Automatic GNSS RFI Classification Challenges

Anja Diez, Aiden Morrison, Nadezda Sokolova (Sustainable Communication Systems, SINTEF Digital)

Abstract

This article describes the real-world challenges that are encountered when trying to categorize and classify the radio frequency interference (RFI) events captured in the GNSS signal bands by an international network of monitoring stations covering all L-band navigation signals. While signals frequently fall into the often-discussed categories such as 'chirp', 'continuous wave', or 'wideband noise' [1]–[3] there are a large and growing number of modulations encountered in reality, both intentional and unintentional [4]. These bear varying degrees of resemblance to the aforementioned traditional categories. Work presented herein focuses on some of the main complications encountered when categorizing multiple years of GNSS RFI event data, and the algorithmic approaches used to proceed with classification in these conditions.

Introduction

Categorization, long term monitoring, and centralized reporting of Radio Frequency Interference (RFI) events in the Global Navigation Satellite System (GNSS) bands can help to inform on the operating parameters of GNSS systems such as availability and continuity at a given site. Further, it can help system manufacturers and researchers to assess the variety and occurrence rates of various types of intentional and unintentional GNSS jamming. Through categorization of jammer events, it might be possible in the future to extract so-called 'fingerprints' [5] of individual jamming devices to help apprehend the operators and, thus, discourage jammer use. An automatic classification algorithm was developed based on data from the Advanced RFI Detection Analysis and Alerting System (ARFIDAAS) and is described in [6]. ARFIDAAS was developed to monitor all GNSS L-band signals and record them when RFI events are detected and issue alerts to the site stakeholders [7], [8]. At present there are 14 ARFIDAAS stations in operation around Europe providing the background collection of difficult to classify events under discussion here. Making use of this raw data, jammer signals are analysed using threshold values for functions and parameters derived from the data's spectrograms in combination with a classification matrix. Classification results, including the derived jammer signal type, centre frequency, bandwidth, and time-modulation frequency of the jammer signals are saved for statistical analysis. Methods for automated jammer classification is an area of active research with recent methods focused on Machine Learning (ML) techniques as presented in [9], however it is believed that all classification families or approaches currently known will experience some difficulty when presented with the non-traditional RFI sources discussed in this work.

The focus of this paper is on the challenge to automatically classify these RFI events. Challenges for the automatic classification algorithm occur for example when signals are aliased, when signals have strong variations in intensity over frequency or time, due to saturation effects, or when multiple jammer signal types occur at the same time. The analysis of the GNSS data is mainly based on the analysis of spectrograms of the data. However, the automatic classification scheme is based on threshold values. By defining threshold values to a range of parameter spaces means that the different jammer signal types are sorted into categories with rigid boundaries. This does not reflect the large variability of jammer signal types we observe in our data set and challenges can, therefore, arise from this chosen classification approach.

ARFIDAAS system, data, and reporting

A large amount of data from a subset of eight stations in the ARFIDAAS monitoring network distributed through Europe, captured during a period of about 1.5 years was used to develop the classification algorithm. The ARFIDAAS units simultaneously record data on four bands. Centre frequencies and bandwidth of these bands can be chosen within constraints defined by the two independent surface acoustic wave (SAW) filters used to isolate the L1 band for channel A, and the L5 through E6 bands for channels B, C and D. However, most of the data were recorded with 60 MHz bandwidth for the following bands: 1555 – 1615 MHz (A), 1249 – 1309 MHz (B), 1203 – 1263 MHz (C), 1162 –

1222 MHz (D). A more detailed description of system components, structure and operation can be found in [7].

The ARFIDAAS system is set up to record data and inform stakeholders when a jammer event is detected. The system is triggered to have received a jammer event when the energy level for one of the bands exceeds the energy level of a recorded environmental baseline. This environmental baseline is recorded once a month in the absence of jammer signals. When data is recorded an initial analysis is carried out to create an event report with information about the jammed bands and the affected frequency ranges. This event report, together with a pdf showing plots of the energy level in comparison to the environmental baseline are immediately send to the side stakeholders by email [7]. Afterwards the recorded event is classified in more detail and the results and parameters of this classification are saved [6]. Finally, the event report, the pdf showing the energy levels, the result of the classification and the event raw data are uploaded to a cloud server. Once a month the saved classification information is accessed to generate a monthly report presenting statistics from the classified jammer events of the previous month for each station in the ARFIDAAS network [6].

Classification

The classification of the jammer signal types relies on properties and parameters derived from the spectrogram of the data (Figure 1A) calculated using the Fast Fourier Transform (FFT) for a moving window with a length of 176 samples (about 3 µs for the 60 MHz sampling rate). Five main functions are derived from the spectrogram, which are afterwards used to derive a range of parameters. Threshold values are defined for these parameters and together with a classification matrix used to classify the jammer events automatically. The classification algorithm, the used functions and derived parameters are in detail explained in [6]. The five main functions used for the classification are:

