Thunderstorms identification tool for LIGO

G. Valdes[1], **B. Martinez**[2], **L. Capistran**[2], **E. Florio**[3], **E. Thompson**[4]

[1]Louisiana State University, Baton Rouge, LA 70803, USA
[2]Center for Gravitational Wave Astronomy, University of Texas Rio Grande Valley, Brownsville, TX 78520, USA
[3]Kenyon College, Gambier, OH 43022, USA
[4] Clemson University, Clemson, SC 29634, USA

E-mail: `guillermo.valdes@ligo.org`

22 October 2020

Abstract. The Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors are sensitive to noise from local environmental sources, such as thunderstorms. We developed a method to identify thunderclaps using feature extraction and machine learning algorithms. In combination with ground motion data, we identify thunderstorm segments. We tested the tool in LIGO-Livingston data and found that, during thunderstorms, this detector experienced losses up to approximately 15% in its sensitivity.

Keywords: detector-characterization, thunderstorms, machine-learning

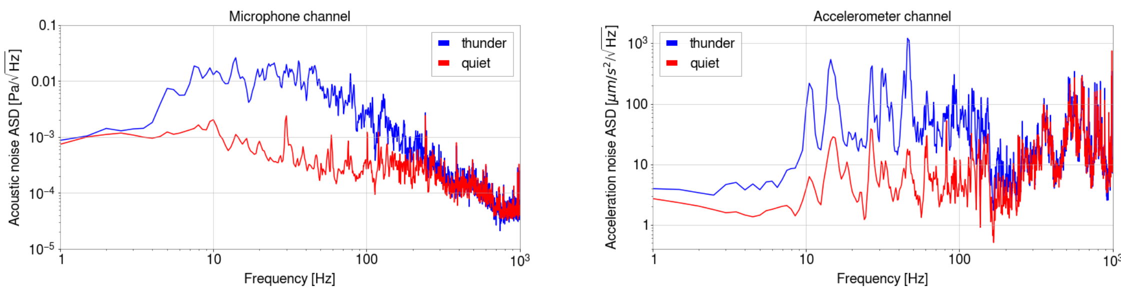
1. Introduction

This document is intended to describe the implementation of the Thunderstorms Identification tool. In the following sections, we report the steps to create the tool and how can we use it to analyze the effect of storms in LIGO binary neutron star (BNS) inspiral range (so-called "BNS range"). LIGO's third observational run (O3) began on April 1, 2019, went through a commissioning break from October 1, 2019, to November 1, 2019, and ended on March 27, 2020. The periods of observations before and after the commissioning break are referred to as O3a and O3b, respectively. We tested the tool in LIGO-Livingston (LLO) data recorded in O3.

2. Methods

2.1. Effect of thunderstorms on sensors

A thunder is accompanied by a thunderclap, detected by the microphones (see Figure 1a), and the rumble, sensed by the ground motion sensors (see Figure 1b), at frequencies between 2 Hz and 200 Hz, with more intensity in the frequency band 10 Hz–100 Hz.



(a) (b)

Figure 1. Amplitude spectral density comparison between a quiet segment (red line) and in the presence of a thunder (blue line) for a microphone (left) and an accelerometer (right). Excess noise is more evident in the frequencies band 10 Hz–100 Hz.

