The Journal of Open Source Software

gemlog: Data Conversion for the Open-Source Gem Infrasound Logger

Jacob F. Anderson [1¶], **Kevin S. Anderson** [2], and **Tamara Beschorner** [3]

1 Department of Geosciences, Boise State University, USA **2** Independent Researcher **3** Oregon Hazards Lab, Earth Science Department, University of Oregon, USA ¶ Corresponding author

DOI: 10.21105/joss.05256

Software

- Review
- Repository
- Archive

Editor: Brian McFee

Reviewers:

- @thelenwes
- @hemmelig

Submitted: 18 November 2022
Published: 15 June 2023

Summary

Infrasound (low-frequency pressure waves in air) is widely used as a geophysical method for monitoring powerful, often hazardous processes like volcanic explosions, nuclear tests, bolides, avalanches, mudflows, and earthquakes. Infrasound can also be used to monitor changes in atmospheric wind fields, and balloon-borne infrasound has been proposed as a means of seismic monitoring on planets like Venus where ground-based monitoring is impractical. In infrasound research campaigns, signals are typically recorded to disks by unsupervised, non-telemetered instruments, and recordings must be converted to standard data formats before analysis and distribution. The `gemlog` package handles data conversion and facilitates data analysis for an open-source instrument, the Gem Infrasound Logger.

Statement of need

The Gem Infrasound Logger (Anderson et al., 2018) is an approach to infrasound recording where the sensor and data logger are built into a single cable-free package that is easy to conceal and permits arbitrary sensor network geometries. Additionally, it has a low cost and weight, runs for months on alkaline batteries, and has a simple, fast installation procedure. These characteristics make it a good choice for temporary infrasound campaigns. By contrast, most campaigns that do not use the Gem use analog infrasound sensors that, via long cables, connect to multichannel data loggers built to record seismometers. This approach yields high-quality data but has several disadvantages: seismic data loggers are expensive, and sensor cables constrain the sensor network's geometry, make a station prone to animal damage and vandalism due to their visibility and exposure, and account for a large share of a recording site's budget for weight, bulk, and setup time. These disadvantages are especially acute for temporary recording campaigns (as opposed to permanent installations), which account for a large share of infrasound research.

Like many geophysical data loggers, the Gem writes data in a non-standard raw format intended to balance firmware simplicity and performance, human readability, and compact file size. Although it is a human-readable text format that an expert can read and understand on a line-by-line basis, data files consist of hundreds of thousands of lines with complicated formatting, meaning that reading it as a spreadsheet or data frame is impractical. Further, operations like clock drift corrections, data decompression, and conversion to standard file formats or classes must be performed to make the data accessible in standard visualization and analysis software. Users often need to convert data from 10 or more infrasound loggers spanning several weeks, meaning that thousands of files including billions of data points must be processed efficiently. The `gemlog` Python package is a cross-platform tool to facilitate data conversion, and is essential for all Gem infrasound logger users.

Some other geophysical data loggers used to record infrasound, for example the DiGOS

DataCube-3 and Reftek RT-130, have non-standard raw data formats that must be converted to standard formats by software distributed by the manufacturer. Like `gemlog`, they conduct data conversion as a simple command line operation. However, to the authors' knowledge, none of those software packages (or the raw data formats they convert) are open-source.

Use in Research

The Gem Infrasound Logger (including `gemlog`) has been included in independent evaluations of infrasound instruments (Kramer et al., 2021; Slad & Merchant, 2021), and has been used in several past and upcoming publications, including the following:

- Volcano monitoring (Bosa et al., 2021; Mock et al., 2020; Rosenblatt et al., 2022)
- Monitoring atmospheric changes using infrasound (Averbuch et al., 2022; Dannemann Dugick & Bowman, 2022)
- Infrasound monitoring from high-altitude balloons (Bowman et al., 2020; Bowman & Albert, 2018; Brissaud et al., 2021; Krishnamoorthy et al., 2020; Young et al., 2018)
- River rapid infrasound monitoring (Gauvain et al., in prep.; Ronan, 2017; Scamfer & Anderson, submitted; Tatum et al., 2023)
- Remotely monitoring earthquake ground shaking (Anderson et al., submitted; Scamfer & Anderson, submitted)

Features

gemlog (GPL-3 license) is a Python library that includes both terminal commands and Python functions. It is installable from PyPI and documented in terminal command help pages, Python function docstrings, and on Read the Docs.