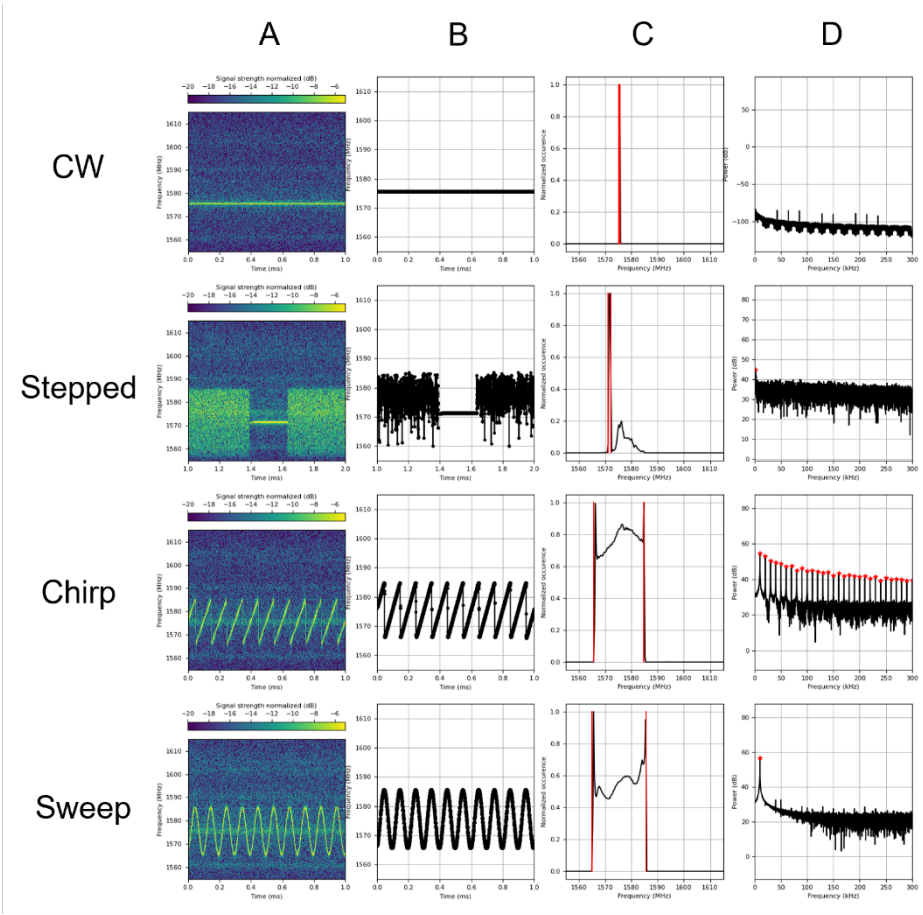

Figure 1: Example for CW, stepped, chirp, and sweep jammer signal, showing (A) the spectrogram, (B) the dominant jammer signal, (C) energy frequency distribution, and (D) the FFT of the jammer signal.

1. **Dominant jammer signal** (Figure 1B): For each time step in the spectrogram the frequency with the highest signal strength is determined. This function will detect the dominating signal in the dataset.
2. **Energy frequency distribution** (Figure 1C): This function gives the frequency distribution of the highest detected energies in the spectrogram. Frequency-time pairs that are within 5 dB of the maximum detected signal strength within the spectrogram are used, their occurrence is summed up for each frequency and normalized. With this function signal parts are detected that are significantly above the noise floor but might be weaker than the dominant jammer signal.
3. **FFT of jammer signal** (Figure 1D): The FFT of the dominant jammer signal (Figure 1B) is calculated. Thus, regular time modulation in the signal can be detected.
4. **Short term-long term (STLT) ratio**: The standard deviation of the mean of the frequency from the dominant jammer signal (Figure 1B) for a long-term window (in this context the entire 17.5 ms data length) and a short time window (0.1 ms) is calculated. Thus, short term variation of the main signal can be detected.
5. **Spectrogram continuity**: The standard deviation of the energy for each frequency is calculated. Signals with a high continuity can be detected.

For the classification of jammer events, we distinguish the three main classes of (i) narrowband, (ii) wideband, and (iii) time-modulated signals with the jammer subclasses of continuous wave (CW) and multi-CW events for narrowband signals, general wideband and stepped events in case of wideband signals, and chirp, multilevel chirp, triangular, sweep, wideband sweep, and low-frequency sweep in case of time-modulated signals. If a jammer signal cannot be characterized in one of these subclasses it is marked as undefined within the relevant main class, e.g., narrowband undefined.

Challenges

A range of challenges arises when automatically classifying jammer events based on threshold values and a classification matrix. Most of the events contain one well defined jammer signal type and are automatically classified correctly. Other jammer events are very difficult to classify due to weak signals, aliased signals or the superposition of multiple jammer signal types and are, therefore, categorized in one of the main classes or are categorized incorrectly. The biggest challenge arises from classifying jammer events that do not fit into one of the predefined categories, bordered by threshold values, because we have not seen this kind of jammer signal type before or because multiple different jammer signal types occur at the same time.

Mistakes in the interpretation of jammers from spectrograms can easily be made by plotting the spectrograms at different resolutions and figure sizes. Figure 2 shows an example of the same data for two different time intervals (2 ms and 8 ms) plotted for different figure sizes. For this example in Figure 2 a chirp signal with a signal repetition frequency of 20 kHz was identified, which is best represented in Figure 2a. Due to variation in power of the chirp signal over frequency and time, patterns appear for the viewer when plotting the data at different resolutions that significantly change the appearance of the signal and might lead to misinterpretation. Hence, events can 'appear' differently in the spectrogram depending on the chosen visualization parameters. When judging the correctness of the classification algorithm it is, therefore, important to investigate the spectrograms with different resolutions and for different time intervals.