2.2. Acoustic features extraction

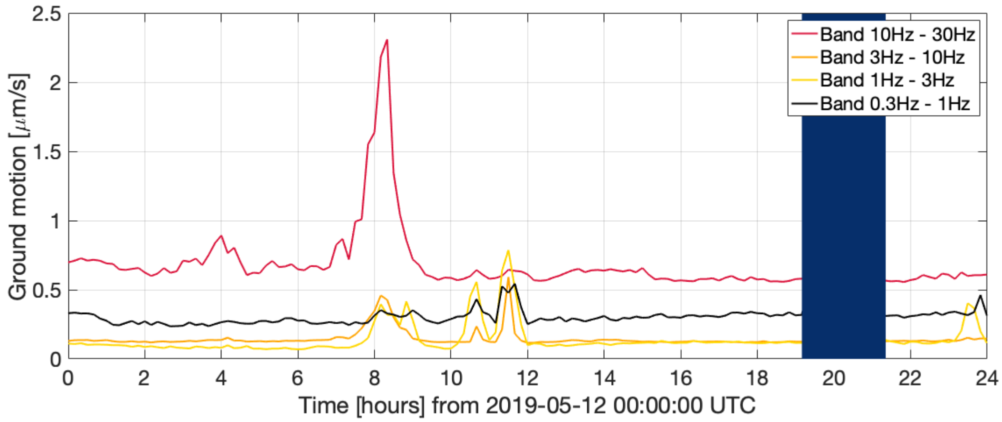
We employ techniques to identify features of the thunderclap while ignoring background noise, such as Mel-frequency cepstral coefficients (MFCC) [1], chroma features, spectral contrast, spectral centroid, spectral flatness, and root-mean-square (RMS) value. We used LibROSA [2], a library for audio analysis for feature extraction, for this purpose. Before the feature extraction, the microphone data is filtered between 2 Hz–200 Hz.

2.3. Classification

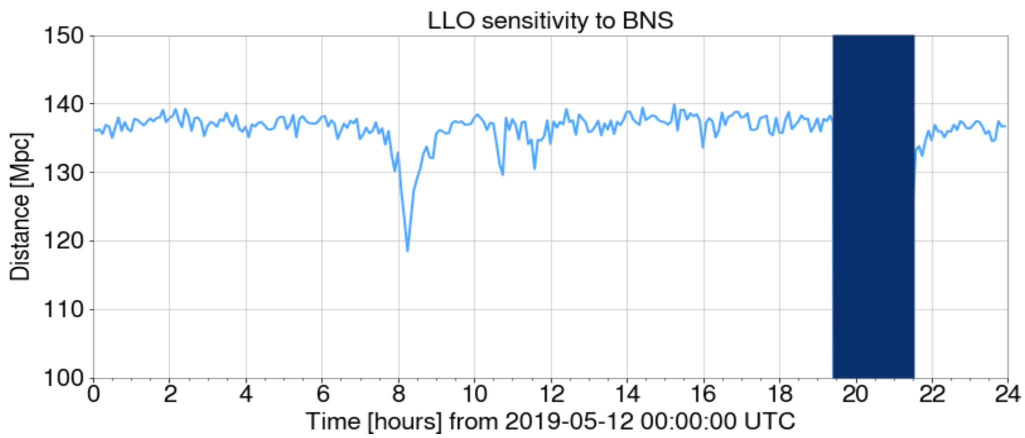
For the categorization, we use a *k-nearest neighbors (KNN)* classifier [3]. In the KNN algorithm, a sample consisting of a set of features is classified by a majority vote of its neighbor samples, with the sample being assigned to the class most common among its k-number nearest neighbors. KNN is a non-parametric method, which means that it does not make any assumptions on the data distributions. KNN is also sensitive to outliers and removing them before tends to improve the results.

2.4. Trigger selection

We use a list of transients (so-called "triggers") in the microphone data identified by Omicron [4, 5]. The information for each trigger includes the time at which it happened, signal-to-noise ratio, and duration. The microphones used are located inside the buildings. For each trigger, we band-pass 10 s of data between 2 Hz–1000 Hz.



(a)



(b)

Figure 2. (a) A thunderstorm on May 12, 2019, 08:15:00 UTC is registered with a higher amplitude by the ground motion in the frequency band 10 Hz–30 Hz (red line). (b) The same thunderstorm is coincident with a drop in LLO's sensitivity from 135 Mpc to 120 Mpc. Blue blocks represent "not observing" periods.

2.5. Band-limited root-mean-square (BLRMS) ground motion

Depending on the frequency band, BLRMS ground motion data, recorded by seismometers, helps to characterize the effect of distinct phenomena in LIGO. For example, we use the frequency band 0.03 Hz–0.1 Hz to describe the earthquake activity, the band 0.1 Hz–0.3 Hz tends to elevate on a high wind day, or 1 Hz–10 Hz relates to anthropogenic events. We found that the band 10 Hz–30 Hz increases more than other bands during thunderstorms, as seen in Figure 2a. At the same time, the LIGO-Livingston's sensitivity dropped from around 135 Mpc to 120 Mpc, as shown in Figure 2b.

