University of Victoria

Department of Physics and Astronomy

Adding Support for Accelerometers and General Seismometer Improvements

Spencer Plovie

spencerwplovie@uvic.ca

In partial fulfillment of the requirements of

the Physics and Astronomy Co-op Program

Summer 2022

Work Term #2 of 4

Junior Scientific Programmer

Confidentiality notice: This report is not confidential and may be shared with other co-op students.

This report will be handled by the UVic Co-op staff and will be read by one assigned report marker who may be a co-op staff member within the Physics and Astronomy Science Co-operative Educative Program, or a UVic faculty member or teaching assistant. The report will be retained and available to the student or, subject to the student's right to appeal a grade, held for one year after which it will be deleted.

I approve the release of this report to the University of Victoria for evaluation purposes only.

Supervisor's Signature: Ben Biffard Position: Sr Scientific Programmer Date: Aug 19, 2022

Supervisor's Name (print): Ben Biffard Email: bbiffard@oceannetworks.ca

Company: Ocean Networks Canada

Contents

I. Introduction

Ocean Networks Canada (ONC) is a data pipeline designed to deliver cabled, mobile, and community-based ocean data to the general public. Established as a not-for-profit in 2007, the data is freely available through ONC's data management system (DMAS), Oceans 3.0[1]. This archived and continuous data acquisition is available in a wide variety of formats, and as new instruments are added to the network, new data products (DPs) are required. The scientific programmer is primarily responsible for creating, maintaining, and improving data products code, and works with the Data Products team in the Software Development Department, which is largely responsible for maintaining the DMAS framework.

Much of the work performed in this position was done using MATLAB, a programming language developed by MathWorks that is widely used by scientists and engineers. When a user goes to the Data Search page on the ONC website, they can select a device they want data from and choose the data product(s) they are interested in, then add this to their cart and 'checkout', where the data is externally retrieved and then processed through the MATLAB data product framework and returned to the user as a compressed file, with a search ID associated with the parameters the user specified. The main function in this process that every type of search uses, *oncmatlabsearch*, calls a wide variety of functions designed to process data from different devices – hydrophones, acoustic doppler current profilers, spectrometers, and much more. One such device, the seismometer, is helpful for detecting earthquakes and other sources of vibrations by recording the motion of the ground. ONC has multiple undersea cabled seismometers in various frequency bands (broadband, short period, etc.), and even seismometers that only specifically measure dynamic accelerating forces, known as an accelerometer. During this term, a sizeable amount of the work completed was improving support for seismometer data products; specifically, the Maris accelerometers, along with the associated documentation.

II. Summary

There were several main points of improvement with the current state of the seismometer DPs: the availability and improvement of seismometer and accelerometer plots in data preview, adding support for non-cardinal directions and accelerometer channel options, and adjusting the acceleration scale factor for the seismometer device attribute.

On the ONC website, one can get a sense of what the data plots for a specific device will look like by going to data preview. For the Maris accelerometers, no plots were generated due to the data product option for the default device parameter being set to select all velocity channel data for a given device, of which an accelerometer would have none. Initially, an override was set in the MATLAB code to include all acceleration channels only for Maris data preview, but the code that sets the data preview options for device types was later changed to select all channels (both velocity and acceleration) so that this override was not needed to help keep the code more streamlined. In addition, the acceleration plots in data preview were showing very little activity, even for time ranges that supposedly capture earthquake events (Figure 1), which suggested that there was a scaling issue with the device attribute associated with the seismometers/accelerometers. To fix this, a test search ID was created and locally run in *oncmatlabsearch*, with this scaling factor manually overridden and adjusted to the proper sensitivity (Figure 2). The device type attribute associated with the Maris accelerometer was then updated in the database to match the adjusted value. A separate adjustment was also made for the Guralp broadband seismometer devices.

