursts. For each frequency channel, the median flux level was adopted as a background reference. Data points with fluxes above the median were corrected by subtracting the median value, while data points with fluxes below the median were scaled down by dividing them by 3. Spectra were then smoothed along the frequency axis with a Gaussian filter (kernel width 4), producing clean, high-contrast DSs suitable for visual identification of bursts. Each DS provided to volunteers covered six hours, with overlapping spectra to prevent losses of features near temporal boundaries. Eight spectra were generated per day, covering four main six-hour intervals (00:00–06:00, 06:00–12:00, 12:00–18:00, and 18:00–24:00), three offset by three hours (03:00–09:00, 09:00–15:00, and 15:00–21:00), and one bridging consecutive days (21:00–03:00). Two representative results of the produced DSs are shown in Fig. 1.

2.2. Volunteers' task

In Zooniverse, a subject refers to a single data item presented to volunteers for classification. In the Solar Radio Burst Tracker project, each subject is a PNG image displaying a six-hour DS (Fig. 1).

Volunteers complete a tutorial, which introduces the characteristic Type III burst morphology, and are provided with a field guide with detailed instructions for common scenarios, including faint, overlapping, or storm events. The training emphasised detecting both strong and weak bursts to ensure high completeness and reliability. After training, volunteers annotate Type III bursts via a web interface by outlining features directly on the image.

Each spectrum image was analysed by five different volunteers, and the subject was retired once this number was reached in order to prevent redundancy. In total, 13 647 spectrum images

---

[1] `https://www.zooniverse.org/projects/xbonnin/solar-radio-burst-tracker`

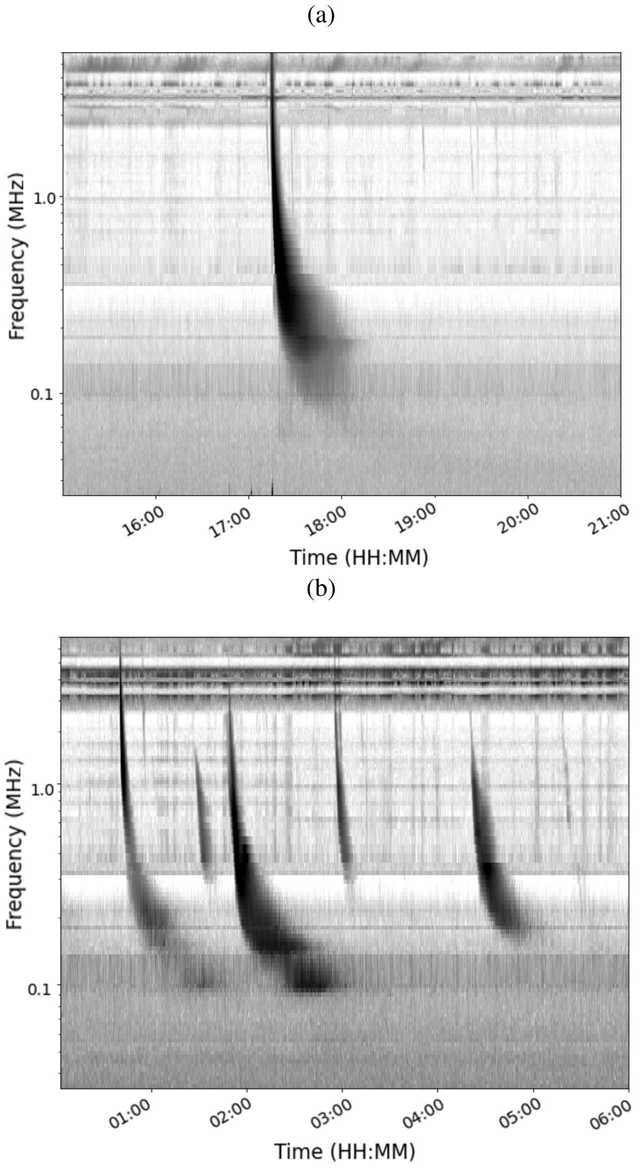


Fig. 1. Examples of the final processed DSs provided to the volunteers of the citizen-science campaign. Panel (a) shows a DS on 9 August 2023, while panel (b) shows a DS on 22 September 2025.

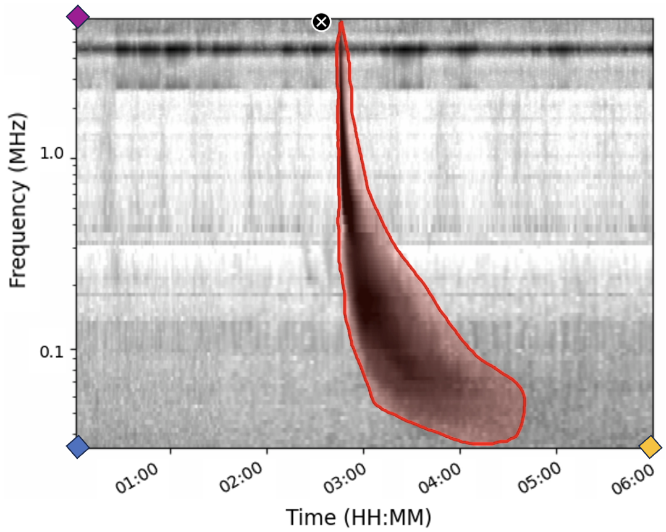


Fig. 2. Schematic of the volunteer annotation task, showing how a Type III burst is outlined (red region) and the three anchor points used to convert pixel coordinates to physical time and frequency: origin (blue), maximum time (yellow), and maximum frequency (purple).

were presented to the volunteers, resulting in 68 235 individual classifications. The total number of volunteers who actively contributed at least one classification was 1952, of which 1201 were unregistered (anonymous) users, and 751 were registered users. The number of subjects analysed per participant exhibits a strongly right-skewed distribution. Most volunteers analysed only a few images, with a median of 9 spectra, while a small fraction of highly active participants analysed up to 2186 spectra. The mean number of analyses per participant was 38, reflecting the contribution of a small number of highly active users. All classifications were completed within 35 days, and each annotation was included in the subsequent analysis.

Each image filename encodes the date and time window, allowing Zooniverse metadata to map annotations precisely to the corresponding spectrum. The user annotations are stored in pixel coordinates (x, y) and converted to physical frequency–time values using three anchor points recorded in a JSON file for each subject:

- the origin (lower-left corner),
- maximum frequency at start time (upper-left corner),
- maximum time at lowest frequency (lower-right corner).

Using these reference points, any pixel coordinate can be mapped to a unique time–frequency location. A schematic of the volunteer task and the reference points is shown in Fig. 2.

2.3. Clustering of multi-user annotations

To analyse volunteer annotations, we first organised the raw data collected from the Zooniverse platform. Each entry corresponds to a user's contribution on a specific image and includes the image filename, username, and annotation data in JSON format. Annotations were grouped by image, including only days with marked Type III activity. Out of 1849 total days, 1449 days (≈78%) contained at least one Type III burst. For these days, the (x, y) coordinates of user-drawn contours were extracted and cleaned by removing self-intersections, resolving inconsistencies, and splitting multi-region annotations into individual components. This resulted in a structured dataset in which each image is associated with a collection of cleaned user-drawn contours.

To identify regions of consensus among volunteers, we applied the Density-Based Clustering Algorithm (DBSCAN) (Ester et al. 1996), adapted for polygon similarity (Appendix A). Clusters were formed only when at least two overlapping shapes were provided by different volunteers. Consequently, isolated shapes that did not overlap with annotations from other volunteers were discarded as unreliable, as they did not reflect consensus. For each resulting cluster, a shape uncertainty (U_s, defined in Appendix A) was computed to quantify the level of agreement among volunteers.