3. Implementation

We created a training set with 148 triggers previously identified as *thunder* or *other* in human evaluation of the microphone recording. We used six neighbors ($k = 6$) to train the KNN algorithm and classified the rest of the microphone triggers. The trained model achieved a 98% true-positive classification. We ran the feature extraction and classification in HTCondor [6]. Of the 1200 loudest Omicron triggers of microphone data during the observation period O3, 345 were categorized as thunder with registered amplitudes ranging from 0.10 Pa up to 3.90 Pa.

We used the *thunder* triggers and the ground motion BLRMS 10 Hz–30 Hz to distinguish thunderstorms segments on July 20–22, 2019 (see Figure 3a). We estimated the difference between the minimum value of sensitivity to BNS reached during the storm and the average sensitivity during quiet-segments of the day (see Figure 3b). These quiet-segments are those times without storms, starting 1 h after the detector entered to observation status and 20 min before any loss of lock of the interferometer. LLO sensitivity to BNS tends to be lower during the daytime. Therefore, we used the average nighttime (00:00:00–12:00:00 UTC) or daytime sensitivity (12:00:00–24:00:00 UTC) depending on the time when the thunderstorm happened.

The percentage of sensitivity loss (R_{loss}) is defined by Equation 1, where R_{storm} is the average sensitivity during the thunderstorm and R_{quiet} is the average sensitivity during the closest quiet-segment.

$$R_{loss} = 100 \times \left[1 - \frac{R_{storm}}{R_{quiet}}\right] \tag{1}$$

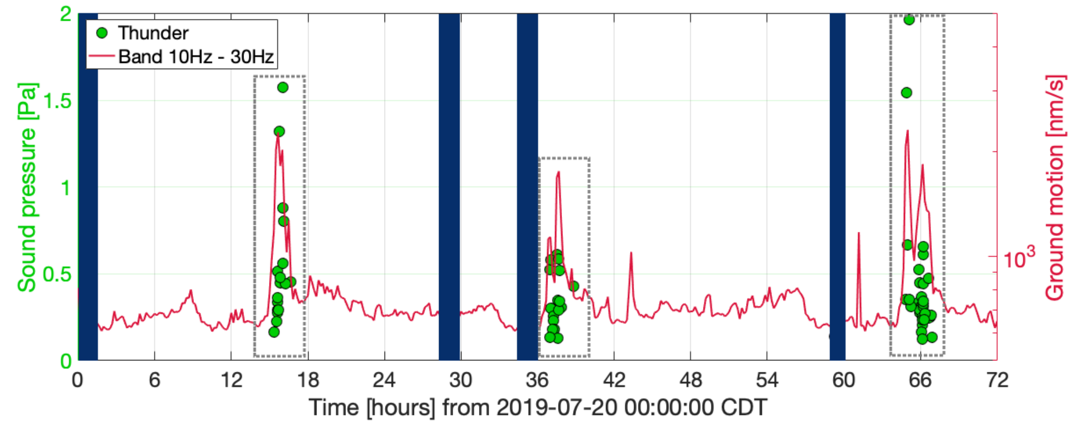
On the other hand, the difference in ground motion (ΔG) is defined as the maximum ground motion value reached during the storm (G_{storm}) minus the average ground motion during the closest quiet-segment (G_{quiet}).

Figure 5 shows the percentage of sensitivity loss versus difference in 10 Hz–30 Hz ground motion, during O3. The Pearson correlation coefficient between these two quantities is 0.63, and there are sensitivity losses up to 14.5%.