While acceleration channels are available for select seismometers through the 'all channels' DP option, for accelerometers it didn't make sense to only be able to select this data along with every other channel and not be able to specify select frequencies or orientations. This prompted improved access to acceleration channels by adding new form fields in the seismometer channel form section in data search (Figure 3). The new channel options included combinations of the different sample rates and orientations available. All velocity, acceleration, and other instrument type data from seismometer devices on the ONC website is retrieved from the Incorporated Research Institutions for Seismology (IRIS) website[3] for the NEPTUNE Canada network, where the metadata can be used to check which channels are available for a device. From this we were able to determine that the Maris accelerometer has two main bands – broadband H and long period M – for its sample rates, and that its data is typically available in the non-cardinal 1/2/3 orientations, rather than the standard N/E/Z (North/East/Vertical). During this process, a broadband C channel was also found for Maris and implemented in the MATLAB code. However, the acceleration scale factor set for the other accelerometer channels was far too sensitive for this new channel and resulted in incredibly noisy data (Figure 4), so a new device attribute was created with a larger scale factor specifically set for the C channel data to produce a more reasonable result (Figure 5).

Other minor improvements included adding a new 'all velocity components' form field option in data search, as well as more informative titles to the seismometer plots, which translates the channel codes and orientations into a more human-readable format. A 'no filter' override was also added for low-rate velocity or acceleration channels in data preview due to the plots showing very noisy data.

To more comprehensively check the stability of these improvements, searches whose parameters overlap with these changes were created in data search. These are known as test cases and are added to a list of many other search ID's created for new data product code changes. These searches are run every night automatically in search automation as code is continually updated.

III. Discussion

As a physics and astronomy student working in backend software development, there have been many opportunities to learn new concepts and techniques, as well as continuing to develop interpersonal skills in a team-oriented environment.

During this work term, programming skills were developed using MATLAB, PostgreSQL, and Git Bash. Having started with a rudimentary understanding of MATLAB, the fundamentals were learned through hands-on practice and online self-paced courses, including using built-in functions and constants, the built-in debugging tool, array manipulation, regexp to search for character patterns, and much more. Further programming techniques in this language will continue to be developed in the second work term. SQL had been worked with minimally in the past to make basic astronomy database queries, but conceptual and technical skills had not been properly built. Using pgAdmin – a graphical user interface tool used to interact with the database management system PostgreSQL – required a moderate understanding of how a database works. Typically, a database is structured in rows and columns that map to a table or multiple different tables[5], not unlike how ONC's database is set up. A query can be made by selecting specific (or all) columns from a table to search for information in the database. For example, if one wanted to see which channel selections are available for the seismometer channel DP option in data search (knowing the form field ID), the following query could be made:

$$select\ *\ from\ formfield\ where\ formfieldid\ =\ 129$$

and then the user could look under the display info column to see the options. Or if the ID is not known, one could make a general search in the form field table (not specifying the form field ID) to see which columns are available, then make a refined search using keywords:

$$select\ *\ from\ formfield\ where\ description\ ilike\ '\%seismometer\%'$$

and obtain the appropriate form field ID from there. Apart from querying, PostgreSQL is also useful for manipulating and defining data (such adding a new device type attribute for the C channel acceleration scale factor previously mentioned in the summary section of this report). Tying together the work done in pgAdmin and MATLAB, a pull request (PR) can then be made to add the code changes to the DP repository, where all the files designed to process data from different devices is stored and available to view. A PR is made to request a change in the remote repository to one (or several) of these files, which can be approved after undergoing review. One method of making a PR is through Git Bash, a command line application that can be used to add, commit, and push local changes to a remote repository. During this work term, a more thorough understanding of merging and rebasing branches, editing and deleting commits, and the many other trials and tribulations of using Git Bash, was obtained.

Some of the other skills that were built upon include problem solving, such as being able to debug code that was written by others. Learning how to effectively read through function comments, understand the purpose of the input/output variables, see how the main function connects to other built-in functions within it, step through the code to the point where it broke and trace it back to the cause of the issue, and then think of a succinct and comprehensive solution, are invaluable skills to have as a developer. Another concept I became familiar with during this work term were daily scrum meetings, where the work completed in the previous day and the plan for the current day are effectively summarized by each team member in a group.

IV. Future Considerations

Current plans for work in the next term will likely be several smaller DP improvements across different devices, along with keeping on top of bug fixes in data search. Larger changes to the seismometer DPs might still occur, such as adding functionality for tiltmeter channel data for certain seismometers. In addition, prototype code for new hydrophone DPs is being created in Python, and after that has been refined a bit further, it will likely be converted over to MATLAB code.

