SANDIA REPORT
SAND2020-5170
Printed Click to enter a date

Sandia National Laboratories

Applying Waveform Correlation to Mining Blasts Using a Global Sparse Network

Amy Sundermier
Rigobert Tibi
Christopher J. Young

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

Agencies that monitor for underground nuclear tests are interested in techniques that automatically characterize mining blasts to reduce the human analyst effort required to produce high-quality event bulletins. Waveform correlation is effective in finding similar waveforms from repeating seismic events, including mining blasts. We report the results of an experiment that uses waveform templates recorded by multiple International Monitoring System stations of the Comprehensive Nuclear-Test-Ban Treaty for up to 10 years prior to detect and identify mining blasts that occur during single weeks of study. We discuss approaches for template selection, threshold setting, and event detection that are specialized for mining blasts and a sparse, global network. We apply the approaches to two different weeks of study for each of two geographic regions, Wyoming and Scandinavia, to evaluate the potential for establishing a set of standards for waveform correlation processing of mining blasts that can be effective for operational monitoring systems with a sparse network. We compare candidate events detected with our processing methods to the Reviewed Event Bulletin of the International Data Centre to develop an intuition about potential reduction in analyst workload.

ACKNOWLEDGEMENTS

This research was initiated by members of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) and experts that attended the October 2018 CTBTO Expert Meeting on Advances in Waveform Processing and Special Studies in Vienna, Austria. The following people attended the initial project launch meeting:

Helmuth Breitenfellner, Steven J. Gibbons, Ivan Kitov, Tormod Kvaerna, Ronan Le Bras, Christos Saragiotis, Megan Slinkard, and Amy Sundermier.

Ronald A. (Chip) Brogan reviewed SeisCorr template libraries and correlation jobs and provided advice on filter bands for signal enhancement.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page left blank

ACRONYMS AND DEFINITIONS

Abbreviation	Definition
CSS	Center for Seismic Studies
CTBT	Comprehensive Nuclear-Test-Ban Treaty
CTBTO	Comprehensive Nuclear-Test-Ban Treaty Organization
EVID	Event Identifier
IDC	International Data Centre
IMS	International Monitoring System
LEB	Late Event Bulletin
LTA	Long-Term Average
NDEF	In the CSS standard, this is defined as the number of locating phases. We use this term here to mean the number of detecting templates in a multistation validation job.
REB	Reviewed Event Bulletin
SEL	Standard Event List
SNL	Sandia National Laboratories
SNR	Signal-to-Noise Ratio
STA	Short-Term Average
STA/LTA	Ratio of Short-Term Average over Long-Term Average

1. INTRODUCTION

Several studies have shown that waveform correlation is effective in detecting similar waveforms from recurring earthquakes and other seismic events, including mining blasts. Mining blasts increase the number of seismic events that are detected on global networks such as the International Monitoring System (IMS), which causes increased analyst workload, even though the mining blasts are not of interest and are considered nuisance events by underground nuclear test monitoring agencies. For that reason, monitoring agencies have shown interest in adopting techniques such as waveform correlation to quickly characterize recurring mining blast events to reduce the amount of effort required by analysts to produce a high-quality event bulletin. Members of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) invited several experts familiar with waveform correlation methods to participate in a study of mining regions that are particularly problematic for the International Data Centre (IDC). The goal of the study was to reduce analyst workload in monitoring system pipelines due to mining blasts. The study was conducted in parallel by different researchers using the same time periods and geographic regions; the CTBTO plans to evaluate the experimental results as part of a prototype pipeline.

This report describes the research conducted at Sandia National Laboratories (SNL) on the mining regions identified by the CTBTO members. The authors used SeisCorr, a software system developed at SNL for waveform correlation event detection that has previously been used for studies of aftershock sequences [1][2][3] and general regional seismicity, including mining regions [4]. For each region, the IDC designated three IMS arrays and two specific week-long timeframes as part of the experimental requirements. Using the SeisCorr system, we create waveform templates from arrivals at the three IMS arrays for historical events in the Reviewed Event Bulletin (REB) to detect and identify events that occur during the study weeks. No limit was set upon the use of historical events, so we used all available arrival waveforms in our waveform archive up to 10 years prior to the week of study for the designated IMS arrays. We did not have 10 years of waveform data in our archive for every array, and for those arrays with less data we used all available waveform history.

Using the SeisCorr system, we customized methods for template selection, threshold setting, and event detection for the mining regions using sparse, global networks. The specialized methods were applied to two separate weeks in the year 2018 for each of two geographic regions, Wyoming (USA) and Scandinavia, to evaluate the potential for establishing a set of standards for mining blast correlation processing that can be effective for operational monitoring systems that use a sparse network such as the IMS. To evaluate the effectiveness of the methods, the bulletin of candidate events detected with our specialized methods are compared to the Reviewed Event Bulletin (REB) to develop a well-informed intuition about potential reduction in analyst workload.

2. EXPERIMENT METHODS AND RESULTS

The research described in this report applies waveform correlation technique to mining regions, and will hereafter be referred to as the mining blast study. The basic experimental design was created by the IDC staff, but the application of SNL's SeisCorr tool required some additional design decisions and technical refinements that are described in detail in this report.

Section 2.1 describes the two geographical regions and timeframes that were chosen by the CTBTO for the study. Section 2.2 briefly describes SNL's application for seismic waveform correlation, including the features of the SeisCorr software. Sections 2.3 through 2.5 provide details about the mining blast study and readers that are familiar with waveform correlation and the mining blast study may choose to begin reading with those sections. Section 2.3 describes the selection of waveform templates and template threshold setting processes to prepare for waveform correlation of each geographical region. Section 2.4 provides an overview of the correlation processing for the mining blast study. Section 2.5 provides an overview on how SeisCorr organizes correlation detections into candidate events and describes the study results.

2.1. Experimental Setup

The CTBTO chose four weeks for study, two weeks for each of the two mining regions that are problematic for the IDC. Mining blasts are nuisance events that raise the event rate and increase the analyst workload to produce the REB, but otherwise are not interesting events for the IDC's mission to monitor for underground nuclear testing. The CTBTO set the geographical boundaries and temporal span for each mining region for the study and provided guidance to use any historical waveform templates and/or waveform templates to detect events that occurred later in time; in other words, to provide results useful for an operational system, template waveforms must chronologically precede the data that are processed for detections. The research described in this report was based on information published by the IDC to a collaboration website after the October 2018 CTBTO Expert Meeting on Advances in Waveform Processing and Special Studies. The collaboration website is access controlled and maintained by the CTBTO, and study participants were provided accounts and access to experimental design requirements for the aftershock [3] and mining blast studies. The maps and geographical regions of the mining blast study were obtained from this collaboration website and are included in this report for the convenience of the reader.

NOTE: To aid in proposing follow-up research, the report includes notes such as this one with descriptions of additional work if more funding becomes available, as well as issues that were discovered during the study.

The two mining regions and two 1-week study periods per region identified by the CTBTO are listed in Table 2-1. The geographical extent of the mining region specifies the location of events that may be used to create historical waveform templates and bounds the area for comparing detected events with REB events. The geographical extents for both the Wyoming, USA and Scandinavia regions are given as a box bounded by latitude/longitude minimum and maximum values. The temporal span is one week for each of four periods of study.

Table 2-1. Mining blast geographical and temporal extents identified by CTBTO.

Region	Temporal Span (UTC)	Temporal span (JDATE)	Geographical Extent (Lat, Lon)
Wyoming, USA	2018-APR-04 to 2018-APR-10	Week 1: 2018094 to 2018100	40°-- 46°N, 110°--100°W
	2018-AUG-08 to 2018-AUG-14	Week 2: 2018220 to 2018226	
Scandinavia	2018-FEB-12 to 2018-FEB-18	Week 1: 2018043 to 2018049	62°--72°N, 16°--37°E
	2018-JUL-06 to 2018-JUL-12	Week 2: 2018187 to 2018193	

The CTBTO provided maps with the event locations for the two mining regions on the collaboration website, which are shown in Figure 2-1 for the convenience of the reader.

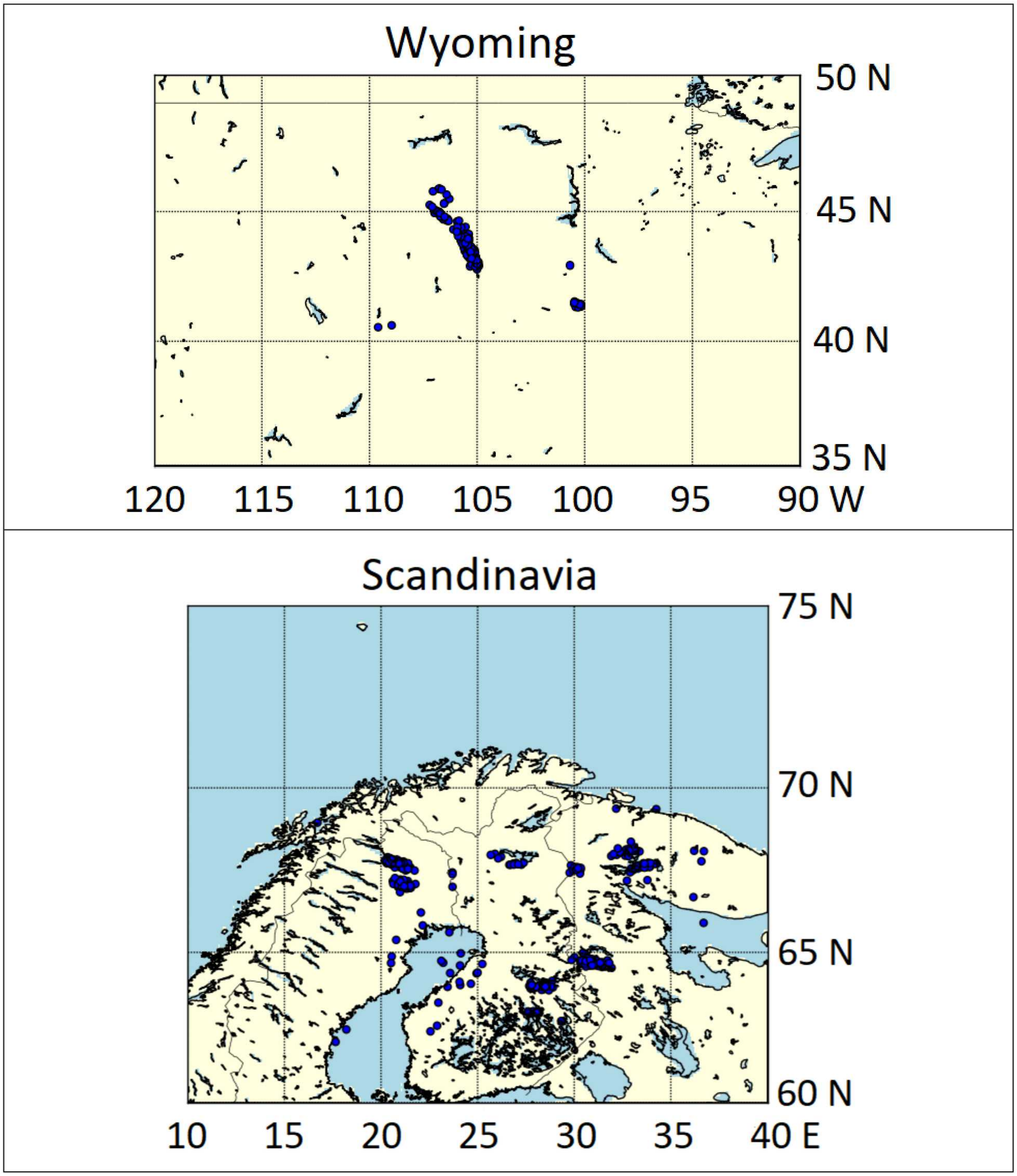

Figure 2-1. CTBTO maps of the two mining regions, showing Wyoming, USA above and Scandinavia below. Blue dots represent REB event locations.

The guidance provided by the CTBTO allowed template waveforms to be created from any events published in the REB bulletin that preceded the timeframe of study. We chose historical REB events to serve as template events that were within the geographical boundaries in Table 2-1 and preceded the start of the week of study by as much as 10 years. We expect that template events from the REB are more likely to have accurate locations and associations than the automated pipeline SEL events because they have been reviewed by analysts. Location accuracy of template events is important because a waveform correlation detection indicates that the unknown event and the template event are co-located or closely located; poorly located template events will result in poorly assigned location for the detections.

It should be noted that the region of study encloses events from multiple source types (e.g., tectonic events, and other anthropogenic sources) that are not desired as templates for the mining blasts. There is no easy way to differentiate source types in the event bulletins when choosing events for mining blast templates, but fortunately for this study we were not required to differentiate source types, so the study includes detection of all seismic events within the geographical and temporal boundaries.

The method chosen for this research is visually described by the timeline shown in Figure 2-2. Templates are created from the REB from up to 10 years of historical events, then the templates are correlated against the one week of continuous data to detect events within the region. A set of template libraries for each week of study was created; thus, although there is substantial overlap between the templates for the two weeks for each region, the libraries are not identical and were created independently.

Figure 2-2. Timeline for template building and event detections.

The method was chosen to be compatible with existing operational pipelines. We desire to limit the number of additional steps that must be added to the operational pipeline, particularly any steps that require human intervention. For example, we use automated methods to generate templates from REB arrivals that do not require hand-picking template waveforms.

Guidance was provided at the project launch meeting that the goal of the research is not to find more events than are listed in the REB, but rather to use waveform correlation to detect events that are included in the REB that are currently built manually or substantially changed by IDC analysts. The main focus of this research project is to find out if waveform correlation can reduce the workload on analysts for mining regions by augmenting the events detected by the existing automatic IDC pipeline.

2.2. Overview of SeisCorr Software

The study was conducted using the SeisCorr software for waveform correlation, developed at SNL. It is important to recognize that SeisCorr has not been implemented for use in a real-time system. It was originally developed as a research tool to rapidly process continuous waveform data recorded by a single station. An additional capability to screen out lower-quality events by performing multistation validation of individual station results was added later. While SeisCorr has an interactive user interface, it is meant to process large amounts of data without human review of the templates prior to their use in correlation jobs.

SeisCorr has been used for regional studies that included mining regions, which demonstrated that waveform correlation works well to detect repeating mining blasts [4]. That study involved IMS arrays in Kazakhstan (BVAR, KURK, MKAR).

SeisCorr supports three major activities for waveform correlation research: 1) template preparation; 2) correlation of template waveforms with continuous waveform data to detect possible events; and 3) candidate event creation from multistation validation. The following subsections provide a high-level overview of these three SeisCorr activities, while specific methods and parameters chosen for this study for the templates and multistation validation are described later in the report.

SeisCorr uses the Center for Seismic Studies (CSS) version 3 database schema [5] for the Origin, Arrival, and Association tables. SeisCorr contains a specialized schema, not described in this report, of database tables for:

- waveform correlation
  - templates, and
  - template matches (i.e., detections), and
- multistation validation
  - candidate origins, and
  - candidate associations.

2.2.1. Template Preparation

SeisCorr contains functionality for the creation of a template library for a given station based on a query of the database arrival table. For the mining blasts study, arrivals were sought for events located within the geographic boundary of the region that occurred from up to 10 years prior to the week of study. For the mining blast study, we chose to use the Pn and Pg or the Lg phase from REB arrivals depending upon the station's distance from the region; the SeisCorr user queries for labeled arrivals of the desired phases at the chosen station. SeisCorr allows the user to filter the waveforms associated with each arrival using a bandpass filter chosen to work well for the epicentral distance of the recording station. Then the SeisCorr user may screen the filtered waveforms based on a specified STA/LTA threshold to eliminate waveforms with low signal-to-noise ratio (SNR), resulting in a set of candidate templates. The candidate templates can be saved as a template library for the station. We use the term "candidate template" to recognize that even after the SNR-screening step, not all arrival waveforms will make acceptable templates for correlation; additional screening steps may be applied to eliminate poor quality templates.

Metadata about the template waveforms are stored in a database table that contains columns such as ORID and ARID that allow the user to discover metadata about the template's corresponding event, such as location, magnitude, phase, and so on.

Prior waveform correlation research at SNL demonstrated the importance of choosing a correlation coefficient threshold that is dependent on the characteristics of the template [4]. In theory, the distribution of correlation values obtained from correlating a template with a continuous data stream can be thought of as the sum of 2 distributions: 1) correlation of the template with noise, and 2) correlation with similar events. SeisCorr implements a sophisticated algorithm for setting template thresholds based upon a false alarm rate obtained from correlation with noise for each correlation threshold. Ordinarily, unless similar events can be identified and cut out first, correlating a template with continuous data will generate both noise and signal correlations. However, time-reversing the template ensures the template has the same time-bandwidth product as the template proper and should yield a very similar distribution of correlation values with noise windows, but also not correlate with similar events even if they are present. Thus, a time-reversed template can be correlated with continuous data without first screening out time windows with similar events to generate a robust distribution of noise correlation values. Using this method, thresholds can be set individually for each template so as to achieve a consistent false alarm rate (FAR) based on noise correlations.

A recent study [6] has shown that for a template waveform with a low time-bandwidth product (e.g. teleseismic P) a threshold based solely on the noise distribution may be too low, because correlation with signals from some non-similar events do not fit within the noise distribution, nor are they part of the target signal distribution. Fortunately, the templates created for the mining blast study were obtained from stations at regional distances and had long time windows ranging from 50s to 160s; thus, we expect a time-bandwidth product high enough to separate the signal characteristics from noise. The template thresholds were set by the time-reverse method implemented in SeisCorr such that the duration of the week-long study period was processed to determine the noise correlation thresholds, i.e. we intentionally chose a time period with many events in it. Thus, the thresholds that were set for this study were based not only on correlation with noise but also with the signals of the numerous mining blasts in the continuous data. More research is needed to develop an optimal approach to setting thresholds for template waveforms for mining regions since the approach used in this study must be modified for an operational system; our approach violated the antecedent constraint by using waveform data throughout the duration of the study period to set template thresholds to detect later events. However, the template threshold statistics are similar across the two study weeks for each region (e.g., Figure 2-9 and Figure 2-15) so it may be as simple as using thresholds set from a week or month prior.

There are several steps required to create a template library in SeisCorr using an event bulletin as a source for template events. This section will describe the steps in general, and later sections will discuss the details for each mining region.

The general steps to create a template library in SeisCorr using a bulletin are:

1. Choose event origins from existing bulletin (e.g., REB Origin table from the IDC). This step creates a set of origin candidates for templates, and for convenience these origins may be stored in an ORIGIN table based on the CSS schema that is specific to the mining region and study period.
2. Create a table from a query of candidate template associations for the origins found in step 1 (i.e., an ASSOC table from the CSS schema).
3. Create a table from a query of candidate template arrivals for the origins and associations found in steps 1 and 2 (i.e., an ARRIVAL table from the CSS schema).
4. At this point in a waveform correlation study, we would typically choose stations for which we want template libraries. In the case of the mining blast study, the CTBTO specified the use of

three stations for each region; PDAR, NVAR and TXAR for Wyoming, and ARCES, FINES, and NOA for Scandinavia. For each station, follow steps 5 through 10 to create a template library for that station.

5. Use SeisCorr to search for arrivals at a station, using the candidate template events and candidate template associations, specifying the phase list and time boxing the arrival time.
6. Select a template window size and offset from the picked arrival time. The template window size will be the same for the entire library in the current version of the software.
7. Choose a filter band based on the epicentral distance and filter the templates to enhance the signal. The filter band will be the same for the entire library in the current version of the software.
8. Run an STA/LTA detector on all the templates in the library. Select an STA/LTA threshold, and discard candidate templates that do not meet the threshold. This step is optional if all the templates appear to have a good SNR.
9. Save candidate templates for the station as a library.
10. Set correlation coefficient threshold using the time-reverse method discussed earlier.

2.2.2. Correlation of Templates with Continuous Data

SeisCorr correlates each template from a template library across continuous waveform data for a specified time period. If the correlation score exceeds the template cross-correlation coefficient threshold, a single-station detection is recorded. Note that a separate correlation job must be run for every template library, and each template library represents waveforms from a single station. For 3 stations, for example, 3 correlation jobs would be required.

As previously mentioned, SeisCorr was implemented for historical data, not designed for use in a real-time system. A real-time waveform correlation system will have to process templates from many stations simultaneously over incoming waveform data; in contrast, SeisCorr is optimized to process many templates from a single station rapidly over historical waveform data. At the project launch meeting the decision was made to bypass the incorporation of SeisCorr algorithms into the real-time pipeline by defining a set of input and output database tables that can act as substitute steps for an algorithm. This approach allows us to use the interactive SeisCorr system on IDC data to develop template libraries, set thresholds and run correlations at SNL in our normal research environment, then deliver the template matches as an output database table that is analogous to the database arrival table for later use in the prototype pipeline being developed by the CTBTO.

2.2.3. Multistation Validation and Candidate Events

SeisCorr declares a detection when the correlation score from the correlation of the template with a waveform exceeds the template correlation coefficient threshold. A list of detections is produced as the template slides along the continuous data stream. We have found that waveform correlation finds many detections that cannot be validated with events in published bulletins. Often the events producing those detections are real, though their magnitudes are below the thresholds captured in the bulletins, but some of the detected events are false, despite the effort we put into screening templates and properly setting correlation thresholds for each template. Thus, SeisCorr includes an additional feature called multistation validation [4] that compares correlation detections across two or more stations to look for detections that are -- in terms of location and time -- consistent with the same causal event. The location of a detection is assumed to be the location of the template event (i.e., the event from which the template waveform was recorded), while the origin time is calculated by subtracting the calculated travel time from the template event to the station. For multistation validation, SeisCorr clusters the detections from each station to improve the estimates of location and origin time for the common events. The SeisCorr user chooses distance and time tolerances to cluster the detections. Once a cluster of detections have been aggregated optimally relative to other clusters, SeisCorr calculates a location and origin time for the candidate event. The candidate events are delivered in the schema of the Origin table for later use in the prototype pipeline.

This subsection completes the overview of the features of SeisCorr. The remainder of the report will describe the specific use of SeisCorr for the mining blast study.

2.3. Template Libraries

A template library is a collection of waveform templates for a single station. For the mining blast study, the CTBTO identified three stations for each mining region for waveform correlation processing, which means that three template libraries are required for each region and timeframe. Twelve template libraries were required to complete the study (2 regions, 2 timeframes, 3 stations).

Each mining region had one station that was actually located within the geographical boundaries of the region within approximately 10 degrees or less of the mines. In the case of Wyoming, the closest station is the PDAR array. In the case of Scandinavia, the closest station is the ARCES array. The closest stations provide valuable waveform template detectors because the waveforms contain high SNR signals with a broad frequency range, resulting in templates with a high time-bandwidth product. For both regions, the other two stations are still within regional distances, which allows the use of Lg templates in the cases where Pn does not propagate with sufficient amplitude to the station.

There are many choices associated with creating a template library. The results of any waveform correlation study depend on both the waveform correlation algorithm for determining waveform similarity and greatly depend on the choice of template waveforms. For this study we chose parameters for defining template waveforms to detect mining blast events that should be useful to analysts; for example, long duration templates for regional signals guarantee a high time-bandwidth product that reduces the overall number of detections but boosts confidence in each detection. Furthermore, these template waveforms have enough complexity that they are unlikely to match signals from a non-collocated geographical location or a dissimilar source type, in contrast with template waveforms derived from teleseismic P (e.g., [6], [7]).

There were some REB events in the group of candidate template events with multiple phase picks for arrivals at the same station. The templates are long enough in duration to include more than one phase, as shown in the example template in Figure 2-3. We chose to make templates from only one phase arrival per event in order to maintain detection independence; in other words, we did not want a Pn template that included the Lg phase and also an Lg template from the same event, since detections from two such template waveforms would overlap. Our preference in this study was to make templates from Pn picks that included later phases; this automatically incorporated correlation with the differential phase travel time and the complexity of different frequency content from each phase. When Pn picks did not meet the STA/LTA threshold test, our second phase choice for the study was a template based on a picked Lg phase; Lg phases have been used successfully for detections within regional distances in prior waveform correlation studies [8][9].

Template windows of different lengths were explored for each station. The final choice for a template window length was based on including as much of the seismic waveform as possible, while excluding trailing noise. The length of waveforms from the nearest to furthest epicentral event-to-station distances were visually compared to make the best compromise for template length. This was done manually, not algorithmically. Once a template window length was chosen for each station for week 1, the same template window length was used for the station for week 2.

NOTE: At the current time the SeisCorr software only supports the choice of one template window length per library, which is a limitation that prevents optimization of template window length per template to include the entire arrival waveform but little additional trailing noise. Custom template lengths could be explored as a future research direction.

The details of the preparation of the template libraries for each region and each week of study is described in the following subsections.

2.3.1. Wyoming Week 1 (April 4–10, 2018)

The template libraries for Week 1 of the Wyoming region were drawn from up to 10 years of historical REB events within the geographical boundaries of the region. In the case of the PDAR and NVAR stations, there were only 7 years of waveform data available in the SNL archive for the same channel labels; in cases where channel labels changed, such as re-instrumentation at the stations, the older waveforms were not included.

Figure 2-3 shows an example Pn template waveform from the PDAR station from an REB event (ORID 14579863) that occurred on 2017-07-03 at 12:09 UTC. Only the waveform for element PD01 and channel SHZ is displayed in this figure even though the template includes waveforms for all elements of the array. The example template is 80s in duration and the windows starts 5s before the picked REB arrival. The filter band used for the PDAR template library is 1.0 – 5.0 Hz. The time-reverse method calculated a custom detection threshold of 0.189 for the template.

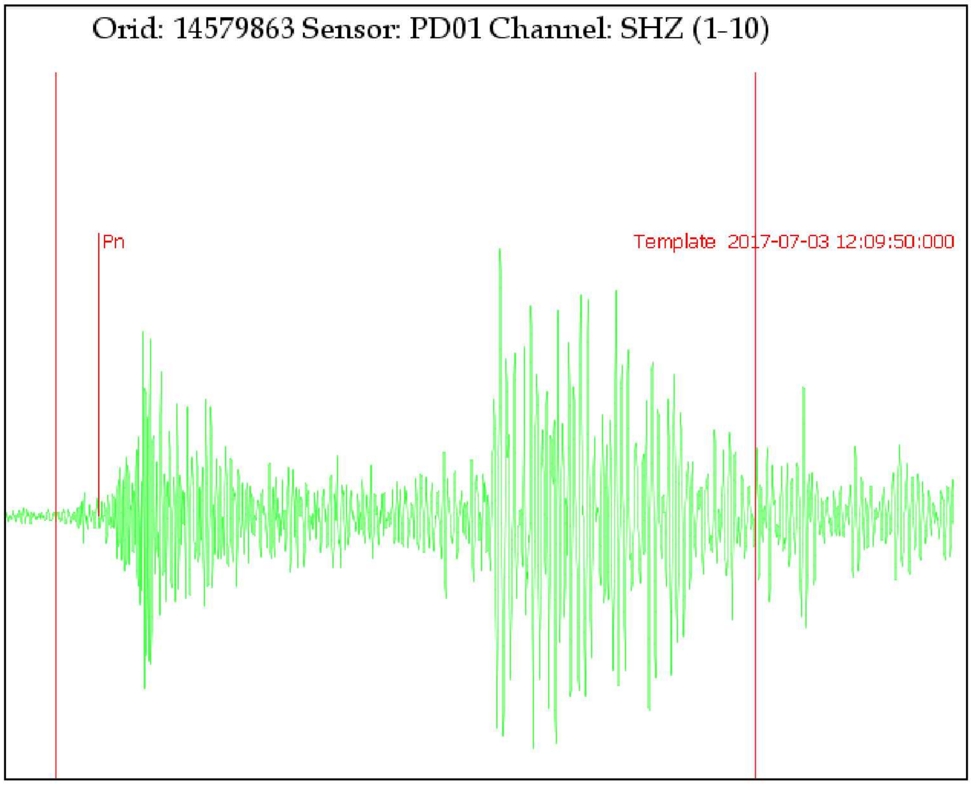

Figure 2-3. An example 80s PDAR Pn template waveform for the PD01 element. The portion of the waveform between the two vertical red lines indicates the portion of the arrival waveform that is used as the template waveform for cross-correlation with continuous data.

Table 2-2 provides an overview of the template library parameters, organized by station, for the template libraries for the Wyoming Week 1 region and timeframe. Later in this section we explain what the parameters are and how they are used for the library creation.

Table 2-2. Template Library Parameters for Wyoming Week 1

Stations	Phase	Avg SNR	Window Length (s)	Filter Band (Hz)	STA/LTA parameters (s)	STA/LTA Threshold	Template Count
PDAR	Pn	51.3	80	1.0-5.0	1/30/0	5.0	1282
NVAR	Pn	10.0	50	1.0-3.5	1/30/0	2.5	119
TXAR	Lg	5.3	60	0.5-5.0	3/30/100	2.0	41

The map in Figure 2-4 shows the station locations, network geometry, geographical boundaries of the study region, and the locations of template events. The color of the template events indicates the number of stations within the network (PDAR, NVAR, TXAR) that have templates from the event. There are many 1-station template events because mining blasts are small, so most of the time only the closest station (PDAR) of the 3 stations in this network detected the event.

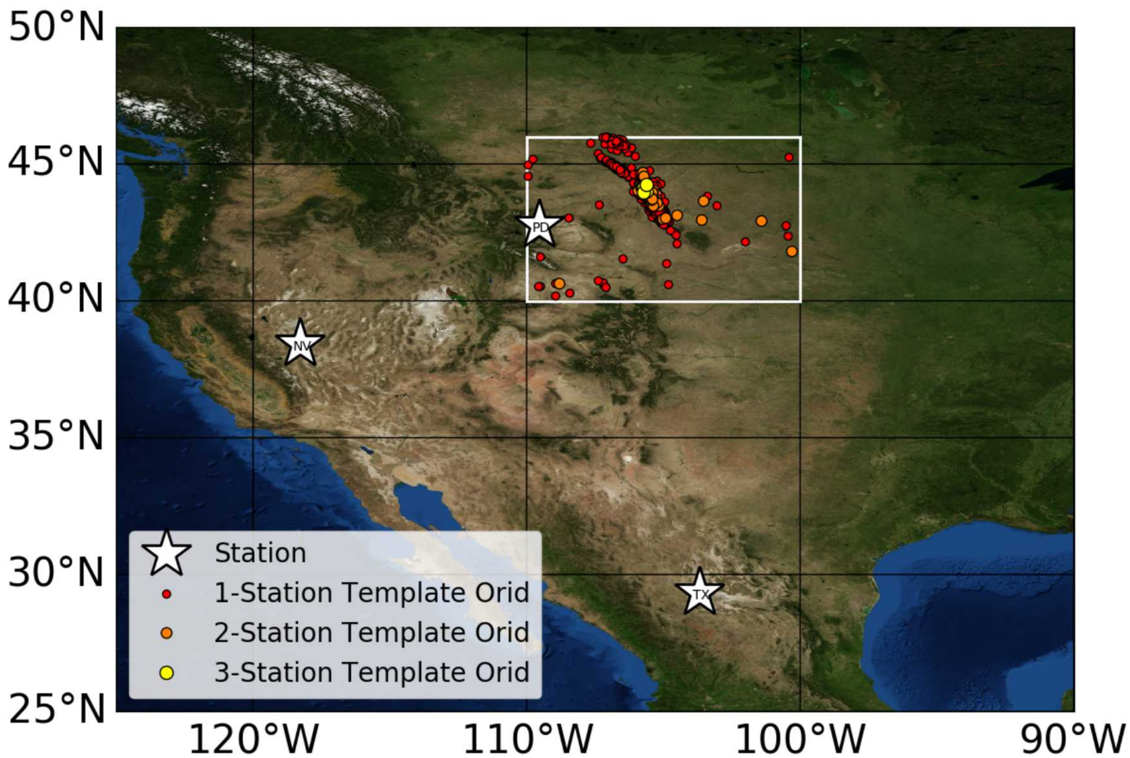

Figure 2-4. Map of the stations PDAR, NVAR, and TXAR, indicating network geometry and distance from the study region. The geographical boundaries of the study region are indicated by a white box, which encloses the Wyoming Week 1 template event locations. The color of the template events indicates the number of stations in the study that detected the template event.

The detailed sequential steps to create the template libraries for Wyoming Week 1 are listed below:

1. The REB origin table was queried for events that occurred within the geographical boundaries from Table 2-1, which are (Lat 40°-- 46°N, Lon 110°--100°W) and the timespan from 2008-04-04 to 2018-04-03. This resulted in 1303 template events with arrivals at all three stations. The locations of the template events are shown in Figure 2-5.