Clustering was first applied on a per-image basis. Since subjects contain overlapping time intervals (Section 2.1), duplicate regions corresponding to the same Type III burst can appear in different files. These duplicates were identified by converting the pixel coordinates of user annotations into physical frequency–time values. The clustering algorithm was then re-applied on a per-day basis in the frequency–time domain, without the user restriction, to merge duplicate shapes arising from overlapping time intervals. When duplicate clusters were merged, the

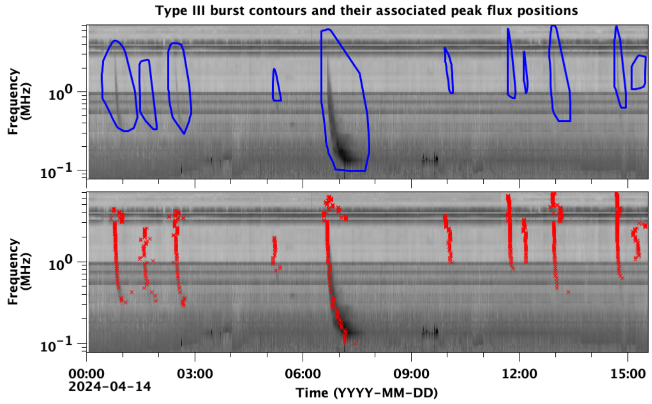


Fig. 3. DSs from 14 April 2024 from both the HFR and TNR receivers, showing the peak-flux positions (red crosses) associated with each selected contour (blue outlines).

number of contributing volunteers was summed, and the shape uncertainty U_s was recalculated as the average of the merged uncertainties. After these steps, each Type III burst is represented by a single, unique region in frequency–time space, accompanied by a shape uncertainty and the total number of volunteers who contributed to its identification.

2.4. Type III burst peak flux and drift rate analysis

After the clustering process, the resulting shapes correspond to unique frequency–time regions potentially containing Type III bursts. These regions were mapped onto the Level 3 data to extract the radio flux enclosed within each contour. Each shape is defined in frequency and time through the transformation of the original (x, y) pixel coordinates. Because the transformed coordinates do not necessarily coincide with the discrete frequency and time grids of the Level 3 data, each point was matched to the nearest available frequency and time sample. In this way, each frequency within a given shape is associated with a corresponding time interval, from which successive measurements provide a light curve at that frequency. For every frequency contained within a shape, the corresponding light curve was extracted from the processed Level 3 data, without applying additional smoothing. From each identified Type III burst, the following physical parameters were derived:

1. the frequency range,
2. the time range,
3. the radio flux values enclosed within the contours, and
4. the peak radio flux.

Figure 3 presents the annotated contours together with the frequency–time peak flux positions of the selected Type III bursts on 14 April 2024, overlaid on the corresponding daily DSs generated with the AutoPlot tool (Faden et al. 2010). The figure highlights 11 Type III bursts.

For erroneous identifications, the peak flux positions do not follow the characteristic frequency–time drift of genuine Type III emission. Therefore, the extracted peak-flux positions serve as a direct indicator of how accurately the annotations trace the expected burst drift.

Type III radio bursts exhibit a characteristic frequency drift, reflecting the decrease of the emission frequency as the electron beams propagate outwards through the solar corona and heliosphere. The dependence of the drift rate on frequency was already recognised in early studies (Wild et al. 1963) and was empirically described by Alvarez & Haddock (1973) as a power law,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = A f^{\alpha}, \tag{1}$$

where $\mathrm{d}f/\mathrm{d}t$ is the frequency drift rate in $MHz \cdot s^{-1}$, f is the emission frequency in MHz, A sets the overall magnitude of the drift rate, and α reflects its scaling with frequency. Based on a statistical analysis of 79 Type III bursts in the 50–350 kHz range, Alvarez & Haddock (1973) reported a power-law index, $\alpha_{\mathrm{AH}} = -1.84 \pm 0.02$, which reflects the rapid frequency decrease as the emission source moves through regions of declining plasma density. This value provides a useful reference for comparison with the drift rates derived from the RPW observations analysed in this work.

Each volunteer-identified contour contains frequency–time points (f, t) corresponding to peak radio flux positions. From these points, the frequency drift rate can be estimated and compared directly with Equation (1). To this end, global time and frequency uncertainties were defined for the HFR and TNR receivers. For each instrument, the smallest time cadence over the period 12 February 2020 to 5 March 2025 was adopted as the effective time resolution, denoted Δt_{HFR} and Δt_{TNR}. An analogous procedure was applied to the frequency sampling, yielding Δf_{HFR} and Δf_{TNR}. Assuming independent errors, the associated uncertainties are given by $\sigma(\mathrm{d}t_i) = \sqrt{(\Delta t_{i+1})^2 + (\Delta t_i)^2}$ and $\sigma(\mathrm{d}f_i) = \sqrt{(\Delta f_{i+1})^2 + (\Delta f_i)^2}$, where Δt and Δf depend on the receiver associated with each point, and the instantaneous drift rate is then computed as $R_i = \mathrm{d}f_i/\mathrm{d}t_i$. For each selected Type III burst, the points $(f_i, \log_{10}(|R_i|))$ were fitted with a linear model. The slope of the fit corresponds to the power-law index (α) in Equation (1). Figure 4 shows two representative examples of drift-rate fits for candidate Type III bursts. The uncertainty in $\log_{10}(|R_i|)$ is obtained through standard error propagation as

$$\Delta \log_{10}(R_i) = \frac{1}{\ln 10} \sqrt{\left(\frac{\Delta f_i}{|\mathrm{d}f_i|}\right)^2 + \left(\frac{\Delta t_i}{|\mathrm{d}t_i|}\right)^2}.$$

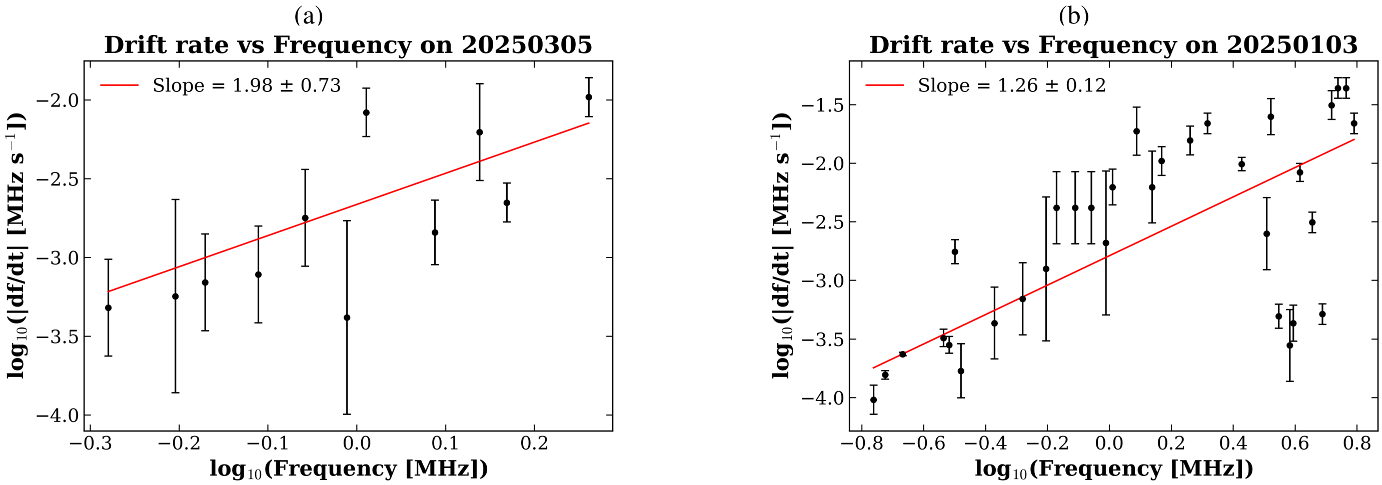
Therefore, for each identification, a linear fit yields a slope corresponding to the drift-rate index (α) and the constant value (A) in Equation (1), together with their associated uncertainty ($\delta\alpha, \delta A$).