Bandwidth & Aliasing

The bandwidth is derived in two ways. One option is to derive the bandwidth as the standard deviation of the mean frequency of the dominant jammer signal function. Another possibility is to derive the bandwidth of a signal from the energy frequency distribution (Figure 1C). Here we define the boundaries of the bandwidth at the location when the energy frequency distribution exceeds 0.2. For the statistical monthly reporting we mainly use the second version of deriving the bandwidth [6]. In case of strong signals like sweep or chip jammer signal types as shown in Figure 1 the derived bandwidth is more correct using energy frequency distribution compared to the standard deviation of the dominant jammer signal. This is also the case for wideband signals that have a peak energy and fade out towards the edges of the wideband jammer signal.

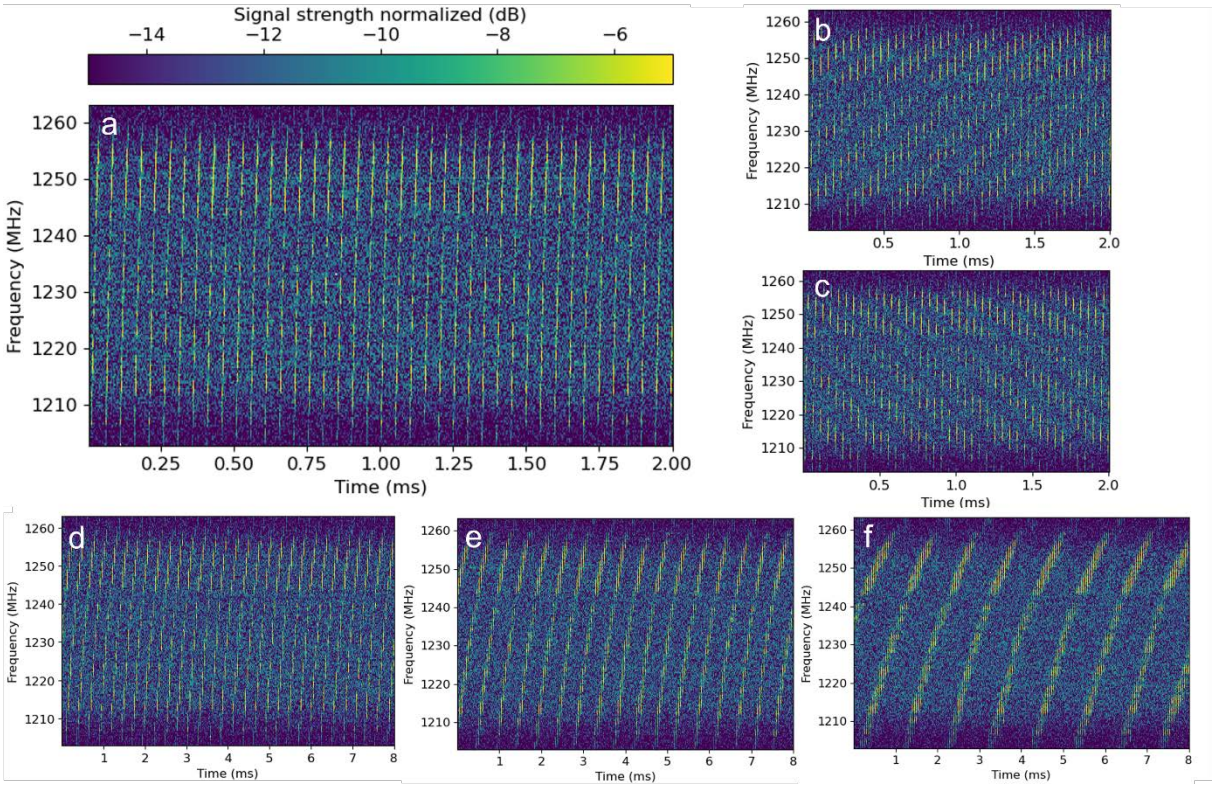

Figure 3: Different visualizations by plotting data on a variation of image sizes and for different time spans. All subplots show the same data!