V. References

[1] Ocean Networks Canada, *About ONC*, accessed 18 August 2022. *Online.* Available: https://www.oceannetworks.ca/about-onc/

[2] Wikipedia, MATLAB, accessed 18 August 2022. *Online*. Available: https://en.wikipedia.org/wiki/MATLAB

[3] IRIS, *MetaData Aggregator NV Network* Summary, accessed 18 August 2022. *Online*. Available: https://ds.iris.edu/mda/NV/

[4] Government of Canada, *Earthquakes Canada*, accessed 19 August 2022. *Online*. Available: https://earthquakescanada.nrcan.gc.ca/index-en.php

[5] Oracle, *What Is a Database?*, accessed 19 August 2022. *Online*. Available: https://www.oracle.com/ca-en/database/what-is-database/

VI. Appendix

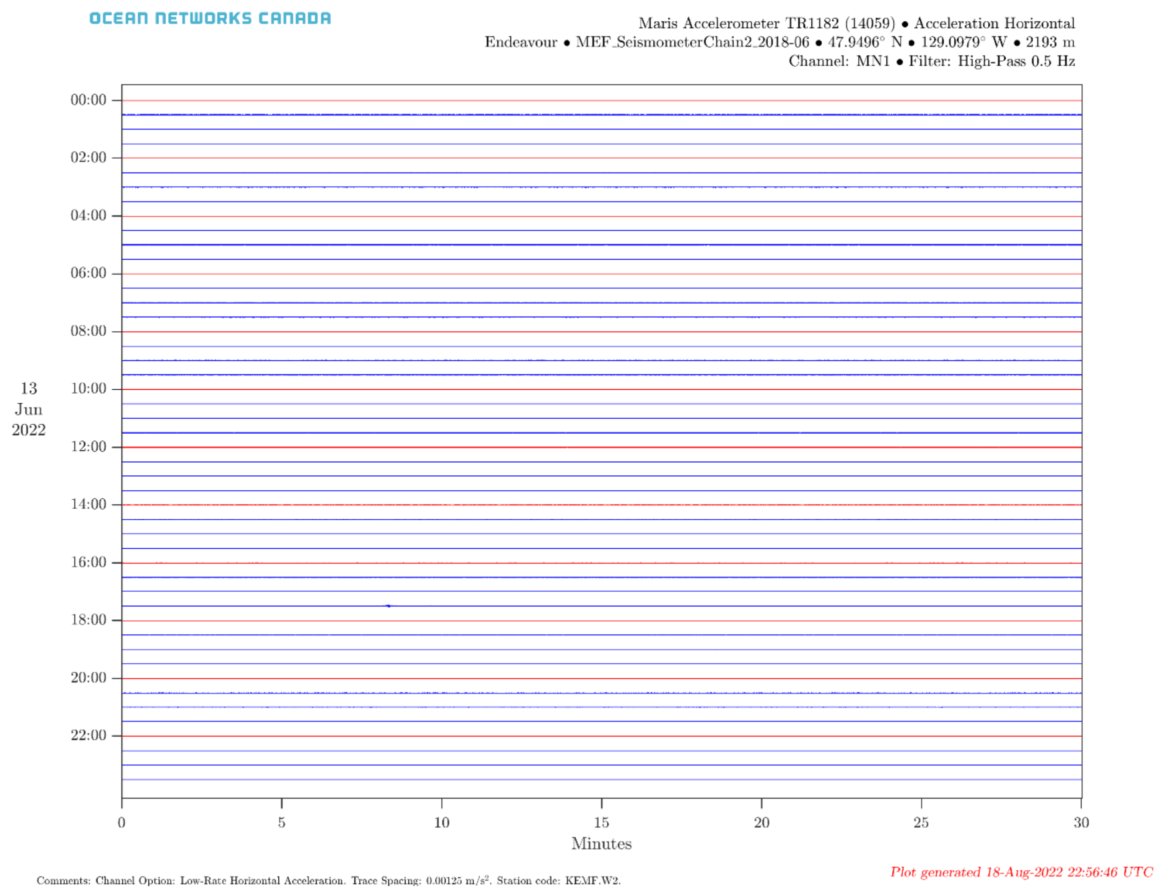

Figure 1. Example of low-rate horizontal acceleration data before the scale factor was adjusted. Around 5:40pm there is a small blip, which is the 5.1 magnitude earthquake off the coast of Port Hardy, BC, found via the Earthquake Canada Database[4].