- Data conversion: Terminal commands `gemconvert` and `gemconvert_single` convert sets of raw files into standard data formats (typically miniSEED). Because infrasound analyses require sample timing to be approximately millisecond-precise, an essential part of this process is correcting clock drift using accurate times provided infrequently by the Gem's on-board GPS. Contiguous blocks of raw data are converted into contiguous blocks of output data, and the software identifies breaks in recording and includes the same breaks in the output. On a typical laptop, conversion may take on the order of 10 wall-clock seconds per day of data for one station (or, a unitless ratio of approximately 10^{-4} between conversion time and data duration).
- Instrument testing: Terminal command `gem_verify_huddle_test` is used to automatically examine waveform, state-of-health, and GPS data from several instruments recording in the same place at the same time (termed a "huddle test" in seismology/infrasound), and verify that all instruments are working properly.
- Data analysis and visualization in Python: several Python functions facilitate working with data, including plotting spectra of the Gem's self-noise and standard environmental noise specs set by the International Monitoring System, and deconvolving the Gem's instrument response from recordings. `gemlog` is well-integrated with the common seismic/infrasound data processing Python package `obspy` (Krischer et al., 2015) and uses its functions and classes when possible.

Demonstrations

`gemlog` includes three demonstrations of its functionality, including example data, example code, and explanations. These demonstrations cover the most common uses of `gemlog`.

- A typical data conversion, pre-processing, and inspection workflow, including using `obspy` tools to process and plot data and metadata (Figure 1).
- Data conversion workflows that can be used for data lacking GPS data (typically recorded at high-altitude, indoors, or underground) in which normal timing corrections are impossible.
- A quality-control workflow to inspect an infrasound dataset from multiple instruments recording simultaneously in the same location, verifying that all instruments appear to work correctly and record similar data.