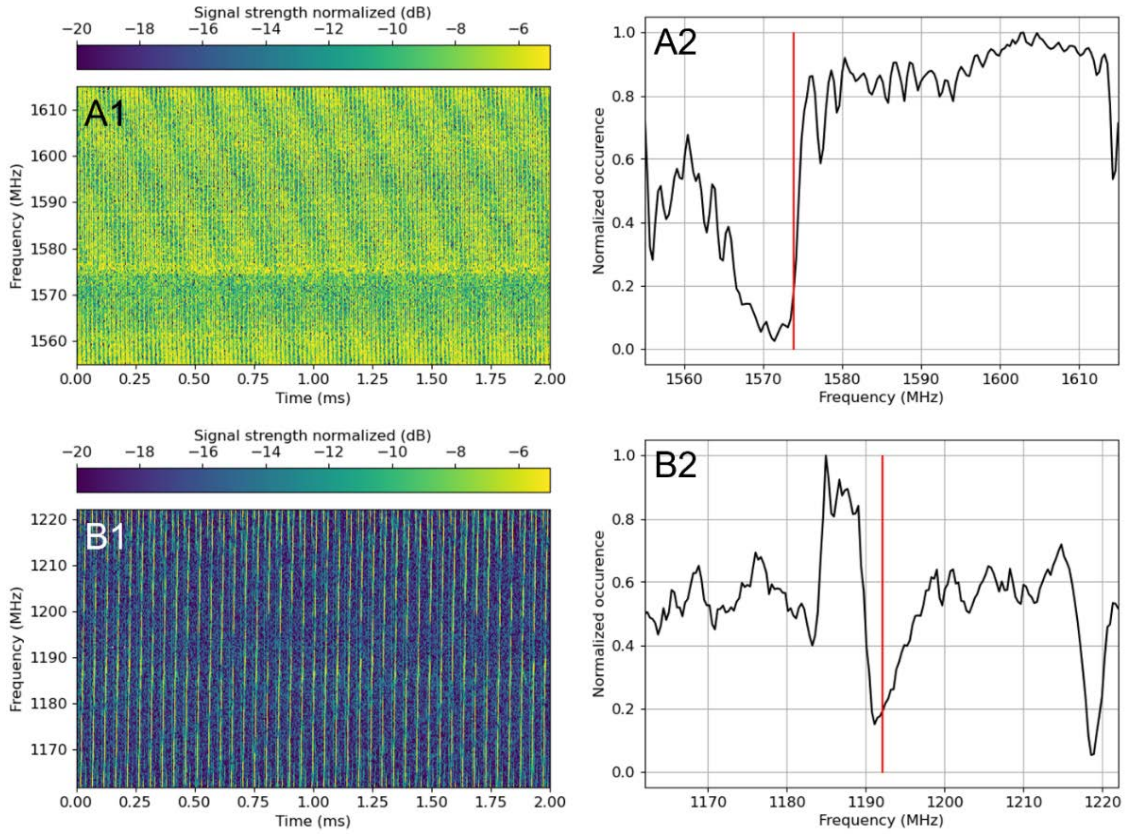

Figure 2: Signals that span the entire bandwidth with areas of weak energy for two examples (A and B) showing the signal as spectrogram (A1 and B1) and the energy frequency distribution (A2 and B2).

In case of an aliased signal, we derive the bandwidth information for an aliased signal form the unaliased and aliased part within the energy frequency distribution (Figure 1C). The centre frequency is then defined at half the bandwidth form the boundary of the bandwidth within the unaliased part of the signal. If the signal is strong over the entire frequency range the entire frequency range is defined as bandwidth. Problems can occur when signals span the complete bandwidth but have a frequency range where the jammer signal is weaker. These signals are not aliased, and are not classified as aliased, but due to the energy distribution and frequencies with low energy level they are handled similar like aliased signal when determining the bandwidth. Examples of this are shown in Figure 3. For the chirp jammer in example A centre frequency and bandwidth of 1600±26 MHz are derived and for example B a centre frequency and bandwidth of 1221±29 MHz from the energy frequency distribution. The derived bandwidth from the dominating jammer signal is 1588±17 MHz for example A and 1189±17 MHz for example B. For both examples a bandwidth that is spanning the entire frequency range with a centre frequency at the centre frequency of the respective band would be the correct result, i.e., 1585±30 MHz and 1192±30 MHz for example A and B, respectively. The derive centre frequency from the energy frequency distribution is more correct than that derived from the dominant jammer signal, however the derived bandwidth is clearly underestimated with 17 MHz instead of 30 MHz and is more correctly derived using the energy frequency distribution. While the centre frequency and bandwidth derived from the energy frequency distribution is preferable for most of the observed jammer events, as the derived bandwidth takes weaker parts of the signal into account, it can lead to incorrect results especially for the centre frequency for some jammer events as presented for the example here.

Signals spanning multiple bands

The bands B to D are partly overlapping with respect to their ranges of digitized spectrum, and while a tuneable IF filter is used to position the passband 3dB attenuation range to nominally eliminate overlap, the jamming signals can be wide enough to cover multiple entire bands simultaneously. Hence, signals can exist that span the complete bandwidth from 1162 MHz (band D) to 1309 MHz (band B). Figure 4 shows a time varying signal that is strongly visible on band B with an aliased portion, well visible on band C with an aliased part and an abrupt stop at about 5 ms and 15 ms, and weakly visible and most likely aliased on band D. The events were classified as wideband general, wideband (aliased), and stepped jammer on band B, C, and D, respectively. At present the analyses of bands B – D is entirely separated from each other to limit the complexity of the analysis process leading to cases where wideband signals that overlap multiple bands may be identified as different jammer signal types.

Signals with strong variation over time

For the analysis of the data, we study 17 ms long segments and assume that the jammer is persistent over the time interval of the recording. However, even within the analysed 17 ms we can observe variations in the signal's appearance, energy, frequency, or bandwidth, which can make it difficult to detect these signals or classify them correctly. Figure 5 shows a range of examples where the signal shows large variations over time. The analysis is not really affected by the increase in energy for the CW jammer shown in Figure 5b. Figure 5d shows an example of two CW events with varying energy over time. The automatic classification does not recognize these as two CW events, due to the variation

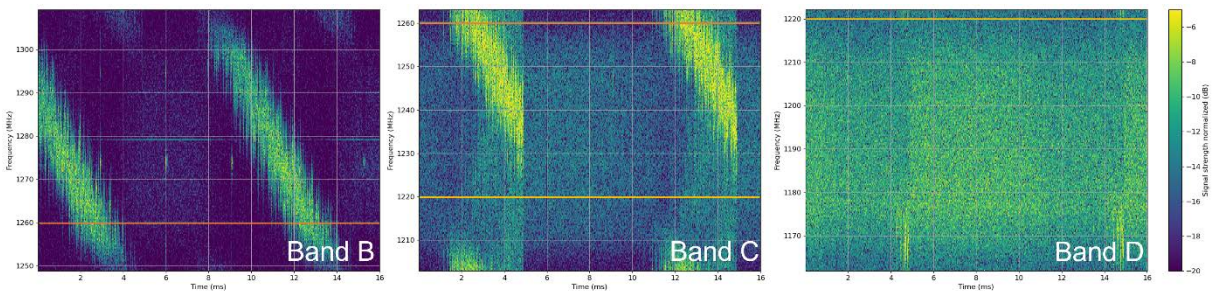

Figure 4: Signals visible on multiple bands. The same frequencies are marked with orange and red lines in the overlapping bands for easier comparison.