2.5. Measure of uncertainty on the Type III detection

The resulting catalogue is based on contributions from 1952 volunteers, and consequently, a fraction of the identified features is expected to be erroneous. For this reason, each volunteer-identified Type III candidate is accompanied by a quantitative uncertainty metric that reflects the confidence of the identification. Two independent sources of uncertainty are considered: one related to the human annotations and one related to the physical behaviour of the derived parameters.

The uncertainty associated with the human selections is quantified through the shape uncertainty U_s, as defined in Appendix A. This quantity measures the similarity of the contours within a cluster, with values close to 0 indicating nearly identical shapes and values approaching 1 indicating low agreement. In addition, the number of volunteers identifying the same region provides an independent measure of reliability, as higher consensus implies greater confidence. These two factors are combined into a human uncertainty term,

$$U_{\mathrm{human}} = \frac{U_s}{\sqrt{N_{\mathrm{users}}}}, \tag{2}$$


(a)

(b)

Fig. 4. Examples of linear fits of $\log_{10}(|df/dt|)$ as a function of frequency for two Type III bursts.

where $N_{\rm users}$ is the number of volunteers contributing to the cluster.

An additional uncertainty component is derived from the frequency drift rates estimated for each candidate Type III burst (Section 2.4). The fitted power-law index (α) was compared to the reference value reported by Alvarez & Haddock (1973). The deviation is quantified as

$$\Delta_\sigma = \frac{|\alpha - \alpha_{\rm AH}|}{\sqrt{(\Delta\alpha)^2 + (\Delta\alpha_{\rm AH})^2}}, \tag{3}$$

where α and $\Delta\alpha$ are the slope and its uncertainty derived from the volunteer-based fit, and $\alpha_{\rm AH}$ and $\Delta\alpha_{\rm AH}$ are the corresponding values from Alvarez & Haddock (1973). To enable direct comparisons across events, the resulting values were normalised using a min–max scaling,

$$U_{{\rm drift},i} = \frac{\Delta_{\sigma,i} - \min(\Delta_\sigma)}{\max(\Delta_\sigma) - \min(\Delta_\sigma)}. \tag{4}$$

The total uncertainty associated with each Type III candidate is then defined as the quadratic mean of the human and drift-related contributions,

$$U_{\rm total} = \sqrt{\frac{U_{\rm human}^2 + U_{\rm drift}^2}{2}}, \tag{5}$$

yielding a normalised confidence metric in the range [0, 1]. Values of $U_{\rm total}$ close to 0 correspond to highly reliable identifications, while values approaching 1 indicate increasing uncertainty. This combined metric enables the catalogue to be filtered or weighted by confidence, allowing robust statistical analyses while retaining faint but physically consistent events. After applying the filtering and merging procedures described in Appendix B, which removes duplicate shapes corresponding to the same Type III event, a final unified catalogue of Type III radio burst features was obtained.

3. Results

3.1. Citizen-science Type III burst (CS-T3) catalogue

The citizen-science project Solar Radio Burst Tracker has produced the first large-scale catalogue of interplanetary Type III solar radio bursts compiled from spacecraft observations. The CS-T3 catalogue is publicly available (Pesini et al. 2025) through the MASER (Measurement, Analysis, and Simulation of Emission in the Radio range) service (Cecconi et al. 2018), hosted at LIRA (Laboratory for Instrumentation and Research in Astrophysics), Paris Observatory. MASER provides long-term access and inter-operability for planetary and solar radio data products, ensuring that the resulting Type III burst catalogue is openly accessible to the scientific community. It contains 15 934 independently identified Type III bursts, representing an extensive statistical sampling of interplanetary radio activity over more than four years of Solar Orbiter RPW measurements.

The CS-T3 catalogue is stored using the time–frequency catalogue (TFCat) format (Cecconi et al. 2022), enabling direct use in multi-mission analysis tools. For each detected Type III burst, the dataset provides:

1. A time–frequency contour outlining the full extent of the emission feature.
2. A frequency-resolved peak-flux curve, providing the time of maximum intensity at each frequency inside the contour.
3. A quantitative uncertainty estimate, derived using Equation (5), which enables reproducible scientific analyses.

3.2. Type III bursts and sunspot activity

To assess the reliability of the CS-T3 catalogue, we compared the temporal occurrence of Type III bursts with a well-established indicator of solar activity: the daily sunspot number. Figure 5 shows the daily number of Type III bursts alongside the corresponding daily sunspot number, covering the period from 12 February 2020 to 5 March 2025. Sunspot data were obtained from the World Data Center SILSO, Royal Observatory of Belgium, Brussels[2] (Clette & Lefèvre 2015). Both time series were smoothed using a Gaussian filter with a window width of four days in order to emphasise long-term trends. Additionally, both series were normalised using min–max scaling. Specifically, for each time series x, the normalised values $x_{\rm norm}$ were computed as $x_{\rm norm} = (x - x_{\rm min})/(x_{\rm max} - x_{\rm min})$, where $x_{\rm min}$ and $x_{\rm max}$ correspond to the minimum and maximum values of each respective dataset over the examined time interval. This procedure rescales both quantities to the range [0, 1], enabling a direct

---

[2] `https://doi.org/10.24414/qnza-ac80`

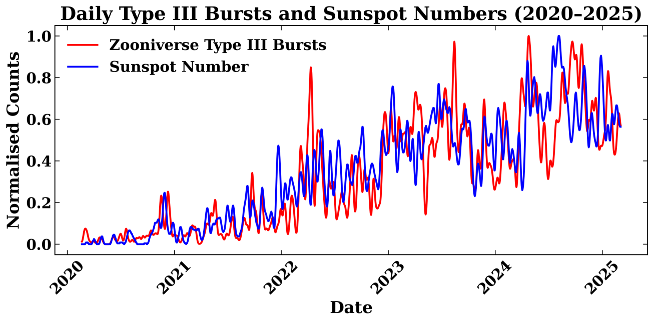


Fig. 5. Daily number of Type III bursts (red) compared with the daily sunspot number (blue) between 12 February 2020 and 5 March 2025.

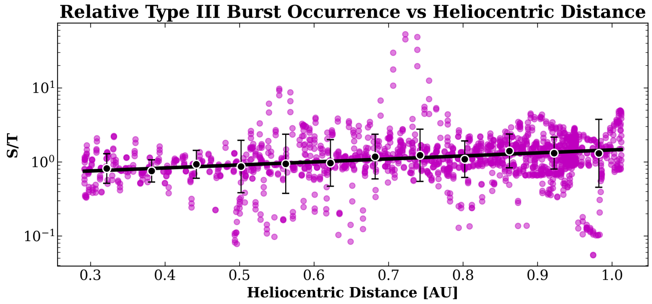


Fig. 6. Scatter plot of the daily S/T ratio as a function of the heliocentric distance of Solar Orbiter between 12 February 2020 and 5 March 2025. Black circles show the median S/T values computed in distance bins, while the vertical error bars correspond to the 1σ scatter within each bin. The solid black line indicates a weighted linear fit performed in logarithmic space.

comparison of their relative temporal evolution independent of their absolute magnitudes.

Both time series exhibit a clear increase towards solar maximum and remain closely aligned throughout the observed period. The Pearson correlation coefficient between the daily Type III burst counts and the sunspot number is $r = 0.813$ with a 95% confidence interval (CI) of 0.797–0.828, indicating a strong positive linear relationship. Despite the overall agreement, deviations on shorter timescales are observed between the two series. These differences can arise for several reasons. The sunspot number is a proxy strictly related to the emergence of active regions on the solar photosphere, whereas Type III radio bursts are more sensitive to changes in the global magnetic field configuration, the properties of the emitting energetic electrons, and the interplanetary plasma environment. Observational effects can also play a role. Sunspots are observed from Earth, providing a view of only one side of the Sun, whereas Solar Orbiter has access to different perspectives. Furthermore, the kilometric frequency range considered here allows Type III bursts to be detected even when the observer is not along the direct line of sight of the burst source.