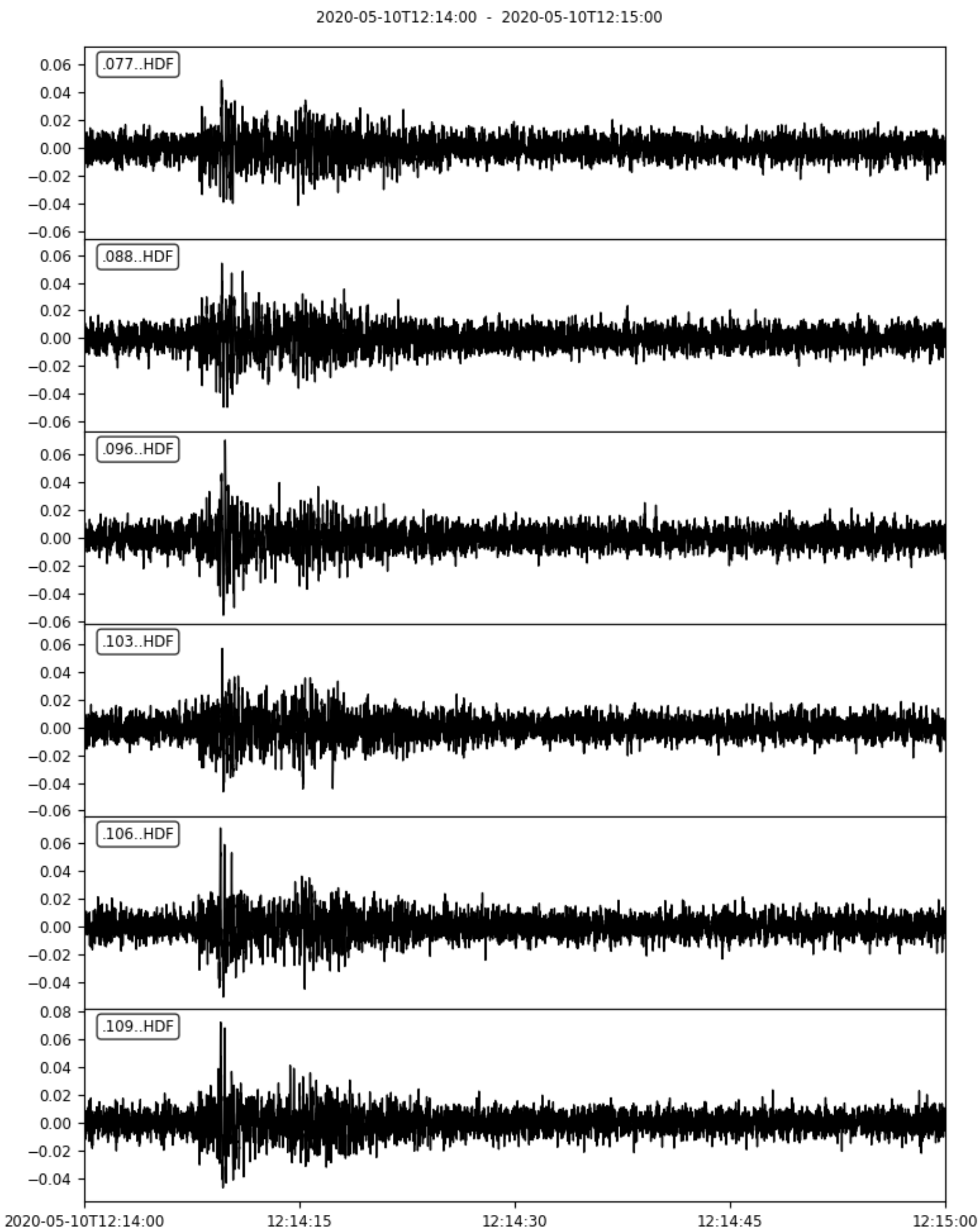

Figure 1: Plot of an obspy.Stream of example infrasound data created by the main data conversion workflow demonstration.

Acknowledgements

This work was funded by NSF award EAR-2122188.

References

Anderson, J. F., Johnson, J. B., Bowman, D. C., & Ronan, T. J. (2018). The Gem infrasound logger and custom-built instrumentation. *Seismological Research Letters*, *89*(1), 153–164. https://doi.org/10.1785/0220170067

Anderson, J. F., Johnson, J. B., Mikesell, T. D., & Liberty, L. M. (submitted). *Remote detection of earthquake ground shaking via high-resolution large-N infrasound beamforming.*

Averbuch, G., Ronac-Giannone, M., Arrowsmith, S., & Anderson, J. F. (2022). Evidence for short temporal atmospheric variations observed by infrasonic signals: 1. The troposphere. *Earth and Space Science*, *9*(3), e2021EA002036. https://doi.org/10.1029/2021ea002036

Bosa, A., Pineda, A., Mock, J., Roca, A., Johnson, J. B., Bejar, G., Waite, G., Gauvain, S. J., & Bartel, B. (2021). Dynamics of rain-triggered lahars inferred from infrasound array and time-lapse camera correlation at Volcan de Fuego, Guatemala. *AGU Fall Meeting Abstracts*, *2021*, V25D–0150.