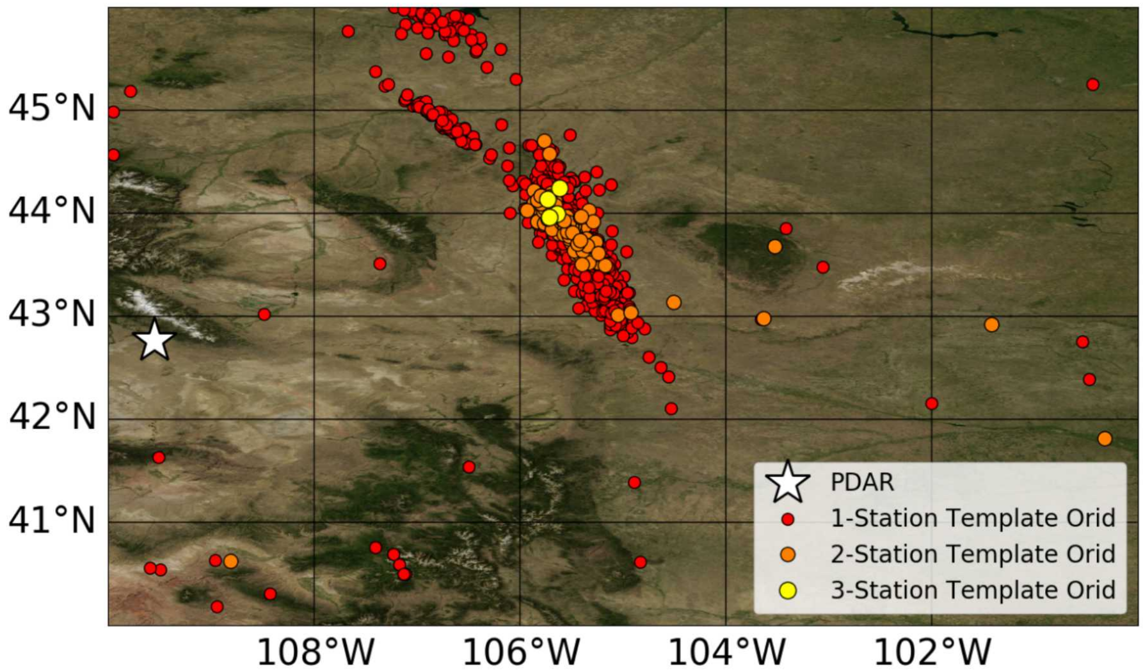

Figure 2-5. Locations of 1303 Wyoming Week 1 template events, where the color and size indicate the number of detecting stations for the event.

2. Create temporary tables with associations and arrivals for the 1303 template events. Queries joined the REB.ORIGINS with the REB.ASSOC and IDCX.ARRIVAL tables to populate the Wyoming Week 1 associations and arrivals tables. SeisCorr uses the temporary tables to identify the arrivals at the stations for the 1303 template events bounded by the study.

The steps 3-8 are performed for each station to create a template library.

3. For each station the phase arrivals were reviewed to determine what phase would make the best template library for the station. We preferred to make templates from the Pn phase pick and include most of the waveform. In the case of station TXAR, the Lg picks offered the best template waveform because the Pn arrivals from the mining region are too small and consequently buried in the noise. Figure 2-6 shows an example Lg template waveform from the TXAR station from an REB template event (ORID 14741815) on 2017-08-15 00:59. Only the waveform for array element TX01, channel SHZ, is displayed in this figure even though the template includes waveforms for all elements of the array. The example template is 60s in duration and is windowed 5s after the picked REB arrival. The filter band used for the TXAR Lg template library is 0.5 – 5.0 Hz. The time-reverse method calculated a custom template threshold of 0.2 for the template.

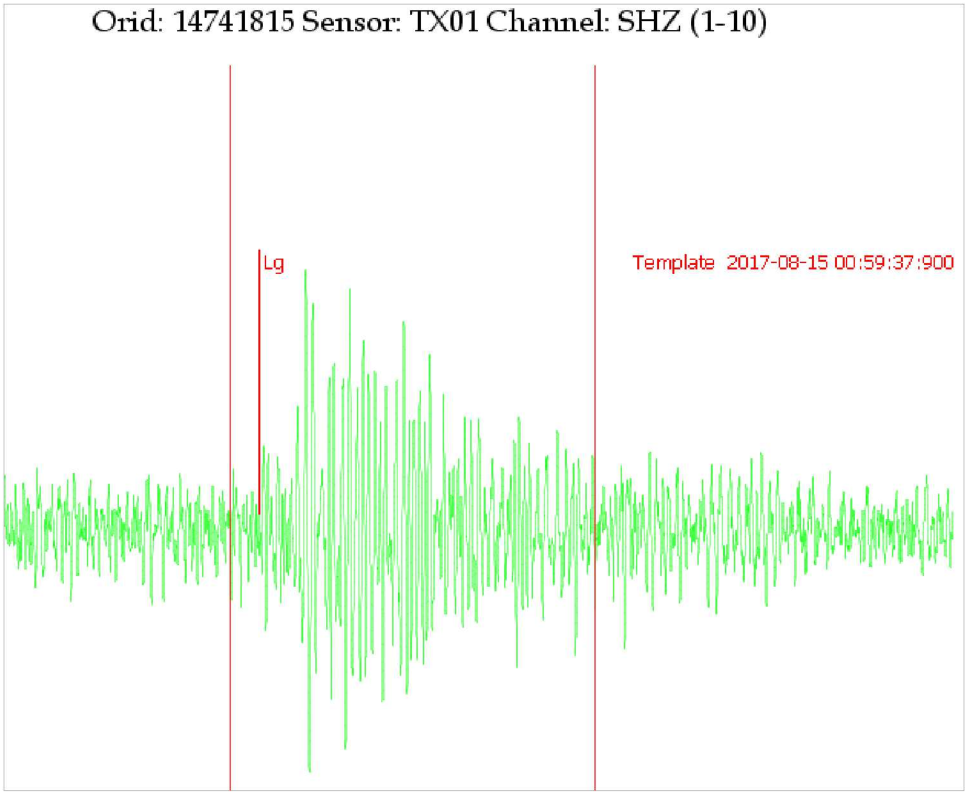

Figure 2-6. An example 60s TXAR Lg template waveform for the element TX01. TXAR templates from the mining region were made from Lg picks because Lg phase arrivals were often the only clearly identifiable signals.

4. Template windows of different lengths were explored for each station. The final choices for template window length for Wyoming Week 1 are shown in Table 2-2.
5. The templates were filtered with a 3-pole Butterworth bandpass filter. The frequency band is shown by station in Table 2-2.
6. STA/LTA threshold screening removed templates that did not have high enough SNR. The parameters for STA/LTA are listed below and are shown in a column of Table 2-2 separated by the "/" character (e.g. "1/30/0"):
   a. Short-term average window
   b. Long-term average window
   c. Gap between windows

   The minimum STA/LTA threshold for each station is also shown in a column in Table 2-2. Note that PDAR has a significantly higher threshold than NVAR or TXAR because the station is closer to the events and the waveforms have higher SNRs. Note that the gap between the STA and LTA is large (100s) for the Lg templates of TXAR so that the LTA is computed prior to the P arrival corresponding to the Lg.
7. The candidate templates that passed the STA/LTA threshold screening were saved in the template library. Figure 2-7 shows the number of templates for each station (i.e., library). The number of templates binned by age, by station, is shown by the histogram in Figure 2-8. We are interested in the performance of templates by age because a mining region changes over time. We wish to know how far back in time the historical waveforms are useful for detection because

waveform correlation is computationally expensive and there are many mining activities taking place around the world. The question we would like to answer, at least for this mine, is whether or not templates from older waveforms are useful.

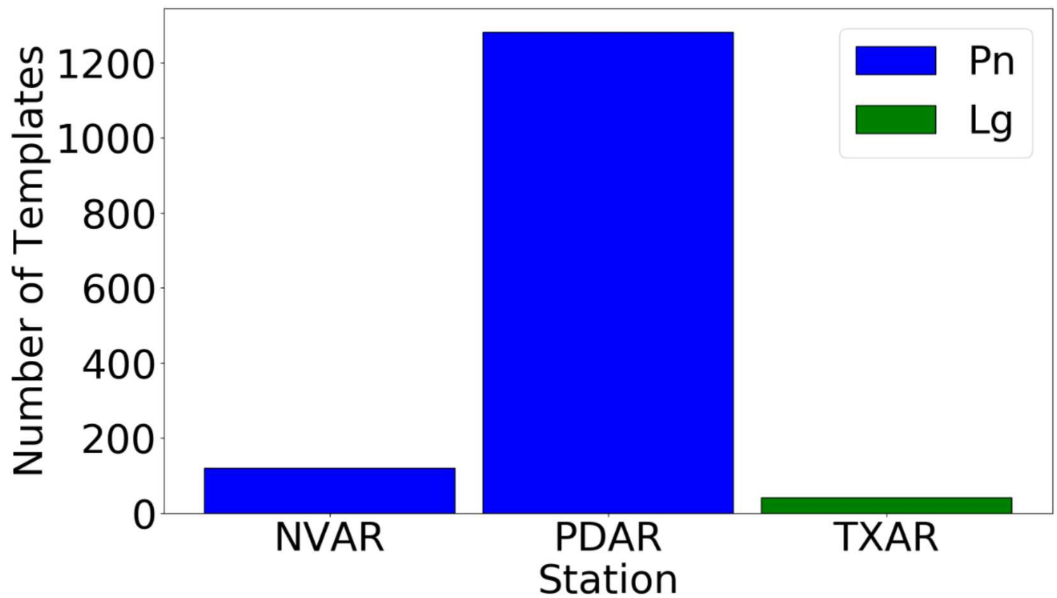

Figure 2-7. Wyoming Week 1 template count by station.

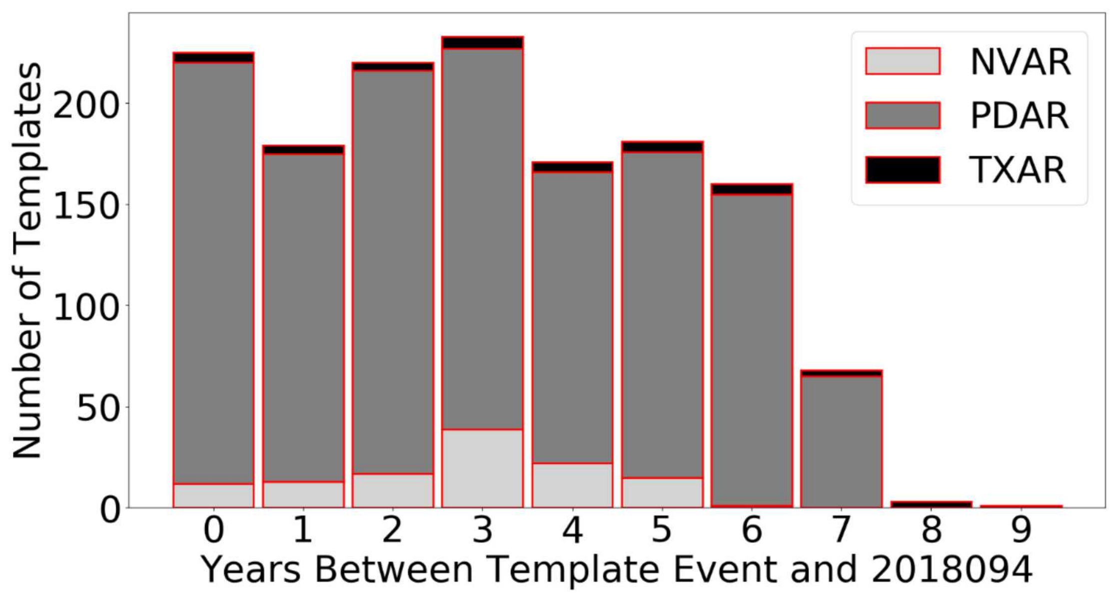

Figure 2-8. Number of templates binned by age for Wyoming Week 1, by station.

8. The template correlation coefficient thresholds were set using the time-reverse method, where the false alarm rate (FAR) was set to 1 FA per year. The templates were reversed and correlated against the waveform data for the week of study from April 4, 2018 through April 10, 2018. After the threshold is calculated, a small offset of 0.05 is added to the threshold value to raise

the value beyond noise correlation values that were observed. The template threshold value statistics by station are shown in Figure 2-9. The variation in custom template thresholds is small for two reasons: 1) the stations are arrays so time-reversing the templates will reverse the moveout across elements; and 2) the templates have a high time-bandwidth product. The result is that the templates generate a low false alarm rate when time-reversed and correlated with continuous data.

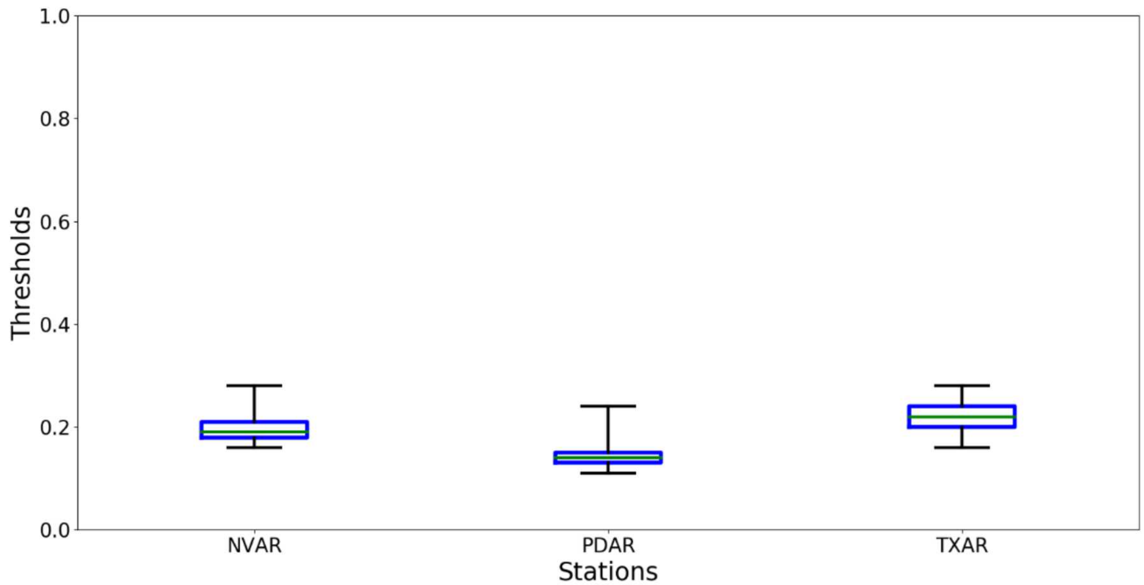

Figure 2-9. Statistics for Wyoming Week 1 template threshold values, alphabetical order by station. The box extends from the lower to upper quartile of the threshold values with a line representing the median threshold value. The minimum (maximum) threshold value is the bottom (top) line of the distribution.

2.3.2. Wyoming Week 2 (August 8–14, 2018)

The template libraries for Week 2 of the Wyoming region were drawn from up to 10 years of historical REB events within the geographical boundaries of the region.

Figure 2-10 shows an example Pn template waveform from the NVAR station from an REB template event (ORID 15211988) on 2017-12-14 22:01 UTC. Only the waveform for element NV01 is displayed in this figure even though the template includes waveforms for all elements of the array. The example template is 50s in duration and is windowed 5s after the picked REB arrival. The filter band used for the NVAR template library is 1.0 – 3.5 Hz. The time-reverse method calculated a custom template threshold of 0.22 for the template.

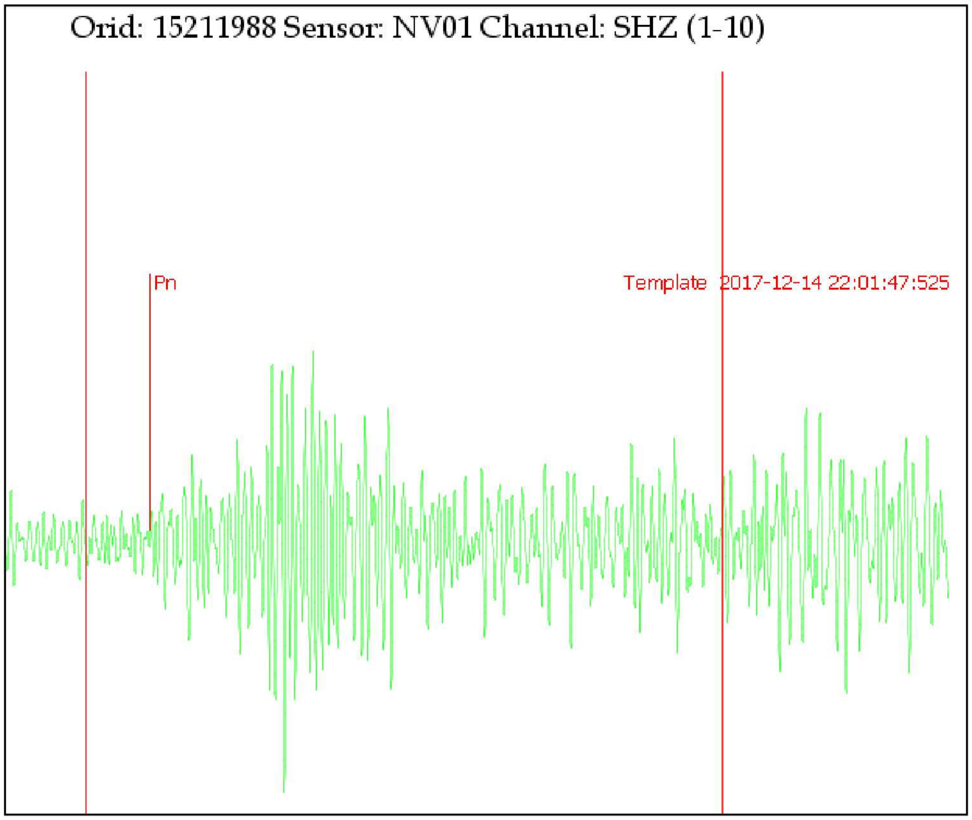

Figure 2-10. An example 50s NVAR Pn template waveform for the NV01 element. The portion of the waveform between the two vertical red lines indicates the portion of the arrival waveform that is used as the template waveform for cross-correlation with continuous data.

Table 2-3 provides an overview of the template library parameters, organized by station, for the template libraries for the Wyoming Week 2 region and timeframe. Later in this section we explain what the parameters are and how they are used for the library creation.

Table 2-3. Template Library Parameters for Wyoming Week 2

Stations	Phase	Avg SNR	Window Length (s)	Filter Band (Hz)	STA/LTA parameters (s)	STA/LTA Threshold	Template Count
PDAR	Pn	50.7	80	1.0-5.0	1/30/0	5.0	1341
NVAR	Pn	9.9	50	1.0-3.5	1/30/0	2.5	122
TXAR	Lg	5.4	60	0.5-5.0	3/30/100	2.0	46

The map in Figure 2-11 shows the station locations, network geometry, geographical boundaries of the study region, and the locations of template events. The color of the template events indicates the number of stations within the network (PDAR, NVAR, TXAR) that have templates from the event.

Because the template libraries for Week 1 and Week 2 are drawn from 10 years of historical waveform data, most of the templates will be the same in the libraries for the two weeks of study. However, there are differences where REB events from 4/10/2018 – 8/7/2018, not present in the template libraries for Week 1, may be included in the template libraries for Week 2. For example, for Week 2 there is a cluster of template events located near (41.5°N, 100.5°E) that were not included in the template library for Week 1, presumably because the events occurred during or after Week 1 (4/10/2018).

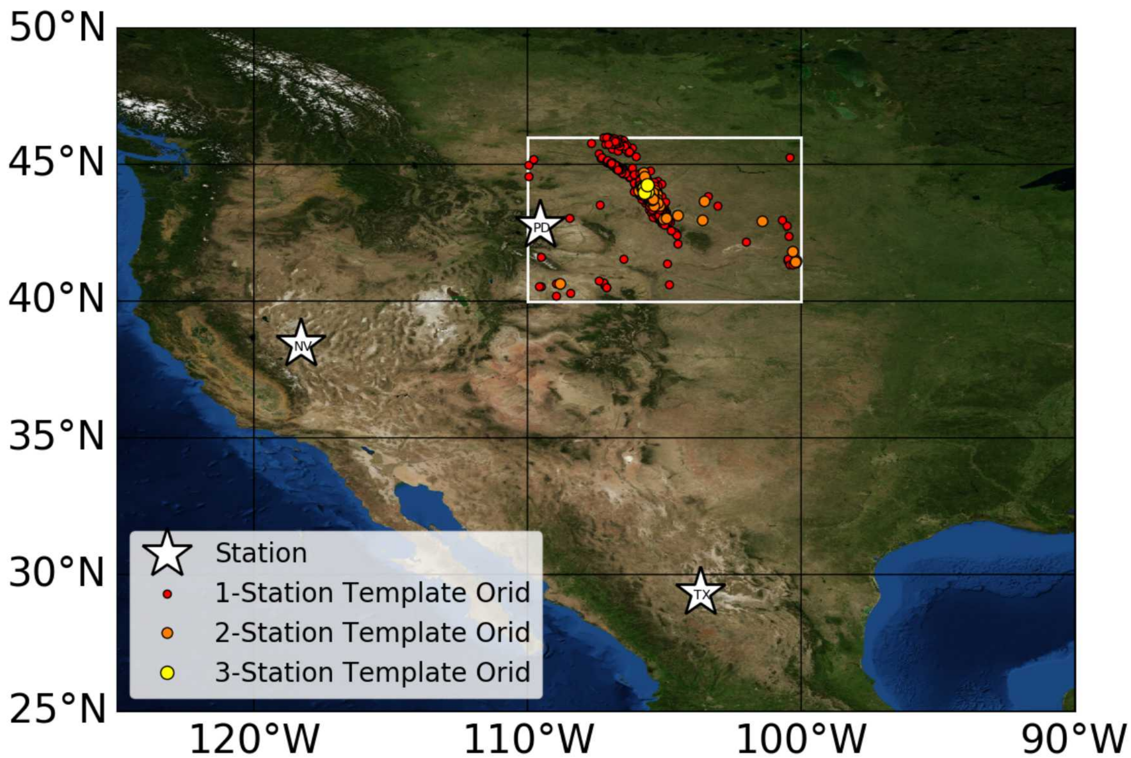

Figure 2-11. Map of the stations PDAR, NVAR, and TXAR, indicating network geometry and distance from the study region. The geographical boundaries of the study region are indicated by a white box, which encloses the Wyoming Week 2 template event locations. The color of the template events indicates the number of stations in the study that detected the template event.

The detailed sequential steps to create the template libraries for Wyoming Week 2 are listed below:

1. The REB origin table was queried for events that occurred within the geographical boundaries from Table 2-1, which are (Lat 40°-- 46°N, Lon 110°--100°W) and the timespan from 2008-08-08 to 2018-08-14. This resulted in 1364 template events from arrivals at all three stations. The locations of the template events are shown in Figure 2-12.

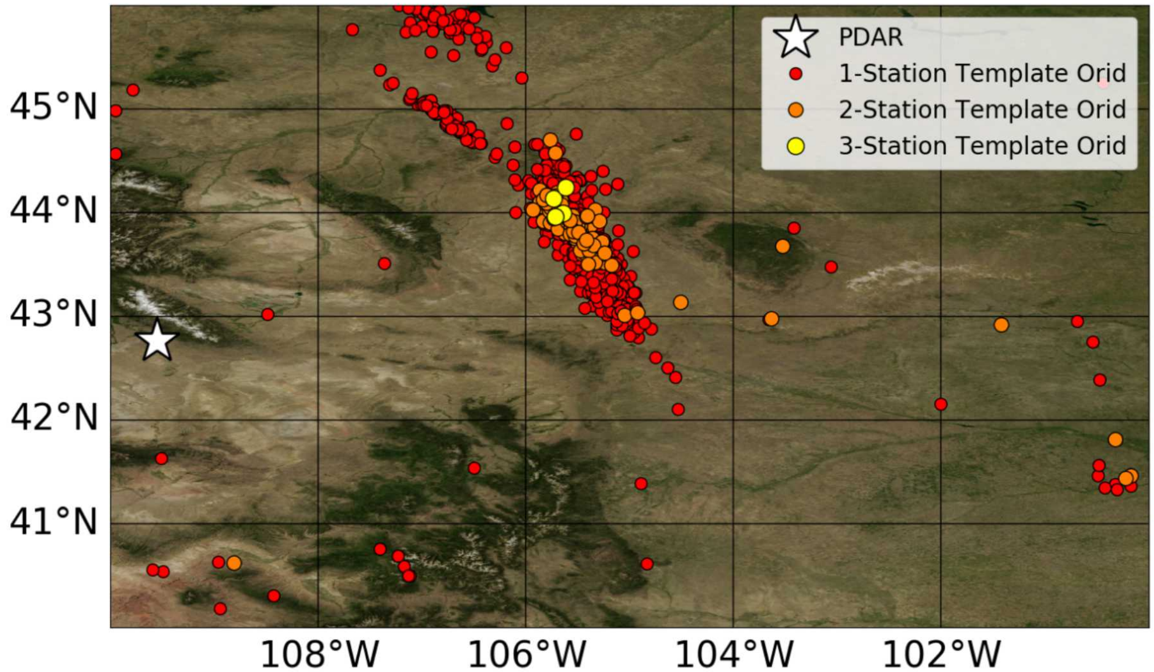

Figure 2-12. Locations of 1364 Wyoming Week 2 template events from REB events, where the color and size indicate the number of stations in the study detecting the event.

2. Create temporary tables with associations and arrivals for the 1364 template events. Queries joined REB.ORIGINS with the REB.ASSOC and IDCX.ARRIVAL tables to populate the Wyoming Week 2 associations and arrivals tables. SeisCorr uses the temporary tables to identify the arrivals at the stations for the 1364 template events bounded by the study.

The steps 3-8 are performed for each station to create a template library are consistent with the process used for Wyoming Week 1.

3. For each station the phase arrivals were reviewed for Week 1 to determine what phase would make the best template library for the station. The same phase arrivals were chosen for Week 2 template libraries because the timeframes of the libraries overlapped.
4. Template windows of different lengths were explored for each station for Week 1. The final choice for a template window length for Wyoming Week 2, shown in Table 2-3, is consistent with the window lengths for Week 1.
5. The templates were filtered with a 3-pole Butterworth bandpass filter. The frequency band used for Week 1 was also adopted for Wyoming Week 2 (Table 2-3).
6. STA/LTA threshold screening removed templates that did not have high enough SNR. The parameters for STA/LTA and the minimum STA/LTA threshold for each station that were chosen for Wyoming Week 1 were also used for Wyoming Week 2 (Table 2-3). Similarly to Week 2, PDAR has a significantly higher threshold than NVAR or TXAR because the station is closer to the events and the waveforms have a higher SNR.

7. The candidate templates that passed the STA/LTA threshold screening were saved as a template library. Figure 2-13 shows the number of templates for each station (i.e., library). The number of templates binned by age, by station, is shown by the histogram in Figure 2-14.

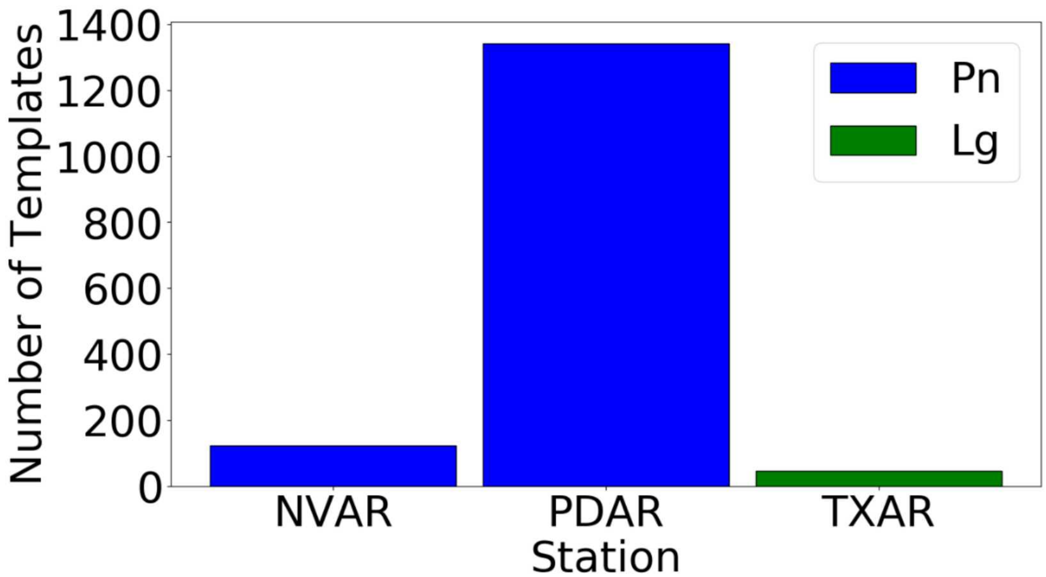

Figure 2-13. Wyoming Week 2 template count by station.

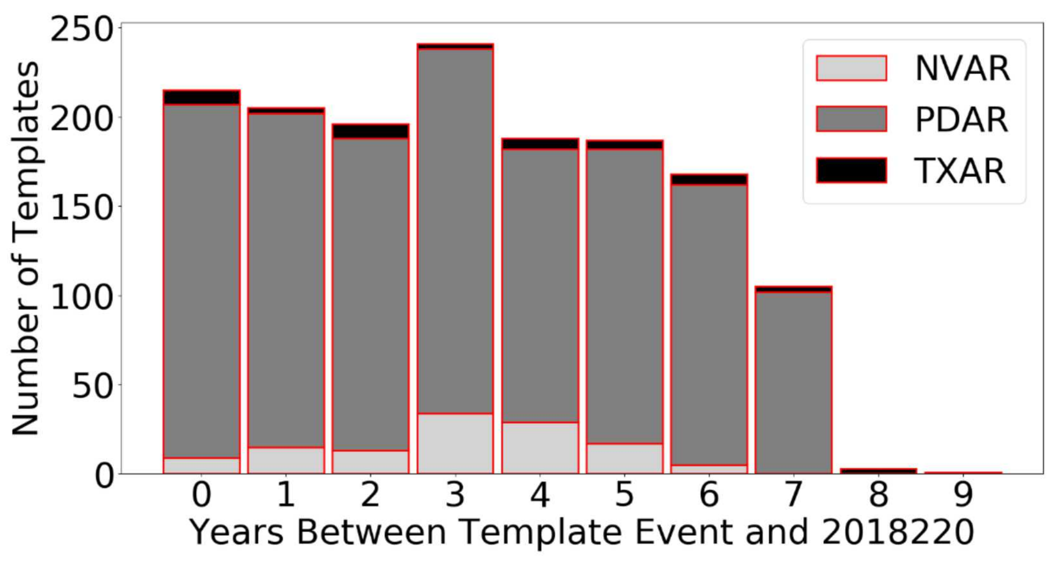

Figure 2-14. Number of templates binned by age for Wyoming Week 2, by station.

8. The template correlation coefficient thresholds were set using the time-reverse method, where the false alarm rate (FAR) was set to 1 FA per year. The templates were reversed and correlated against the waveform data for the week of study from August 8, 2018 through August 14, 2018. The template threshold value statistics by station are shown in Figure 2-15. As we would expect, the custom template thresholds are similar to the values obtained for the Week 1 template library

because the templates included within both libraries is almost identical. Nevertheless, it is reassuring to know that the noise characteristics between Week 1 and Week 2 result in similar custom template cross-correlation threshold values.

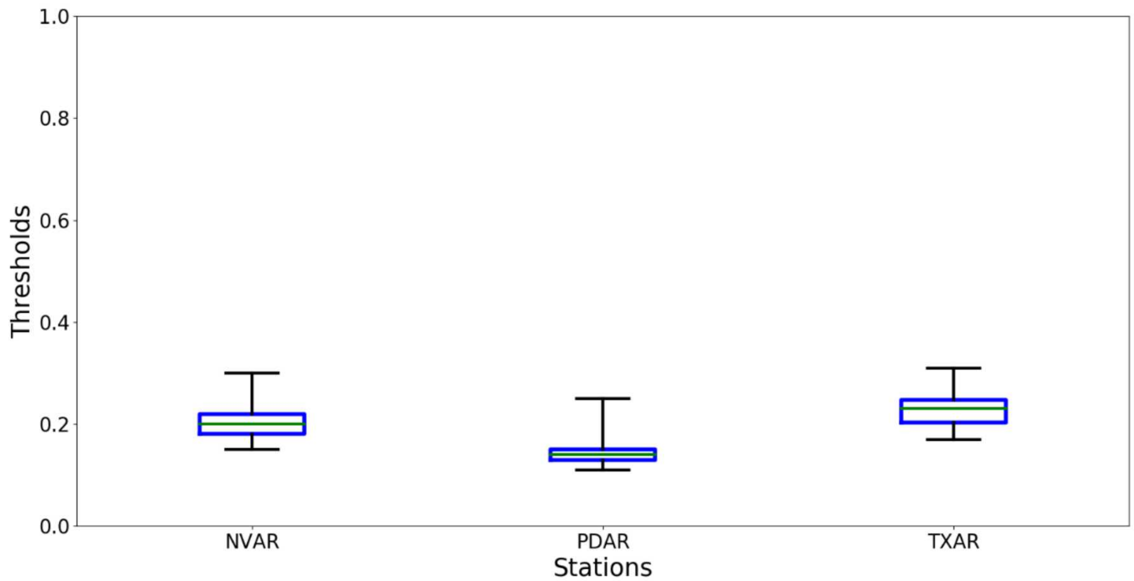

Figure 2-15. Statistics for Wyoming Week 2 template threshold values, alphabetical order by station. The box extends from the lower to upper quartile of the threshold values with a line representing the median threshold value. The minimum (maximum) threshold value is the bottom (top) line of the distribution.

2.3.3. *Scandinavia Week 1 (February 12–18, 2018)*

The template libraries for Week 1 of the Scandinavia region were drawn from up to 10 years of historical REB events within the geographical boundaries of the region. In the case of the ARCES and NOA stations, there were only 4 and 6 years of waveform data, respectively, available in the SNL archive for the same channel labels; in cases where channel labels changed, such as re-instrumentation at the stations, the older waveforms were not included.

There were issues with creating templates from ARCES, which is within the study region. There were 739 arrivals that were potential templates. However, SeisCorr failed to calculate a valid STA/LTA for 129 of those arrivals. This may have been due to signal quality on one or more array elements. Fortunately, the ARCES templates have a high time-bandwidth product and SNR, so the remaining 610 template waveforms are high quality for the purpose of detections with subsequent events.

There were also issues with creating the template library for NOA. The long template window in combination with the full array size of 31 elements resulted in failed correlation jobs that seemed to be due to computer memory limitations. As a result, we chose to use a subarray to reduce the number of elements in the templates. This simplification allowed the creation of a template library that enabled the completion of the study. We were able to use all 31 elements of NOA in templates for the aftershock study where the template window length was only 15 seconds. For this reason, we believe that the strain on computer memory resources was due to the template window length of 50 seconds in combination with a large number of array elements.