To further study if there is an observational impact on the Type III measurements, we compared the ratio of the normalised number of sunspots to the normalised number of Type III bursts (S/T) with the heliocentric distance of Solar Orbiter over the same period. The heliocentric distances of Solar Orbiter were obtained from SPICE (European Space Agency 2024). Figure 6 illustrates the variation of the S/T as a function of heliocentric distance. A slight positive trend indicates that the ratio increases as the spacecraft moves farther from the Sun. Since sunspot numbers are derived from observations made at Earth, they are not affected by the heliocentric distance of the spacecraft, unlike the Type III burst measurements. Consequently, the observed variation in the ratio is primarily driven by the number of Type III bursts detected by Solar Orbiter, indicating that fewer bursts are observed at larger heliocentric distances. Additionally, the density of points varies along the trajectory. When Solar Orbiter is closer to the Sun, it moves faster along its orbit, resulting in fewer observations per unit distance, while at larger distances the spacecraft moves more slowly, leading to a higher density of points.

3.3. Type III burst peak flux distribution

The human-based campaign proved effective in identifying both intense and faint Type III bursts, as reflected in the distributions of their peak flux values. We analysed the peak fluxes of all Type III bursts in the catalogue at 23 clean frequencies unaffected by spacecraft-related noise (Maksimovic et al. 2021). Peak flux values are provided by the Level 3 data products described in Section 2.1. To account for variations in heliocentric distance, all fluxes were scaled to a reference distance of 1 AU according to $F_{1\,\mathrm{AU}} = F_{\mathrm{obs}}(R/1\,\mathrm{AU})^2$, where R is the heliocentric distance of Solar Orbiter at the time of observation. Figure 7 shows representative peak-flux distributions at four frequencies. At each frequency, the distributions follow a power law of the form $(\mathrm{d}N/\mathrm{d}S_\nu) \propto S_\nu^\beta$, where N is the number of detected Type III bursts and S_ν is the flux density ($\mathrm{W\,m^{-2}\,Hz^{-1}}$) at observing frequency ν (in MHz).

At low fluxes, the distributions approach the RPW sensitivity limit ($\sim 10^{-20}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$; Vecchio et al. 2021), demonstrating that the volunteers' identifications include very faint Type III bursts, as instructed. The weighted average power-law index over the full frequency range is $\langle\beta\rangle = -1.56 \pm 0.004$, where the uncertainty corresponds to the weighted standard error derived from the individual slope uncertainties.

3.3.1. Power-law slope as a function of frequency

To investigate whether the power-law slope exhibits a systematic dependence on frequency, we show in Fig. 8 the derived power-law indices as a function of observing frequency. For comparison, we also include in the same figure the slopes reported by Fitzenreiter et al. (1976), which were obtained from a substantially smaller sample of Type III bursts (94–530 bursts per frequency) and limited to ten discrete frequencies. The present study includes 3811–14 724 bursts across 23 frequencies, significantly improving the statistical robustness.

The slopes exhibit a decreasing trend with increasing frequency, indicating a progressive steepening of the peak-flux distributions in both studies. At lower frequencies, the distributions appear flatter, indicating a relative deficit of weak Type III bursts compared to higher frequencies. Early studies have shown that the low-frequency limit of Type III bursts strongly depends on burst intensity (Leblanc et al. 1995, 1996; Dulk et al. 1996), with weaker bursts ceasing at higher frequencies. Numerical simulations by Reid & Kontar (2015) further show that the frequency below which the burst is no longer detectable depends on the injected beam density and the spectral index of the electron energy distribution, with lower densities and higher spectral indices leading to higher stopping frequencies. Our results are consistent with this behaviour. Another possible explanation is a higher background level at lower frequencies (Dulk et al. 2001), which could reduce the detectability of weak events.

At much higher frequencies in the GHz range, slopes of $\beta \sim 1.7$–1.8 have been reported (Saint-Hilaire et al. 2013; Nita et al. 2002), in agreement with the slopes estimated above 5 MHz in

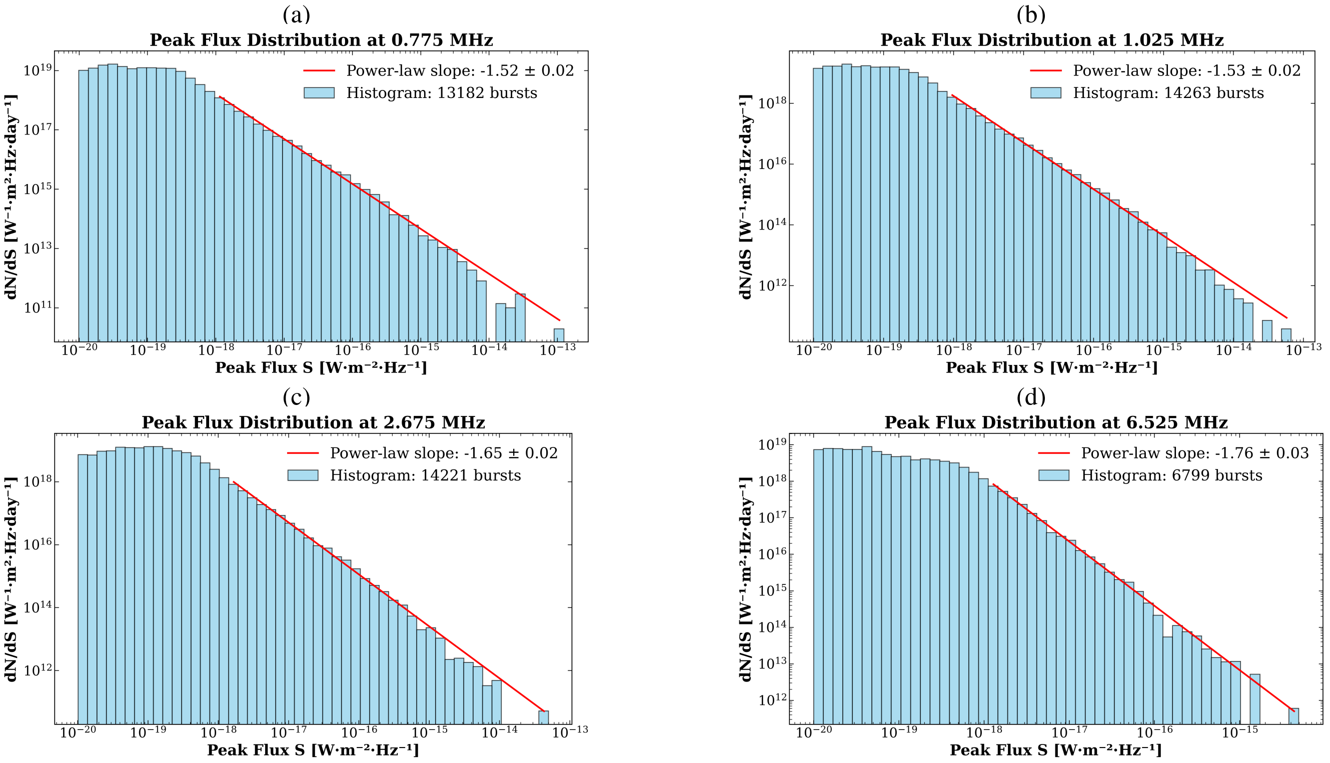


Fig. 7. Peak flux distributions of Type III radio bursts at four representative frequencies. All fluxes are scaled to a reference distance of 1 AU. The red lines indicate power-law fits applied to the decaying (high-flux) part of each distribution.