Bowman, D. C., & Albert, S. A. (2018). Acoustic event location and background noise characterization on a free flying infrasound sensor network in the stratosphere. *Geophysical Journal International*, *213*(3), 1524–1535. https://doi.org/10.1093/gji/ggy069

Bowman, D. C., Norman, P. E., Pauken, M. T., Albert, S. A., Dexheimer, D., Yang, X., Krishnamoorthy, S., Komjathy, A., & Cutts, J. A. (2020). Multihour stratospheric flights with the heliotrope solar hot-air balloon. *Journal of Atmospheric and Oceanic Technology*, *37*(6), 1051–1066. https://doi.org/10.1175/jtech-d-19-0175.1

Brissaud, Q., Krishnamoorthy, S., Jackson, J. M., Bowman, D. C., Komjathy, A., Cutts, J. A., Zhan, Z., Pauken, M. T., Izraelevitz, J. S., & Walsh, G. J. (2021). The first detection of an earthquake from a balloon using its acoustic signature. *Geophysical Research Letters*, *48*(12), e2021GL093013. https://doi.org/10.1029/2021gl093013

Dannemann Dugick, F., & Bowman, D. (2022). *Data report: TurboWave I and II data release.* Sandia National Lab.(SNL-NM), Albuquerque, NM (United States). https://doi.org/10.2172/1863279

Gauvain, S. J., Anderson, J. F., Yager, E. M., & McNamara, J. P. (in prep.). *Effects of discharge and morphology on fluvial sound.*

Kramer, A., Marty, J., & Doury, B. (2021). Current PTS activities related to low-cost infrasound sensors. *SnT 2021 CTBT Science and Technology Conference*, *2021*, T3.1–221.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., & Wassermann, J. (2015). ObsPy: A bridge for seismology into the scientific python ecosystem. *Computational Science & Discovery*, *8*(1), 014003. https://doi.org/10.1088/1749-4699/8/1/014003

Krishnamoorthy, S., Bowman, D. C., Komjathy, A., Pauken, M. T., Cutts, J. A., Carlo, M. D., Ardhuin, F., & Le Pichon, A. (2020). *Infrasound measurements in the Arctic on a long-duration balloon*. https://hdl.handle.net/2014/52499

Mock, J., Anderson, J. F., Johnson, J. B., Rosenblatt, B., & Gauvain, S. J. (2020). Topographic effects of infrasound wave propagation. *AGU Fall Meeting Abstracts*, *2020*, S001–0002.

Ronan, T. J. (2017). *Fluvial seismo-acoustics: Characterizing temporally dependent river conditions through near river seismic and acoustic signals* [PhD thesis, The University of North Carolina at Chapel Hill]. https://doi.org/10.17615/nseg-wy10

Rosenblatt, B. B., Johnson, J. B., Anderson, J. F., Kim, K., & Gauvain, S. J. (2022). Controls on the frequency content of near-source infrasound at open-vent volcanoes: A case study from Volcán Villarrica, Chile. *Bulletin of Volcanology*, *84*(12), 1–17. https://doi.org/10.1007/s00445-022-01607-y

Scamfer, L. T., & Anderson, J. F. (submitted). *Exploring background noise with a large-N infrasound array: Waterfalls, thunderstorms, and earthquakes*.

Slad, G., & Merchant, B. (2021). *Evaluation of low cost infrasound sensor packages.* Sandia National Lab.(SNL-NM), Albuquerque, NM (United States). https://doi.org/10.2172/1829264

Tatum, T. A., Anderson, J. F., & Ronan, T. J. (2023). Whitewater sound dependence on discharge and wave configuration at an adjustable wave feature. *ESS Open Archive*. https://doi.org/10.22541/essoar.167397475.53275976/v1

Young, E. F., Bowman, D. C., Lees, J. M., Klein, V., Arrowsmith, S. J., & Ballard, C. (2018). Explosion-generated infrasound recorded on ground and airborne microbarometers at regional distances. *Seismological Research Letters*, *89*(4), 1497–1506. https://doi.org/10.1785/0220180038