

WALDO! A Massive Public Repository of Global ELF/VLF Radio Data

Morris Cohen¹, Mark Golkowski², Umran Inan³, and John DeSilva⁴

¹Georgia Institute of Technology Main Campus

²University of Colorado Denver

³Koç University

⁴Stanford University

November 22, 2022

Abstract

Observations of radio waves in the Extremely Low Frequency and Very Low Frequency band (ELF/VLF, 0.3-30 kHz) have a host of geophysical uses, including lightning detection and characterization, D-region ionosphere remote sensing, detection of solar flares and geomagnetic storms, gravity waves, gamma-ray burst detection, observations of whistlers, chorus and hiss, to infer wave-particle interactions in the magnetosphere, plasmaspheric state. It's been looked at for earthquake forecasting and also has commercial uses like submarine communications and subterranean prospecting. For many years ELF/VLF data have been collected at various locations and by various groups around the world for a variety of scientific purposes, but most of this data is not available publicly. We introduce the World Archive of Low-frequency Data and Observations (WALDO), a repository of ELF/VLF data from recordings taken over the past two decades by Stanford University and subsequently by Georgia Tech and University of Colorado Denver. The locations of the recordings are all around the world, including Alaska, Antarctica, Australia, and many low and mid latitude stations. Some sites were more consistent than others but there's a lot of untapped value in this dataset. Funding for these recordings came from many years of funding from NSF, NASA, DoD, and others, on various basic science projects, and we feel a responsibility to make sure the datasets are now preserved. We are in the process of transferring many 100s of TBs of data and sharing every raw bit for anyone to download and analyze. This includes both "broadband" data that includes the entire spectrum from 500 Hz – 50 kHz, and "narrowband" data corresponding to amplitudes and phases of specific transmitting beacons. We are also including automatically generated summary plots, and a host of basic analysis tools to allow anyone to download and analyze the data. We will announce and present WALDO, update its status and timeline for full deployment, and detail some of the uses of ELF/VLF data, with the goal of enabling its use by anyone interested. We will not be finished by the Fall meeting (ripping 80,000 DVDs can take a while) but whatever we finished will be public and hopefully we will be far along by then. Finally, we will have the answer to the age-old question... "Where's WALDO?"

A Massive Public Repository of Global ELF/VLF Radio Data

Morris Cohen¹, Mark Golkowski², Umran Inan^{3,4}, John DeSilva³



You found WALDO!

- 1 Electrical and Computer Engineering, Georgia Tech
- 2 Electrical Engineering, University of Colorado-Denver
- 3 Electrical Engineering, Stanford University
- 4 Electrical Engineering, Koc University

AE33B-3156

We are proud today to announce WALDO: the Worldwide Archive of Low-frequency Radio Data. WALDO is the largest repository of publicly available ELF/VLF radio data. It includes decades of historical data captured around the world by VLF research groups.



http://waldo.world



What Data is on WALDO?

For decades, the Stanford VLF group collected data for all kinds of geophysical experiments. Much of it piled up in DVDs, or even analog tapes, as the group wound down. Now, they are all being placed in the cloud. With big data and machine learning spreading, we hope to catalyze new discoveries.

Data that are or will be on WALDO include

- Antarctic recordings at Palmer Station, South Pole
- Siple Station Antarctica recordings from 1974-1986
- Alaska recordings during HAARP experiments
- Data from the 2017 Great American Solar Eclipse
- Data from global IHY/ISWI/AWESOME receivers
- Recordings preceding the 2011 Tohoku earthquake

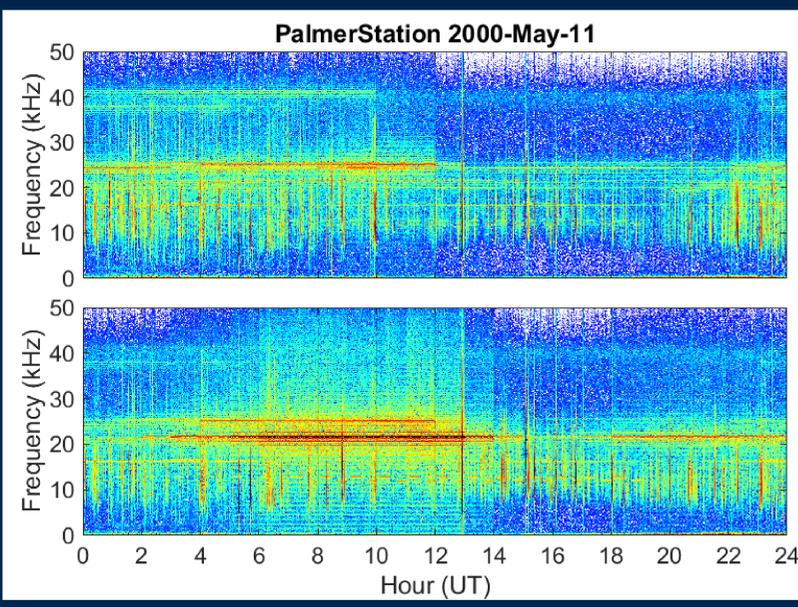
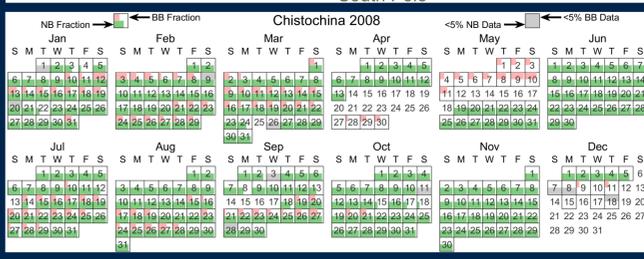
In addition to ~80,000 DVDs and ~250 TB of old Stanford VLF group recording, we are also using it to deposit more recent recordings (2015 and later) from Georgia Tech and the University of Colorado Denver. So WALDO will also be a living data repository.



How Do I Navigate WALDO?

WALDO is populated with several automatically-generated maps, calendars, quick-look plots, to help you identify what is already on the server. See examples below. There is also a detailed description of the format and sample MATLAB scripts to enable anyone to read the data with ease. You can then navigate directly to whatever site and date you want, and download entire files or folders.

Remember, we are anticipating ~1000 TB of data eventually, and so far have only transferred ~150 TB, so there's a lot more on the way in the coming months. Check with us if you want to know what we've got coming.



Are There Any Strings Attached?

No. These data were collected under various US government grants (NSF, ONR, DARPA, AFRL) over many years, and belong now to the entire scientific enterprise. Anyone can download the data, no permission, no cost, no account. We request only an acknowledgement in any future publication that uses data from WALDO.

If extensive amounts of WALDO data are used, the curators request, but do not require, to be contacted to discuss the possibility of joint authorship, with the WALDO curators providing help analyzing and interpreting the large dataset.

If more direct and extensive access is needed for large-scale studies, please contact the curators, we may be able to provide a direct Google Drive account.

morris.cohen@waldo.world, mark.golkowski@waldo.world

But Where's WALDO?

Haha! In the top right corner. But can you find *Odlaw*?