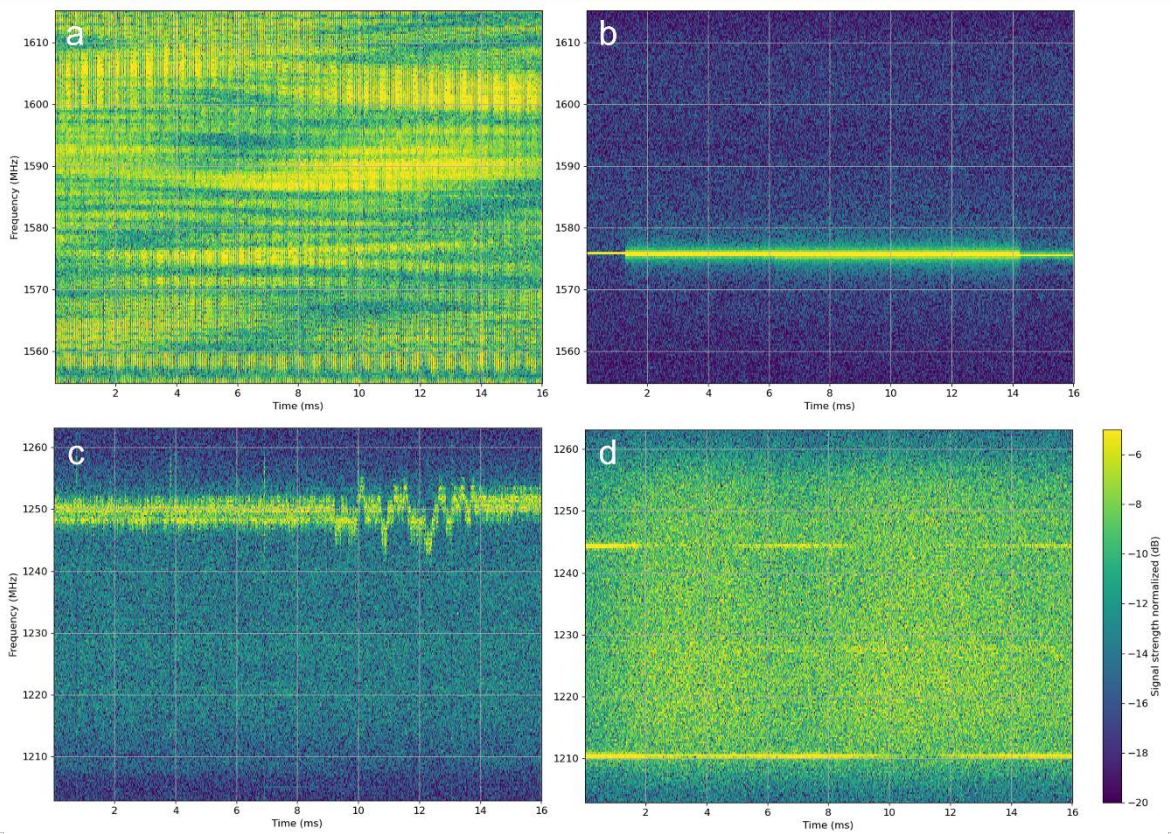

Figure 5*:* Variation in signal frequency and strength over time for four examples: (a) chirp jammer with strong variations in energy, (b) CW jammer with varying signal strength, (c) classified as multi-level chirp, a jammer with multiple time varying signals and (d) a jammer classified as wideband, with two CW jammers with varying energy level over time.

in energy over time and the event is not identified as multi-CW but is classified as wideband signal instead.

Figure 5a shows the spectrogram of a multilevel chirp signal spanning the entire bandwidth of band A with random variations in energy over time and frequency. The strong and random variations in energy the bandwidth and centre frequency result in centre frequencies and bandwidth derivation problems as described before for the examples in Figure 3. Figure 5c shows a wideband signal with sudden changes in centre frequency, which is automatically classified as multilevel chirp due to the variation of peaks in the FFT of the jammer signal. Those variabilities over time can make it very challenging to sort the event into the class that is correct for the majority of the jammer signal, when it is altered by some shorter time variations. It would potentially be more appropriate to define this specific case as a time-modulated undefined jammer signal.

Weak signals

If jammer signals are weak compared to the noise floor it can be challenging to identify the jammer signal type, centre frequency, and bandwidth correctly. This can be a problem when the aim of the classification is to identify the jammer signal type for potential signal localization and identification. However, the aim can also be to identify the disturbed part of the GNSS band not identifying the correct jammer signal type is less of a problem. The largest challenge is then to correctly determine centre frequency and bandwidth of the disturbed part of the GNSS band investigated. Here it can happen that the energy of parts of the background noise is so similar to the energy level of the weak jammer signal that both parts are analysed together and conflated as a jammer, resulting in a centre frequency and bandwidth that does not reflect the frequency range that is affected by the jammer signal.