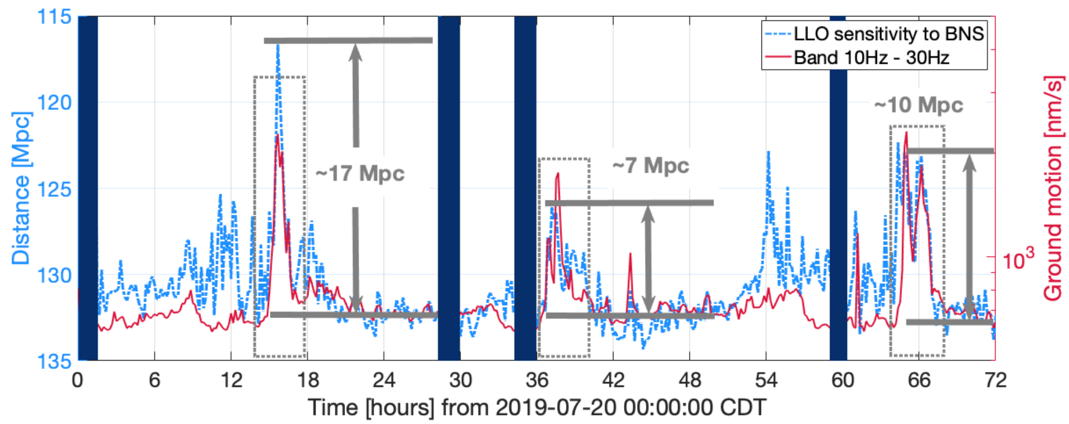
4. Summary

The sensitivity of LIGO detectors is affected by thunderstorms. We have developed a method using Omicron triggers in microphone channels, acoustic feature extraction, and a machine learning classifier to identify thunderclaps. Together with 10 Hz–30 Hz ground motion data sensing the rumble of thunder, we identify thunderstorm segments.

Having identified thunderstorm segments, we quantified its effect on the LIGO-Livingston sensitivity to binary neutron star systems and found that it decreased by up to approximately 15%, during O3.



(a)



(b)

Figure 3. Ground motion and LLO sensitivity during thunderstorms on July 20–22, 2019, UTC. (a) Peaks in the ground motion (solid line) at frequencies between 10 Hz–30 Hz and loud noises in the microphones, identified as thunder (solid circles), characterize thunderstorm segments (dashed rectangles). (b) Decrease in the LLO sensitivity (dashed line) coincide with thunderstorms (dashed rectangles). We flipped the LLO sensitivity for better appreciation. Dark blocks represent "not observing" periods.

References

[1] Oppenheim A V and Schafer R W 2004 *IEEE Signal Processing Magazine* **21** 95–106

[2] McFee B *et al.* 2019 librosa/librosa: 0.7.1 URL `https://doi.org/10.5281/zenodo.3478579`

[3] Tan P N, Steinbach M and Kumar V 2005 *Introduction to Data Mining, (First Edition)* (USA: Addison-Wesley Longman Publishing Co., Inc.) ISBN 0321321367

[4] Robinet F 2015 Omicron: An algorithm to detect and characterize transient noise in gravitational-wave detectors Tech. rep. Virgo URL `https://tds.ego-gw.it/ql/?c=10651`

[5] Chatterji S *et al.* 2004 *Class. Quant. Grav.* **21** S1809–S1818 (*Preprint* `gr-qc/0412119`)

[6] Thain D, Tannenbaum T and Livny M 2005 *Concurrency - Practice and Experience* **17** 323–356

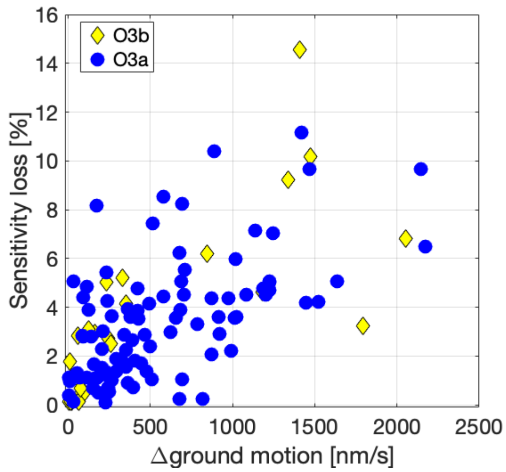


Figure 4. Scatter plot of sensitivity loss versus the difference in 10 Hz–30 Hz ground motion. The Pearson correlation coefficient between these two quantities is 0.63. Blue dots represent storms in O3a, and yellow diamonds represent storms in O3b.