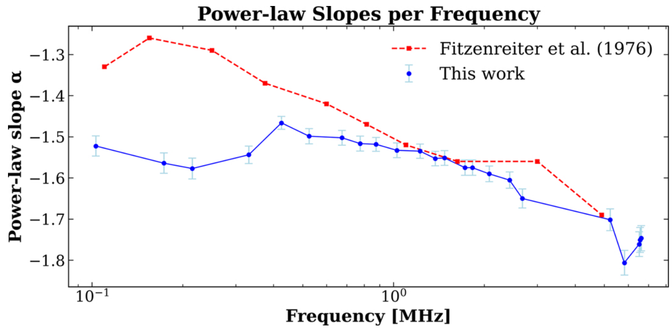


Fig. 8. Power-law indices derived from the peak flux distributions as a function of observing frequency, shown together with slope values reported in previous studies.

this work. The observed steepening of the power-law slope with increasing frequency, followed by an apparent saturation beyond a critical frequency, may reflect the behaviour expected from self-organised critical systems, in which scale-invariant statistics emerge above a characteristic threshold. In these systems, the critical slope of avalanche-like events is continuously self-adjusted: instabilities initially grow exponentially once a critical threshold is exceeded, and subsequently saturate following a logistic-like evolution (Aschwanden 2022).

3.3.2. Number of Type III bursts as a function of frequency

For each frequency at which the peak flux distribution was calculated, we also determined the number of detected Type III bursts. Figure 9 shows the frequency dependence of the number of bursts. The number of detected bursts peaks at 14 724 at 1.725 MHz and decreases to 3811 at 0.103 MHz, which corresponds to the lowest value within the examined frequency range.

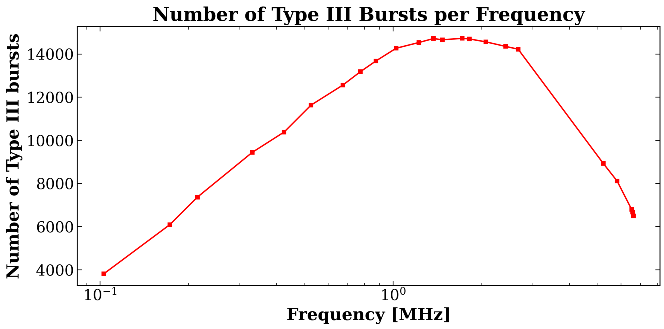


Fig. 9. Number of Type III bursts as a function of frequency.

The peak coincides with the frequency at which Type III radio flux is known to reach its maximum (Dulk et al. 2001; Sasikumar Raja et al. 2024), indicating that 1–2 MHz is the most efficient frequency range for Type III bursts. This maximum likely reflects efficient electron-beam transport and beam–plasma interactions in the inner heliosphere, and can be further compared with the advection–non-linear-diffusion model of flare-accelerated electron transport developed by Kontar et al. (2024).

3.4. Frequency drift rate of Type III bursts

From the resulting catalogue, we derived the frequency drift rate–frequency relation from Equation (1), with the drift rate expressed in $MHz \cdot s^{-1}$. For a total of 13 074 Type III bursts observed in the frequency range 4 kHz–7 MHz, we found the following equation:

$$\frac{df}{dt} = -0.001\, f^{1.37}. \tag{6}$$

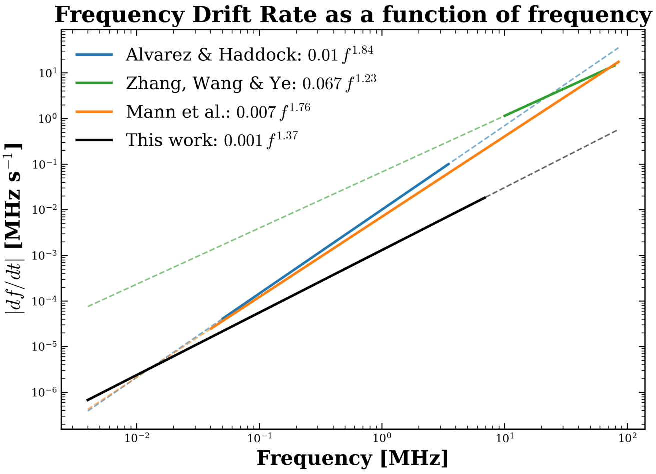


Fig. 10. Comparison of Type III burst drift rates. The dashed line shows the relation derived in this work from RPW data (4 kHz–7 MHz) from 13 074 Type III bursts, while the solid lines show previously reported relations (Alvarez & Haddock (1973), Zhang et al. (2018), Mann et al. (1999)).

From the full sample of Type III bursts, we retained only those selections with $U_{\text{total}} \leq 0.5$, thereby keeping the bursts for which the shape agreement among volunteers is higher. This filtering resulted in 13 074 selected Type III bursts. From these Type III bursts, we extracted the parameters A and α (see Section 2.4) together with their individual uncertainties δA and $\delta\alpha$. To account for the uncertainties of the fitted parameters in the CS-T3 catalogue, we calculated the weighted mean of each parameter, using weights defined as the inverse squared uncertainties ($w_A = 1/\delta A^2$, $w_\alpha = 1/\delta\alpha^2$). The uncertainty of the mean was estimated from the standard deviation of the distributions divided by the square root of the number of selected Type III bursts, yielding $\alpha = 1.374 \pm 0.005$ and $A = (1.30 \pm 0.01) \times 10^{-3}$. Figure 10 presents a comparison of our derived drift-rate relation and previously published relations, which were obtained over different frequency ranges.

Our comparison shows that previously reported relations differ in both amplitude and slope, reflecting the respective frequency ranges and observational regimes. Zhang et al. (2018) exhibits the highest $|\mathrm{d}f/\mathrm{d}t|$ at high frequencies due to its large amplitude ($A = 0.067$), as it refers to the 10–80 MHz range where the highest frequency drift rates are expected according to the Type III emission mechanism. Alvarez & Haddock (1973) and Mann et al. (1999) show intermediate $|\mathrm{d}f/\mathrm{d}t|$ values, with steeper slopes ($\alpha = 1.84$ and 1.76) causing the drift rate to increase rapidly with frequency. Our work, covering 4 kHz–7 MHz, exhibits the lowest $|\mathrm{d}f/\mathrm{d}t|$ and a moderate slope ($\alpha = 1.37$), consistent with the lower drift rates expected in the interplanetary regime. The frequency–drift slopes differ across studies and frequency ranges, essentially reflecting the varying speeds of the electron beams that generate the Type III radio bursts. In particular, our work and that of Zhang et al. (2018) show the lowest slopes, corresponding to the lowest and highest frequency ranges, respectively.

4. Conclusions and discussion

A citizen-science campaign with the primary goal of producing the first extensive catalogue of interplanetary Type III radio bursts identified through human participation was undertaken. Following the completion of the campaign, a dedicated processing pipeline was implemented to combine and consolidate all volunteer annotations into a unified catalogue comprising 15 934 individual Type III burst identifications. Each entry provides the spatial location of a Type III burst as well as the corresponding peak radio flux position at each observing frequency. All entries are accompanied by quantitative uncertainty estimates, derived from both the level of agreement among volunteer selections and the properties of the retrieved frequency drift rates. This uncertainty characterisation allows the catalogue to be flexibly filtered according to the requirements of specific scientific applications, enabling users to balance completeness against reliability.

The first statistical analysis of the catalogue shows that the occurrence of Type III bursts is strongly correlated with the sunspot number, following the expected evolution of the solar cycle. This agreement supports the reliability of the volunteer-based identifications and indicates that the catalogue captures the underlying physical processes governing solar radio emission. Deviations on shorter timescales are nevertheless present, as the two phenomena are governed by different physical mechanisms and are therefore not expected to be perfectly aligned. Additional discrepancies arise from observational effects: sunspots are observed from Earth through optical measurements, whereas Type III bursts are detected by Solar Orbiter in the kilometric radio domain. Although the number of measurements from Solar Orbiter increases at larger heliocentric distances, the observed number of Type III bursts decreases. This behaviour can be explained by the propagation of the radio emission. As the radio waves propagate away from the source, their intensity decreases due to geometrical spreading, approximately following a $1/r^2$ dependence. Consequently, weaker Type III bursts can fall below the instrument's sensitivity threshold before reaching the spacecraft when it is located farther from the source. As a result, only the stronger events remain detectable at larger heliocentric distances. A deeper analysis could extend the study of Type III bursts in correlation with sunspots by including other indicators of solar activity.