Figure 6 shows an example with a very weak chirp signal on band C (A1-A4) and D (B1-B4). Both these examples are from the same event. For this event we identify a multilevel chirp with a bandwidth of 9 MHz and a repetition frequency of 116 kHz on band A. On band C, and D the automatic classification identifies a wideband sinus (Figure 6). On band D (Figure 6, B1-B4) the chirp is well

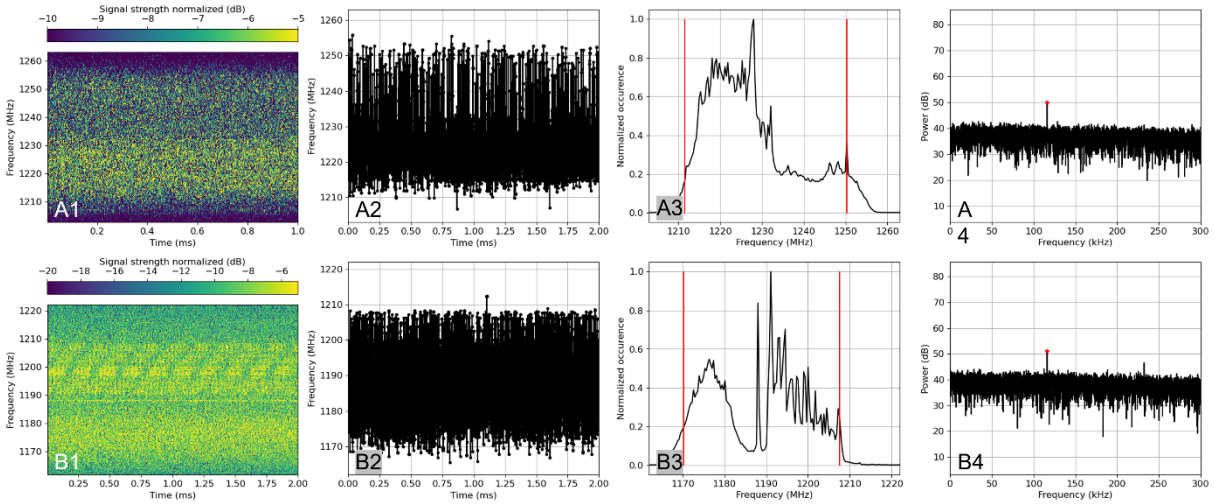

Figure 6: Example of a weak signal on band C (A1-A4) and band D (B1-B4) both automatically classified as wideband sinus. Shown are (1) the spectrogram, (2) the dominating jammer signal, (3) the energy frequency distribution, and (4) the FFT of the dominating jammer signal.

visible with a centre frequency and bandwidth of about 1199±9 MHz. The automatic classification derives this to be a wideband sinus signal with centre frequency and bandwidth of 1189±19 MHz, merging the bandwidth of the chirp between 1190 and 1210 MHz and the higher energy level of the environmental baseline between 1170 and 1183 MHz. On band C (Figure 6, A1-A4) the chirp is barely visible for frequencies below 1210 MHz, while the observed energy between 1210 and 1250 MHz is really the energy level of the environmental baseline. Even though the signals are weak on band C and D the same repetition frequency of 116 kHz as that of the chirp observed on band A is derived. Hence, weak signals might be detected but due to the energy level being similar to the environmental baseline it can be difficult to extract the correct jammer signal type, centre frequency and bandwidth.

Multiple signals at the same time

Significant challenges arise for the automatic classification when multiple jammer signal types occur at the same time. Figure 7 show four examples from data with multiple jammer signal types, with different energy levels occurring at the same time.

The first example (Figure 7, A1-A4) shows a wideband signal with some random variation of the centre frequency over time and a stepped signal at the same time. While the dominant jammer signal function (A2) is mainly picking up the wideband signal, the energy frequency distribution is dominated by the stepped signal (A3). The bandwidth derived from the energy frequency distribution results in a centre frequency and bandwidth of 1274±1 MHz, while we derive 1289±4 MHz from the dominating jammer signal, which much better reflects the disturbed frequency range in this case. The automatic classification algorithm is not able to fit the signal into one of the subcategories and defines this signal as time-modulated und.

The second example (Figure 7, B1-B4) shows a weak triangular signal on a station where a constant CW signal can be observed. Hence, the CW signal is part of the environmental baseline. In the presence of strong jammer signals this CW signal has no influence as the jammer will dominate the analysis. However, if the jammer signal is weak as in this example the analysis is a combination of the CW signal and the jammer. Here, the dominating jammer signal picks up the CW event most of the time except when the triangular is at the maximum frequency. This leads to a signal picked up by the dominating jammer signal function that does not reflect the actual signals present. At the same time is the energy frequency distribution dominated by the CW signal. We derive a centre frequency and bandwidth of 1590±2 MHz and 1591±0.2 MHz from the dominating jammer signal and the energy frequency distribution, respectively. Both do not reflect the centre frequency and bandwidth of the jammed area which is rather 1570±15 MHz. The automatic classification falsely derives this signal to be a chirp signal instead of a triangular signal.