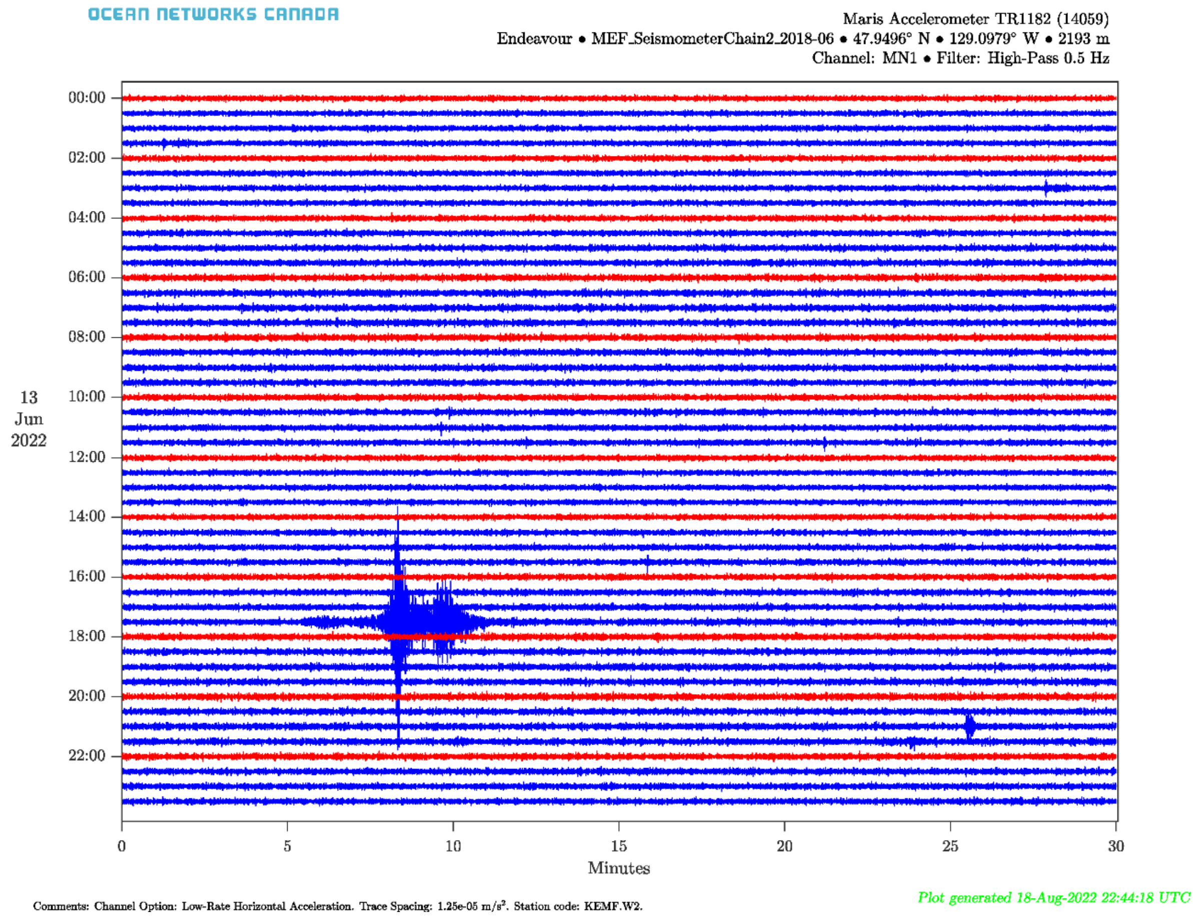

Figure 2. Example of low-rate horizontal acceleration plot after the scale factor was adjusted. Most noticeable in the plot is a 5.1 magnitude earthquake off the coast of Port Hardy, BC found via the Earthquake Canada Database[4]. It is much more prominent than in Figure 1, with other small tremors at earlier and later times of day also noticeable.