The analysis also demonstrates that peak radio flux distributions are frequency-dependent. At lower frequencies, the measured fluxes (scaled to 1 AU) exhibit a relative deficit of weak events. This behaviour is likely driven by the properties of the electron beam, such as the injected density and energy of the electrons, and may also be influenced by higher background levels at lower frequencies, which reduce the detectability of weak events. Comparison with previous studies suggests that the peak flux distributions tend towards a saturated power-law slope, a behaviour reminiscent of self-organised critical systems.

In addition, the number of detected Type III bursts is frequency-dependent, reaching a maximum around 1–2 MHz, consistent with previous studies of Type III burst flux. This suggests that the 1–2 MHz range represents the most efficient frequencies for Type III bursts, likely reflecting effective electron-beam transport and beam–plasma interactions in the inner heliosphere.

Another finding concerns the frequency drift rates of 13 074 bursts selected from the CS-T3 catalogue based on their associated uncertainty. The drift rates follow Equation (6), being lower at lower frequencies, in agreement with the expected behaviour from the Type III emission mechanism. This highlights the potential of the catalogue to be used to investigate particle velocities and plasma density structures across different frequency bands.

Overall, this work presents an innovative and effective approach to creating an extensive Type III burst catalogue and provides a first preliminary statistical analysis. The results reveal correlations between Type III bursts and sunspot activity, describe the frequency dependence of peak radio flux and burst occurrence, and characterise the frequency drift rate relation. This catalogue enables more robust statistical studies of Type III bursts, which can enhance research on solar activity, radio wave propagation, and electron acceleration mechanisms. Additionally, it offers a high-quality training dataset for optimising machine-learning algorithms for future Type III observations in similar frequency ranges.

Data availability

The CS-T3 catalogue (Pesini et al. 2025) is publicly accessible at `https://doi.org/10.25935/bync-2m17`

For this work, we used RPW Level 3 preliminary data products, derived from calibrated Level 2 data, which are publicly available at `https://rpw-lira.obspm.fr/roc/data/pub/solo/rpw/data/L2/thr/`

Acknowledgements. This work was supported by the European Space Agency (ESA) through the IDeas programme (IDEAS: I-2024-05402), which partially funded A.P.'s PhD research. We sincerely thank all volunteers whose dedicated efforts in identifying Type III bursts made this citizen-science campaign possible. Solar Orbiter is a mission of international cooperation between ESA and NASA, operated by ESA. The RPW instrument was designed and funded by CNES, CNRS, the Paris Observatory, the Swedish National Space Agency, ESA-PRODEX, and all participating institutes. This publication uses data generated via the Zooniverse.org platform, whose development is supported by generous contributions, including a Global Impact Award from Google and a grant from the Alfred P. Sloan Foundation.

References

Alvarez, H., & Haddock, F. T. 1973, Sol. Phys., 29, 197
Aschwanden, M. J. 2022, ApJ, 934, 33
Bai, L., Hancock, E. R., & Han, L. 2013, in International Conference on Computer Analysis of Images and Patterns (Springer), 102
Benz, A. O. 2017, Liv. Rev. Sol. Phys., 14, 2
Bussons Gordo, J., Fernández Ruiz, M., Prieto Mateo, M., et al. 2023, Sol. Phys., 298, 82
Cecconi, B., Le Sidaner, P., Savalle, R., et al. 2018, in EGU General Assembly Conference Abstracts, 6207
Cecconi, B., Louis, C. K., Bonnin, X., Loh, A., & Taylor, M. B. 2022, Front. Astron. Space Sci., 9, 368
Chen, P. F. 2011, Liv. Rev. Sol. Phys., 8, 1
Clette, F., & Lefèvre, L. 2015, SILSO Sunspot Number V2.0, `https://doi.org/10.24414/qnza-ac80` (WDC SILSO – Royal Observatory of Belgium (ROB))
Dulk, G. A. 1985, ARA&A, 23, 169
Dulk, G. A., Leblanc, Y., Bougeret, J.-L., & Hoang, S. 1996, Geophys. Res. Lett., 23, 1203
Dulk, G. A., Erickson, W. C., Manning, R., & Bougeret, J.-L. 2001, A&A, 365, 294
Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD) (AAAI Press), 226
European Space Agency. 2024, Solar Orbiter Operational SPICE Kernel Dataset, ESA SPICE Service
European Space Agency. 2024, Solar Orbiter RPW Level 2 Data, `https://doi.org/10.57780/esa-3xcjd4w`, accessed from the ESA Space Science Data Centre (ESDC)
Faden, J., Weigel, R. S., Merka, J., & Friedel, R. H. W. 2010, Earth Sci. Inform., 3, 41
Fitzenreiter, R. J., Fainberg, J., & Bundy, R. B. 1976, Sol. Phys., 46, 465
Fletcher, L., Dennis, B. R., Hudson, H. S., et al. 2011, Space Sci. Rev., 159, 19
Ginzburg, V. L., & Zhelezniakov, V. V. 1958, Soviet Astron., 2, 653
Hathaway, D. H. 2015, Liv. Rev. Sol. Phys., 12, 4
Henri, P., Briand, C., Mangeney, A., et al. 2009, J. Geophys. Res. (Space Phys.), 114, A03103
Klein, K.-L. 2021, Front. Astron. Space Sci., 7, 105
Kontar, E. P., Azzollini, F., & Lyubchyk, O. 2024, ApJ, 976, 233
Kullback, S., & Leibler, R. A. 1951, Ann. Math. Statist., 22, 79
le Roux, H., Steyn, R., Strauss, D. T., et al. 2025, Sol. Phys., 300, 179
Leblanc, Y., Dulk, G. A., & Hoang, S. 1995, Geophys. Res. Lett., 22, 3429
Leblanc, Y., Dulk, G. A., Hoang, S., Bougeret, J.-L., & Robinson, P. A. 1996, A&A, 316, 406
Lobzin, V. V., Cairns, I. H., Robinson, P. A., Steward, G., & Patterson, G. 2009, Space Weather, 7, S04002
Lobzin, V. V., Cairns, I. H., & Zaslavsky, A. 2014, J. Geophys. Res. (Space Phys.), 119, 742
Maksimovic, M., Bale, S. D., Chust, T., et al. 2020, A&A, 642, A12
Maksimovic, M., Souček, J., Chust, T., et al. 2021, A&A, 656, A41
Mann, G., Jansen, F., MacDowall, R. J., Kaiser, M. L., & Stone, R. G. 1999, A&A, 348, 614
Mann, G., Melnik, V. N., Rucker, H. O., Konovalenko, A. A., & Brazhenko, A. I. 2018, A&A, 609, A41
Melrose, D. B. 1980, Space Sci. Rev., 26, 3
Müller, D., St. Cyr, O. C., Zouganelis, I., et al. 2020, A&A, 642, A1
Nita, G. M., Gary, D. E., Lanzerotti, L. J., & Thomson, D. J. 2002, ApJ, 570, 423
Paipa-Leon, D., Vilmer, N., Maksimovic, M., Krupar, V., & Vecchio, A. 2025, A&A, 694, A111
Pesini, K., Vecchio, A., Bonnin, X., et al. 2025, Catalogue of Solar Type-III Radio Bursts Identified in the Solar Orbiter/RPW Observations (Version 1.0), Public dataset, PADC/MASER, [Data set]
Priest, E., & Forbes, T. 2000, Magnetic Reconnection: MHD Theory and Applications (Cambridge, UK: Cambridge University Press)
Raouafi, N. E., Patsourakos, S., Pariat, E., et al. 2016, Space Sci. Rev., 201, 1
Reid, H. A. S., & Kontar, E. P. 2015, A&A, 577, A124
Saint-Hilaire, P., Vilmer, N., & Kerdraon, A. 2013, ApJ, 762, 60
Salmane, H., Weber, R., Abed-Meraim, K., Klein, K.-L., & Bonnin, X. 2018, J. Space Weather Space Climate, 8, A43
Sarkar, S., Paul, S., & Denra, R. 2015, Phys. Plasmas, 22, 102109
Sasikumar Raja, K., Maksimovic, M., Kontar, E. P., et al. 2024, in 42nd Meeting of the Astronomical Society of India (ASI), 42, 251
Scully, J., Flynn, R., Carley, E., Gallagher, P., & Daly, M. 2021, in 2021 32nd Irish Signals and Systems Conference (ISSC), 1
Shibata, K., & Magara, T. 2011, Liv. Rev. Sol. Phys., 8, 6
Solanki, S. K. 2003, A&A Rev., 11, 153
Suzuki, S., & Dulk, G. A. 1985, in Solar Radiophysics: Studies of Emission from the Sun at Metre Wavelengths, eds. D. J. McLean, & N. R. Labrum, 289
Vecchio, A., Maksimovic, M., Krupar, V., et al. 2021, A&A, 656, A33
Wild, J. P. 1950, Austr. J. Sci. Res. A Phys. Sci., 3, 541
Wild, J. P., Smerd, S. F., & Weiss, A. A. 1963, ARA&A, 1, 291
Zhang, P. J., Wang, C. B., & Ye, L. 2018, A&A, 618, A165
Zheleznyakov, V. V., & Zaitsev, V. V. 1970, Soviet Ast., 14, 47