The third example (Figure 7, C1-C4) shows a combination of a triangular signal, a chirp signal, and a CW signal. The CW signal is again part of the environmental baseline and is here weak compared to the

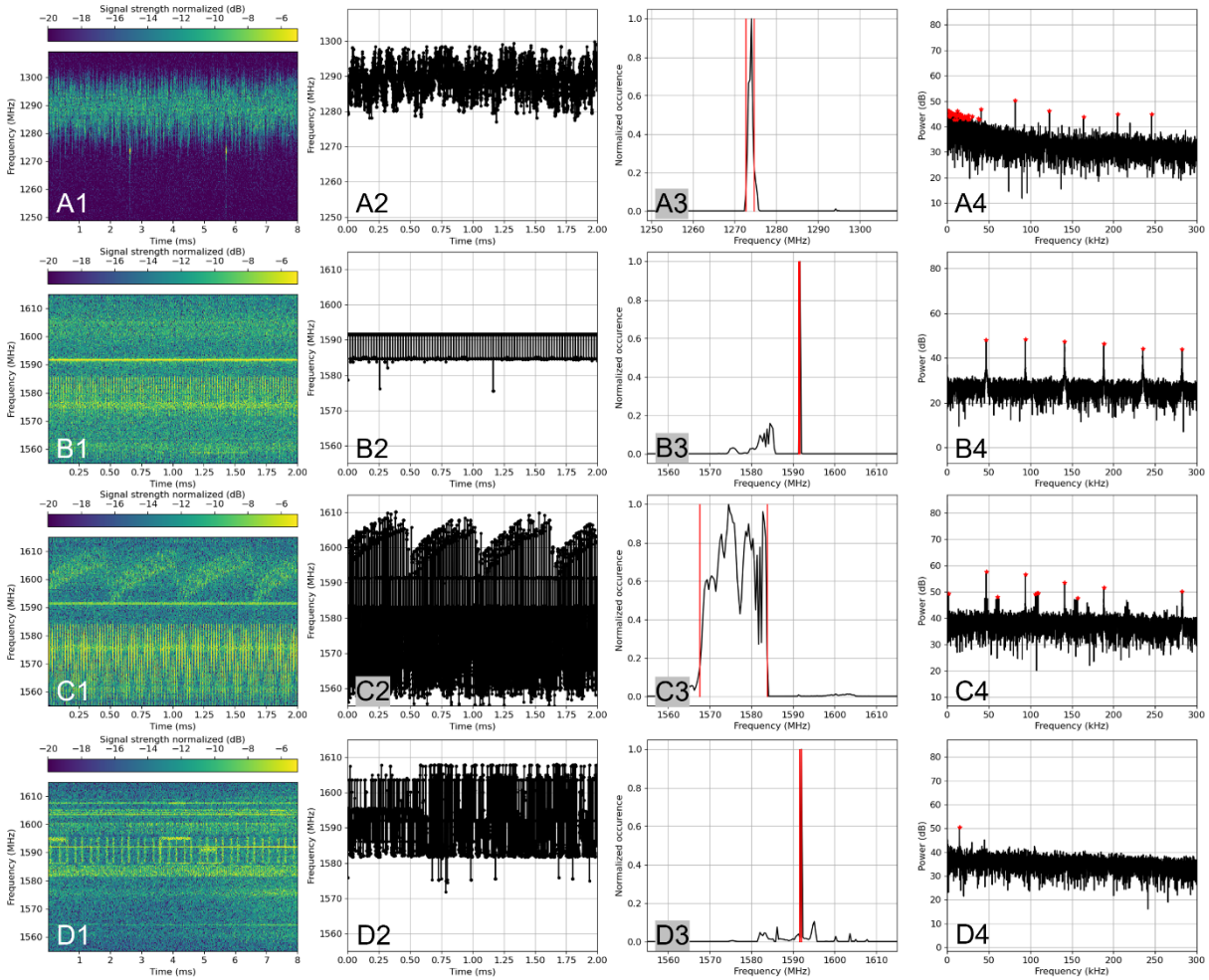

Figure 7: Examples where multiple different jammers occur at the same time. The different effects are discussed in the text. Shown are (1) the spectrogram, (2) the dominating jammer signal, (3) the energy frequency distribution, and (4) the FFT of the dominating jammer signal.

other signals and, therefore, not picked up as dominating signal. The triangular signal observed below 1565 MHz is most likely the same signal as the one observed in example B (same station, events observed on day apart). The chirp signal has an actual chirp over 8 MHz, while at the same time changing the centre frequency of this chirp form about 1595 to 1608 MHz within 0.6 ms. For this event we derive a centre frequency and bandwidth of 1581±14 MHz and 1575±8 MHz from the dominating jammer signal and the energy frequency distribution, respectively. The jammer is classified as multilevel chirp, while it should rather be defined as time-modulated undefined.

The fourth example (Figure 7, D1-D4) shows a whole range of different signals. We observe CW signals, at 1603 and 1595 MHz and weaker CW events at 1608 and 1605 MHz, and an interrupted CW jammer at 1596 MHz, whereby the signal at 1595 MHz is part of the environmental baseline. Wideband signal between 1581 and 1586 MHz are observed, which might be time modulated, as well as wideband signals between 1596 to 1600 MHz, wideband bursts of energy between 1586 and 1596 MHz with an alternating repetition after 0.3 and 0.2 ms, blocks of energy between 1594 and 1596 MHz with a length of 1 ms, and between 1590 and 1591 MHz of about 0.6 ms length. For this event we derive a centre frequency and bandwidth of 1591±5 MHz and 1591±0.2 MHz from the dominating jammer signal and the energy frequency distribution, respectively. The area affected by this jammer is however 1591±17 MHz. The jammer is classified as wideband sinus, while a classification as wideband undefined or time-modulated undefined would be more appropriate.

With a combination of different jammers, it is difficult to classify the dominating jammer signal type. Nevertheless, we derive a jammer signal type that is reasonable in most of these cases. More problematic is again the correct derivation of the centre frequency and bandwidth. Especially in the presence of a

CW event in combination with a time-modulated jammer signal, we observe that the CW event can dominate the energy frequency distribution, resulting in a strong underestimation of the bandwidth.