We used Lg picks for NOA because it was the most prominent phase for the epicentral distances. The travel time of Lg across the region of study made it difficult to choose a template window that would capture the Lg phase for all epicentral distances. Instead, we created two NOA Lg libraries. The first Lg library contains Lg template events at epicentral distances up to 9°, and has a template window length of 50 seconds. The second Lg library contains Lg template events at an epicentral distance from 9° to 12°, and has a template window length of 60 seconds. Beyond 12° the SNRs of the waveforms were too low for useful templates. Even the 9° – 12° library yielded few templates and very few correlation detections. NOA did not contribute templates that covered the entire geographical span of the region. We recommend further study to choose NOA templates for this region.

NOTE: The NOA template libraries were problematic for this study and the template events did not extend throughout the entire region. The authors recommend that additional effort be devoted to developing useful NOA template libraries for the Scandinavia mining region or investigating the use of other stations for waveform templates.

Figure 2-16 shows an example Pn template waveform from the ARCES station from an REB template event (ORID 15424672) on 2018-02-04 20:37 UTC. Only the waveform for array element ARA0 and channel BHZ is displayed in this figure even though the template includes waveforms for all elements of the array. The example template is 140s in duration and is windowed 5s before the picked REB arrival. The filter band used for the ARCES template library is 2.0 – 10.0 Hz. The time-reverse method calculated a custom template threshold of 0.119 for the template. This particular template correlated with six events during Week 1, with the correlation scores ranging from 0.6231 to 0.3630. Two of the six detected events have picked arrivals in the REB arrival table (this example is revisited in Section 2.4 and Figure 2-30).

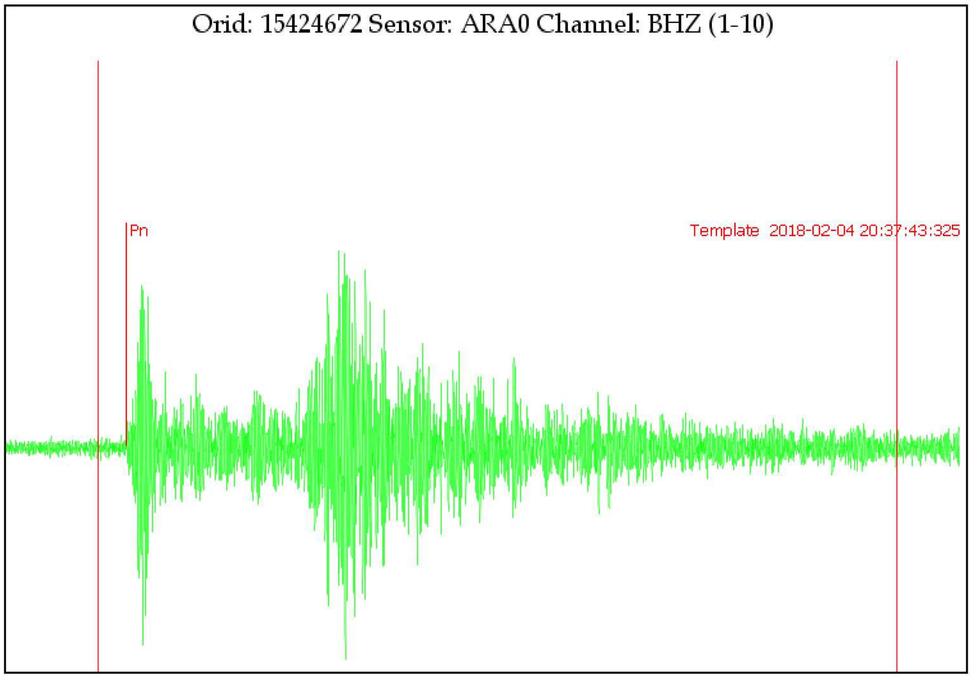

Figure 2-16. An example 140s ARCES Pn template waveform for the ARA0 element. The portion of the waveform between the two vertical red lines indicates the portion of the arrival waveform that is used as the template waveform for cross-correlation with continuous data.

Table 2-4 provides an overview of the template library parameters, organized by station, for the template libraries for the Scandinavia Week 1 region and timeframe. As previously described, SeisCorr failed to calculated STA/LTA values for many ARCES arrivals so longer STA/LTA timeframes of 2 s and 60 s were used for to increase the percentage of ARCES arrivals that were eligible for template waveforms. Note that ARCES and FINES waveforms had a higher average SNR than NOA, thus, the STA/LTA threshold for these two stations ensures a higher quality signal-to-noise ratio than for the NOA template waveforms. Moreover, ARCES and FINES contribute hundreds more template waveforms than NOA for this study.

Table 2-4. Template Library Parameters for Scandinavia Week 1

Stations	Phase	Avg SNR	Window Length (s)	Filter Band (Hz)	STA/LTA parameters (s)	STA/LTA Threshold	Template Count
ARCES	Pn	88.4	140	2.0-10.0	2/60/0	5.0	610
FINES	Pn	20.1	160	2.0-10.0	1/30/0	4.0	336
NOA 0-9°	Lg	5.5	50	2.0-6.0	3/30/100	3.0	90
NOA 9-12°	Lg	4.6	60	1.0-4.0	3/30/100	2.0	45

The map in Figure 2-17 shows the station locations, network geometry, geographical boundaries of the study region, and the locations of template events. The color of the template events indicates the number of stations within the network (ARCES, FINES, NOA) that have templates from the event. There are many 1-station template events because mining blasts are small so only each one of the closest two stations, ARCES and FINES, detected the event. This mining region has two stations that contributed many good quality template waveforms for the study and more template events were detected by 2 of the 3 stations (shown as orange circles on the map) in contrast with the Wyoming region. This is important because waveform correlation only detects events where a template from a prior event exists. We must have template events with multiple detecting stations to expect to detect recurring events with multiple stations.

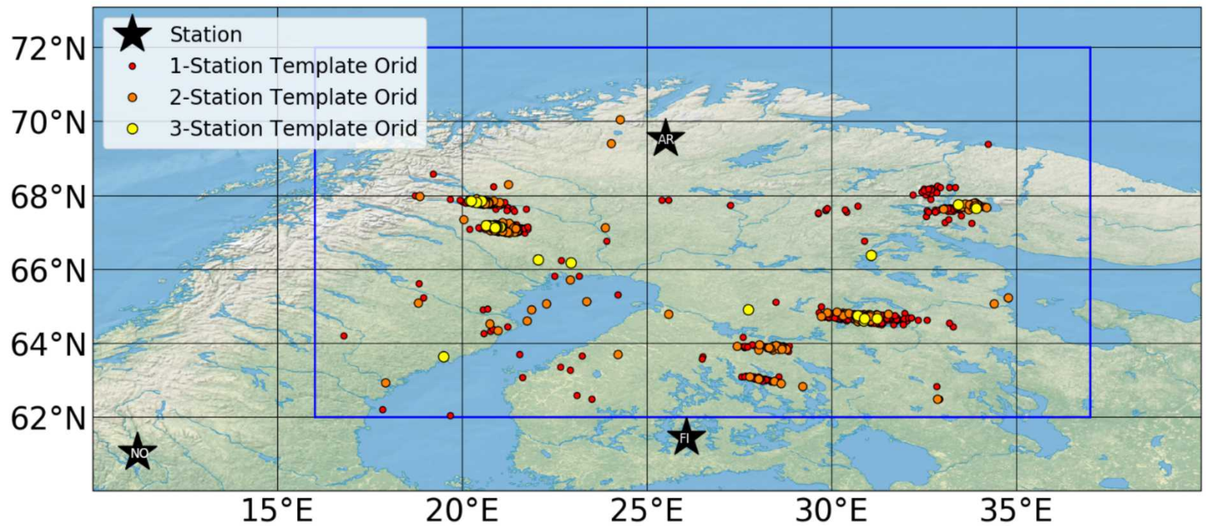

Figure 2-17. Map of the stations ARCES, FINES, and NOA, indicating network geometry and distance from the study region. The geographical boundaries of the study region are indicated by a dark blue box, which encloses the Scandinavia Week 1 template event locations. The color of the template events indicates the number of stations in the study that detected the template event.

The detailed sequential steps to create the template libraries for Scandinavia Week 1 are listed below:

1. The REB origin table was queried for events that occurred within the geographical boundaries from Table 2-1, which are (Lat 62°-- 72°N, Lon 16°--37°E) and the timespan from 2008-02-12 to 2018-02-09. This resulted in 970 template events from arrivals at all three stations. The locations of the template events are shown in Figure 2-18.

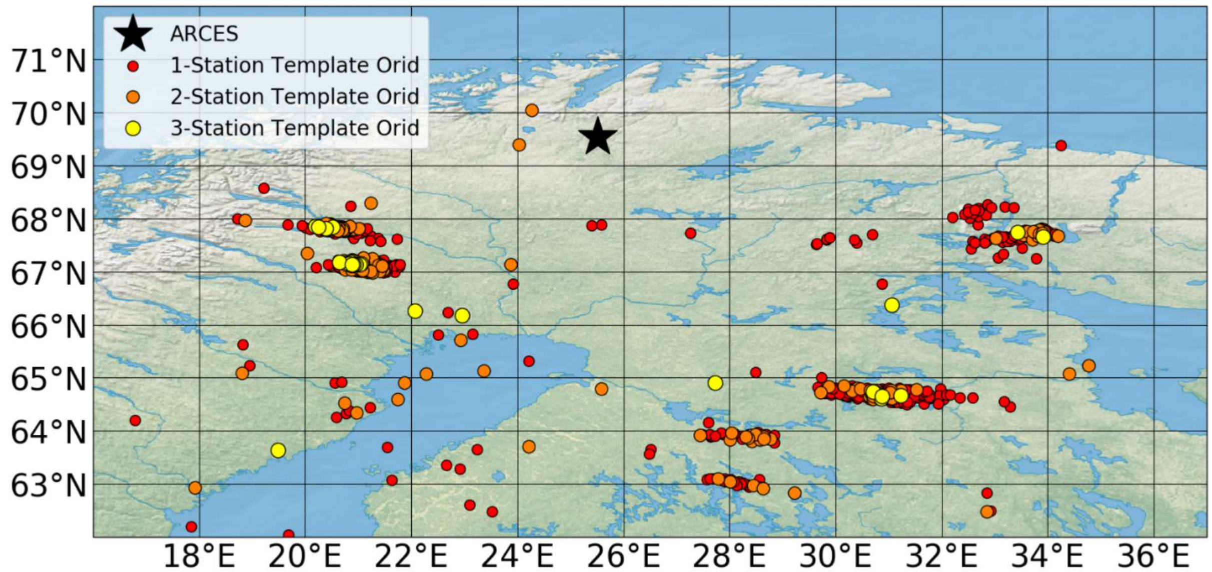

Figure 2-18. Locations of 970 Scandinavia Week 1 template events from REB events, where the color and size indicate the number of stations in the study detecting the event.

2. Create temporary tables with associations and arrivals for the 970 template events. Queries joined REB.ORIGINS with the REB.ASSOC and IDCX.ARRIVAL tables to populate the Scandinavia Week 1 associations and arrivals tables. SeisCorr uses temporary tables to identify the arrivals at the stations for the 970 template events bounded by the study.

The steps 3-8 are performed for each station to create a template library.

3. For each station the phase arrivals were reviewed to determine what phase would make the best template library for the station. We preferred to make templates from the Pn phase pick that included most of the full seismic waveform. In the case of station NOA, the Lg picks offered the best template waveform because the Lg arrivals from the mining region had the highest amplitude (Figure 2-19). The start of the template window is 5 s prior to the arrival pick for all the stations.

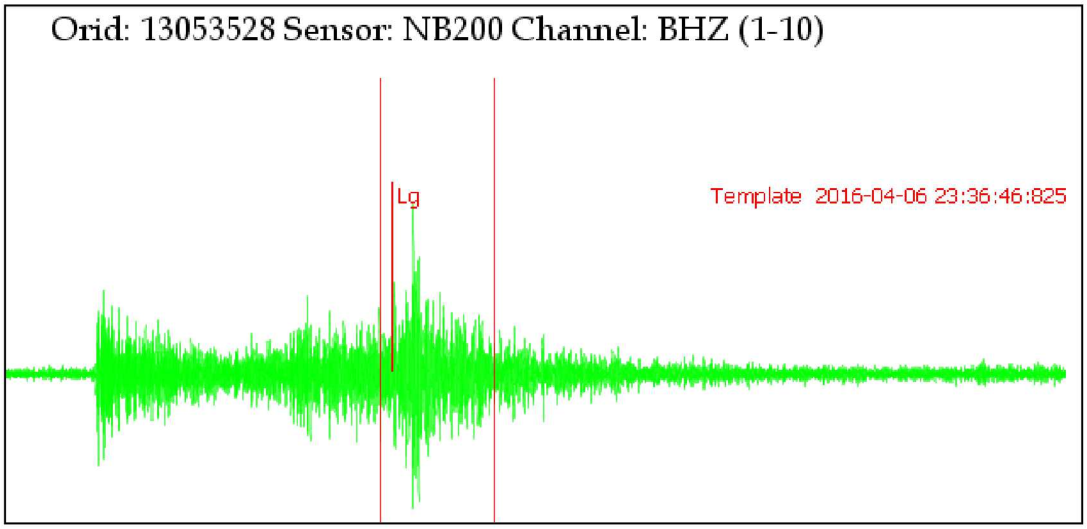

Figure 2-19. An example 60s NOA Lg template waveform for the NB200 element from an REB event located at (67.855°, 20.197°) on 2016-04-06 23:36 UTC.

4. Template windows of different lengths were explored for each station. The final choice for a template window length for Scandinavia Week 1 are shown in Table 2-4.
5. The templates were filtered with a 3-pole Butterworth bandpass filter. The frequency band is shown by station in Table 2-4.
6. STA/LTA threshold screening removed templates that did not have high enough SNR. The parameters for STA/LTA calculation and the STA/LTA threshold are shown in Table 2-4. Note that ARCES and FINES waveforms had a higher STA/LTA threshold than NOA, resulting in higher quality templates.
7. The candidate templates that passed the STA/LTA threshold screening were saved as a template library. Figure 2-20 shows the number of templates for each station and library. The number of templates binned by age, by station, is shown by the histogram in Figure 2-21. Although ARCES contributed the highest number of templates, those templates were drawn from only 4 years of history; in contrast, FINES contributed templates for the entire 10 years.

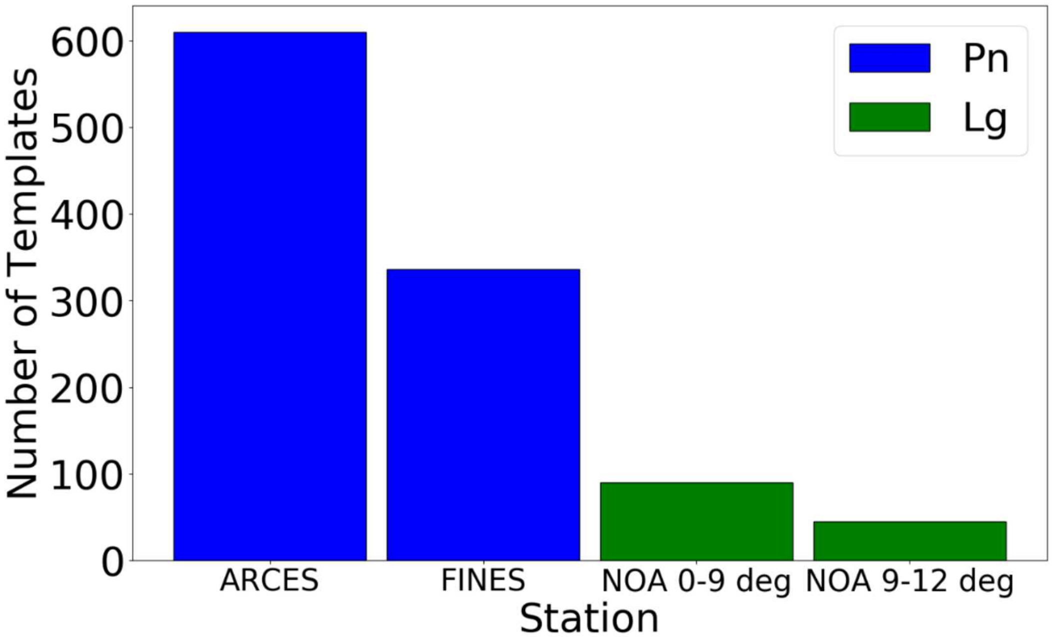

Figure 2-20. Scandinavia Week 1 template count by station, and in the case of NOA, by library.

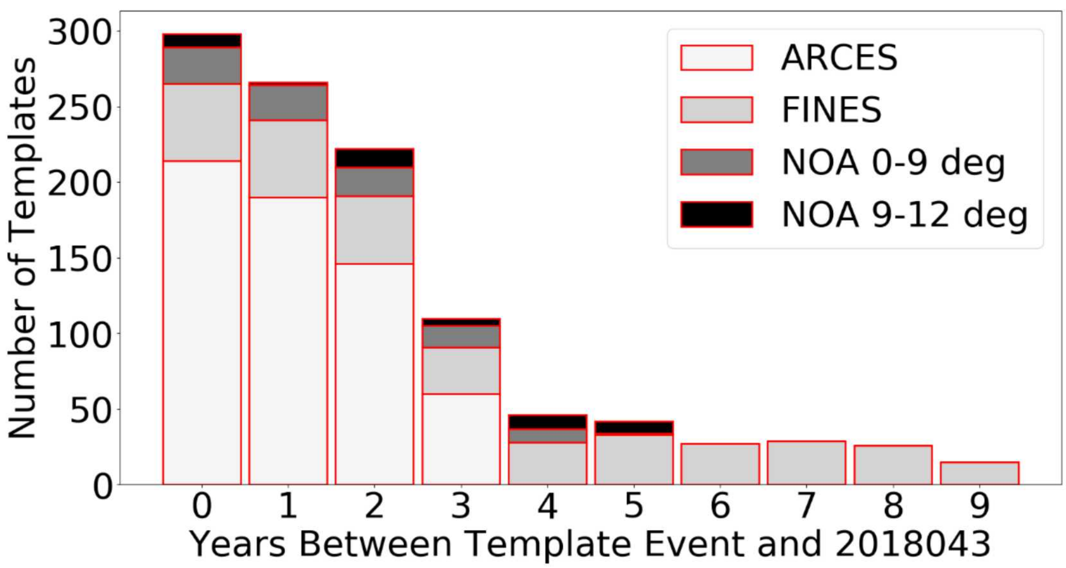

Figure 2-21. Number of templates binned by age for Scandinavia Week 1, by station, and in the case of NOA, by library.

8. The template correlation coefficient thresholds were set using the time-reverse method, where the false alarm rate (FAR) was set to 1 FA per year. The templates were reversed and correlated against the waveform data for the week of study from February 12, 2018 through February 18, 2018. After the threshold is calculated, a small offset of 0.05 is added to the threshold value to raise the value beyond noise correlation values that were observed. The template threshold value statistics by station are shown in Figure 2-22. The variation in custom template thresholds is small for two reasons: 1) the stations are arrays so time-reversing the templates will reverse

the moveout across elements; and 2) the templates have a high time-bandwidth product. Note that the NOA templates have a lower time-bandwidth product than the ARCES or FINES templates because the template length is shorter and the frequency band is narrower, thus the median cross-correlation threshold value is higher. However, all of the templates generate a low false alarm rate when time-reversed and correlated with continuous data so the cross-correlation thresholds are low.

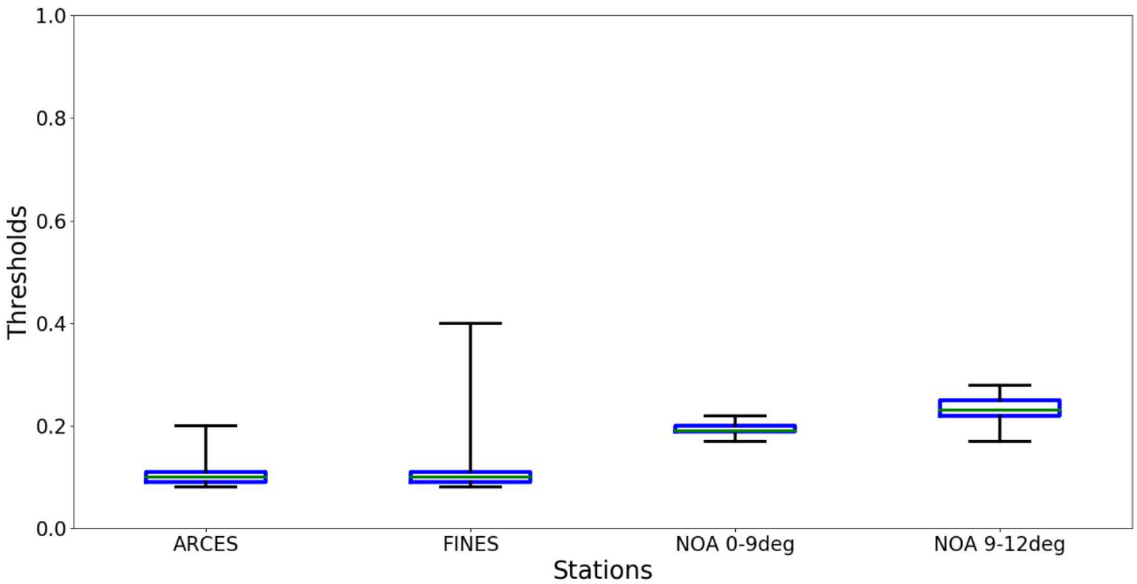

Figure 2-22. Statistics for Scandinavia Week 1 template threshold values, alphabetical order by station. The box extends from the lower to upper quartile of the threshold values with a line representing the median threshold value. The minimum (maximum) threshold value is the bottom (top) line of the distribution.

2.3.4. Scandinavia Week 2 (July 6–12, 2018)

The template libraries for Week 2 of the Scandinavia region were drawn from up to 10 years of historical REB events within the geographical boundaries of the region.

Figure 2-23 shows an example Pn template waveform from the FINES station from an REB event (ORID 11997624) on 2015-05-13 22:21 UTC. Only the waveform for element FIA0, channel SHZ, is displayed in this figure even though the template includes waveforms for all elements of the array. The example template is 160s in duration and is windowed 5s before the picked REB arrival. The filter band used for the FINES template library is 2.0 – 10.0 Hz. The time-reverse method calculated a custom template threshold of 0.16 for the template.

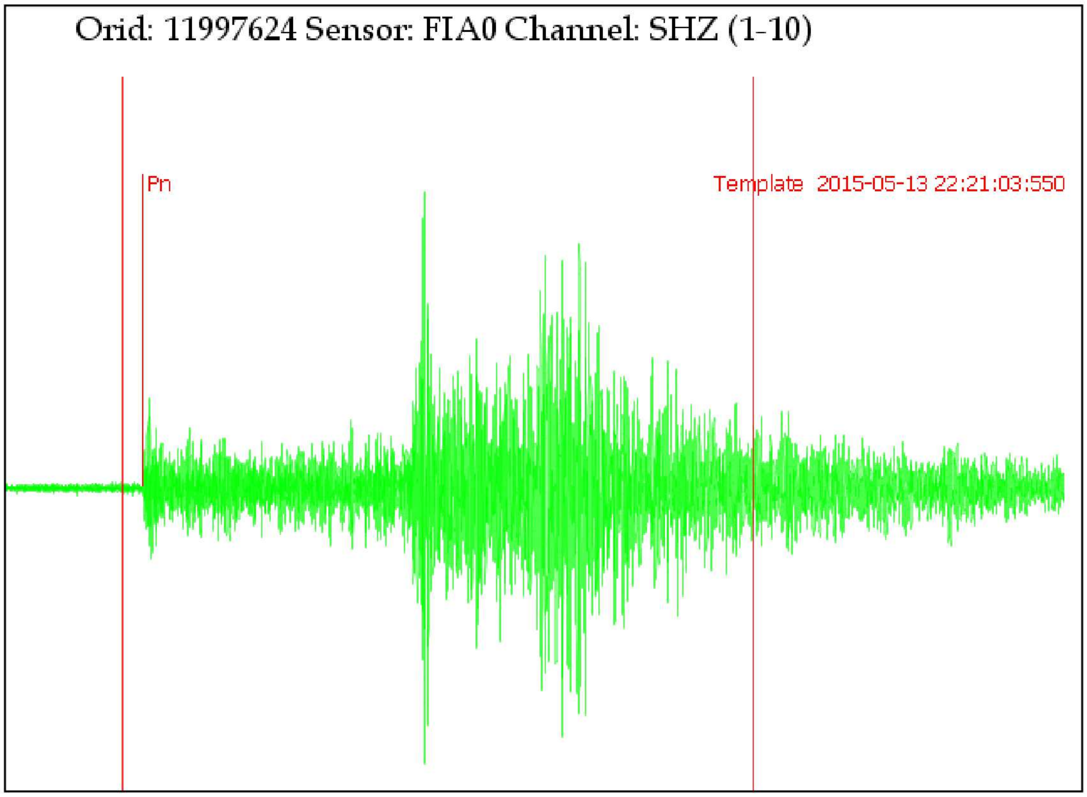

Figure 2-23. An example 160s FINES Pn template waveform for the FIA0 element from an REB event located at (67.209°, 20.939°) on 2015-05-13 22:21 UTC. The portion of the waveform between the two vertical red lines indicates the portion of the arrival waveform that is used as the template waveform for cross-correlation with continuous data.

Table 2-5 provides an overview of the template library parameters, organized by station, for the template libraries for the Scandinavia Week 2 region and timeframe. For Week 2, the STA/LTA thresholds for ARCES and NOA 0-9° were lowered to include more candidate arrivals as templates, which resulted in a lower average SNR than the template libraries for Week 1. In the case of ARCES, this was done because SeisCorr was failing to calculate an STA/LTA for a large number of arrivals that appeared to have good waveforms, so this change in STA/LTA threshold was chosen as a software workaround. In the case of NOA, the lower STA/LTA threshold was intentionally chosen to include marginal templates that could increase the number of 3-station template events for Week 2.

Table 2-5. Template Library Parameters for Scandinavia Week 2

Stations	Phase	Avg SNR	Window Length (s)	Filter Band (Hz)	STA/LTA parameters (s)	STA/LTA Threshold	Template Count
ARCES	Pn	73.2	140	2.0-10.0	2/60/0	3.0	862
FINES	Pn	20.0	160	2.0-10.0	1/30/0	4.0	352
NOA 0-9°	Lg	5.3	50	2.0-6.0	3/30/100	2.0	145
NOA 9-12°	Lg	4.7	60	1.0-4.0	3/30/100	2.0	44

The map in Figure 2-24 shows the station locations, network geometry, geographical boundaries of the study region, and the locations of template events. The color of the template events indicates the number of stations within the network (ARCES, FINES, NOA) that have templates from the event. There are many 1-station template events because mining blasts are small so only the closest station, PDAR, of the 3 stations in this network detected the event.

Because the template libraries for Week 1 and Week 2 are drawn from 10 years of historical waveform data, most of the templates will be the same in the libraries for the two weeks of study. However, due to different STA/LTA thresholds, additional arrival waveforms are included in the template library for Week 2.

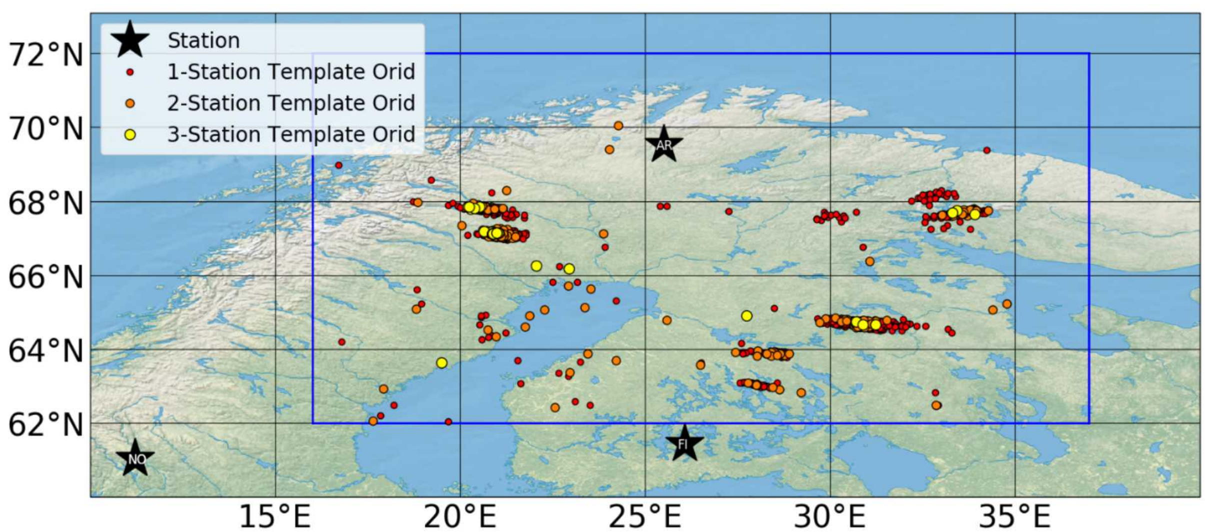

Figure 2-24. Map of the stations ARCES, FINES, and NOA, indicating network geometry and distance from the study region. The geographical boundaries of the study region are indicated by a blue box, which encloses the Scandinavia Week 2 template event locations. The color of the template events indicates the number of stations in the study that detected the template event.

The detailed sequential steps to create the template libraries for Scandinavia Week 2 are listed below:

1. The REB.ORIGINS table was queried for events that occurred within the geographical boundaries from Table 2-1, which are (Lat 62°-- 72°N, Lon 16°-- 37°E) and the timespan from 2008-07-05 to 2018-07-05. This resulted in 1140 template events from arrivals at all three stations. The locations of the template events are shown in Figure 2-25.

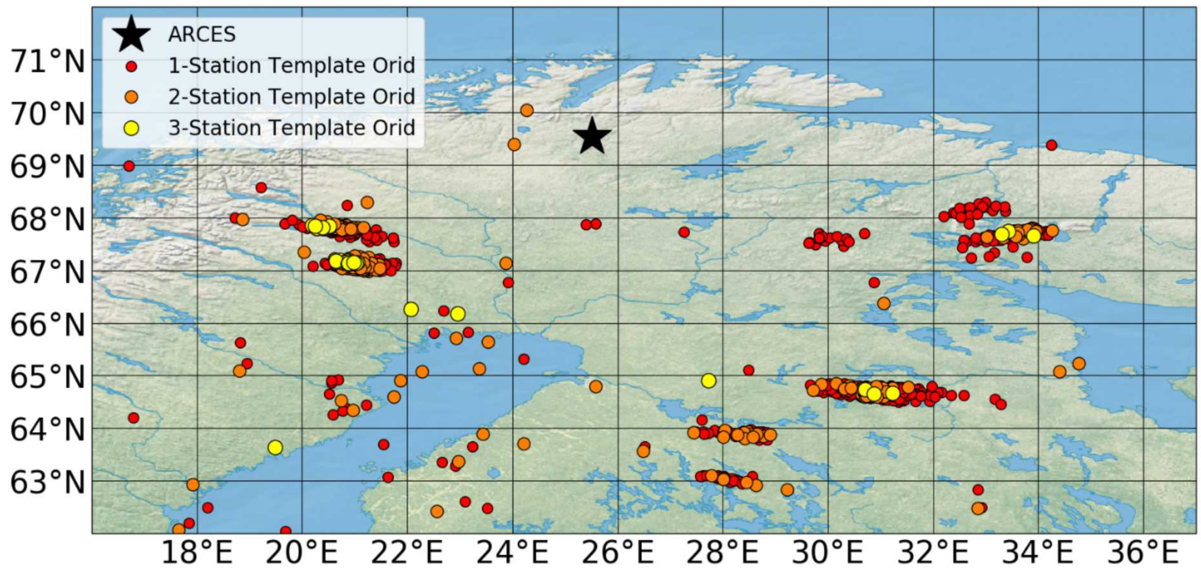

Figure 2-25. Locations of 1140 Scandinavia Week 2 template events from REB events, where the color and size indicate the number of stations in the study detecting the event.

2. Create temporary tables with associations and arrivals for the 1140 template events. Queries joined REB.ORIGINS with the REB.ASSOC and IDCX.ARRIVAL tables to populate the Scandinavia Week 2 associations and arrivals tables. SeisCorr uses the temporary tables to identify the arrivals at the stations for the 1140 template events bounded by the study.

The steps 3-8 are performed for each station to create a template library are consistent with the process used for Scandinavia Week 1.

3. For each station the phase arrivals were reviewed for Week 1 to determine what phase would make the best template library for the station. The same phase arrivals were chosen for Week 2 template libraries because the timeframes of the libraries overlapped.
4. Template windows of different lengths were explored for each station for Week 1. The final choice for a template window length for Scandinavia Week 2, shown in Table 2-5, is consistent with the window lengths for Week 1.
5. The templates were filtered with a 3-pole Butterworth bandpass filter. The frequency band used for Week 1 was also adopted for Wyoming Week 2 (Table 2-5).
6. STA/LTA threshold screening removed templates that did not have high enough SNR. The parameters for STA/LTA and the minimum STA/LTA threshold for each station for Scandinavia Week 2 are shown in Table 2-5.

The candidate templates that passed the STA/LTA threshold screening were saved as a template library. Figure 2-26 shows the number of templates for each station and library. The number of templates binned by age, by station, is shown by the histogram in Figure 2-27.