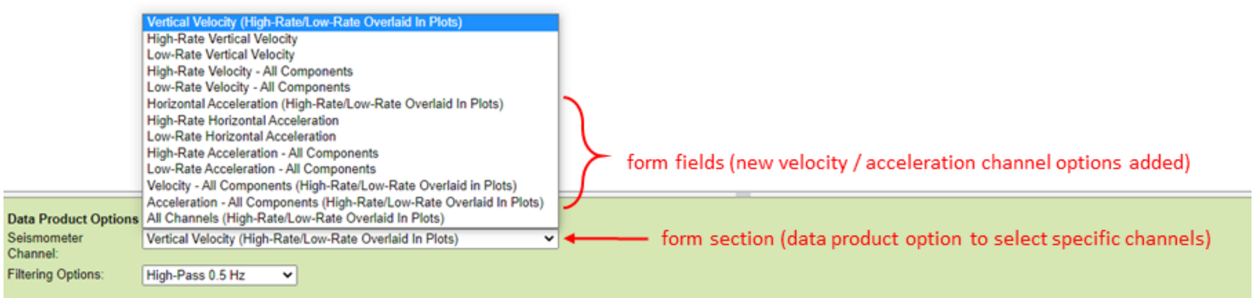

Figure 3. Demonstrating how multiple form fields can map to a single form section for the new acceleration and velocity channel options added.

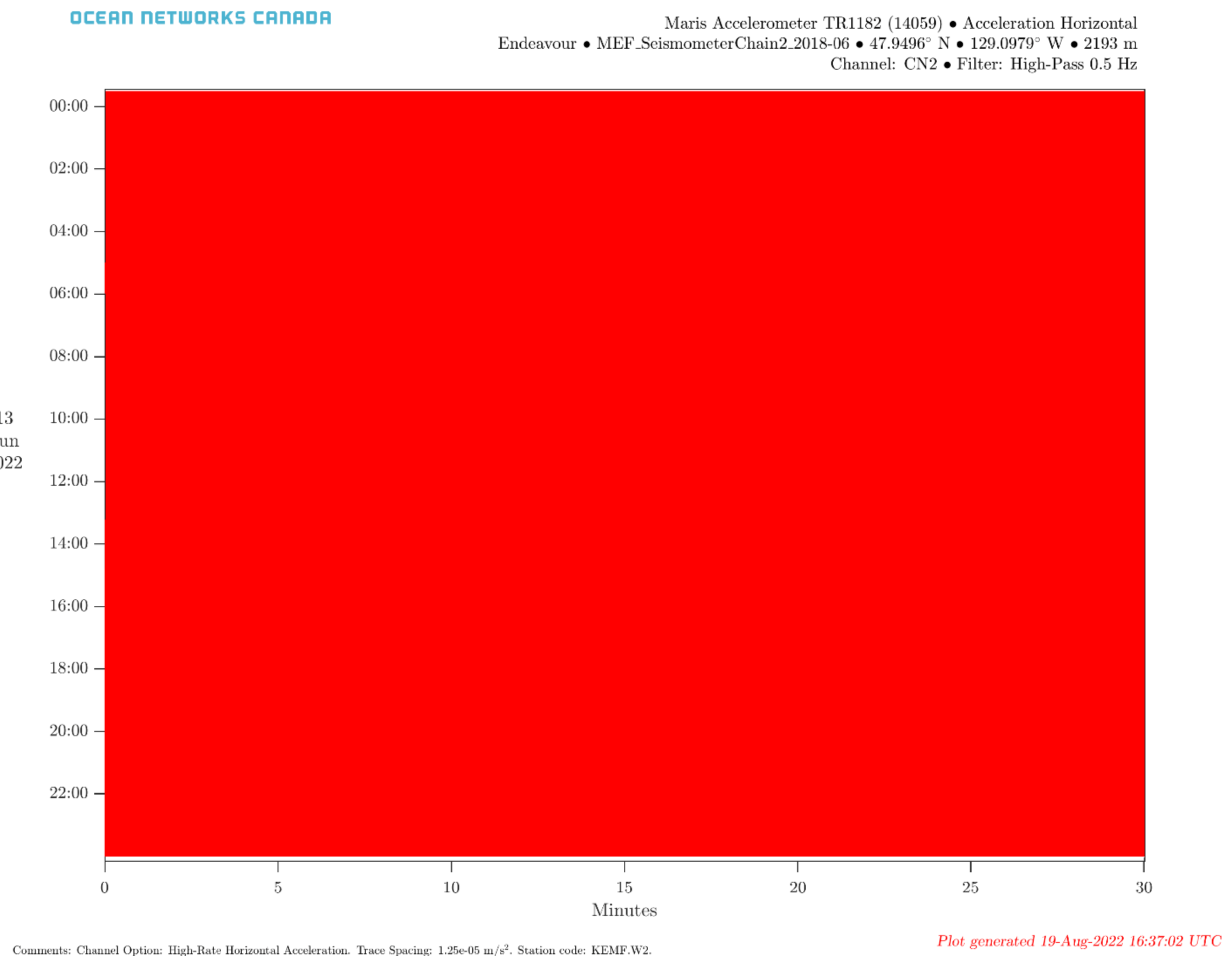

Figure 4. Example of high-rate horizontal acceleration C-channel data before a separate scale factor was created for this channel. As seen, the noise level is significantly greater than the line spacing in the plot, and so we get a red plot of death.