Appendix A: Density-Based Clustering Algorithm

DBSCAN groups together items that are closely packed in the input space, based on two parameters: a distance threshold ε, and a minimum number of samples required to form a cluster. DBSCAN was adapted for this use case as follows:

- The intersection over union (IoU) is first computed between all pairs of annotated polygons, but only when the polygons originate from different users. The IoU is defined as:

$$\mathrm{IoU}(A, B) = \frac{\mathrm{Area}(A \cap B)}{\mathrm{Area}(A \cup B)}, \tag{A.1}$$

  where A and B are two polygons from different users.
- The IoU values are converted into distances using:

$$d = 1 - \mathrm{IoU}, \tag{A.2}$$

  where d close to 0 indicates high similarity, and values near 1 suggest minimal overlap.
- DBSCAN is applied to this distance matrix, with the constraint that no cluster can include more than one polygon from the same user. This ensures clusters reflect consensus across multiple individuals, rather than the internal consistency of a single user's annotations.
- Each polygon receives a cluster label, which is propagated back to the original dataset for traceability.
- For every resulting cluster, we computed the geometric union of all polygons it contains. This union represents the spatial envelope of the shared annotation—the agreed-upon location and shape of a Type III burst as independently identified by multiple users.
- To quantify intra-cluster agreement, we defined the shape uncertainty (U_s) as the mean of all pairwise distances ($1 - \mathrm{IoU}$) between polygons within a cluster. Lower values indicate stronger consensus in shape and location.

The output of this procedure is a set of spatial clusters for each image, where each cluster corresponds to a possible Type III radio burst independently identified by multiple volunteers. Each cluster is associated with a union shape and an uncertainty measure that quantifies the degree of agreement among users' annotations. Figure A.1a shows the raw shapes received from the users, while Fig. A.1b presents the results produced by the described clustering algorithm after excluding the shapes that, according to ε, are associated with only a single user.

After some testing, the clustering parameter ε was set to 0.8, as it provided the best performance. Values of the clustering parameter ε between 0.5 and 0.9 were evaluated. Volunteer annotations were abstract due to the free-drawing interface provided by Zooniverse. Using a larger ε enabled the algorithm to accommodate small variations in shape, allowing similar shapes to be grouped together more effectively. This corresponds to a minimum overlap of approximately 20% between volunteer annotations, ensuring that shapes included in a cluster are supported by multiple users. The minimum number of samples was set to 2, which was chosen to ensure that at least two users cross-checked a given region, providing confidence that it corresponds to a potential Type III burst. This also represents the minimum number of shapes required to form a cluster, meaning that a shape is included in a cluster only if it is defined by at least two users.

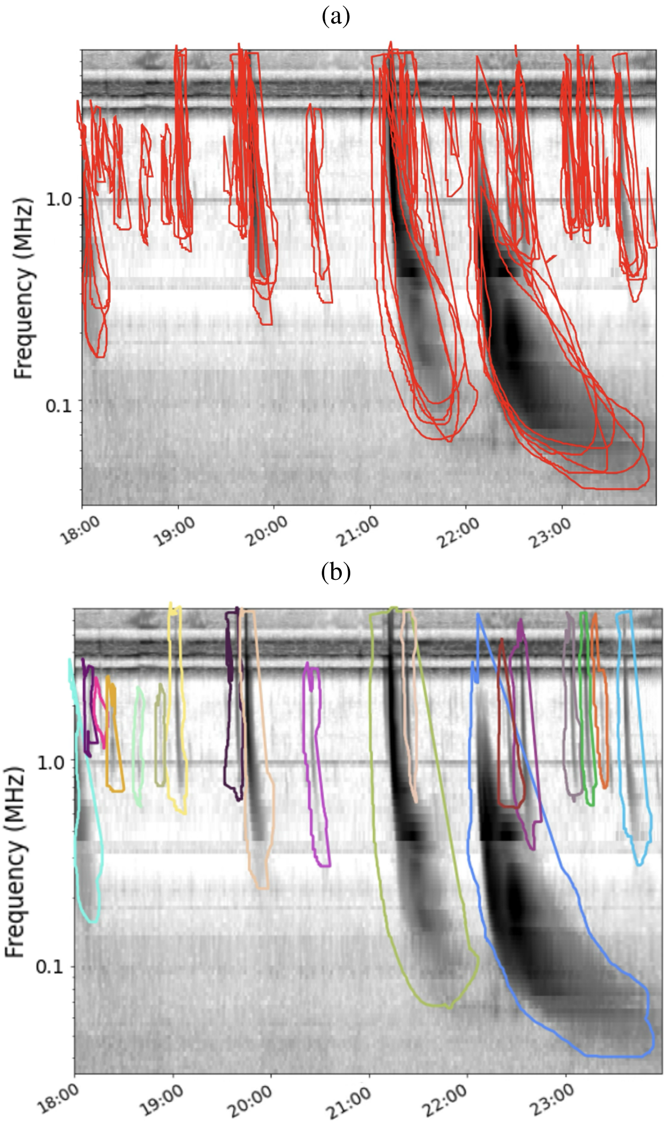


Fig. A.1. Volunteer selections for the same subject on 2 February 2025 before (a) and after (b) applying the clustering algorithm.