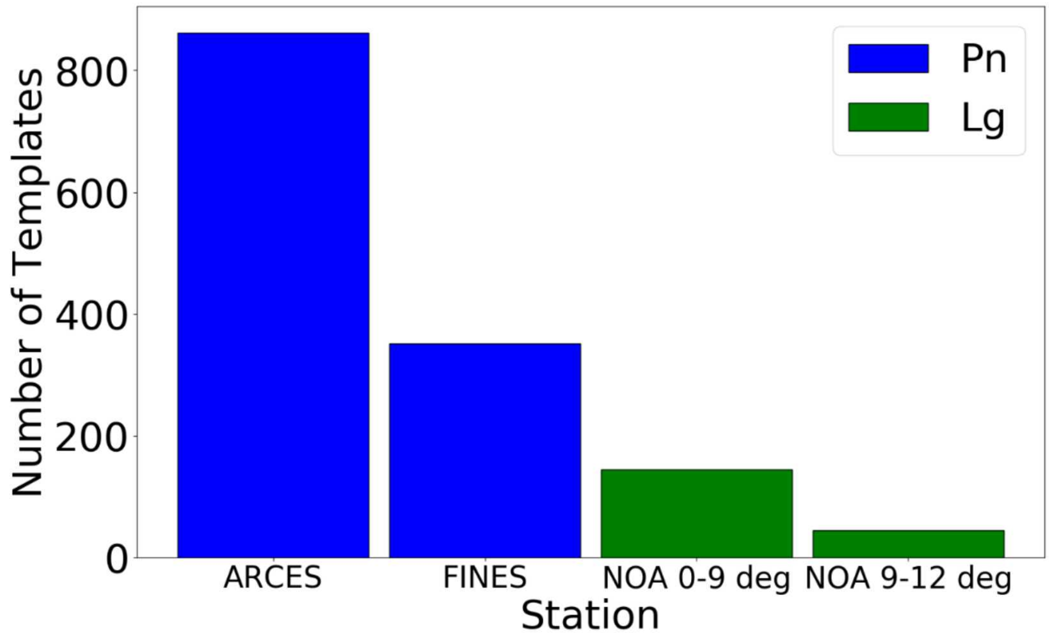

Figure 2-26. Scandinavia Week 2 template count by station, and for NOA, by library.

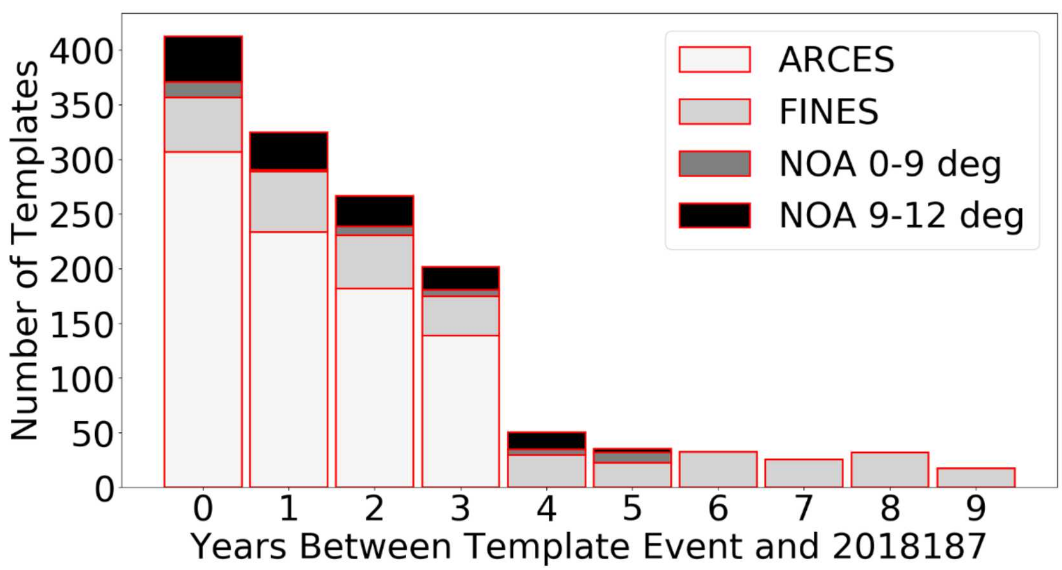

Figure 2-27. Number of templates binned by age for Scandinavia Week 2, by station, and for NOA, by library.

7. The template correlation coefficient thresholds were set using the time-reverse method, where the false alarm rate (FAR) was set to 1 FA per year. The templates were reversed and correlated against the waveform data for the week of study from July 6, 2018 through July 12, 2018. The template threshold value statistics by station are shown in Figure 2-28. We would expect the

custom template threshold statistics for Week 1 (Figure 2-22) and Week 2 to be similar. However, Week 2 shows a wider quartile spread in the custom template threshold values, which is particularly noticeable for FINES. It is possible that the noise characteristics between Week 1 and Week 2 are slightly different, resulting in variation in the statistics for custom template cross-correlation threshold values.

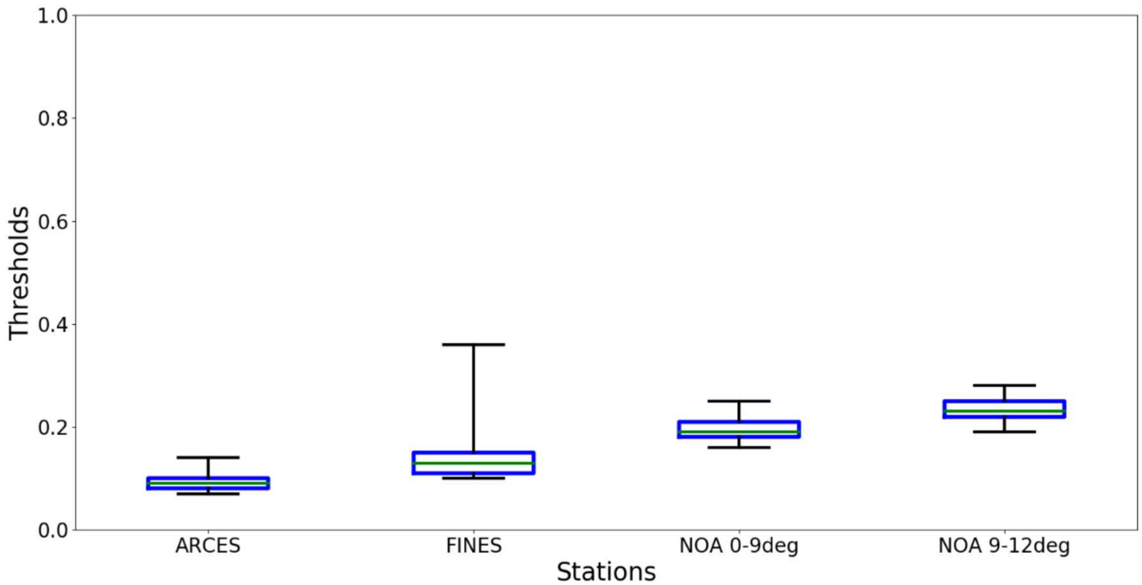

Figure 2-28. Statistics for Scandinavia Week 2 template threshold values, alphabetical order by station. The box extends from the lower to upper quartile of the threshold values with a line representing the median threshold value. The minimum (maximum) threshold value is the bottom (top) line of the distribution.

2.4. Correlation Jobs

For each region and week of study, template waveforms in the template libraries are correlated against continuous waveform data from the same station to detect similar waveforms from mining blasts. Detections are recorded if the correlation score at a point in time exceeds the template correlation coefficient threshold.

Table 2-6 shows a summary of the correlation jobs that were executed for each mining region and week of study, indicating the overall number of templates and detections.

Table 2-6. Summary of correlation jobs by mining region and week of study. All start and end times are 00:00:00 UTC.

Mining Region	Week	Correlation Start	Correlation End	# Templates	# Detections
Wyoming	1	April 4, 2018	April 11, 2018	1441	148
Wyoming	2	August 8, 2018	August 14, 2018	1509	123
Scandinavia	1	February 12, 2018	February 19, 2018	1081	447
Scandinavia	2	July 6, 2018	July 12, 2018	1403	342

This table documents an error in the correlation jobs for Week 2 of both regions that was not discovered until evaluation of the study results. The end times for all runs is 00:00:00 UTC. The end date for Wyoming Week 2 was incorrectly entered as 2018-08-14 rather than 2018-08-15, thus day 7 of Week 2 was not processed. Similarly, the end date for Scandinavia Week 2 was incorrectly entered as 2018-07-12 rather than 2018-07-13, also not processing day 7. As a result of this error, events that occurred in the last day of the Week 2 for both regions were not detected. Later in the report (Section 3.2), we discuss the impact of the missing events when comparing waveform correlation detections with detections by the IDC automated pipeline and the REB. This error is unfortunate but not meaningful, as it does not affect the study conclusions.

NOTE: Any use or evaluation of the dataset accompanying this report should adjust Wyoming Week 2 and Scandinavia Week 2 to six days to avoid incorrect conclusions.

The rest of this section includes some example detections from the stations for both mining regions. Some of these examples were detected by templates that were previously presented in figures in Section 2.3 Template Libraries.

The PDAR Pn template (Figure 2-3) is shown with a detection in Figure 2-29. This is an example of a good quality detection. The template has a complex signal with multiple phases and is clearly distinguishable from the noise. The detected signal follows the amplitude and frequency characteristics of the template; the phase differential consistency builds confidence in the detection. The cross-correlation score is 0.4837 and the relative magnitude of the detected signal is 0.48, so the detected event is smaller than the template event. The location of the template event is (44.245, -105.532). The cross-correlation score between waveforms will decline if the events are not exactly collocated, so it seems likely that the detected event is near the template event but not at exactly the same location. It is also possible that the mine blasting and excavation between the template event (2017-07-03) and the detected event (2018-04-10) changed the physical structure, resulting in signal decorrelation for collocated events. Another possibility is that the blasting techniques are slightly different between the two events.

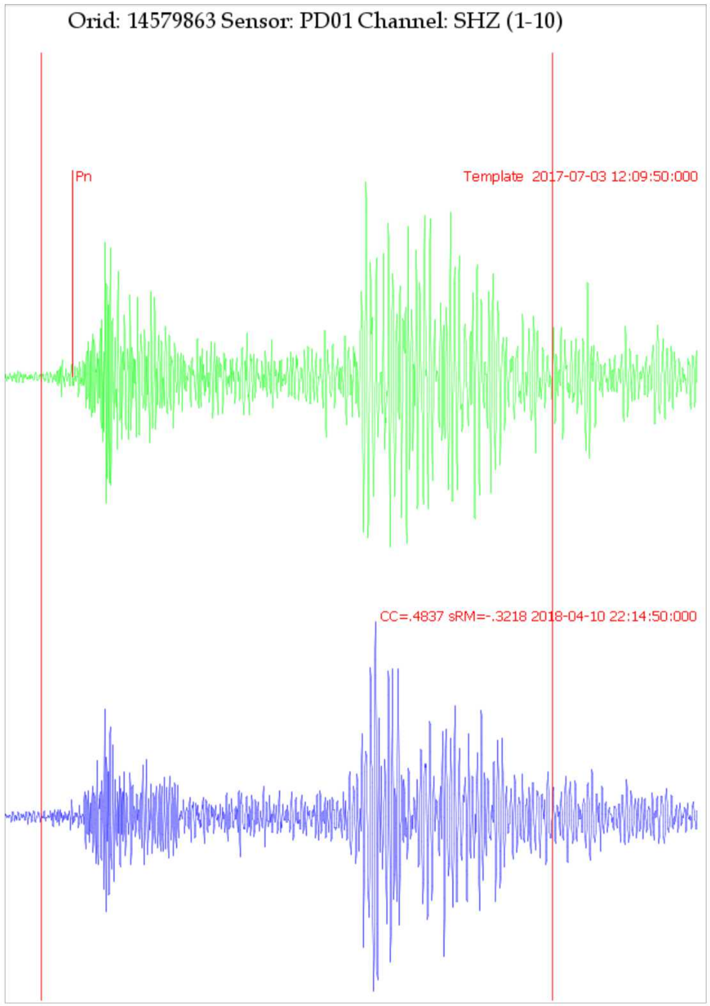

Figure 2-29. PDAR Pn template with detection. The detection has a cross-correlation score of 0.4837. The template has a cross-correlation threshold of 0.189. The template had two detections in the Wyoming region during Week 1.

The ARCES Pn template (Figure 2-16) is shown with a series of detections ordered by cross-correlation score in Figure 2-30. This example was chosen to indicate the potential of waveform correlation to improve the detection of recurring mining blasts while reducing analyst workload to produce the bulletin. The template has a complex signal with multiple phases and is clearly distinguishable from the noise; the template cross-correlation threshold is 0.119. The detected signals follow the amplitude and frequency characteristics of the template and the phase differential consistency builds confidence in the detections. The cross-correlation scores range from 0.6231 and 0.3630 but all look highly similar. The location of the template event is (67.784, 20.683), which is within an area of active mine blasts. Moreover, two of the six detections were picked as REB arrivals. This example also shows that waveform correlation detects likely real events that are not in the bulletin; thus, it can be difficult to get a realistic number for false positives when comparing with a bulletin.

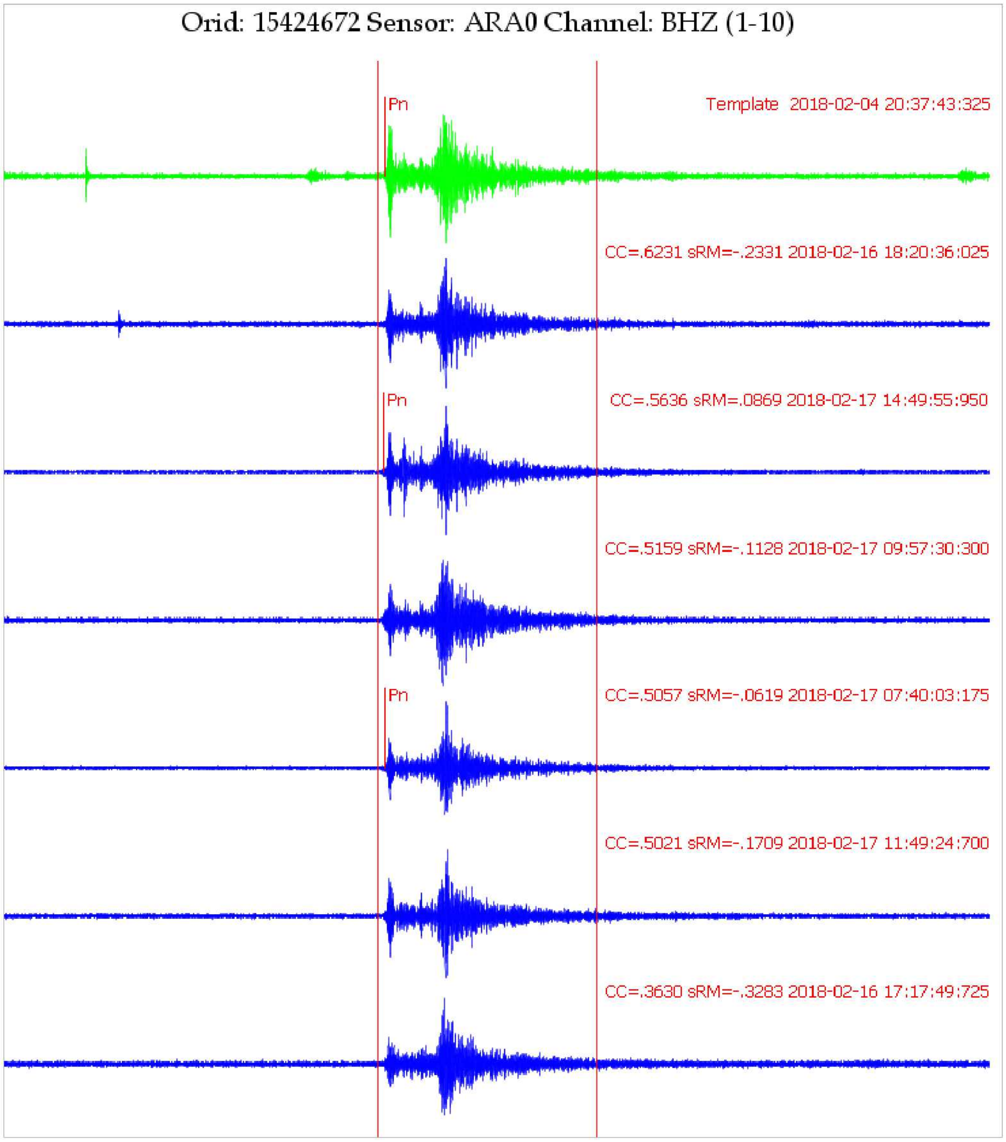

Figure 2-30. ARCES Pn template with a series of detections ordered by cross-correlation score. Two of the six detections were picked arrivals for the REB bulletin (the ones with labelled Pn picks).

The TXAR Lg template (Figure 2-6) is shown with a detection in Figure 2-31. This is an example of a marginal detection. The template has a low SNR, and the amplitudes of the other phases do not exceed the noise level. The detected signal seems likely to be an event, but the low cross-correlation score of 0.2064 is only slightly above the template's cross-correlation threshold of 0.2; it is possible that the template detected an actual, but non-collocated event. This is an example of a detection from the most distant station in our network of 3 stations that does not inspire confidence alone but may combine with detections from the two closer stations to corroborate an event. More experimentation with waveform correlation in an operational pipeline will be required to determine the potential usefulness of detections such as this example.

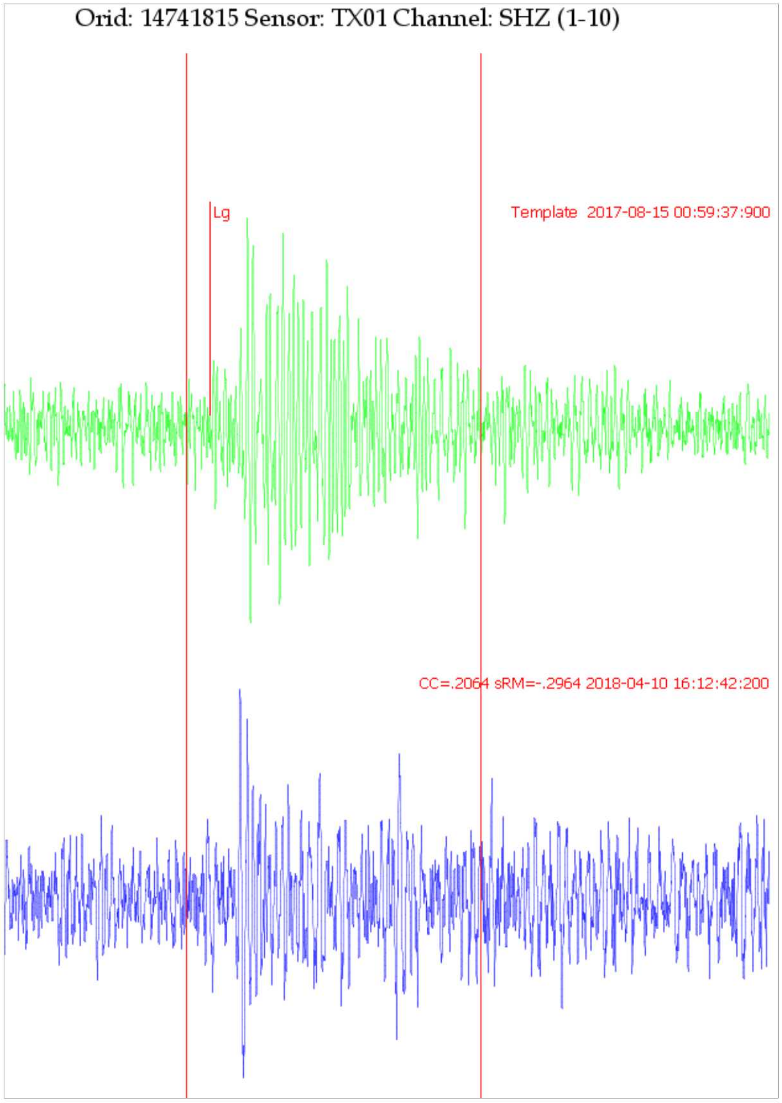

Figure 2-31. TXAR Lg template with detection. The detection has a cross-correlation score of 0.2064. The template has a cross-correlation threshold of 0.2. The template had two detections in the Wyoming region during Week 1, both with cross-correlation scores barely above the threshold.

The next template with set of detections is included as a counterexample to raise awareness of the potential risks in incorporating automated waveform correlation in an operational pipeline. As mentioned earlier, the SeisCorr creation of template libraries for this study used REB picked arrivals to identify potential templates. Visual inspection of selected arrivals chose the template window and bandpass filters for the library. The selection included some randomly chosen arrivals, arrivals from large and small magnitude events, and arrivals from both the nearest and furthest events to evaluate the duration of the waveform for the entire study region.

The example NVAR Pn template (Figure 2-32) shows five incorrect detections. The template is based on a Pn arrival, but the template window included a portion of an unrelated event. The high amplitude arrival of the second event within the template window allowed the candidate Pn arrival to pass STA/LTA threshold screening and become a template. The correlation detections are matching the high amplitude arrival from the second event, not the Pn arrival from the first event. This template should not be used, but it is included to provide an example of what can go wrong in an automated approach to template selection. We recommend research to develop a more rigorous automated template selection method, especially in the application of waveform correlation to mining regions where events may be frequent and overlap in long template windows.

Furthermore, it is important to recognize that the detected events in this example will be erroneously located where the template is located, potentially affecting the downstream association process. The metadata associated with the template event (i.e., the Pn arrival's associated event) is used to calculate information about the detected events, such as location and relative magnitude, that will be completely inaccurate in this example. We will review the risk associated with poor template selections further in the section on multistation validation when we can present examples of the effect on association of detections into events.

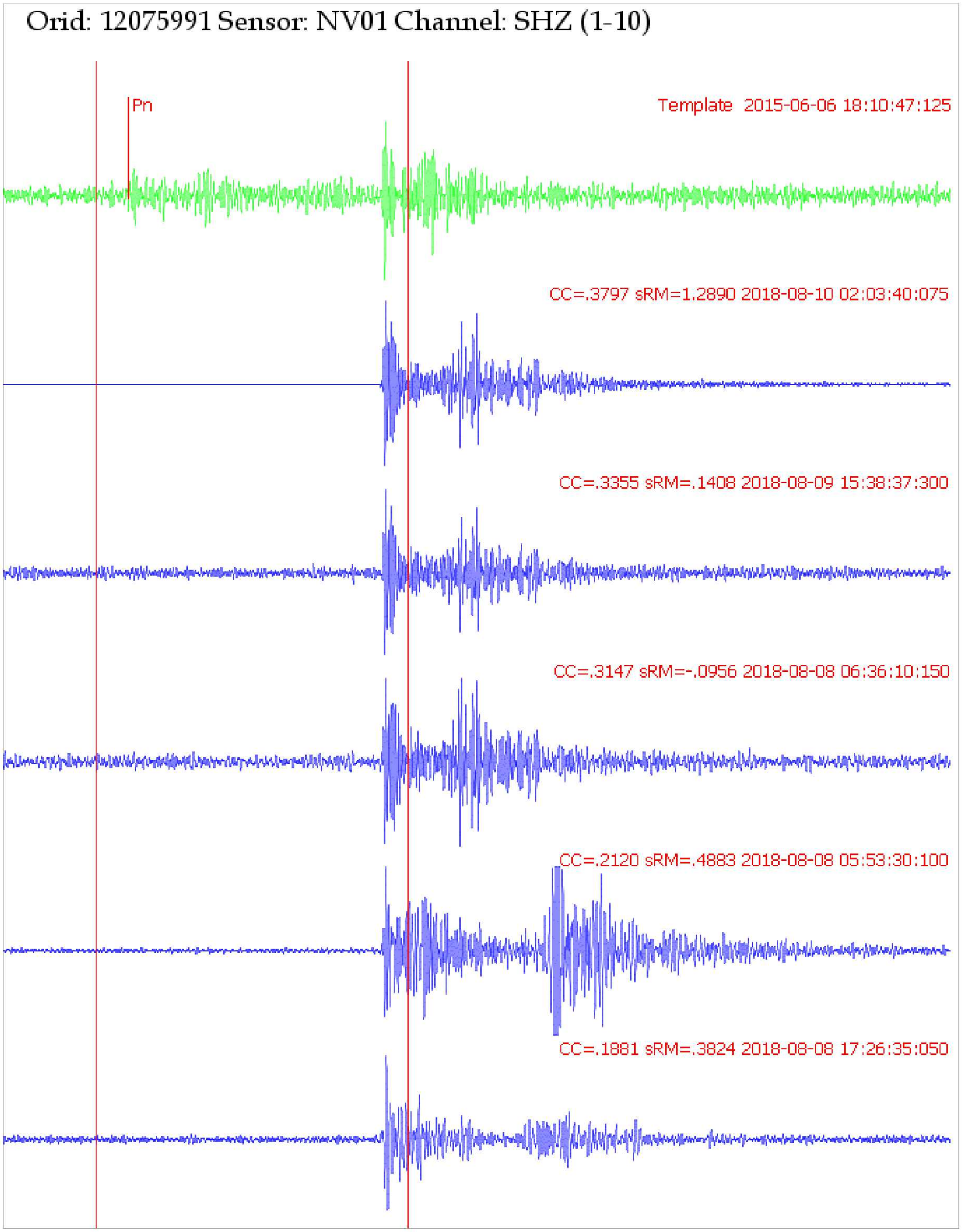

Figure 2-32. Example NVAR Pn template with detections. This example shows the problems that can arise when a template is incorrectly windowed to include a portion of an unrelated arrival. In this case, the detections are correlating with the high amplitude unrelated arrival at the end of the template window, resulting in inaccurate detections.

The last example that we present in this section shows the power of waveform correlation to detect events in noisy background. The FINES Pn template (Figure 2-23) is shown with a detection in Figure 2-33. This is an example of a marginal detection, yet the detected arrival is also a picked REB arrival (ORID 15997904 at 2018-07-06 02:37). The template has a complex signal with multiple phases that are clearly distinguishable from the noise. The cross-correlation score is 0.1742 while the template cross-correlation threshold is 0.16. The relative magnitude of the detected signal is 0.056, so the detected event is much smaller than the template event.

Previously we explained that the cross-correlation score between waveforms will decline if the events are not exactly collocated and this example gives us a good opportunity to show the effect of such decorrelation because we have the bulletin locations of the template and detected events. The REB location of the template event is (67.209°N, 20.939°E) while the REB location of the detected event is (67.118°N, 21.183°E). Both locations are within an active area of mining in the Scandinavia region, so it is likely they are from the same mine or mines that are nearby. Despite the low cross-correlation score due to location differences and signal magnitude differences, we conclude that this is a good waveform correlation detection that has been validated by the analyst picks. We believe this example also shows that waveform correlation detections can lead the analyst to a likely location for marginal arrivals, decreasing the time to search for corroborating arrivals for events.

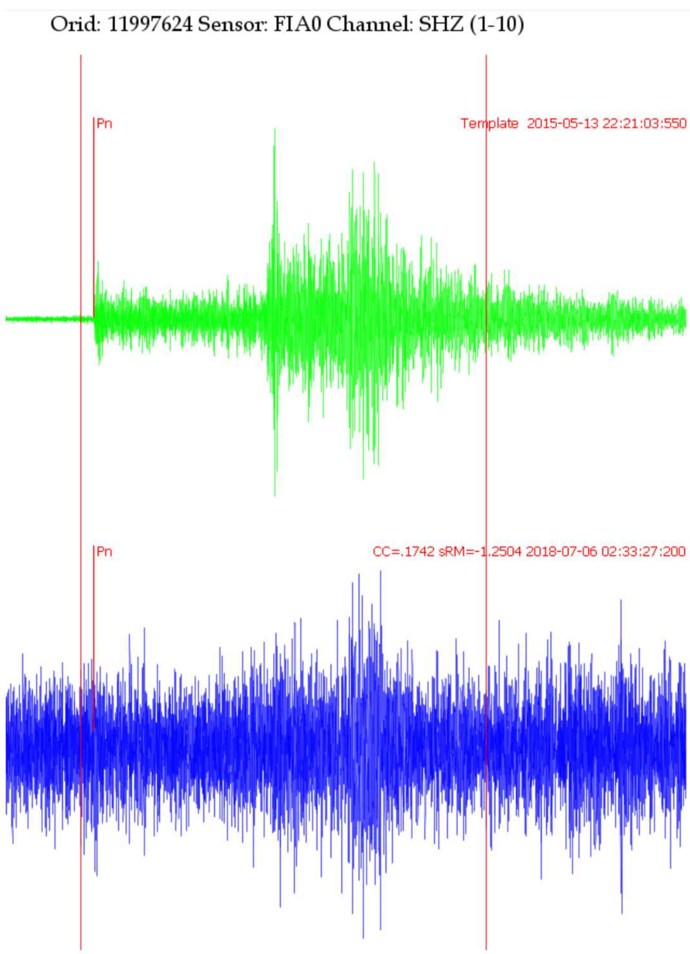

Figure 2-33. Example FINES Pn template with detection. The template had only one detection in the Scandinavia mining region during Week 2, and no detections during Week 1. The Pn arrival for the detected event was picked for the REB bulletin.

2.5. Multistation Validation

Multistation validation is a feature in SeisCorr that joins correlation detections that are coincident in origin time and location -- within user-specified tolerances -- to form candidate events. Studies over broad areas under non-aftershock conditions can use tolerances on the order of 30-60 seconds and 150 km because seismicity is typically sparse in location and time. For a network, multistation validation can significantly reduce the overall false alarm rate under the requirement that an event is declared only if it is detected by multiple stations in the network.

Yet small events, such as mining blasts, pose a particular challenge because the events may not be detected at multiple stations. The problem is exacerbated if the network is global and sparse, like the IMS, because the mining blast signals may attenuate below background noise levels before arriving at a second or third station. However, waveform correlation can detect signals near the noise level, allowing more distant stations to corroborate small mining blast events at stations that are closer to the mines.

This study used the multistation validation method with different tolerances to form candidate events. Then we compared the SeisCorr events with the REB events for the weeks of study to choose the best tolerances for time and distance to match the known REB events. Applying those tolerances to additional detections not in the REB gives insight into how many events waveform correlation could detect if added to the operational pipeline. Once the tolerances were determined for Week 1 of each mining region, the same tolerances were chosen for Week 2 of each mining region. Note that we found that different tolerances were needed for Wyoming and Scandinavia.

NOTE: Future research could investigate whether every mining region requires analysis to set tolerances for multistation validation, or if the tolerances can be predicted based on associator parameters, station sensitivity, network geometry, underlying geology and signal attenuation.

Section 2.5.1 presents some example candidate events produced by multistation validation to develop intuition about the SeisCorr method. Section 2.5.2 describes the parameterization study of the Week 1 mining blast detections to select the tolerance values. The following sections will analyze the candidate events for each of the two mining blast regions.

2.5.1. Example candidate events

The purpose of this section is to develop an intuition about the way that SeisCorr uses multistation validation to join correlation detections into candidate events. The correlation jobs for all the stations must be completed prior to running a multistation validation job. The user specifies what correlation jobs to include in a multistation validation job and what distance and time tolerances to use to search for candidate events.

Figure 2-34 shows a map of the Wyoming study region with the Week 1 template events. The white square delineates a subregion that contains example SeisCorr multistation validated events that will be displayed in Figure 2-35 through Figure 2-37.

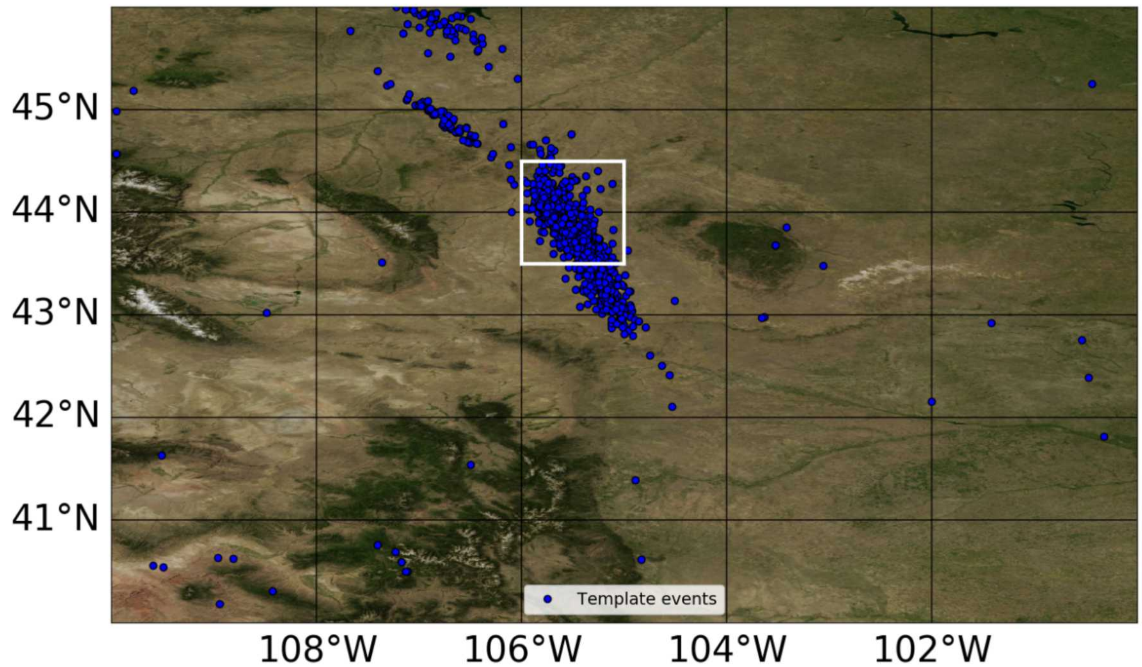

Figure 2-34. Map of Wyoming study region with the Week 1 template event locations (blue). The square (white) outlines the boundaries of the maps shown in Figure 2-35 and Figure 2-37.