Conclusion

We observe a large range of different jammer signals that are difficult to sort into a clearly defined category. When analysing the spectrograms manually we are able to distinguish the different jammer signal types occurring at the same time or signals that are weak compared to the noise floor. This separation is possible by analysing the spectrograms at different resolutions and over different time intervals. However, our automatic classification algorithm is not thus far capable of separating these jammer signal types. Nevertheless, we are often able to derive a jammer class that is reasonable. The biggest challenge with this kind of data is the derivation of a correct bandwidth.

The derivation of the correct bandwidth is also a challenge when the jammer signal is weak compared to the environmental baseline [8], when the data is aliased, or for jammer signals that span the entire width of the recorded GNSS band, especially if the signal has frequencies where the jammer signal energy is weak.

A misclassification that often occurs is that of classifying stepped signals as time-modulate undefined signals. In the case of a very regular repetition rate of a stepped signal with short bursts of energy the result of the FFT of the dominant jammer signal has a peak at this repetition frequency resulting in this misclassification. Another challenge is the classification of one jammer spanning multiple bands, which is identified as different jammer classes. This can make it difficult to recognize that this is indeed the same jammer at the different bands. Here, the classification could in future try a more advanced approach analysing the overlapping bands together. However, this is challenging if the signal strength changes between bands, and the signals are aliased on some or multiple of the bands.

For many of the observed jammer signals in our data the classification and derivation of centre frequency and bandwidth works well as shown in [6]. Here we investigated the most challenging jammer signals observed and show that the largest challenge for the classification algorithm is the derivation of the bandwidth in case of weak jammers, strongly varying energy of jammers over time, or multiple signals occurring at the same time. Improvements of the classification algorithm in future should, therefore, concentrate on improving the derivation of centre frequency and bandwidth for challenging jammer signals. Further possibilities for an improved classification might be in the application of machine learning algorithms which could allow for a less rigid classification system than the one based on categories defined by threshold values and a classification matrix which we presented here.

We show that a large variety of jammer signals and combinations of jammer signal can be observed that go far beyond the simple derivation between CW, sweep, and chirp events. Analysis of jammer signals in future should take these variations and possibilities of complicated multi-band jammer signal types into account.

References

[1] T. Kraus, R. Bauernfeind, and B. Eissfeller, ‘Survey of In-Car Jammers - Analysis and Modeling of the RF Signals and IF Samples (Suitable for Active Signal Cancelation)’, p. 7, 2011.

[2] J. C. Grabowski, ‘Field Observations of Personal Privacy Devices’, *Proceedings of the 2012 International Technical Meeting of The Institute of Navigation*, pp. 689–741, Jan. 2012.

[3] N. Gerrard, A. Rødningsby, A. Morrison, N. Sokolova, and C. Rost, ‘GNSS RFI Monitoring and Classification on Norwegian Highways – An Authority Perspective’, St. Louis, Missouri, Oct. 2021, pp. 864–878. doi: 10.33012/2021.17952.

[4] N. Gerrard, A. Rødningsby, A. Morrison, N. Sokolova, and C. Rost, ‘Exploration of unintentional GNSS RFI sources’, *proceedings of the 35th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2022)*, 2022.

[5] R. Morales-Ferre, W. Wang, A. Sanz-Abia, and E.-S. Lohan, ‘Identifying GNSS Signals Based on Their Radio Frequency (RF) Features—A Dataset with GNSS Raw Signals Based on Roof Antennas and Spectracom Generator’, *Data*, vol. 5, no. 1, p. 18, Feb. 2020, doi: 10.3390/data5010018.

[6] A. Diez, A. Morrison, and N. Sokolova, ‘Automatic classification of RFI events from a multi-band multi-site GNSS monitoring network’, *proceedings of the 35th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2022)*, Sep. 2022.
[7] A. Morrison, ‘NAVISP Element 3, ARFIDAAS, User Manual V1’, SINTEF, 2020.
[8] A. Morrison, N. Sokolova, J. E. Hakegard, T. H. Bryne, and L. Ruotsalainen, ‘A Multi-Site Quad-Band Radio Frequency Interference Monitoring Alerting and Reporting System’, in *2020 European Navigation Conference (ENC)*, Dresden, Germany, Nov. 2020, pp. 1–10. doi: 10.23919/ENC48637.2020.9317522.
[9] R. Morales Ferre, A. de la Fuente, and E. S. Lohan, ‘Jammer Classification in GNSS Bands Via Machine Learning Algorithms’, *Sensors*, vol. 19, no. 22, p. 4841, Nov. 2019, doi: 10.3390/s19224841.

Biographies

Anja Diez received a PhD (Dr. rer. nat.) from the Karlsruhe Institute of Technology. During her PhD and PostDoc time she worked on signal processing of seismic and radar data applied to glaciological application. Currently, she works as a researcher in the acoustics group at SINTEF. Her work concentrates on signal processing of ultrasound data.

Aiden Morrison received his PhD degree in 2010 from the University of Calgary, where he worked on ionospheric phase scintillation characterization using multi frequency civil GNSS signals. Currently, he works as a senior research scientist at SINTEF. His main research interests are in the areas of GNSS and multi-user collaborative navigation systems and GNSS RFI monitoring and analysis.

Nadezda Sokolova received her PhD degree in 2011 from Norwegian University of Science and Technology (NTNU), where she worked on weak GNSS signal tracking and use of GNSS for precise velocity and acceleration determination. She is now working as a senior research scientist at SINTEF, and adjunct associate professor at the Engineering Cybernetics Department, NTNU, focusing on GNSS integrity and multi-sensor navigation for autonomous system operations.