Appendix B: Filtering and merging of Type III features

After extracting the peak radio flux and drift rates (Section 2.4), the catalogue was checked for two points of concern. The first arises when, after the clustering algorithm, duplicate features remain. This can occur because some of the user-drawn shapes are too dissimilar in morphology to be grouped into the same cluster, resulting in multiple contours representing what is in fact the same Type III burst. Since the peak flux has been calculated for all features (see Fig. 3), intersections are identified by comparing the peak flux frequency and time positions. The shapes that share the same peak flux positions are further filtered to determine whether they correspond to the same Type III burst or to two overlapping bursts. For each pair of polygons sharing peak fluxes, we checked if one polygon is contained within the other. A polygon was considered to contain another if it fully covers the other polygon, or if the overlapping area relative to the smaller polygon is at least 90%. This empirical threshold was determined after testing several values and visually inspecting the resulting shapes, where a 90% overlap reliably indicated full coverage rather than a partial intersection. These cases are then flagged and stored for further analysis. For these features, shared points are examined, and duplicates are resolved using the Jensen-Shannon divergence (JSD) metric (Bai et al. 2013). The JSD measures the similarity between two discrete probability distributions P and Q, and is defined as

$$\mathrm{JSD}(P\|Q) = \frac{1}{2}D_{\mathrm{KL}}(P\|M) + \frac{1}{2}D_{\mathrm{KL}}(Q\|M), \tag{B.1}$$

$$M = \frac{1}{2}(P+Q), \tag{B.2}$$

$$D_{\mathrm{KL}}(P\|Q) = \sum_i p_i \log\frac{p_i}{q_i}, \tag{B.3}$$

where D_{KL} is the Kullback–Leibler divergence, which intuitively measures how much information is lost if Q is used to approximate P (Kullback & Leibler 1951). We chose the JSD because it is symmetric, always finite, and provides a bounded measure of similarity between discrete distributions, making it well suited for comparing our peak flux distributions. Lower values of JSD indicate that the two distributions are more similar.

For each group of overlapping features, the set of common points is first identified, and their time and frequency coordinates are extracted. Two-dimensional histograms of time and frequency are then constructed, with the histogram of the common points used as the reference distribution. The JSD was computed separately for the time and frequency distributions:

$$\begin{aligned} JS_{\mathrm{time}} &= JSD\left(P_{\mathrm{feature,time}} \,\|\, P_{\mathrm{common,time}}\right), \\ JS_{\mathrm{freq}} &= JSD\left(P_{\mathrm{feature,freq}} \,\|\, P_{\mathrm{common,freq}}\right), \end{aligned} \tag{B.4}$$

where P_{feature} denotes the normalised distribution of a feature and P_{common} the distribution of the common points. The overall divergence is defined as the arithmetic mean of the two measures:

$$JS_{\mathrm{total}} = \frac{JS_{\mathrm{time}} + JS_{\mathrm{freq}}}{2}. \tag{B.5}$$

This metric quantifies how closely a feature's frequency–time distribution is to the shared peak flux points and provides a measure of which feature in a pair more accurately represents the shared peak flux positions. The JSD ranges from 0 to 1, where 0 indicates identical distributions and 1 indicates completely dissimilar distributions. For pairs of overlapping features, the one with the smallest JS_{total} is retained in the catalogue, while the other is discarded. The bottom panel of Fig. B.1 illustrates an example of this procedure. Initially, three Type III bursts were represented by duplicate contours. After applying the JSD-based filtering, only unique shapes corresponding to distinct Type III burst areas remain.

The second case that requires re-evaluation occurs when users identify Type III structures that are not fully contained within a single spectrum, as illustrated in the upper panel of Fig. B.2. Some Type III bursts are expected to span consecutive days. Because the re-clustering procedure described in Section 2.3 is performed on a daily basis, clusters belonging to the same physical Type III event can be separated into two consecutive shapes, each corresponding to the portion of the burst appearing in the daily spectra. This situation naturally produces two adjacent structures that refer to the same Type III burst. To identify such cases, we searched for features occurring at the end of one day (23:00–00:00) and at the start of the following day (00:00–01:00). Consecutive features were flagged based on the continuity of their peak-flux positions. More specifically, the start and end times and frequencies of the found shapes were calculated so that we could compare how far they are from each other and whether the stopping frequency of one corresponds to the starting frequency of the other. Consecutive pairs were identified based on the time gap and frequency tolerance criteria.

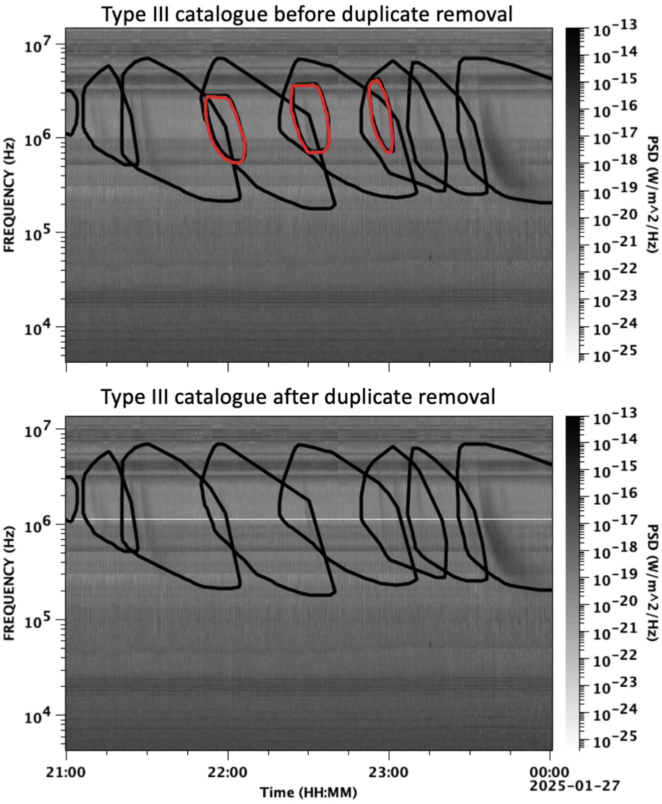


Fig. B.1. Example of duplicate features (red shapes) before and after JSD-based filtering on 26 January 2025.

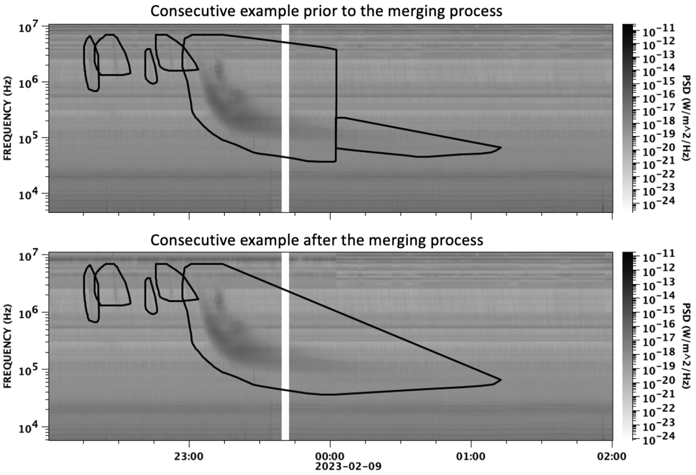
A pair of features is considered consecutive if it satisfies the following criteria:

1. The time gap between the end of the first feature and the start of the second feature lies between −60 s and 100 s, allowing for slight temporal overlaps or moderate separations.
2. The difference between the starting frequency of the second feature and the ending frequency of the first feature does not exceed 0.2 MHz, consistent with the expected smooth downward drift of Type III bursts.

When both criteria are satisfied, the features are flagged as potential continuations of the same Type III burst. All such cases are subsequently examined through visual inspection of the corresponding DSs in order to determine whether the features represent segments of the same physical burst rather than two distinct consecutive events. In total, 65 cases satisfied the above criteria. Among these, 25 corresponded to shapes appearing on consecutive days, consistent with bursts that extend across daily data boundaries. The final merged features are constructed by defining the contour as the union of the individual shapes and by concatenating the coordinates of the peak flux lines. Since each feature is associated with an uncertainty estimate, the merged feature is assigned the minimum of the two uncertainty values.

This approach ensures that the final representation retains the most reliable characteristics while preserving the spatial extent and temporal evolution of the original Type III structures.

Figure B.2 illustrates an example of consecutive features before and after the merging process.



Consecutive example prior to the merging process

Consecutive example after the merging process

Fig. B.2. Example of a consecutive feature from 8-9 February 2023, shown before and after the merging process.

After completing the entire filtering, merging, and cross-day linking procedure, we obtained the final version of the CS-T3 catalogue.