Figure 2-35 shows an example multistation validated candidate event from the Wyoming Week 1 study period that was detected by templates from all three stations, superimposed over the map of the template events. The underlying map of template events is shown here to indicate the locations of the detecting template events as well as the locations of all the template events, and to provide the scale for the clustered detections. This is an active mining area and the templates are dense. The candidate event is based on coincident detections with the tolerances of 100 km and 60 s. SeisCorr does not require that the detecting templates all be based upon the same template event; thus, we make use of the metadata from every detecting template to calculate information about the candidate event.

The legend indicates information about both the template event and the detection for each detecting template. In the case of this example, all three templates are from different template events. For example, the detecting TXAR template (#1) is based on an event from 2012-01-24, while the PDAR template (#2) is based on event from 2016-11-23. The detection at PDAR was on 2018-04-04 21:35:16, that at NVAR on 2018-04-04 21:37:06, and that at TXAR on 2018-04-04 21:42:06.

In SeisCorr, the location of the candidate event is estimated using an algorithm that averages over the locations of the events for the detecting templates. The location of the candidate event is indicated in Figure 2-35 by a white star, and the calculated origin time of the candidate event is indicated in the legend as 2018-04-04 21:34:28. The location calculated for the candidate event is heavily weighted towards the location of the majority of the detecting templates for this event, which places it near the location of the PDAR template in this example.

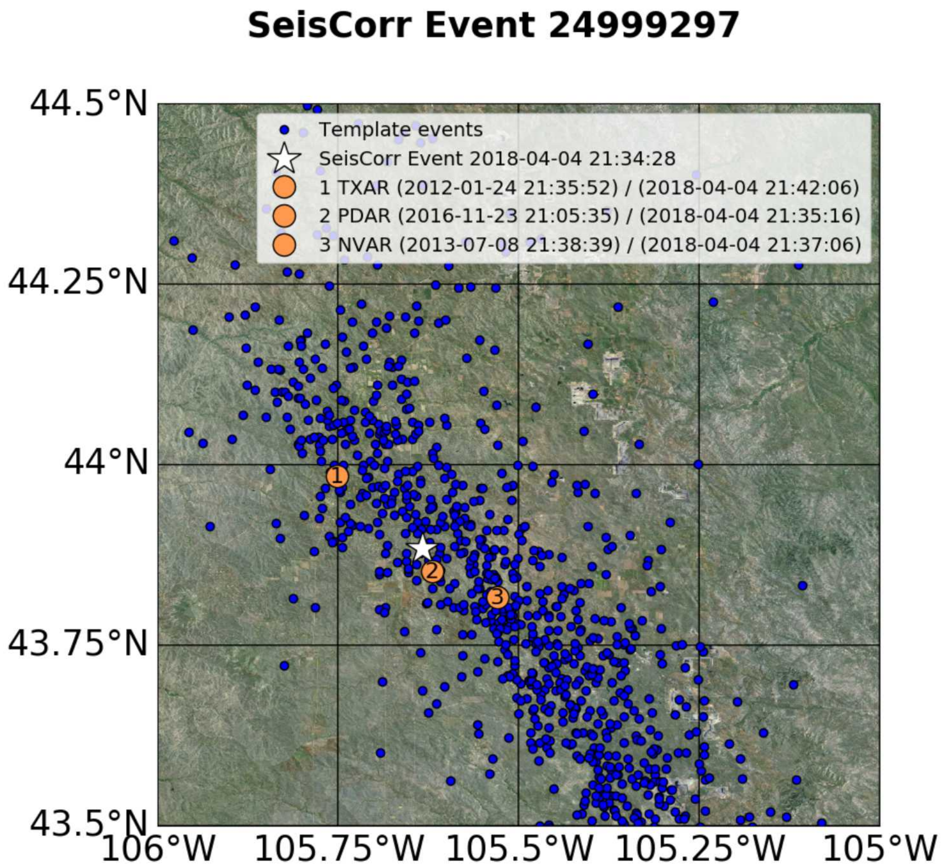

Figure 2-35. Example SeisCorr multistation validated candidate event 24999297 (white star) from Wyoming Week 1 coincident detections using tolerances of (100 km, 60 s). The three orange circles indicate the locations of detecting templates. The legend indicates the template station, template event date/time (UTC), detection date/time (UTC), and a number chosen for convenience to label the location of the template on the map.

The waveforms associated with SeisCorr event 24999297 are shown in Figure 2-36 to expand on the example. The template waveforms are shown in red, with the detected waveforms shown in blue. The array elements are lined up in red/blue pairs to aid with visual comparison, and the cross-correlation score is indicated with the label "Score". The cross-correlation score is an average of the cross-correlation scores over all the elements of the array.

Let us assume that we agree with the detections shown in the waveform plots. If the template events and the detected events have similar source mechanisms, then we would expect decorrelation of the waveform with increasing distance from the source. In other words, if the candidate event is located optimally relative to the template locations shown in Figure 2-35 then the nearest template location (PDAR) should have the highest correlation score (0.479) and the furthest template (TXAR) should have the lowest correlation score (0.398), while the remaining template (NVAR) has a correlation score between the other two (0.414). The cross-correlation scores vary with distance as expected, which gives more confidence in the detection and in the calculated multistation validated event location. Note that this decorrelation with distance argument only works for a similar source mechanism so it is particularly suited for an active mining region, although we acknowledge that reality is more complicated than our simplified discussion.

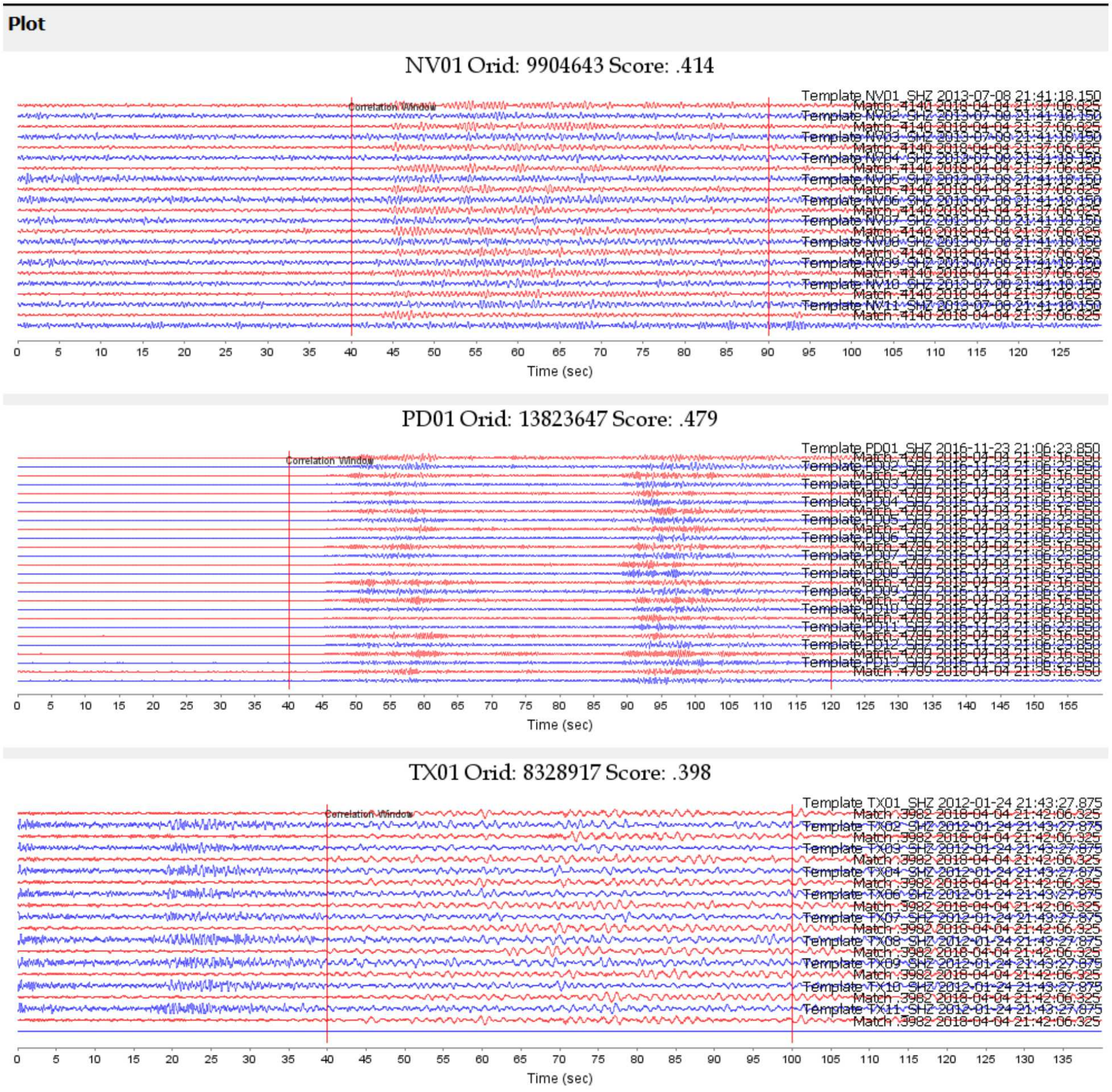

Figure 2-36. Waveforms from the templates and detections from Figure 2-35. The template waveforms are shown in red, while the detected waveforms are shown in blue. The top group of waveforms are from array NVAR. The middle group is from array PDAR. The bottom group is from array TXAR. The cross-correlation score is indicated for each station.

A second example SeisCorr candidate event 24999401 is shown in Figure 2-37. This example candidate event is interesting because the NVAR (#2) and PDAR (#3) templates are from the same template event (same location and event date/time). The orange circles for NVAR and PDAR overlap on the map, and the legend shows the template event date/time of 2015-01-11 21:06:17. The fact that two stations detected this recurring event lends credibility to the detection.

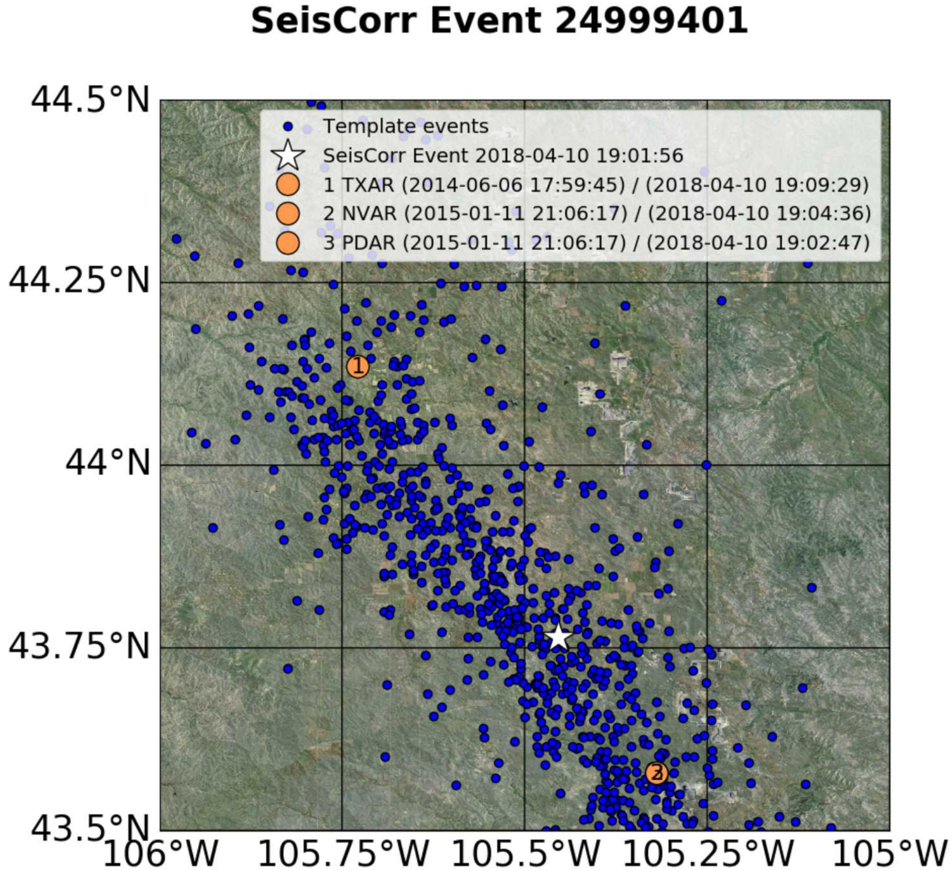

Figure 2-37. Example SeisCorr multistation validated candidate event 24999401 (white star) from Wyoming Week 1 coincident detections using tolerances of (100 km, 60 s). The orange circles indicate the locations of detecting templates. The legend indicates the template station, template event date/time (UTC), detection date/time (UTC) , and a number chosen for convenience to label the location of the template on the map.

Before leaving examples from the Wyoming region, let's recall the example NVAR template discussed in section 2.4 that was poorly windowed and which resulted in 5 false detections (Figure 2-32). The detections shown in that figure were not combined with any other detections within a tolerance of (100 km, 60 s) to result in a candidate event with NDEF > 1. In this case, events were sparse enough in distance and time to prevent an inaccurate corroboration of the false detections, showing the usefulness of multistation validation. Nevertheless, SeisCorr creates an event for every detection, so there are 5 single-station events at the location of the poorly windowed template with NDEF = 1 for the 5 false detections.

We prefer events that have been detected at multiple stations, but mining blasts are sometimes too small to be detected at multiple stations. However, in considering this example with 5 false detections, we would not expect to have detections at NVAR without a corroborating detection at PDAR since PDAR is much nearer than NVAR to the (false) detected event location. Looking towards the future implementation of waveform correlation in an operational pipeline, associators will be required to make intelligent decisions about the combination of waveform correlation detections with other corroborating signals. This will be easier if the location of the detected events, a crucial advantage of waveform correlation detections, can be trusted; thus, we reiterate that template selection will be critical to minimize the possibility of correlations with non-collocated events.

NOTE: The IDC plans to incorporate waveform correlation into a prototype automated pipeline that generates correlation detections above thresholds, but an associator (e.g., NET-VISA [11]) will build events by associating the correlation detections to the events from the standard event lists. The authors encourage the IDC to consider a display of correlation templates and detections in a manner that will allow analysts to easily include those detections when they are credible or discard them when they are dubious.

2.5.2. Time/location tolerances for multistation validation

SeisCorr has been used to study global and regional seismicity, where events are sparse in distance and time, and for aftershocks where the events are so close in distance and time that it is difficult to separate them. This study of mining regions with IMS stations alone offers an opportunity to study recurring events that are close in distance but sufficiently time-separated to use long duration templates that produce credible detections. Because the events are sufficiently time-separated, the study of time/location tolerances for multistation validation was simpler than the parameterization study employed in the aftershock study [3]. We used knowledge of the REB events to establish tolerances that did not split (tolerances too tight) or combine (tolerances too loose) mining events. We established the tolerances for Week 1 of each mining region, then applied the same tolerances to Week 2 of the same mining region after verifying similar results.

Using the REB events to establish tolerances seemed to be an acceptable compromise for this study to quickly determine the best multistation validation tolerances. The mining region's operation is likely to be consistent from one week to the next, so in an operational setting it would be feasible to use prior week's events to establish tolerances for the following week of detections. Finally, we do not suggest that our approach to establishing these tolerances are research results, but rather are a step in our process that must be explained and documented; thus, this section is not meant to belabor a straightforward technical decision. The process is enumerated below:

1. After running all the correlation jobs for a mining region and week of study, multistation validation jobs were run with different tolerances for the same set of correlation detections. These jobs resulted in events with different grouped detections. For example, for Wyoming Week 1, a multistation validation job with tolerances of (100 km, 60 s) resulted in 9 events with NDEF > 1, while a multistation validation job with tolerances of (80 km, 50 s) resulted in 8 events with NDEF > 1. Decreasing the tolerances decreased the number of events with NDEF > 1 from 9 to 8. In contrast, a multistation validation job with increased distance tolerances of (150 km, 60 s) also resulted in 9 events with NDEF > 1. These three examples suggest that (150 km, 60 s) is an adequate tolerance to avoid combining events, while (80 km, 50 s) split at least one event.

2. The next step involved comparison with the REB. For example, for Wyoming Week 1 there were 13 known REB events. Each of the multistation validation jobs' events were compared with the known REB events to determine which comparison tolerances in a range of 1.5° -- 0.5° distance and ±30 s -- ±10 s best reproduced the 10 REB events that were detected by waveform correlation. Using a multistation validation tolerance of (50 km, 15 s) split

waveform correlation events, and this became obvious because 2 or 3 waveform correlation events would be likely to match the same REB event. The goal of this process step was to choose both a multistation validation tolerance and REB comparison tolerance to allow every REB event that was detected by waveform correlation to match one and only one waveform correlation event. The final tolerances, which are specific to each mining region, are shown in Table 2-7.

Table 2-7. Tolerances for Multistation Validation and REB Comparison.

Region/Timeframe	Multistation Validation Tolerances (km, s)	REB Comparison Tolerances (deg, s)
Wyoming Week 1	(100, 60)	(1.0, ±15)
Wyoming Week 2	(100, 60)	(1.0, ±15)
Scandinavia Week 1	(50, 15)	(1.0, ±15)
Scandinavia Week 2	(50, 15)	(1.0, ±15)

NOTE: Further research may be required in this area to develop an approach to optimize time and distance tolerances for combining detections coincident in distance and time to detect mining events with waveform correlation in the context of a global network. As we can see from this study, each mining region required a different set of tolerances to optimally detect the known REB events. In an operational pipeline, the presence of a sophisticated associator may eliminate this issue.

2.5.3. Wyoming Week 1 (April 4–10, 2018) Events

The locations of events detected by waveform correlation for Wyoming Week 1 are shown in the map (Figure 2-38) with superimposed REB events. There were 136 distinct events confirmed by multistation validation using the tolerances of 100 km and 60 s for the three IMS stations (PDAR, NVAR, TXAR); only 9 of the 136 events were detected by two or more stations. 10 of the 13 REB events were detected by waveform correlation events within tolerances of 1° and ±15 s. The three REB events that were not detected by waveform correlation are in an area with only one template event (Figure 2-5) and waveform correlation can only detect recurring events; thus, it is not surprising that these events went undetected.

We do not have ground truth for the majority of the waveform correlation detections at the time this report was written. This map includes single station events, which make up most of the detections; a single-station detection means that a detection was made by one template (station) but there were no other validating templates within the chosen multistation validation tolerance. In the case of a single station event, the location is assumed to be the location of the template event. Most single station events were detected by the nearest station, PDAR.

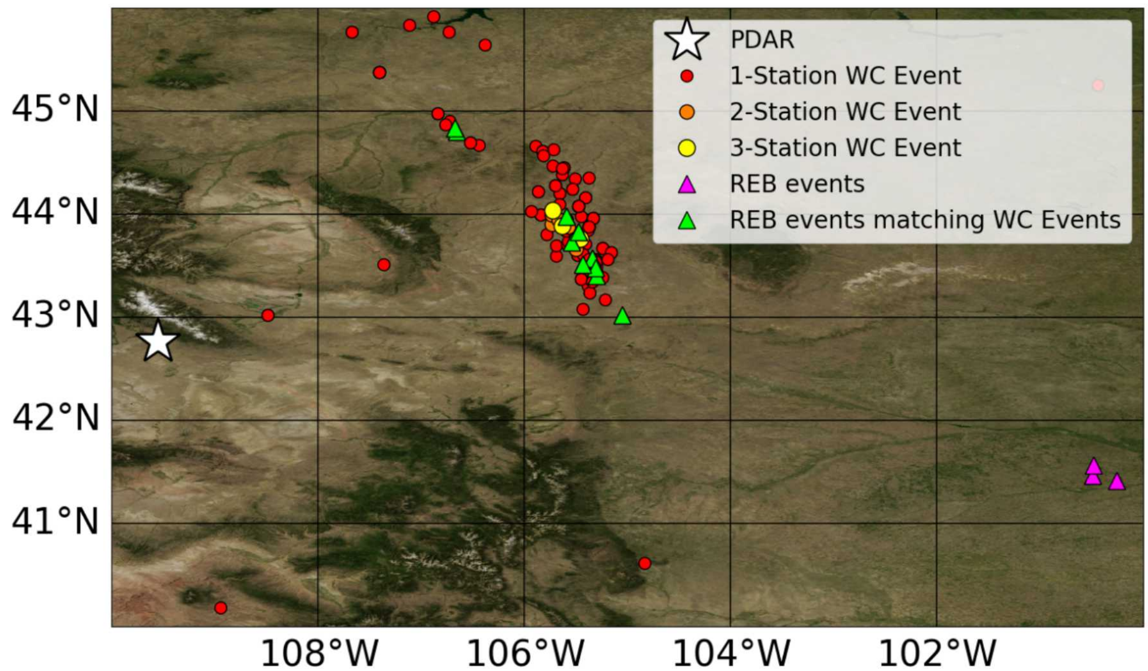

Figure 2-38. Map showing the locations of 136 waveform correlation events (circles) versus 13 REB events for Wyoming Week 1. The size and color of the circle indicates the number of detecting stations. 10 REB events were detected by waveform correlation (green triangles), while 3 REB events were not detected by waveform correlation (magenta triangles).

The bar chart (Figure 2-39) shows the number of Wyoming Week 1 template detections by template age, and further divided by station. The age of a template is the difference between the date of the template event and the date of the detected event. This bar chart shows that most of the detections were made by templates from events within the last 3 years; moreover, most of the detections were made by station PDAR in all years for which PDAR had templates. Sandia's waveform archive did not have waveform data for PDAR for years 8 or 9 so there are no PDAR templates for those years (Figure 2-8).

We expect that in a mining region the efficiency of a template will decline over time as the structure of the mine changes. In general, the data in the bar chart confirms that expectation. However, we see that some older templates also detect events; we do not have any analysis at this time to determine whether the older templates are from tectonic events (i.e. not mining blasts) that recurred in the region. Prior regional studies using waveform correlation have suggested that seismic faults can be detected with old templates when they reactivate [10][10]. The map of events (Figure 2-38) shows some scattered single-station events that may not be mining blasts. Further analysis beyond the scope of this report would be required to investigate whether older templates are detecting events from source types that are not mining blasts.

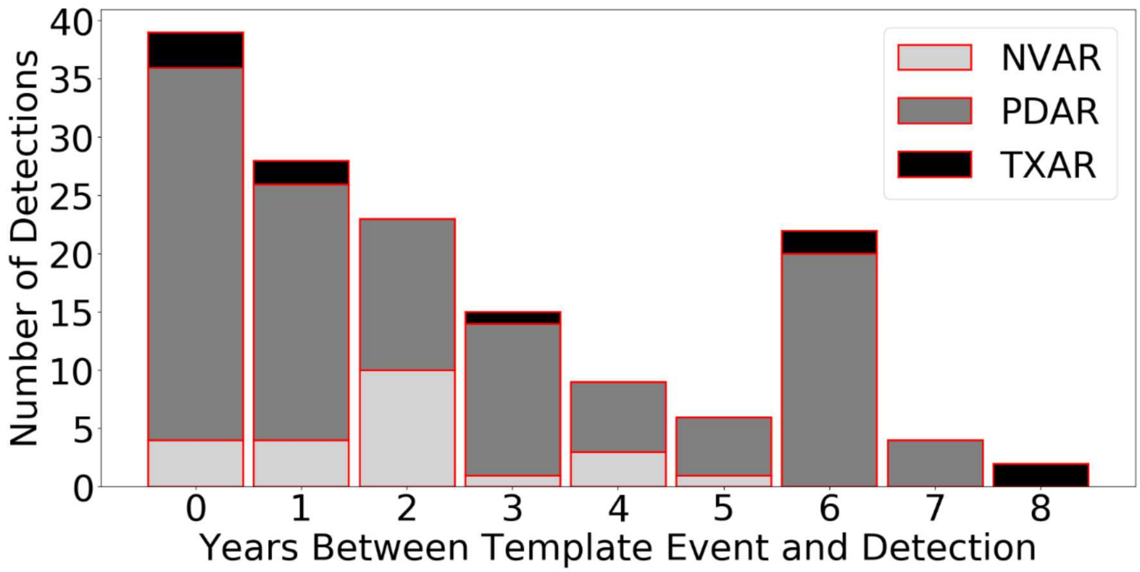

Figure 2-39. The number of Wyoming Week 1 detections by age of template, and by station.

The 10 waveform correlation events that match the REB events are mapped in Figure 2-40, along with the REB events. The purpose of this map is to identify which waveform correlation events were matched to the REB events, and to provide a visual indication of the proximity of the matched events.

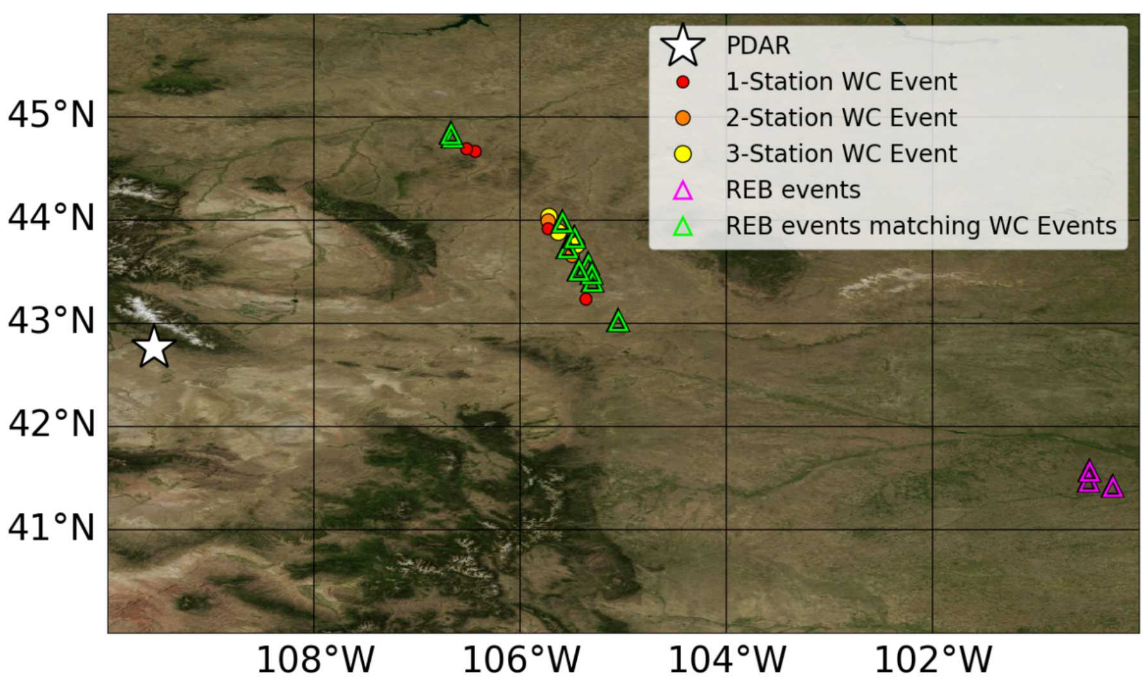

Figure 2-40. Map showing the locations of the 10 waveform correlation events (circles) that matched REB events within the comparison tolerance of (1°, ±15 s) versus the 13 REB events for Wyoming Week 1. The size and color of the circle indicates the number of detecting stations. 10 REB events were detected by waveform correlation (green triangles), while 3 REB events were not detected by waveform correlation (magenta triangles).

The number of waveform correlation detections that match REB events is shown in the bar chart (Figure 2-41). It is clear from this chart that older templates detect bulletin events and are therefore important to analysts. We recommend that template libraries include historical templates as well as recent templates until there is a more complete understanding of the useful longevity of waveform templates in a waveform correlation system.

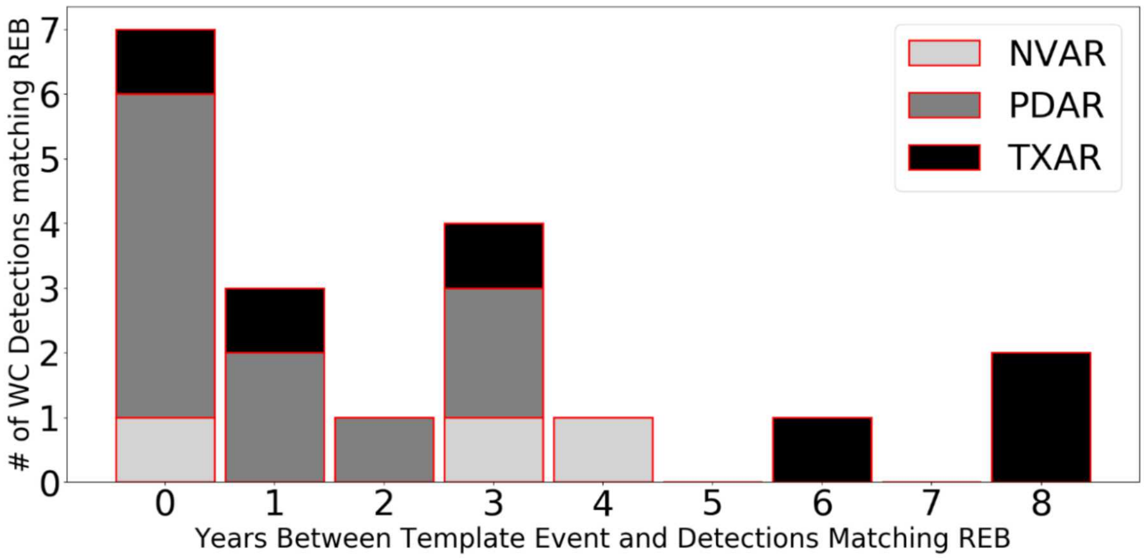

Figure 2-41. Number of Wyoming Week 1 waveform detections matching the REB, by age and by station.

2.5.4. Wyoming Week 2 (August 8–14, 2018) Events

The locations of events detected by waveform correlation for Wyoming Week 2 are shown in the map (Figure 2-42) with superimposed REB events. There were 112 distinct events confirmed by multistation validation using the tolerances of 100 km and 60 s for the three IMS stations (PDAR, NVAR, TXAR); only 9 of the 112 events were detected by two or more stations. 9 of the 10 REB events were detected by waveform correlation events within tolerances of 1° and ±15 s. The single REB event not detected by waveform correlation is in a dense cluster of template events (Figure 2-12) and other waveform correlation detections, which makes this an interesting anomaly that was investigated further.

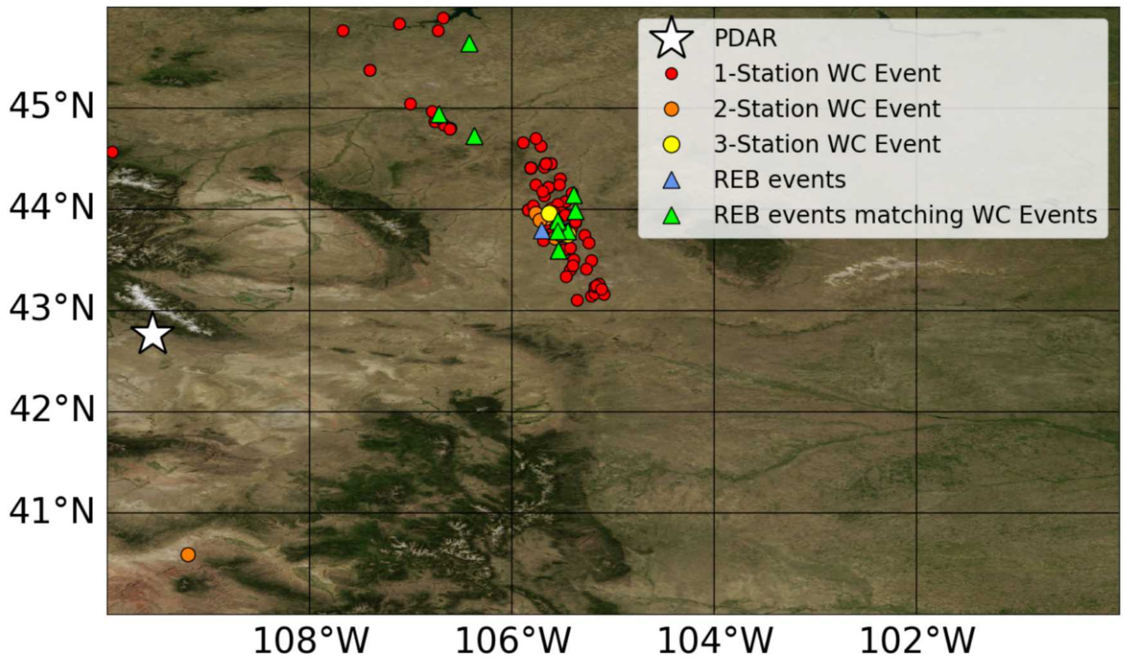

Figure 2-42. Map showing the locations of the 112 waveform correlation events (circles) versus the 10 REB events for Wyoming Week 2. The size and color of the circle indicates the number of detecting stations. 9 REB events were detected by waveform correlation (green triangles), while 1 REB event was not detected by waveform correlation (blue triangle).

Figure 2-43 shows the waveforms for the undetected REB event at station PDAR. The signals are clearly visible and there is no immediately obvious reason why this event should not have been detected by waveform correlation.

The reason for the failure to detect this REB event is quite mundane once investigated. The researcher (the first author claims the credit for the error) accidently entered 2018-08-14 00:00:00 as the end date for the correlation job instead of 2018-08-15 00:00:00, entirely missing the final day of the study period during the detection runs. This undetected event was on 2018-08-14 20:22, and the continuous data was excluded from the correlation jobs that form the basis of the analysis for Wyoming Week 2.