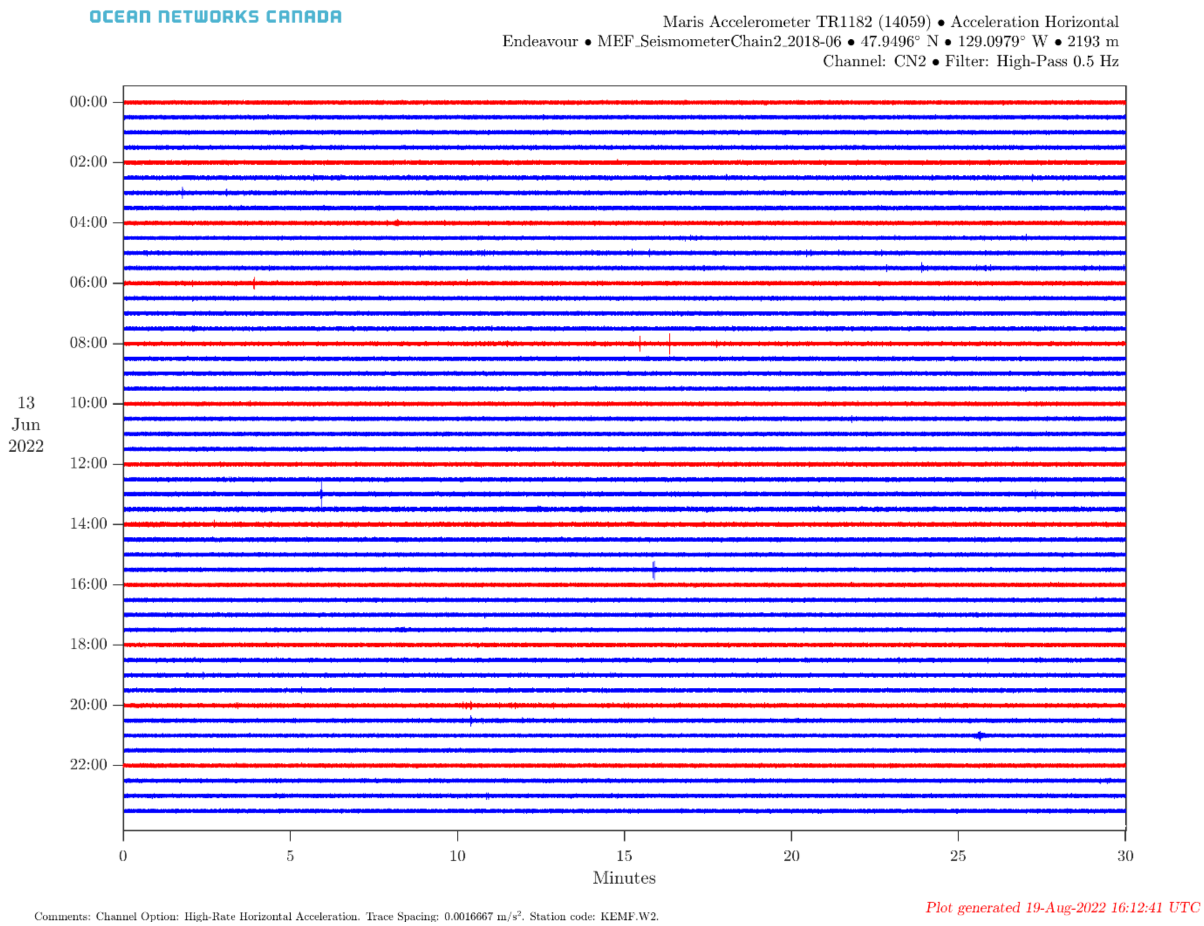

Figure 5. Example of high-rate horizontal acceleration C-channel data after a separate scale factor was created for this channel.