There was inadequate time to run all the correlation jobs, multistation validation jobs, and re-run the analysis to make the figures in this report prior to the funding deadline. However, running a PDAR correlation job for the single day of 2018-08-14 resulted in an additional 8 detections, including the previously missed REB event (Figure 2-44).

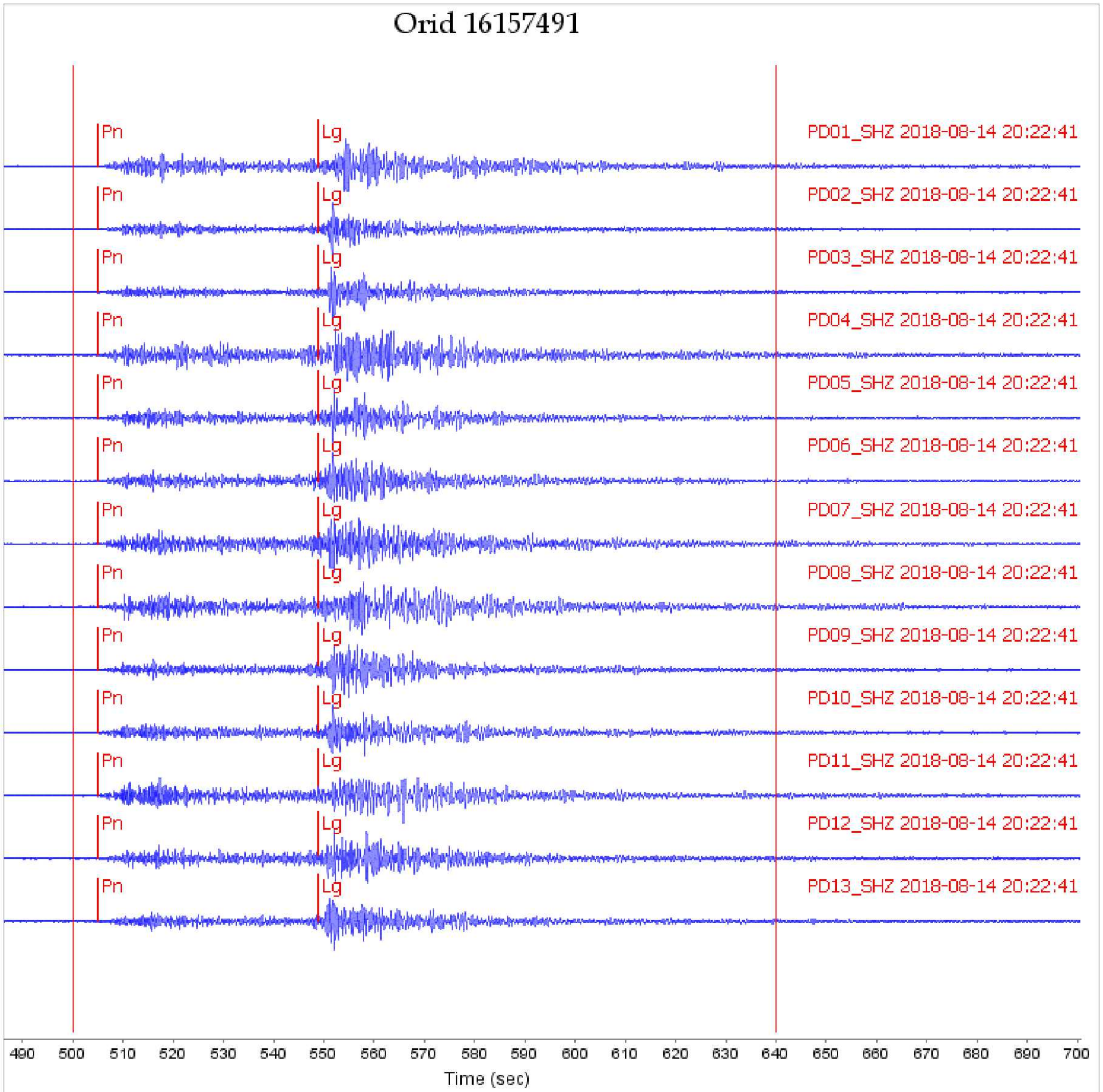

Figure 2-43. Waveform plot of Wyoming Week 2 undetected REB event on 2018-08-14.

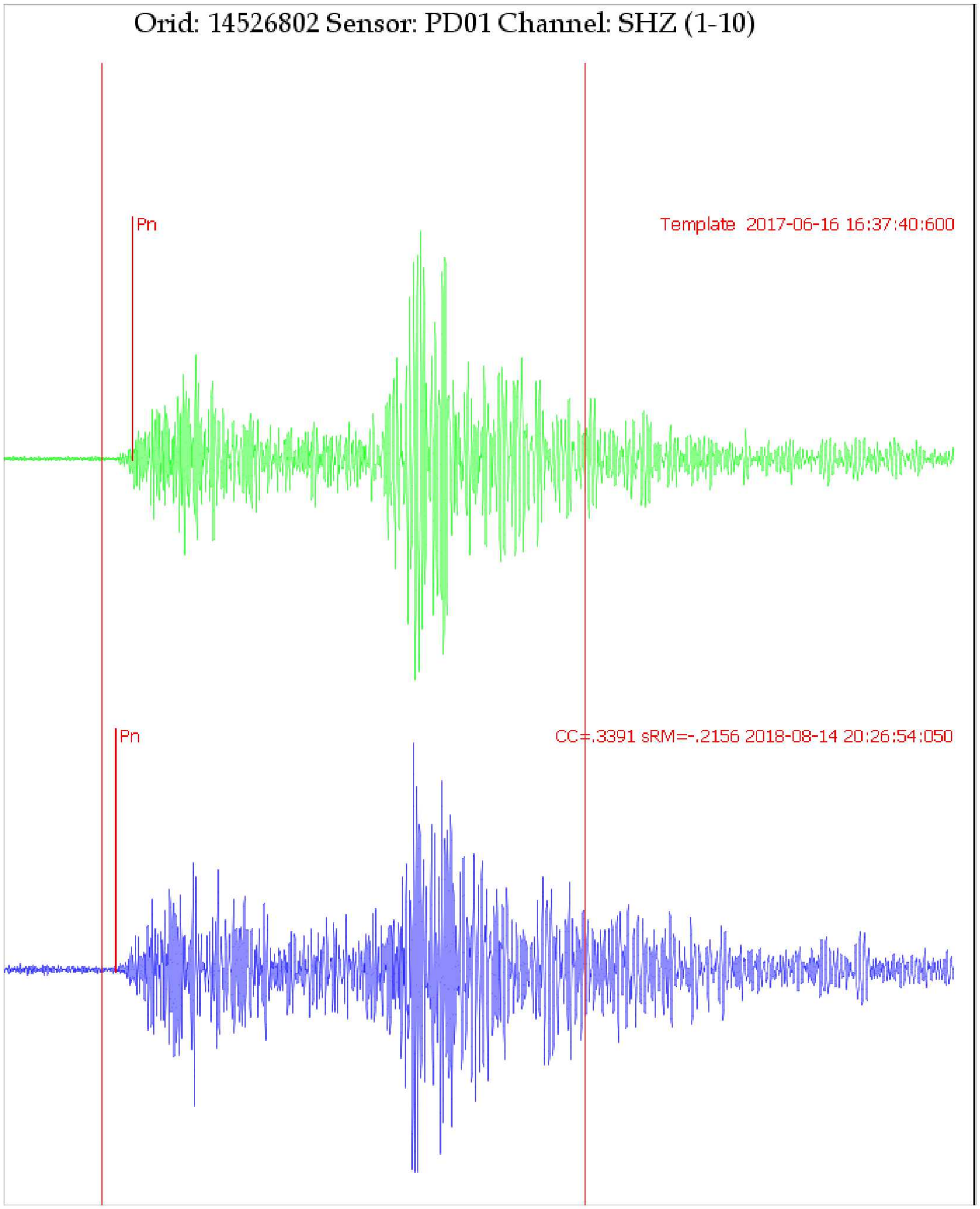

Figure 2-44. Wyoming Week 2 detection of missing REB event on 2018-08-14 with additional PDAR correlation job.

The bar chart (Figure 2-45) shows the number of Wyoming Week 2 template detections by template age, and further divided by station. Similarly to the analogous bar chart for Week 1, the chart for Week 2 shows that most of the detections were made by templates from events within the last 3 years and that most of the detections were made by station PDAR in all years for which PDAR had templates. Similarly to Week 1, we recommend further analysis to investigate whether older templates are detecting events from source types that are not mining blasts.

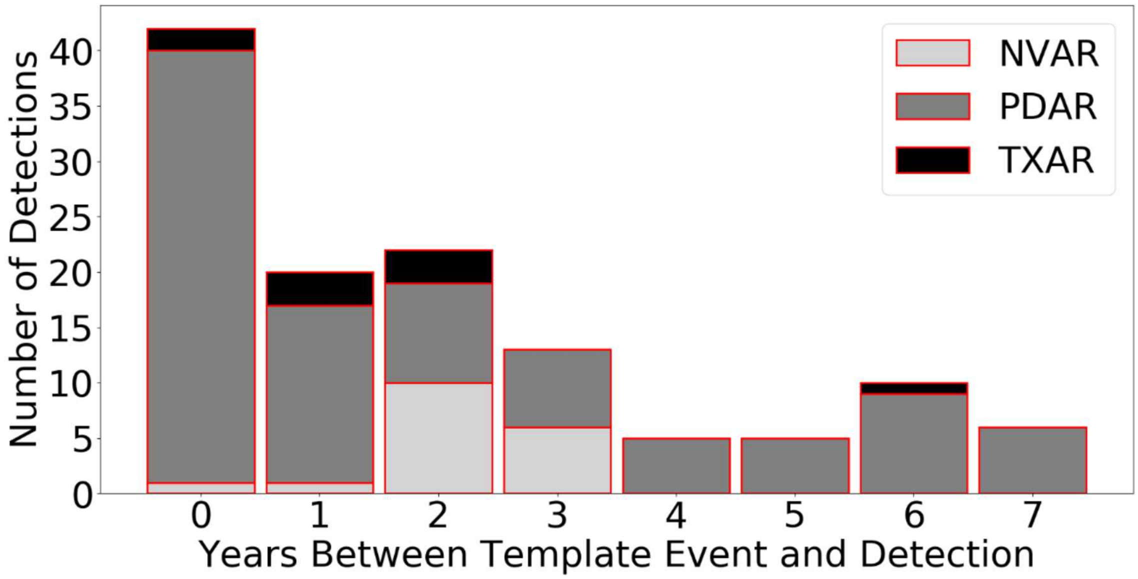

Figure 2-45. The number of Wyoming Week 2 detections by age of template, and by station.

The 9 waveform correlation events that match the REB events are mapped in Figure 2-46, along with the single REB event that was not detected due to a processing error. The purpose of this map is to identify which waveform correlation events were matched to the REB events, and to provide a visual indication of the proximity of the matched events.

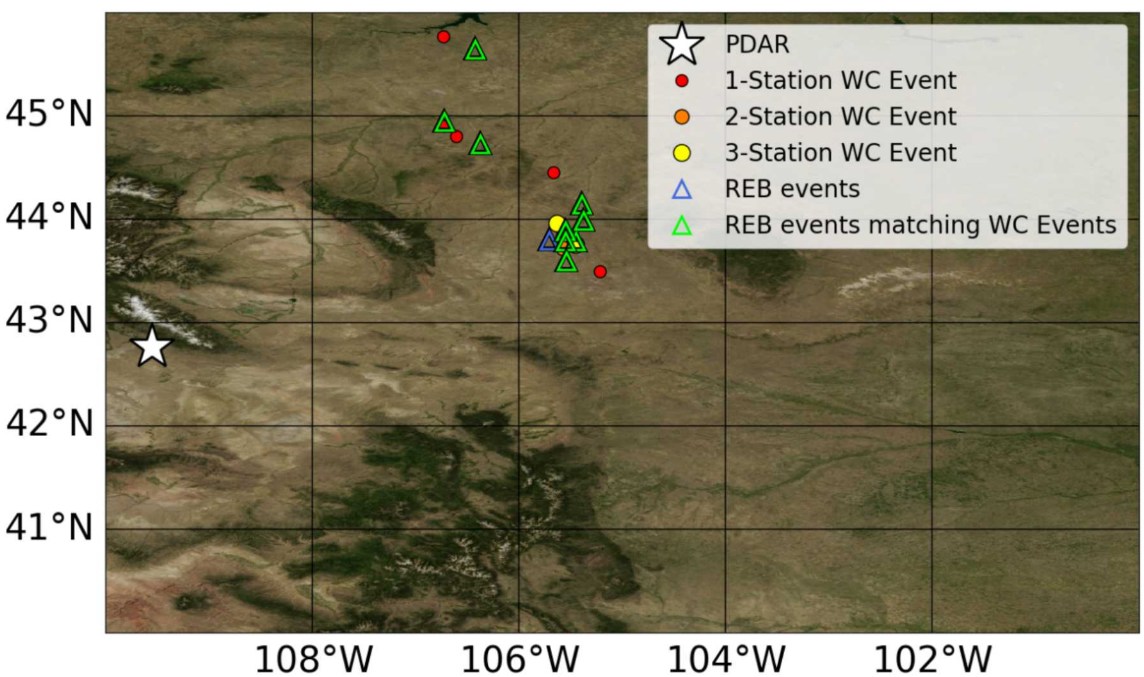

Figure 2-46. Map showing the locations of the 9 waveform correlation events (circles) that matched REB events within the comparison tolerance of (1°, ±15 s) for Wyoming Week 1. The size and color of the circle indicates the number of detecting stations. 9 REB events were detected by waveform correlation (green triangles), while 1 REB event was not detected by waveform correlation (blue triangle).

The number of Wyoming Week 2 waveform correlation detections that match REB events is shown in the bar chart (Figure 2-47). Unlike the Week 1 results, older templates were less effective in detecting bulletin events for Week 2, which may indicate that the sources are mines. We still recommend a diverse template library, including historical templates as well as recent templates, until template effectiveness by age is better understood.

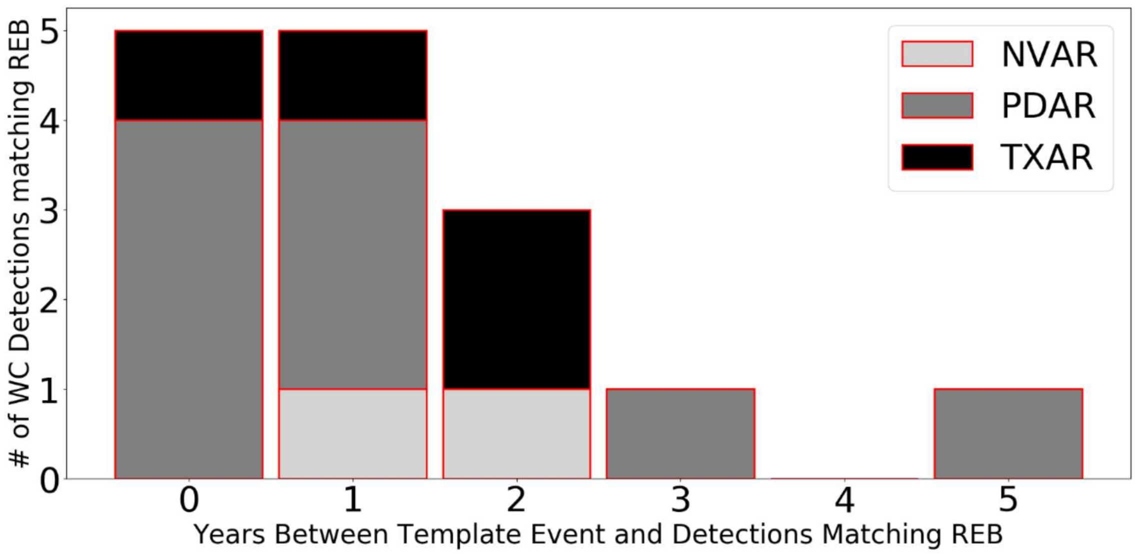

Figure 2-47. Number of Wyoming Week 2 waveform detections matching the REB, by template age and by station.

2.5.5. Scandinavia Week 1 (February 12–18, 2018) Events

The locations of events detected by waveform correlation for Scandinavia Week 1 are shown in the map (Figure 2-48) with superimposed REB events. There were 390 distinct waveform correlation events. Using multistation validation tolerances of 50 km and 15 s for the three IMS stations (ARCES, FINES, NOA), 49 of the 390 events were detected by two or more stations. 14 of the 19 REB events were detected by waveform correlation events within tolerances of 1° and ±15 s.

Three of the five REB events that were not detected by waveform correlation were located near (68°N, 37°E) without nearby template events (Figure 2-18) and waveform correlation can only detect recurring template events; thus, it is not surprising that these events went undetected. Moreover, investigation into these three events showed that they are infrasound only arrivals in the REB association/arrival tables with no associated seismic arrivals for the events.

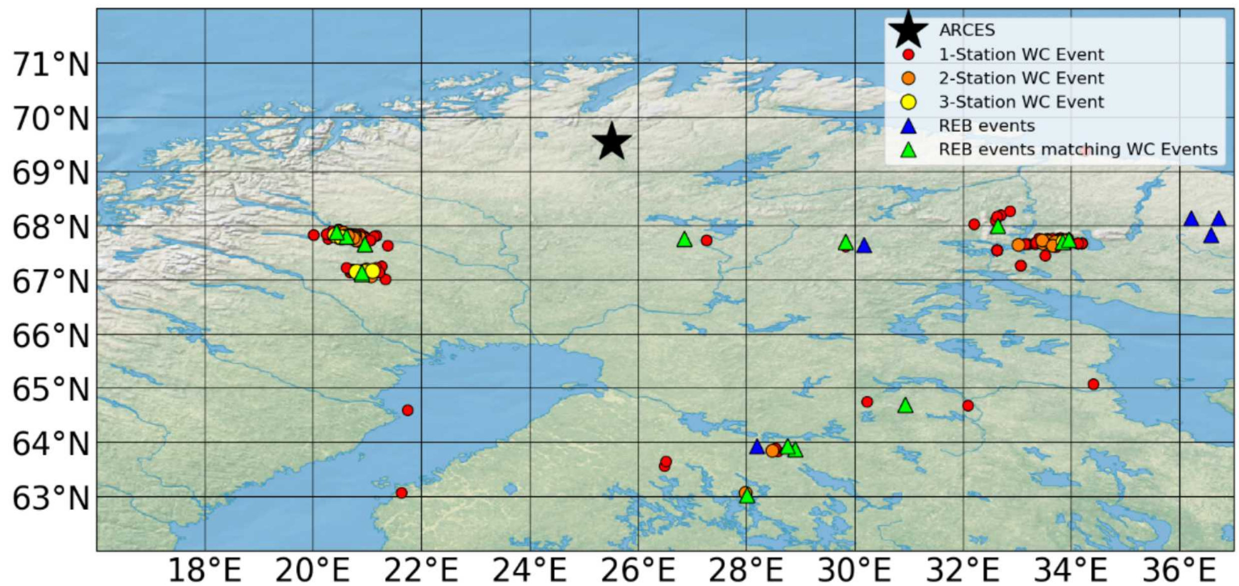

Figure 2-48. Map showing the locations of the 390 waveform correlation events (circles) versus the 19 REB events for Scandinavia Week 1. The size and color of the circle indicates the number of detecting stations. 14 REB events were detected by waveform correlation (green triangles), while 5 REB events were not detected by waveform correlation (blue triangles).

The remaining two REB events that were not detected by waveform correlation were investigated in more detail. The REB ORID 15464786 (67.649872°N, 30.164704°E, 17-Feb-18 11:27:18) was not detected by waveform correlation in this study. In the REB, the event was detected by only one seismic station, ARCES, and the infrasound station I37NO. The ARCES waveform plots of the arrivals (Figure 2-49) show additional unpicked arrivals at about 570 and 610 seconds, still within the template window indicated by the vertical red lines. Waveform correlation detection can be defeated by overlapping events within the template window, which may have been the reason this event was not detected by a waveform template from ARCES.

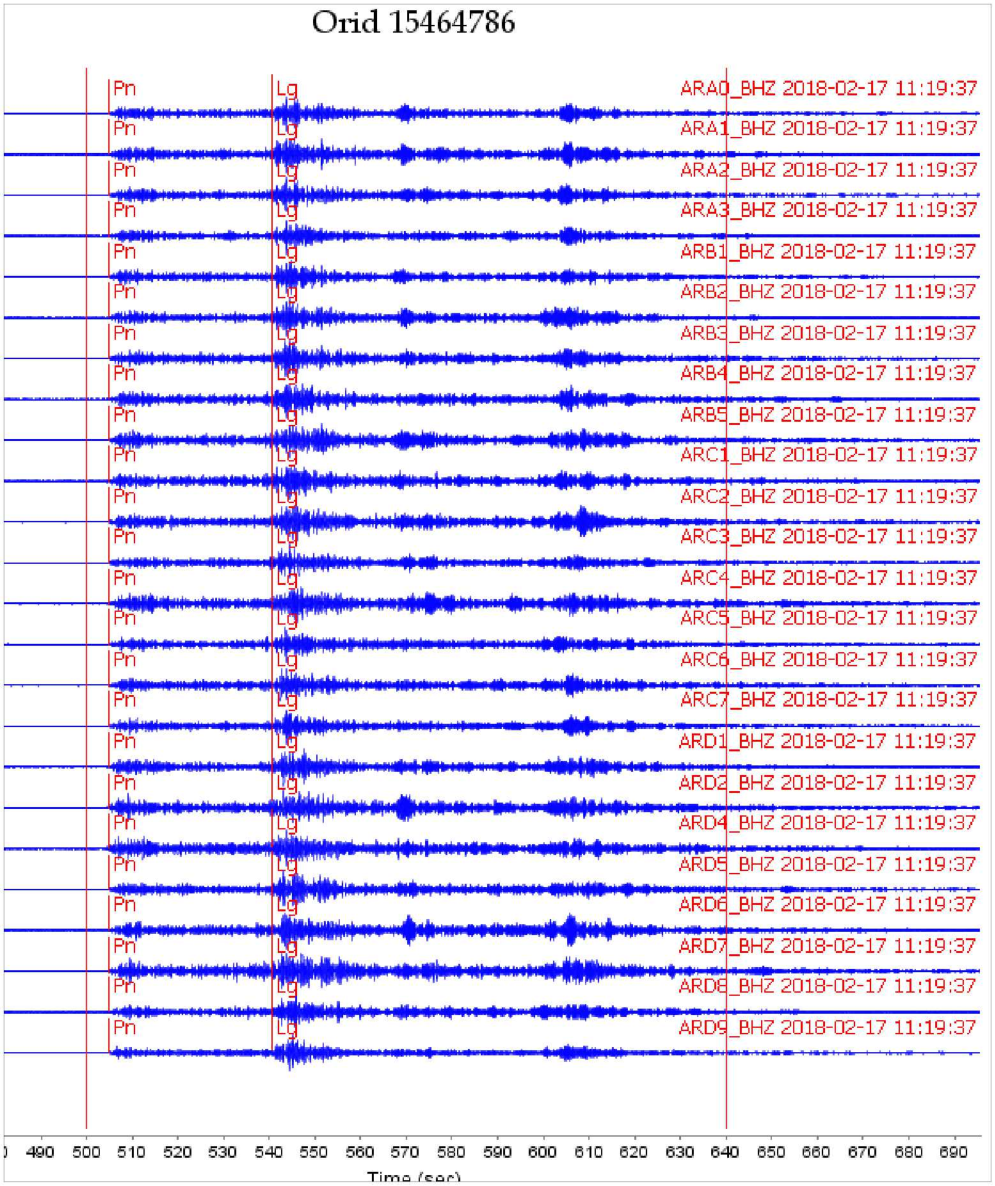

Figure 2-49. ARCES waveform plots of Scandinavia Week 1 undetected REB event on 2018-02-17.

The REB ORID 15460810 (63.933705°N, 28.196828°E, 12-Feb-18 08:00:05) was not detected by waveform correlation in this study. In the REB, the event was detected only by one seismic station, FINES, and the infrasound station I37NO. The FINES waveform plots of the arrivals (Figure 2-50) show picked Pn and Sn arrivals, but the Pn phase amplitude is close to the noise level. A FINES template for element FIA0 from REB ORID 12934868 (63.962, 28.036, 04-Mar-16 12:27:50) is shown for comparison (Figure 2-51) because the distance from the station and locations of the two events are similar; the Pn in the template waveform is above the noise floor. The example template is typical of the REB events in this area, so our conclusion is that the missed event is marginal and correlation with the waveform templates from the immediate vicinity failed to exceed the template cross-correlation thresholds.

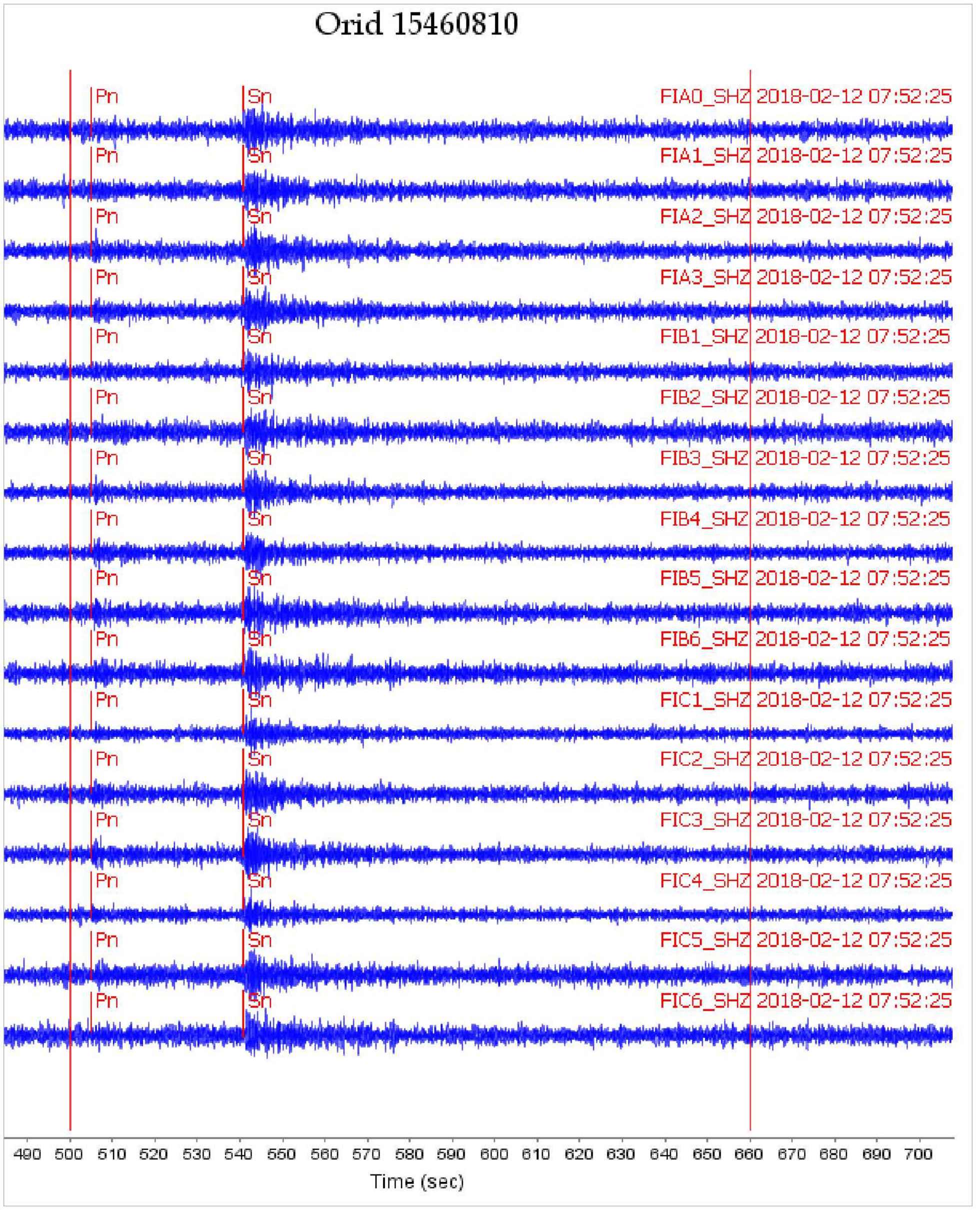

Figure 2-50. FINES waveform plots of Scandinavia Week 1 undetected REB event on 2018-02-12.

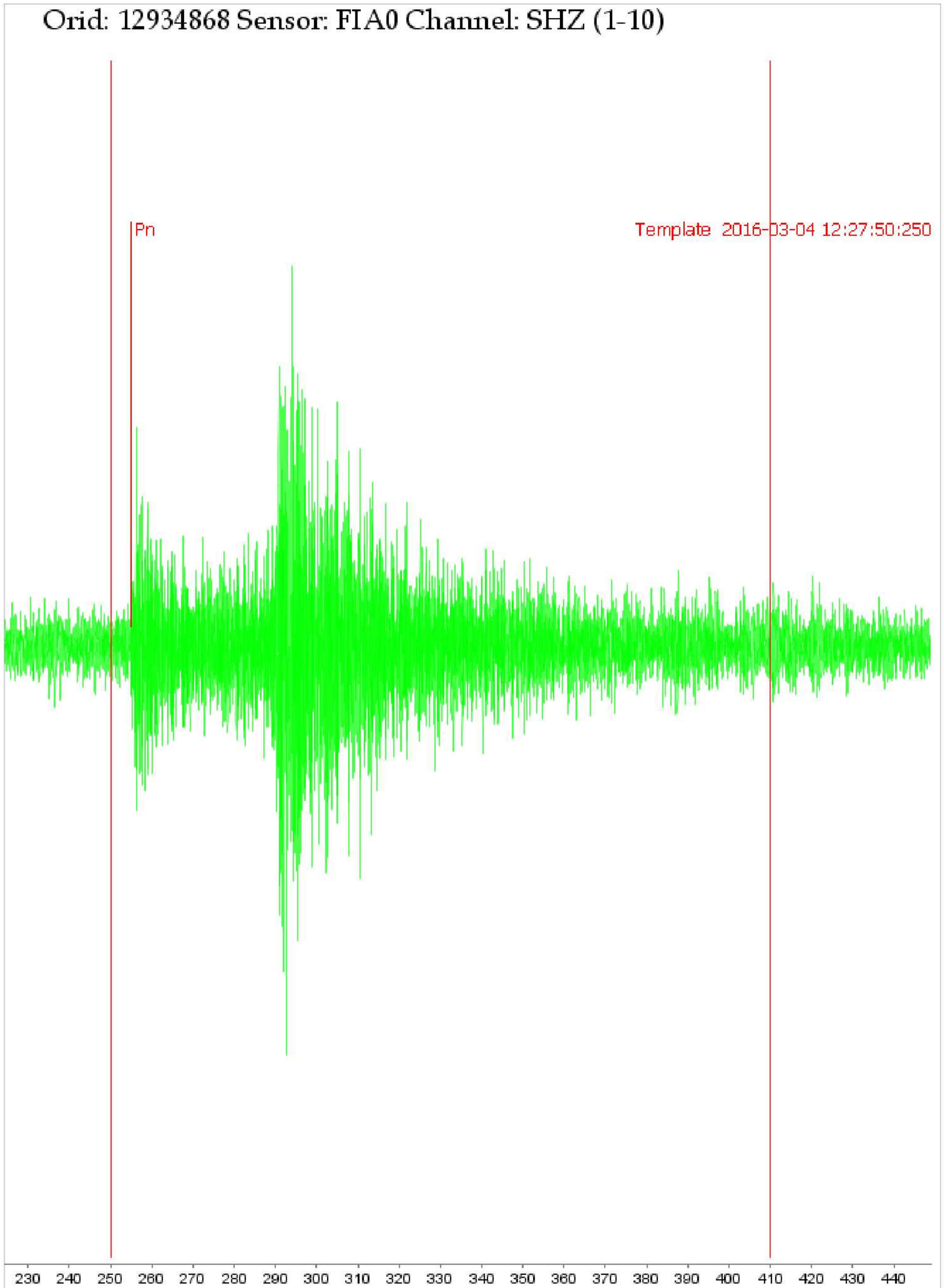

Figure 2-51. Example FINES template from REB ORID 12934868 (63.962, 28.036, 04-Mar-16 12:27:50) shown for comparison with missed REB event in Figure 2-50.

The bar chart (Figure 2-52) shows the number of Scandinavia Week 1 template detections by template age, and further divided by station. This bar chart shows that most of the detections were made by templates from events within the last 3 years for this region; moreover, most of the detections were made by station ARCES in all years for which ARCES had templates. Sandia's waveform archive did not have waveform data for ARCES beyond year 3 so there are no ARCES templates for those years (Figure 2-21). Our historical archive was most complete for station FINES.

We expect that in a mining region the efficiency of a template will decline over time as the structure of the mine changes; for this region the expected trend is shown in the bar chart. The effectiveness of the templates drops dramatically with time; for example, the number of detections from templates less than 1 year of age were almost twice the number of detections from templates that are 1 year old.

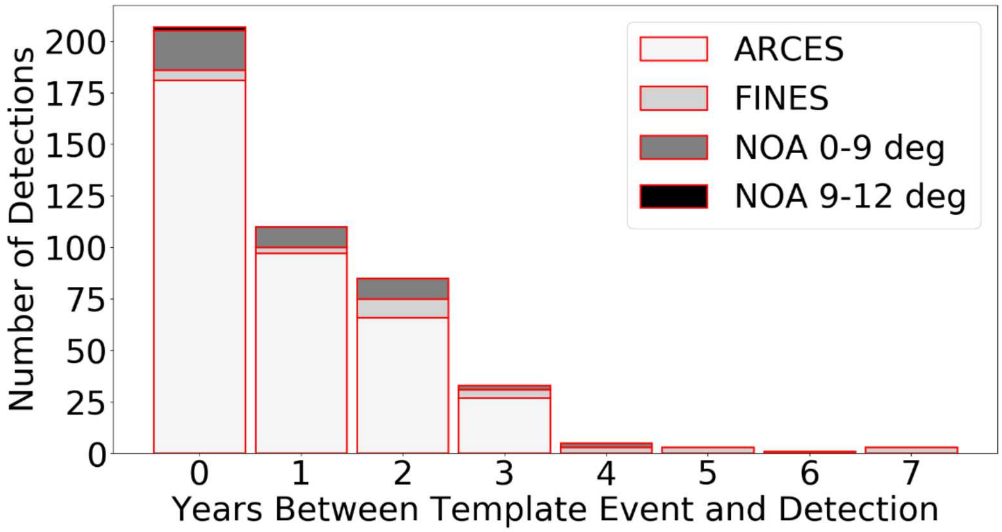

Figure 2-52. The number of Scandinavia Week 1 detections by age of template, and by station.

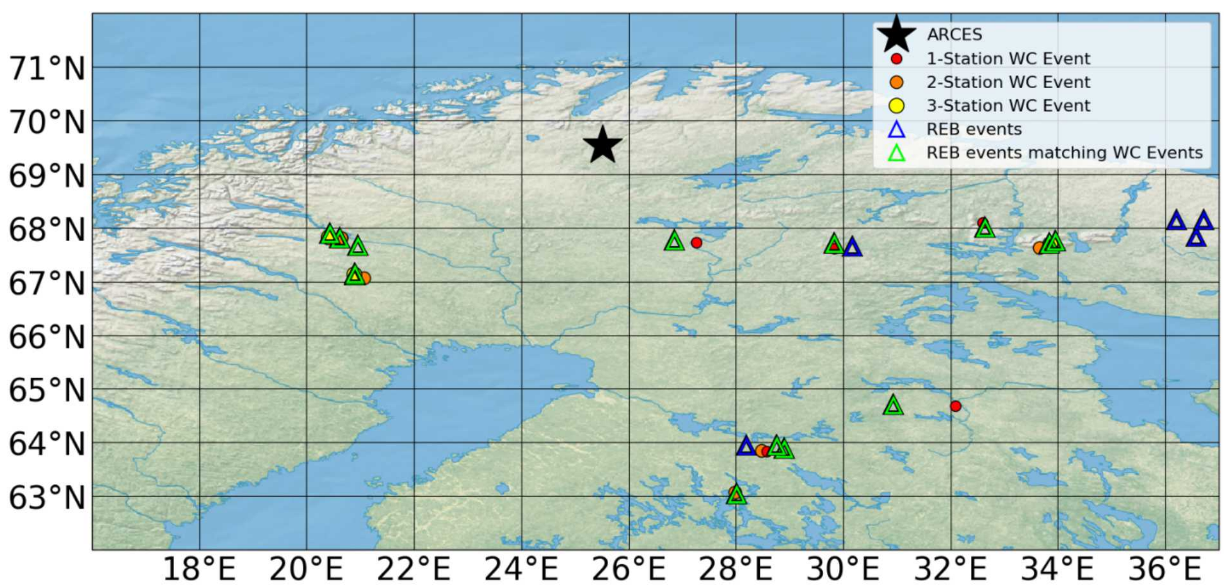

Figure 2-53. Map showing the locations of 14 waveform correlation events (circles) that matched REB events within the comparison tolerance of (1°, ±15 s) versus 19 REB events for Scandinavia Week 1. The size and color of the circle indicates the number of detecting stations. 14 REB events were detected by waveform correlation (green triangles), while 5 REB events were not detected by waveform correlation (blue triangles).

The 14 waveform correlation events that match the REB events are mapped in Figure 2-53, along with the REB events. The purpose of this map is to identify which waveform correlation events were matched to the REB events, and to provide a visual indication of the proximity of the matched events.

The number of waveform correlation detections that match REB events for Scandinavia Week 1 is shown in the bar chart (Figure 2-54). For this mining region, older templates are not contributing detections that are useful for building bulletin events. This contrasts with the Wyoming region, where older templates were effective detectors. This result is in conflict with other studies[10] [10], and we recommend including older templates in the template library for the region unless this result is confirmed over longer time periods.

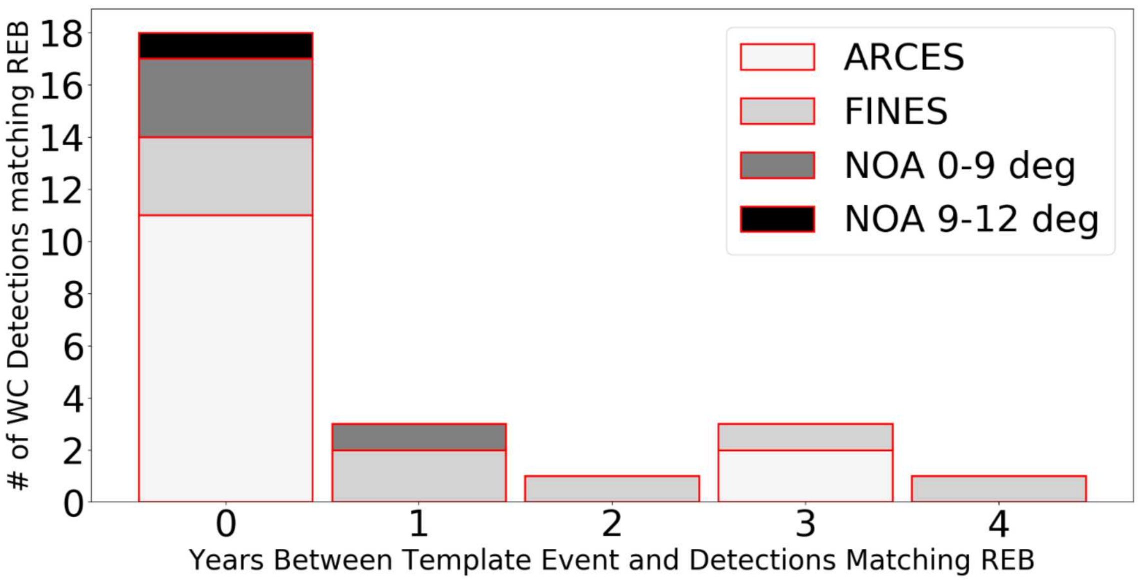

Figure 2-54. Number of Scandinavia Week 1 waveform detections matching the REB events, by age and by station.

2.5.6. Scandinavia Week 2 (July 6–12, 2018) Events

The locations of events detected by waveform correlation for Scandinavia Week 2 are shown in the map (Figure 2-55) with superimposed REB events. There were 291 distinct waveform correlation events. Using multistation validation tolerances of 50 km and 15 s for the three IMS stations (ARCES, FINES, NOA), 39 of the 291 events were detected by two or more stations. 9 of the 17 REB events were detected by waveform correlation events within tolerances of 1° and ±15 s.

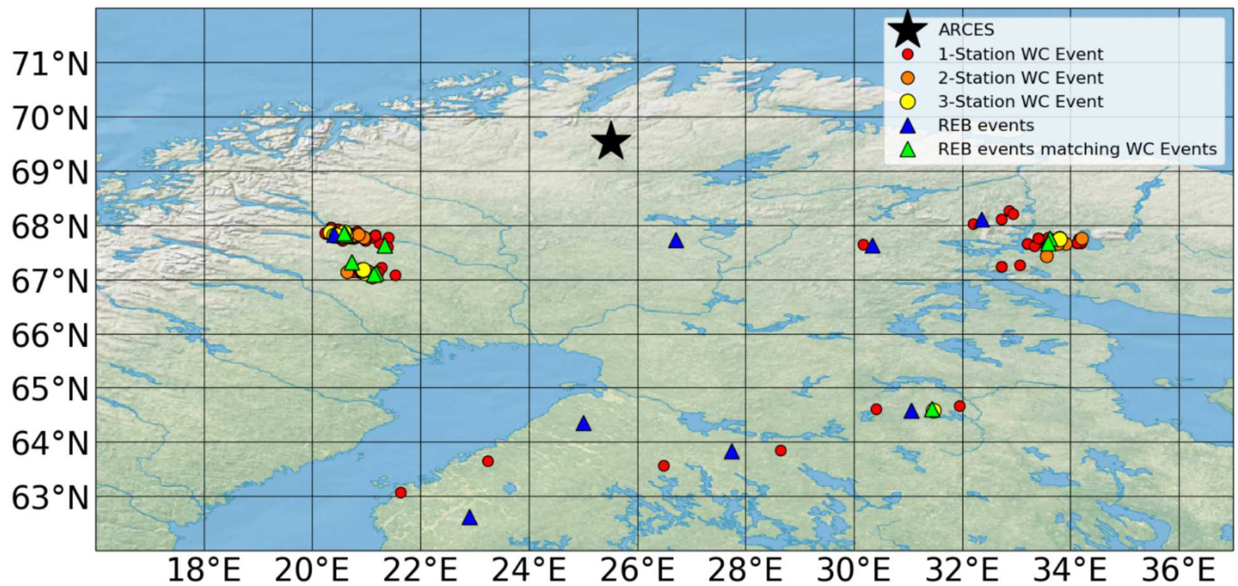

Figure 2-55. Map showing the locations of 291 waveform correlation events (circles) versus 17 REB events for Scandinavia Week 2. The size and color of the circle indicates the number of detecting stations. 9 REB events were detected by waveform correlation (green triangles), while 8 REB events were not detected by waveform correlation (blue triangles).

There were 8 REB events that were not detected by waveform correlation. Three of those events were not detected because the correlation job ended on 12-JUL-2018 00:00 instead of 13-JUL-2018 00:00 due to user parameter entry error. Table 2-8 contains information about the three events that were not detected due to the correlation job ending a day early. All three events are in proximity to mines with large numbers of template events (Figure 2-25). When the error was discovered, an additional correlation job for the single day of 12-JUL-2018 on station ARCES detected an additional 41 detections, bring the total waveform correlation events for Week 2 to 332. The figures referenced in the table show the template waveform and detected ARCES arrival waveform.

Table 2-8. Three REB Events not detected because the correlation job ended too early. The events were later detected by a correlation job for station ARCES for 12-JUL-18.

LAT	LON	ORID	Date/Time	Figure
67.823522	20.407853	16031667	12-JUL-18 11.21.43.760910000 PM	2-56
68.105414	32.356038	16026586	12-JUL-18 12.39.44.253060000 PM	2-57
64.584342	31.064614	16032016	12-JUL-18 10.11.16.659690000 AM	2-58

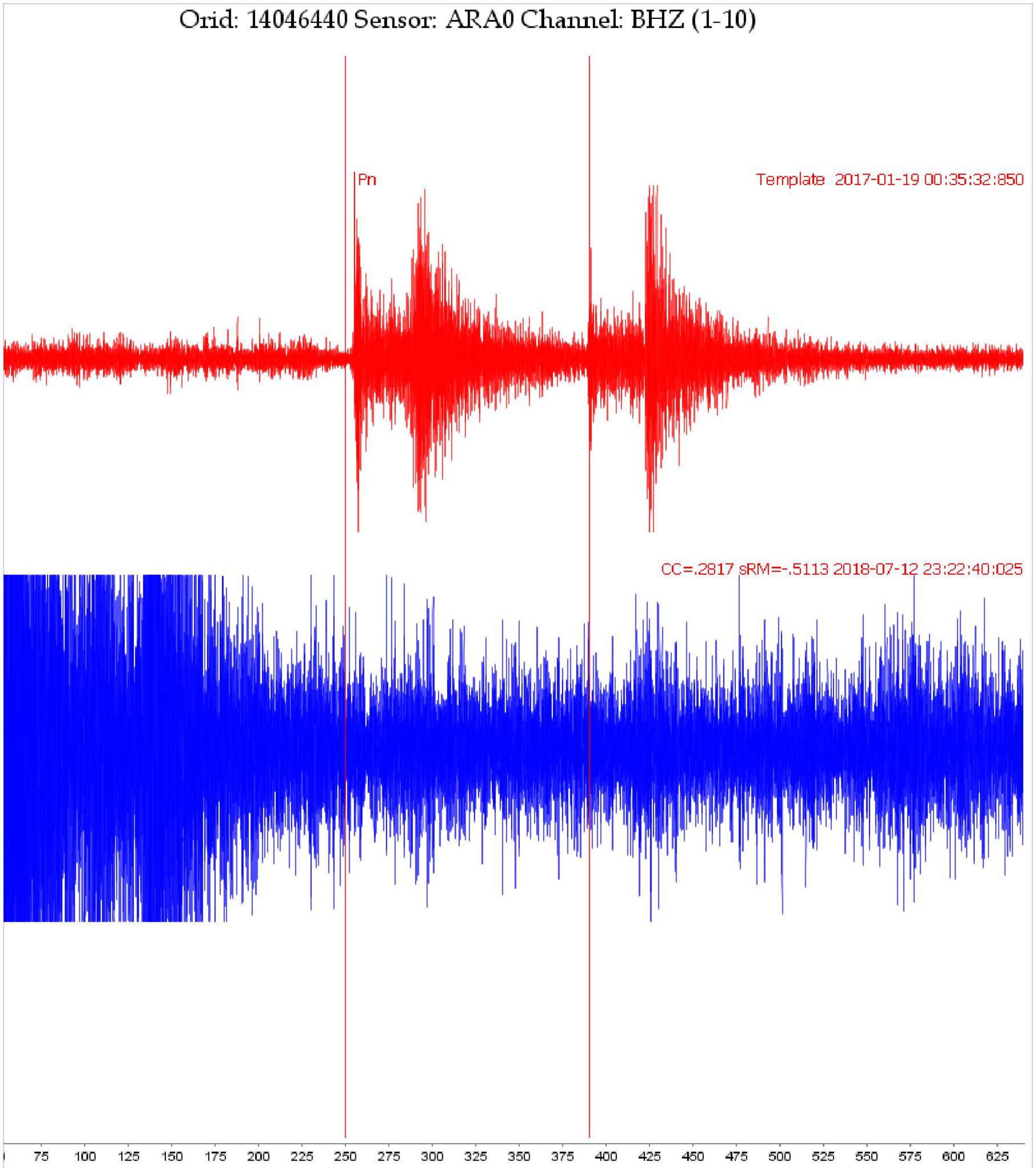

Figure 2-56. The ARCES arrival waveform for REB ORID 16031667 at (67.824°N, 20.408°E) (Table 2-8) was detected by a template event located at (67.831°N, 20.438°E). This arrival was preceded by a much larger event (blue waveform).

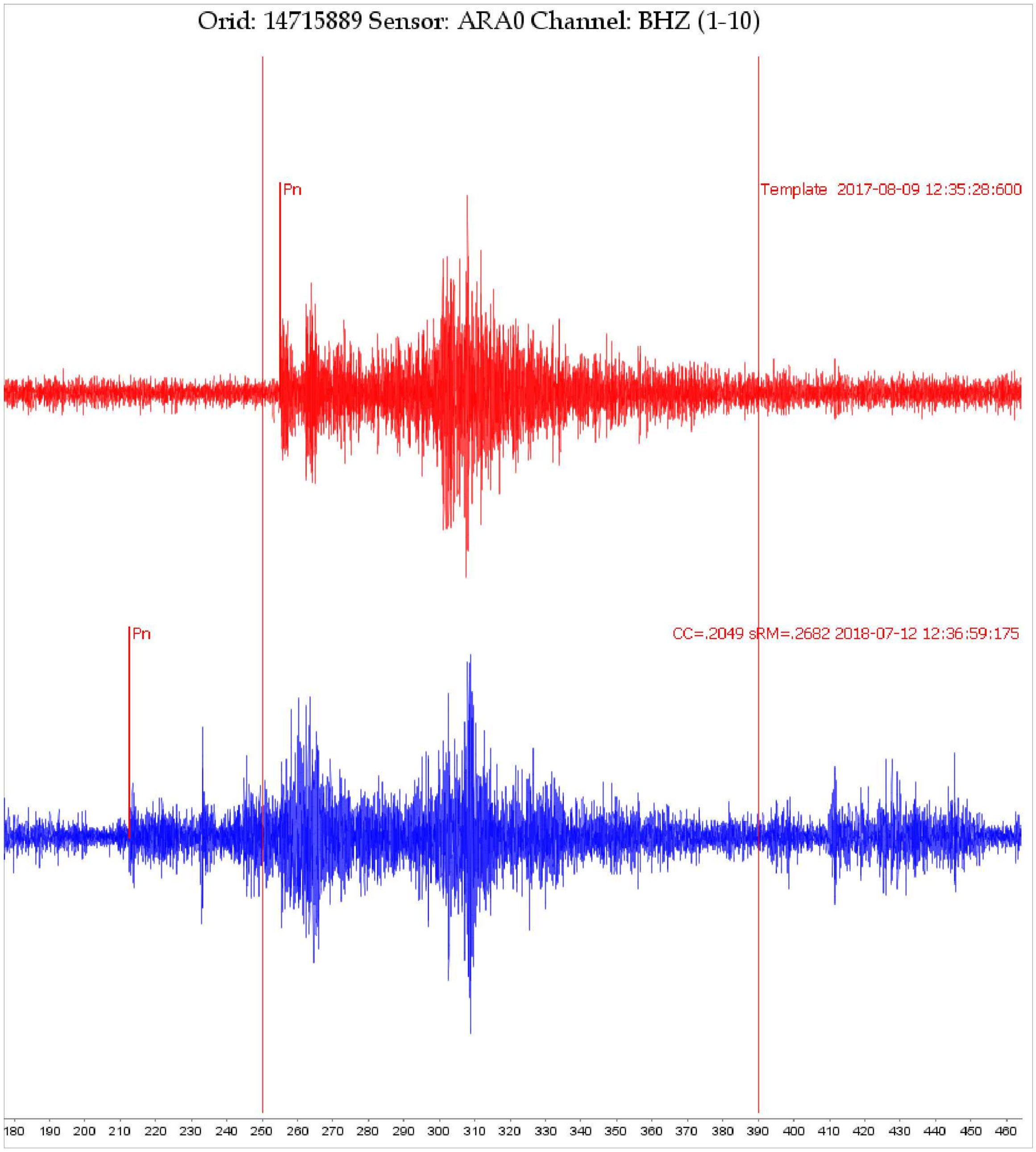

Figure 2-57. The ARCES arrival waveform for REB ORID 16026586 at (68.105°N, 32.356°E) (Table 2-8) was detected by a template event located at (68.129, 32.5).

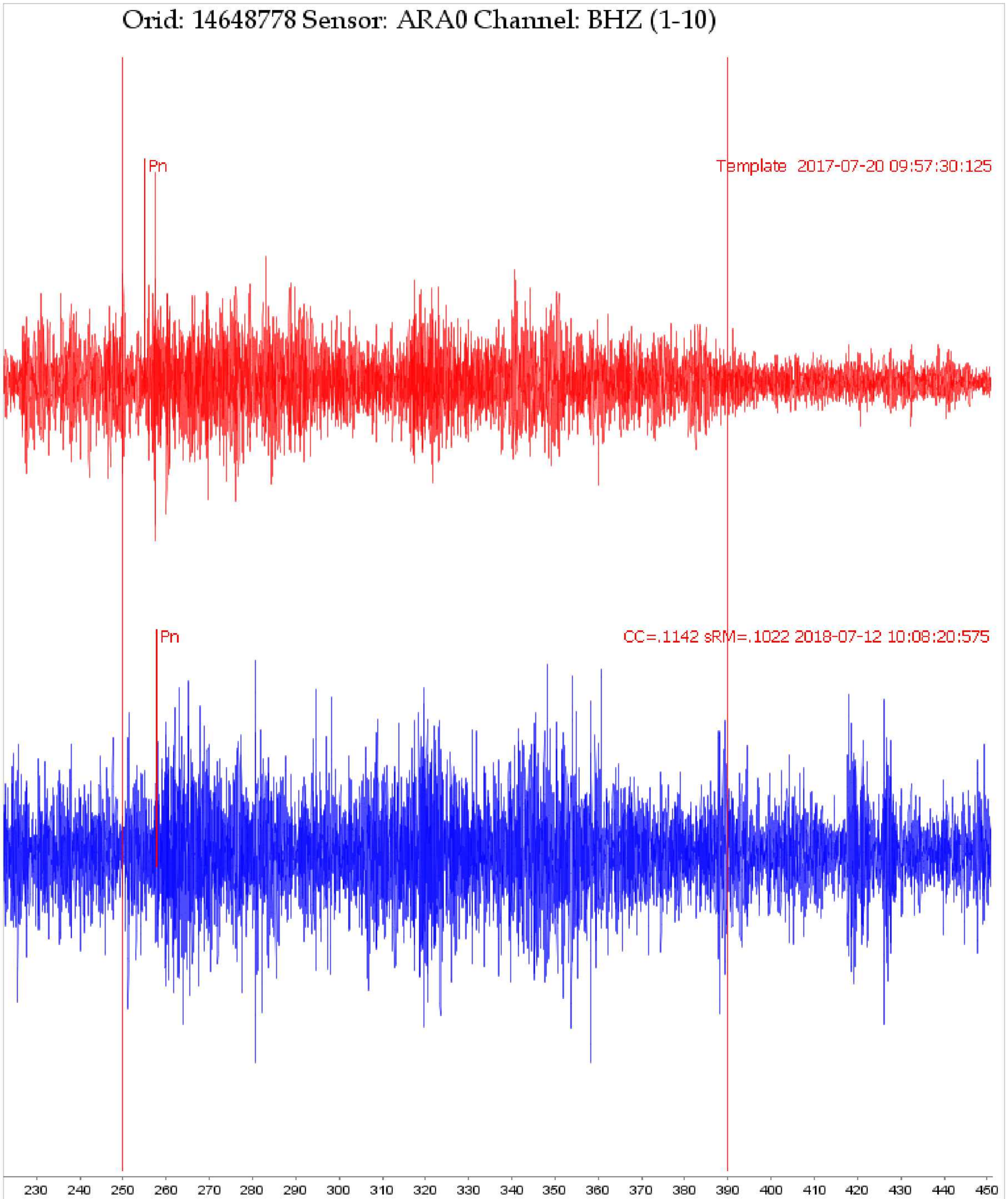

Figure 2-58. The ARCES arrival waveform for REB ORID 16032016 at (64.584°N, 31.065°E) (Table 2-8) was detected by a template event located at (64.707°N, 30.236°E).

The remaining five REB events that were not detected by waveform correlation are listed in Table 2-9. All five of these events had associated IDCX arrivals from a detection by infrasound station I37NO. The seismic waveform plots showed an arrival, but in several cases, there were additional signals within the correlation window that may have resulted in a decorrelation with the template waveform (arrival and template waveforms not shown). Moreover, the events are located in areas of sparse template events so there may not have been a template event within comparison tolerances with the REB events. Similarly to the events for Scandinavia Week 1, the single-station IDCX seismic detections associated with infrasound detections were not detected by waveform correlation. This may require further investigation, but is beyond the scope of this report.

Table 2-9. Five REB events with arrivals not detected by waveform correlation.

LAT	LON	ORID	EVENT DATE/TIME	STATION	ARID	PHASE
67.724	26.711	16001788	06-JUL-18 10.09.43.22 AM	ARCES	133642004	Pg
				ARCES	133642005	Lg
				I37NO	133642748	I
67.628	30.339	16006190	07-JUL-18 11.07.47.50 AM	ARCES	133664488	Pn
				ARCES	133664490	Sn
				I37NO	133665095	I
63.837	27.742	16010674	06-JUL-18 11.29.21.43 AM	FINES	133643774	Pn
				FINES	133643777	Lg
				I37NO	133644391	I
64.351	24.997	16011281	09-JUL-18 09.06.42.69 AM	ARCES	133705513	Pn
				ARCES	133705515	Sn
				I37NO	133707082	I
62.613	22.903	16015125	10-JUL-18 02.33.17.27 PM	FINES	133738775	Sn
				HFS	133739540	Lg
				I37NO	133740366	I

The bar chart (Figure 2-59) shows the number of Scandinavia Week 2 template detections by template age, and further divided by station. This bar chart shows that most of the detections are made by templates from events within the last 3 years for this region; moreover, most of the detections were made by station ARCES.

Similar to the result for Week 1 (Figure 2-52), for Week 2 the effectiveness of the templates drop dramatically with time; for example, the templates less than 1 year of age produced more than three times the number of detections as templates that are 1 year old.

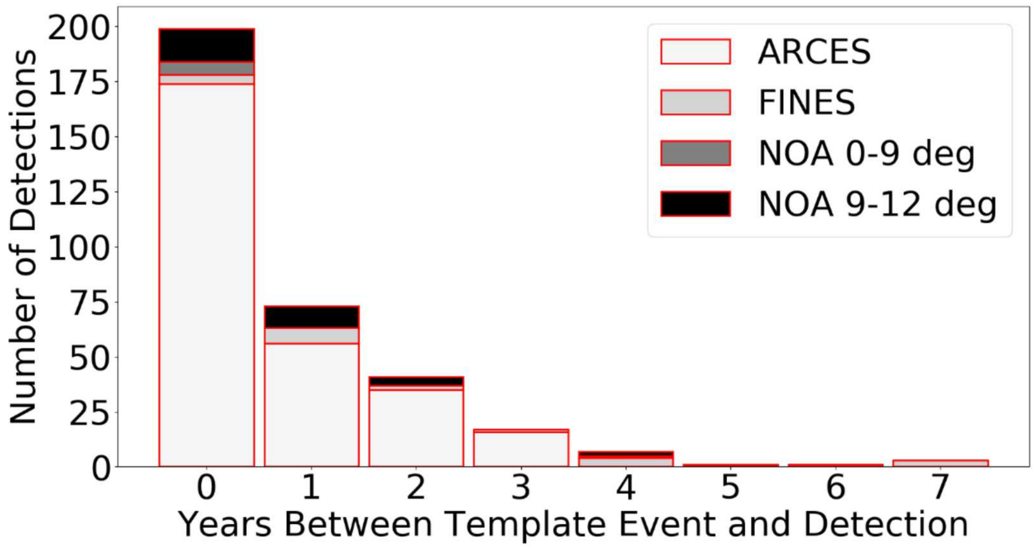

Figure 2-59. The number of Scandinavia Week 2 detections by age of template, and by station.

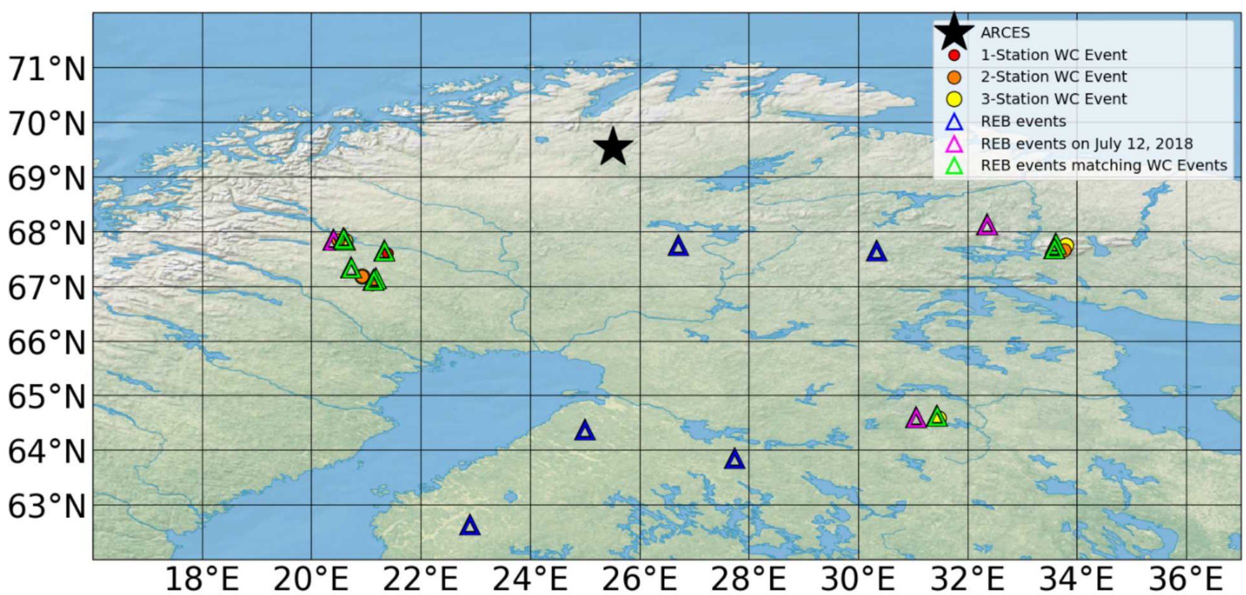

Figure 2-60. Map showing the locations of the waveform correlation events (circles) that matched REB events within the comparison tolerance of (1°, ±15 s) versus the total 17 REB events for Scandinavia Week 2. The size and color of the circle indicates the number of detecting stations. 9 REB events were detected by waveform correlation (green triangles), 3 REB events were excluded from processing on July 12, 2018 (magenta triangles), and 5 REB events were not detected by waveform correlation (blue triangles).

The waveform correlation events that match the Week 2 REB events are mapped in Figure 2-60, along with the REB events. The purpose of this map is to identify which waveform correlation events were matched to the REB events, and to provide a visual indication of the proximity of the matched events. The REB events detected by waveform correlation within a tolerance of (1°, ±15s) are shown as green triangles. The REB events on July 12, 2018 that were not detected because the data was not processed for that date are shown as magenta triangles on this map because these events are in areas of dense template events and would have been detected without the processing error; from this point forward in the report, these events will be identified only as missed REB events. The REB events that were not detected by waveform correlation within the tolerances are shown as blue triangles (Table 2-9).

The number of waveform correlation detections that match REB events for Scandinavia Week 2 is shown in the bar chart (Figure 2-61). For this week of data, the older templates are contributing detections that are useful for building bulletin events; despite fewer detections by older templates, those detections are proportionately more valuable. The detections that match REB events by station and age are more diverse for Week 2 than Week 1, signifying that template diversity is important for effective detectors.

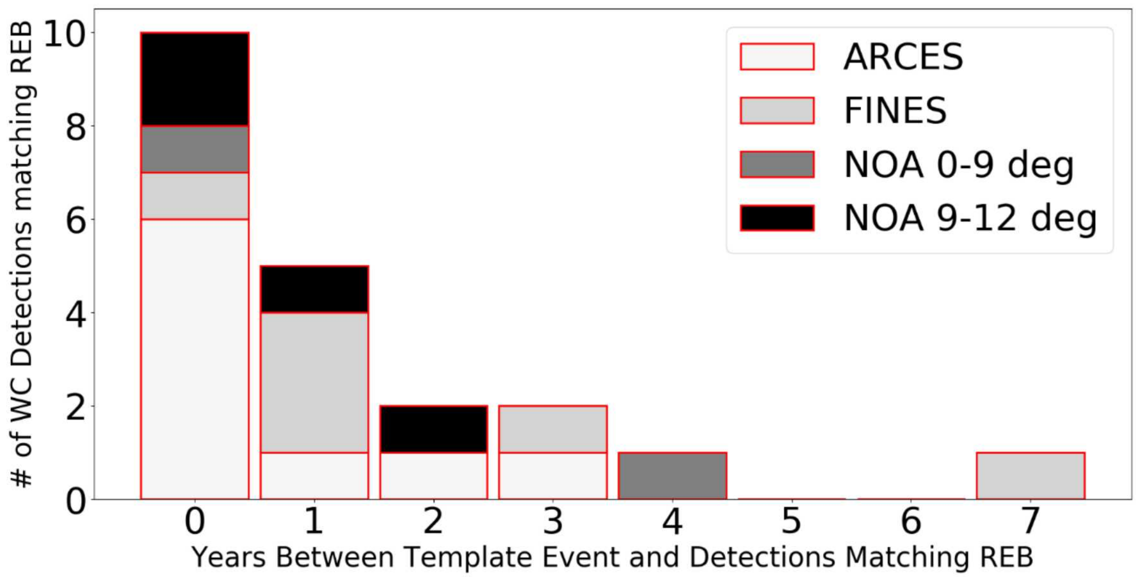

Figure 2-61. Number of Scandinavia Week 2 waveform detections matching the REB events, by age and by station.

3. DISCUSSION

The focus of this research study was to apply waveform correlation to events from two mining regions deemed particularly problematic for the IDC to find out if the technique could reduce analyst workload to produce the REB. Two weeks for each of two mining regions identified by the IDC were studied using SNL's SeisCorr application for waveform correlation. In all cases, SeisCorr detected candidate events that matched REB events within a tolerance of 1 degree and ±15 s and that have an initial event location that is better than the event location in the SEL3 bulletin. We offer this as evidence that waveform correlation has the potential to reduce analyst workload for mining blast events but recommend additional research to quantify the workload reduction accurately. Waveform correlation detects events that are not in the REB; those additional events increase analyst workload when they must be built manually.

The experimental setup previously described emphasized that operational constraints must be adhered to when creating template waveforms. Previous studies of mining regions often used waveforms from all catalog events to detect additional mining blast events because the studies are done on historical data. Those studies show impressive results but sometimes do not address the operational realities of using only historical templates that precede the study period. This study offers insight into the choices available to an operational monitoring agency. The template waveforms were chosen from historical events from the REB. We are not aware of any limit to the use of historical waveforms; thus, we chose to use up to 10 years of data, where available, for the stations chosen by the IDC.

3.1. Waveform Templates

SeisCorr automatically uses arrivals from bulletin events to window waveform templates according to user-defined parameters such as total template duration and the time window preceding the arrival; for this study we chose long duration template windows to increase the time-bandwidth product, with 5 seconds preceding the arrival. The actual length of the template window was chosen depending on each station after reviewing the variety of arrival waveforms for the study region. Typically, a user-defined STA/LTA threshold is applied to guarantee that a signal is present within the window and such user-defined parameters were chosen per template library to vary by station and/or phase.

As described below, there are several types of problems that can be encountered with automatically created SeisCorr templates. Although the templates are screened to pass an STA/LTA threshold, it is possible that an unrelated signal arrives within the template window, which meets the threshold criteria but which does not accurately match the expected signal when correlated with continuous data (see example NVAR template, Figure 2-32). The templates for the mining blast study were not reviewed and pruned manually, and examples of poorly windowed arrivals have been observed in the template libraries. Poorly windowed templates may detect valid events, but the detections may be assigned the incorrect location because the template event is not truly the recurring event. An operational pipeline may require algorithms to screen templates for poorly windowed templates. For arrays, it may be possible to develop an algorithm that includes only templates that demonstrate a moveout across elements that matches the location of the template event, thus enforcing a location-specific template filter. More research on the selection of waveform templates for automated pipelines is needed.

This study created template libraries for three IMS stations for each mining blast region; the number of stations was limited by the CTBTO study. In an automated pipeline, it would make sense to include all the stations of the primary IMS network in the process; thus, with more template libraries and more detecting stations, it seems reasonable to expect that the results of waveform correlation using IMS network would be better than those reported in this study.

3.2. Comparison with Automated Pipeline

SeisCorr includes the multistation validation feature to group detections into candidate events; thus, our results can compare candidate events from waveform correlation detections with events from the IDC automated pipeline that generates the SEL3 as well as the analyst-reviewed REB. A candidate event detected by multiple stations is generally considered more reliable than a single station detection. For the mining blast study, we chose to include results from candidate events based on any number of detecting stations, including only one station. We included single-station events, despite them being sometimes uncertain, because a waveform correlation detection assigns the location of the template event to the detected event, so that these events can be compared with REB and SEL3 events. Moreover, we chose not to exclude detections from single stations that may be grouped with other detections using an associator (e.g., NET-VISA [11][11]), even though they were unassociated after multistation validation. Furthermore, this study is based on waveform correlation with templates from only three stations for each mining region, and the single-station events from waveform correlation may be associated with detections from other IMS stations not included in this study.

The automated pipeline of the IDC produces the SEL3 list of events that are then reviewed and adjusted by analysts to produce downstream bulletins (LEB, and REB); the analysts often improve the formation of events by adding arrivals and adjusting their locations. Within this discussion, we first compare SEL3 events with REB events to provide insight into the analyst effort for each week of study. We compare the locations of the SEL3 events with the REB events for Wyoming Week 1. Where the analyst has modified an event, we trace the evolution of a SEL3 event to REB event through the event identifier (EVID) that uniquely specifies an event through modifications. An REB event that was manually built by an analyst will not have an EVID that corresponds to an original SEL3 event, and detection of such events by waveform correlation is particularly interesting because manually building an event requires a significant amount of analyst effort.

Then we compare the waveform correlation events to the SEL3 events and REB events to look for evidence of improvement. Our analysis for this report is based primarily on location comparison because we believe that waveform correlation can improve the initial starting location for events due to the fact that the locations are assigned based on template events.

We anticipate that the IDC will do a more complete and thorough analysis of results from all study participants (i.e., IDC, NORSAR, SNL) that includes a more comprehensive event evolution analysis through association of arrivals. We consider such an analysis to be beyond the scope of this report.

3.2.1. Wyoming Week 1 Event Comparison

The SEL3 and REB events for Wyoming Week 1 are shown in the map (Figure 3-1). The symbol colors in the map indicate which SEL3 events were traceable to an REB event, which SEL3 events were discarded, and which REB events were manually built. The SEL3 event locations are scattered relative to the final, high-quality REB locations, which cluster around the mines.

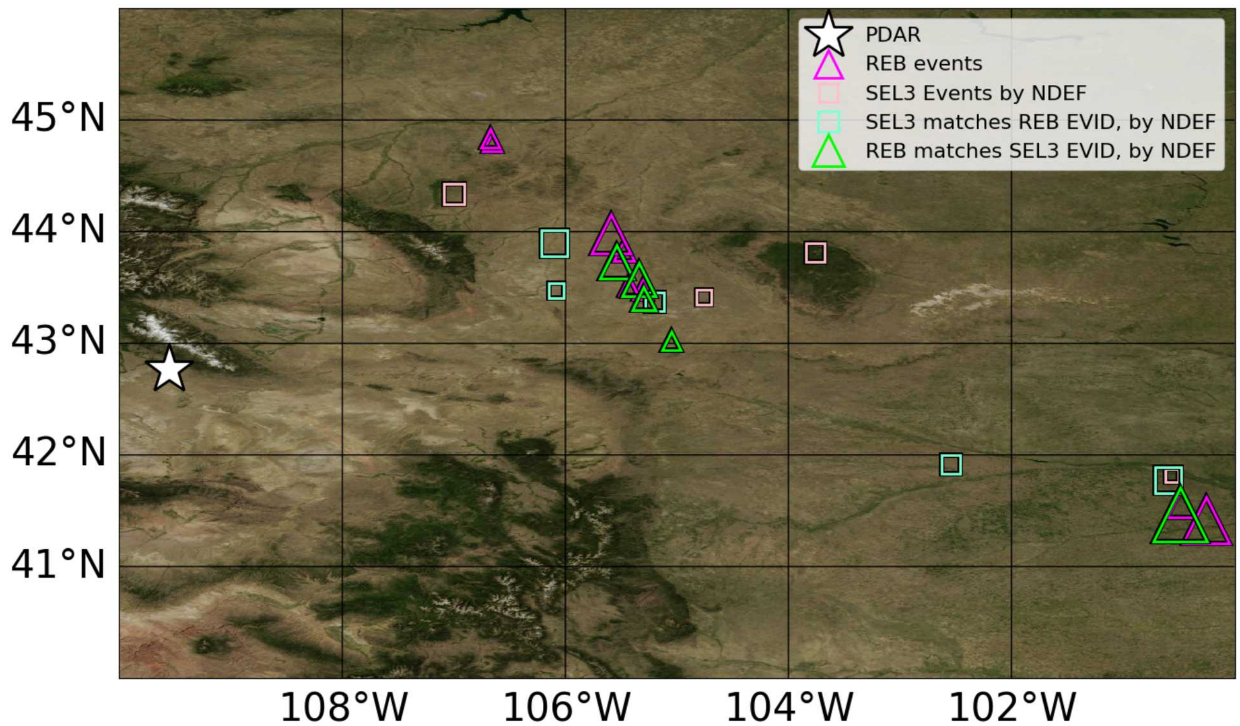

Figure 3-1. Map of SEL3 and REB events for Wyoming Week 1, where SEL3 (square) and REB (triangle) event symbol size indicates the number of associated phases. The colors show pairs of 5 SEL3 (sea green) events that were modified to become REB (green) events (i.e., the EVIDs are the same), while 4 SEL3 (pink) events were discarded and 8 REB (magenta) events were manually built.

The bar graph (Figure 3-2) provides another view of the event data, where the colors match the categories on the map. There was a total of 13 REB events in Wyoming Week 1. Only five REB events are traceable to SEL3 events by EVID. This suggests that four SEL3 events were discarded by analysts, or were modified so substantially that they are no longer located in the study area. Furthermore, eight REB events were manually built by analysts, or were derived from SEL3 events that were located outside the study area. Thus, only 38% of REB events (5/13) are clearly traceable to the automatic pipeline (SEL3) for Week 1 in this study region.

The paired SEL3-REB events are mapped (Figure 3-3) with connecting arrows to show the difference between the locations produced by the automatic pipeline and the final analyst-determined location. This map provides perspective on the location inaccuracy even for the best SEL3 events that are retained by analysts.

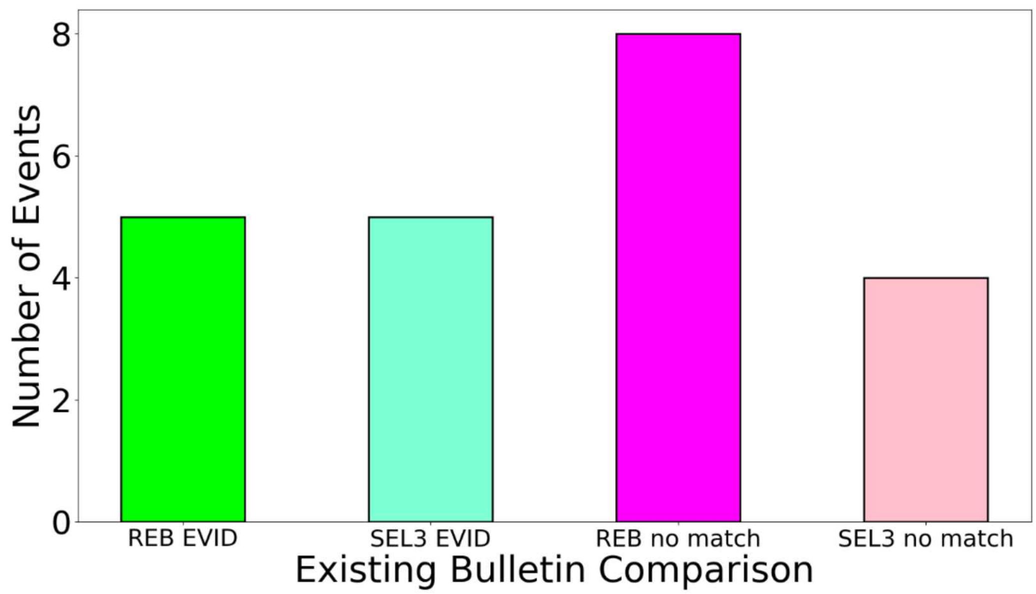

Figure 3-2. Number of Wyoming Week 1 events that are categorized in Figure 3-1 (same colors used on map). The first two columns represent the 5 SEL3 (sea green) events that were modified to become 5 REB (green) events (same EVID). 8 REB events that were manually built and 4 SEL3 events were discarded by analysts.

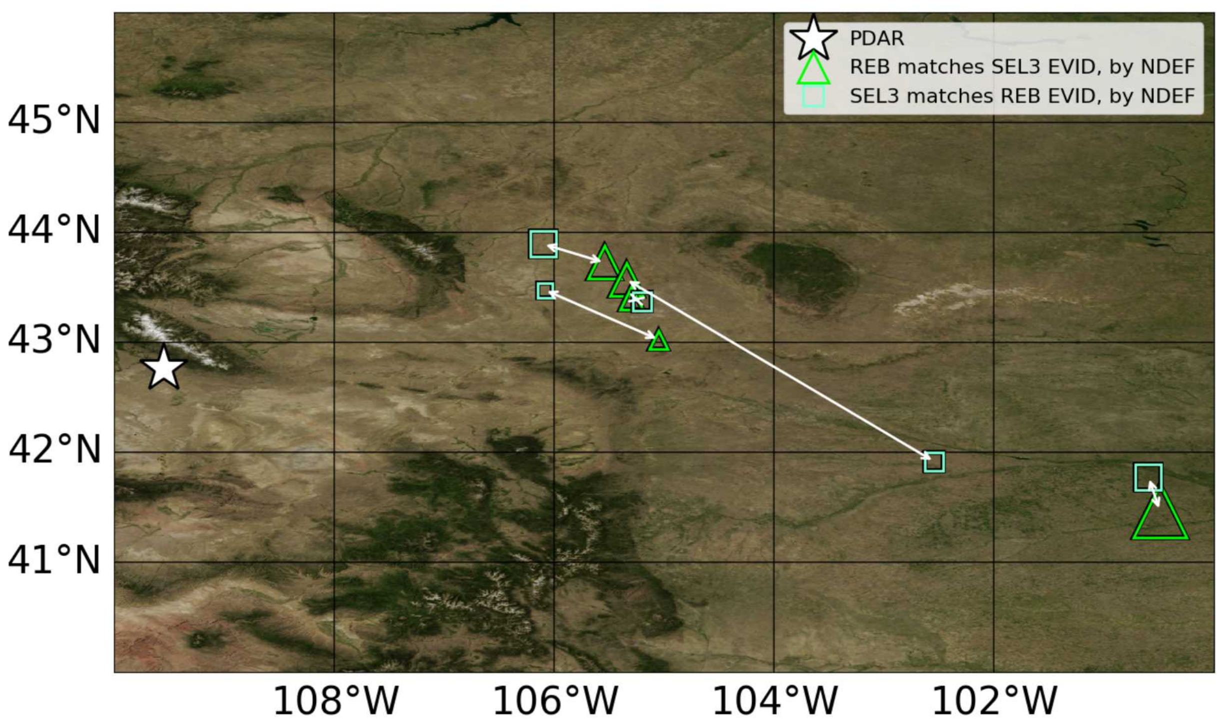

Figure 3-3. Map of 5 SEL3-REB event pairs for Wyoming Week 1, with a white arrow to link each pair and show the analyst relocation of the SEL3 event to the final REB event location.

The waveform correlation (WC) events are mapped with the SEL3 and REB events (Figure 3-4) for Wyoming Week 1. On this map, the REB events are matched to WC events within distance and time tolerances of (1°, ±15s). The WC events are adjacent to the REB events in comparison with the SEL3 events, indicating the potential improvement in initial location. However, we must also point out that three REB events in the SE corner of the map were not matched by WC, presumably because there were no templates in this area, but do have SEL3 events in proximity. As previously explained, waveform correlation can only find recurring events for which template waveforms have been created. This small group of undetected events shows that WC can improve, but not replace, the existing SEL3 pipeline.

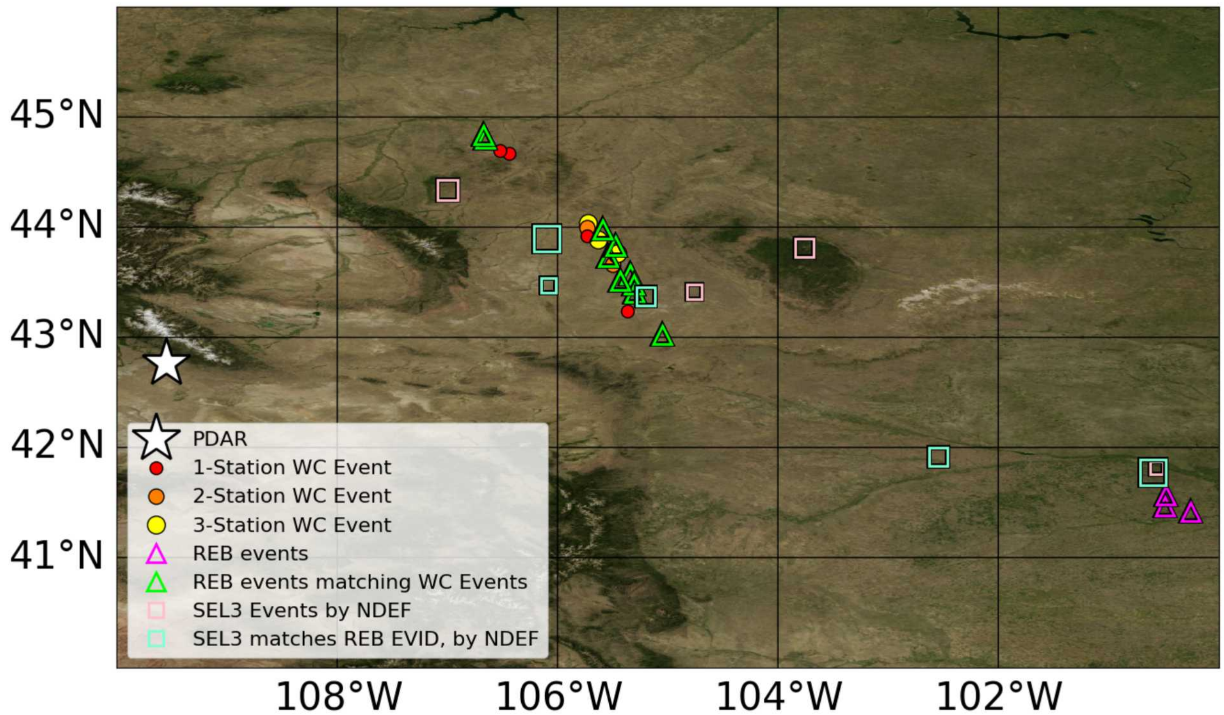

Figure 3-4. Map of SEL3, REB, and waveform correlation (WC) events for Wyoming Week 1. The waveform correlation events are color-coded by number of detecting stations (PDAR, NVAR, TXAR). 10 REB (green triangle) events match WC events within tolerances of (1°, ±15s), while 3 REB (magenta triangle) events do not match within the tolerances. 5 SEL3 (sea green square) events were modified to become an REB event (i.e., the EVIDs are the same), where the size of the square indicates the number of associated phases. 4 SEL3 (pink square) events were discarded by analysts.

The bar graph (Figure 3-5) provides another view of the event data by category counts, where the colors match the categories on the map. There was a total of 13 REB events in Wyoming Week 1, and 10 of the REB events match waveform correlation (WC) events within a tolerance of (1°, ±15s), which is a tight enough tolerance to ensure that one and only one WC event matches each of the REB events. The numbers of matching WC events are further divided by number of detecting stations in the second column, showing that six of 10 REB events were detected by more than one station using waveform correlation for Wyoming Week 1. In comparison, only five REB events were traceable to SEL3 events by EVID. Three REB events were not detected by waveform correlation because there were no templates from recurring events in this area. Four SEL3 events were discarded by analysts.

Previously we pointed out that only 38% of REB events are traceable to the automatic pipeline for Week 1 in this study region. In contrast, 77% of REB events were detected by WC within the specified tolerances.

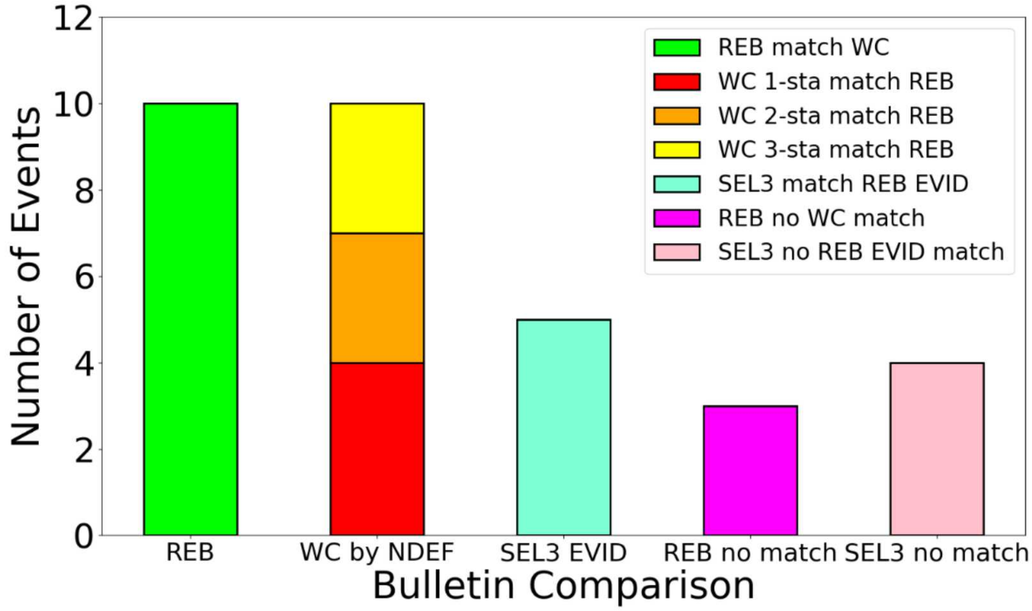

Figure 3-5. Number of Wyoming Week 1 events that are categorized in Figure 3-4 (same colors used on map). The first green column represents 10 REB events that matched waveform correlation (WC) events within tolerances of (1°, ±15s). The second column represents 10 WC events that matched the REB events, color-coded by number of detecting stations (PDAR, NVAR, TXAR). The third column represents 5 SEL3 events that were modified to become REB events (same EVID). The magenta column represents 3 REB events that did not match WC events within the tolerances. The pink column represents 4 SEL3 events that were discarded.

3.2.2. Wyoming Week 2 Event Comparison

The SEL3 and REB events for Wyoming Week 2 were mapped (Figure 3-6) with the same color scheme discussed above for Week 1. For this week, nine of ten REB events were manually built, and three out of four SEL3 events were discarded.

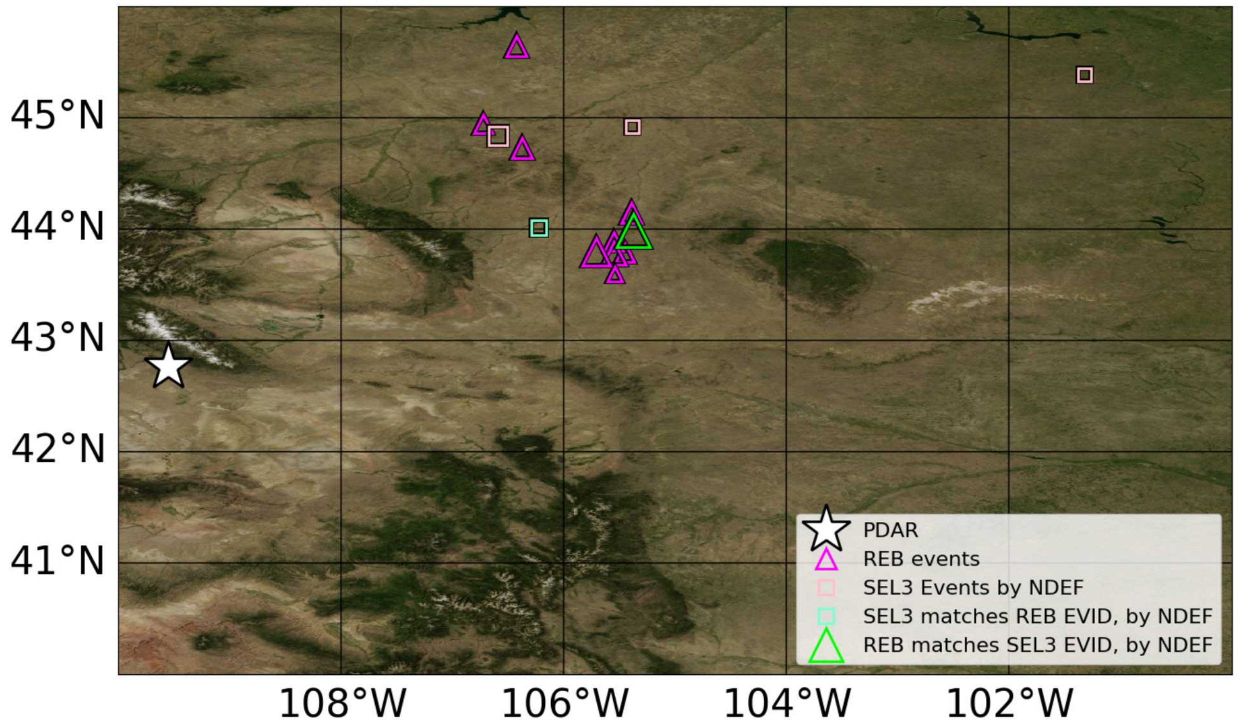

Figure 3-6. Map of SEL3 and REB events for Wyoming Week 2, where SEL3 (square) and REB (triangle) event symbol size indicates the number of associated phases. The colors show the 1 paired SEL3 (sea green) event that was modified to become an REB (green) event (i.e., the EVIDs are the same). 3 SEL3 (pink) events were discarded and 9 REB (magenta) events were manually built by analysts.

The bar graph (Figure 3-7) provides another view of the event data, where the colors again match the categories on the map. There was a total of 10 REB events in Wyoming Week 2. Only one REB event is traceable to a SEL3 event by EVID. The remaining three SEL3 events were discarded by analysts. Nine of the 10 REB events were manually built by analysts. Only 10% of REB events (1/10) are traceable to the automatic pipeline for Week 2 in this study region.

The paired SEL3-REB event is mapped (Figure 3-8) with a connecting arrow to provide perspective on the location inaccuracy for the only SEL3 event that was retained by analysts.

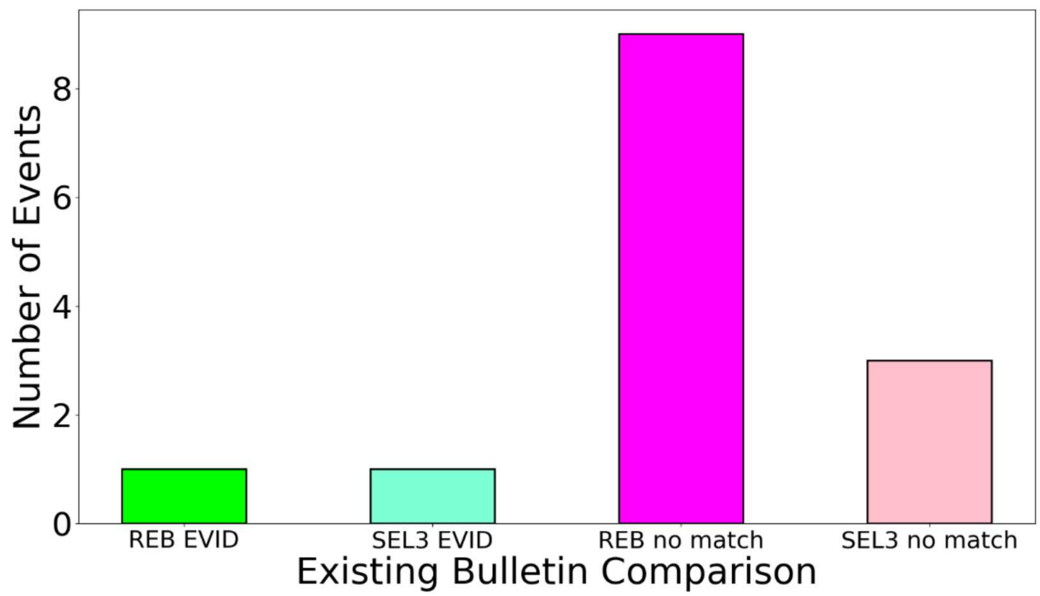

Figure 3-7. Number of Wyoming Week 2 events that are categorized in Figure 3-6 (same colors used on map). The first two columns represent 1 SEL3 (sea green) event that was modified to become an REB (green) event (same EVID). 9 REB (magenta) events were manually built and 3 SEL3 (pink) events were discarded by analysts.

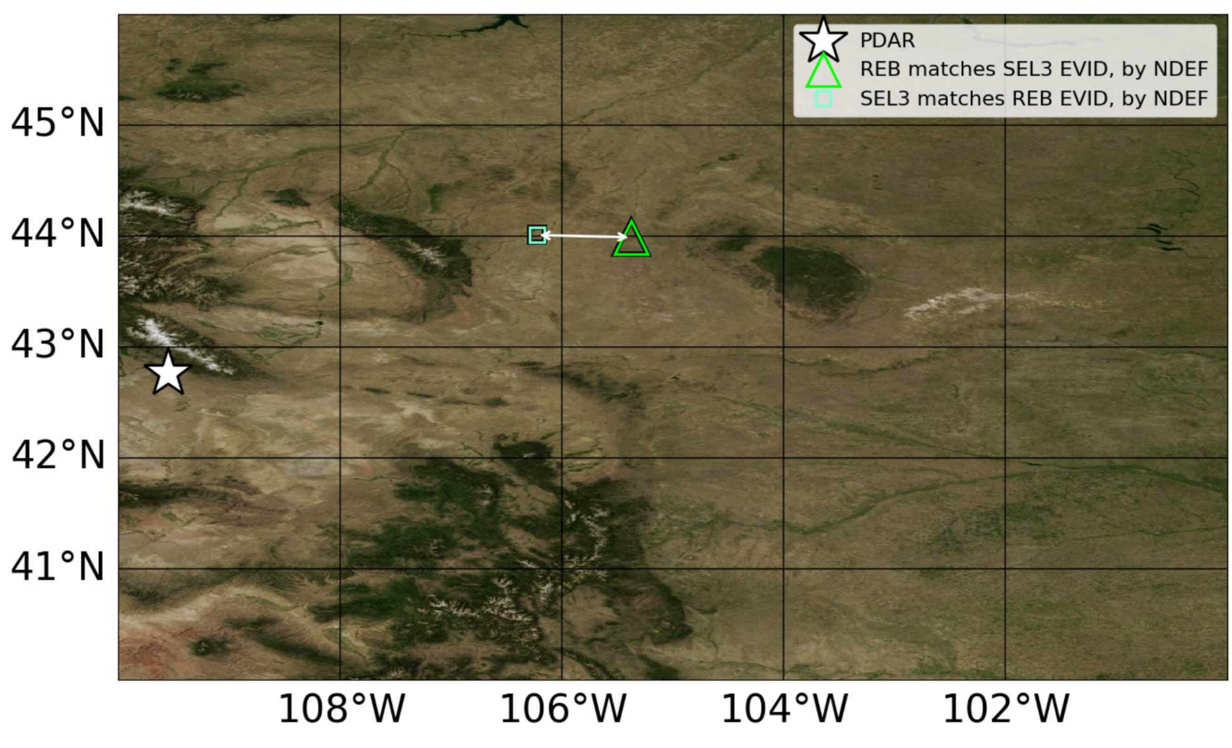

Figure 3-8. Map of SEL3-REB event pair for Wyoming Week 2, with a white arrow to link the pair and show the analyst relocation of the SEL3 event to the final REB event location.

The waveform correlation (WC) events for Wyoming Week 2 are shown with the SEL3 and REB events (Figure 3-9). On this map, the REB events are matched to WC events within distance and time tolerances of (1°, ±15s). Waveform correlation detected nine of the 10 REB events for Week 2; recall that the 10th event was not detected because only 6 of 7 days of Week 2 were processed. Similarly to Week 1, the WC events are adjacent to the REB events in comparison with the SEL3 events, indicating potential improvements in both detection and initial location for events.

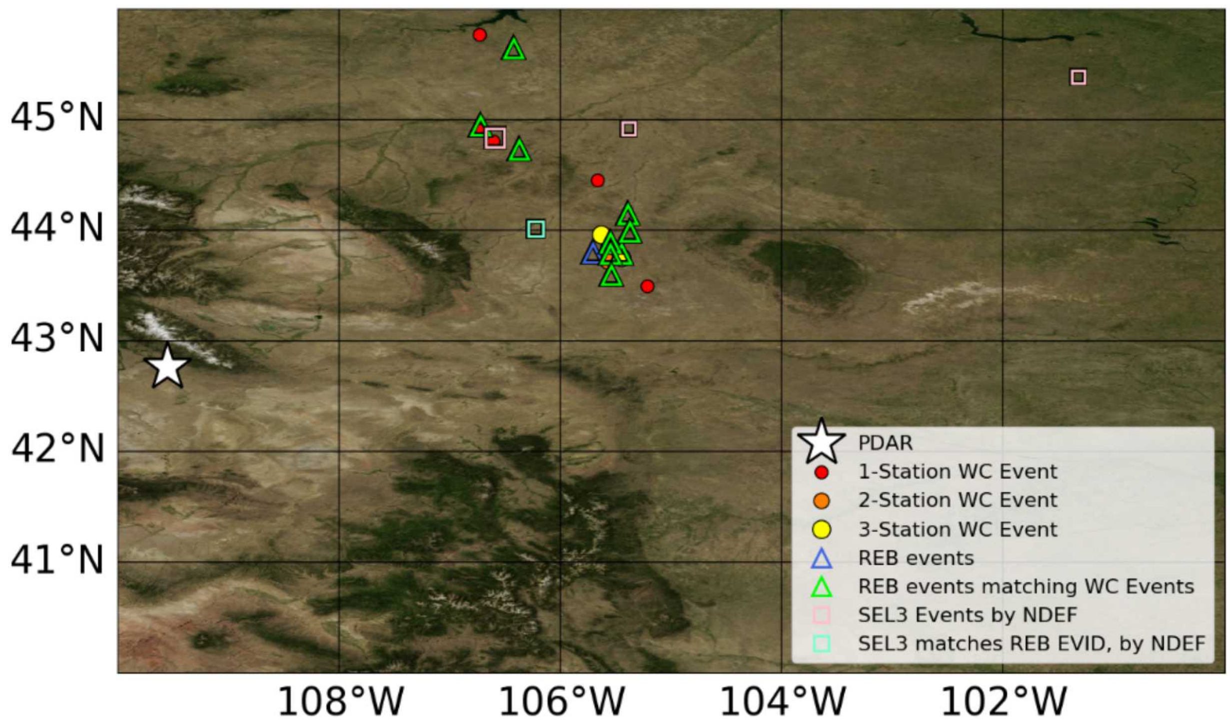

Figure 3-9. Map of SEL3, REB, and waveform correlation (WC) events for Wyoming Week 2. The waveform correlation events are color-coded by number of detecting stations (PDAR, NVAR, TXAR). 9 REB (green triangle) events match WC events within tolerances of (1°, ±15s), while 1 REB (blue triangle) event does not match within the tolerances. 1 SEL3 (sea green) event was modified to become an REB event (i.e., the EVIDs are the same), where the size of the square indicates the number of associated phases. 3 SEL3 (pink) events were discarded by analysts.

The bar graph (Figure 3-10) provides another view of the event data by category counts, where the colors match the categories on the map. There was a total of 10 REB events in Wyoming Week 2, and nine of the REB events match waveform correlation (WC) events within a tolerance of (1°, ±15s). The numbers of matching WC events are further divided by number of detecting stations in the second column, showing that four of nine REB events were detected by more than one station using waveform correlation for Wyoming Week 2. In comparison, only one REB event was traceable to a SEL3 event by EVID. One REB event was not detected by waveform correlation because that day of waveform data was not processed due to an error. Three SEL3 events were discarded by analysts.

Previously we pointed out that only 10% of REB events are traceable to the automatic pipeline for Week 2 in this study region. In contrast, 90% of REB events were detected by WC within the specified tolerances.

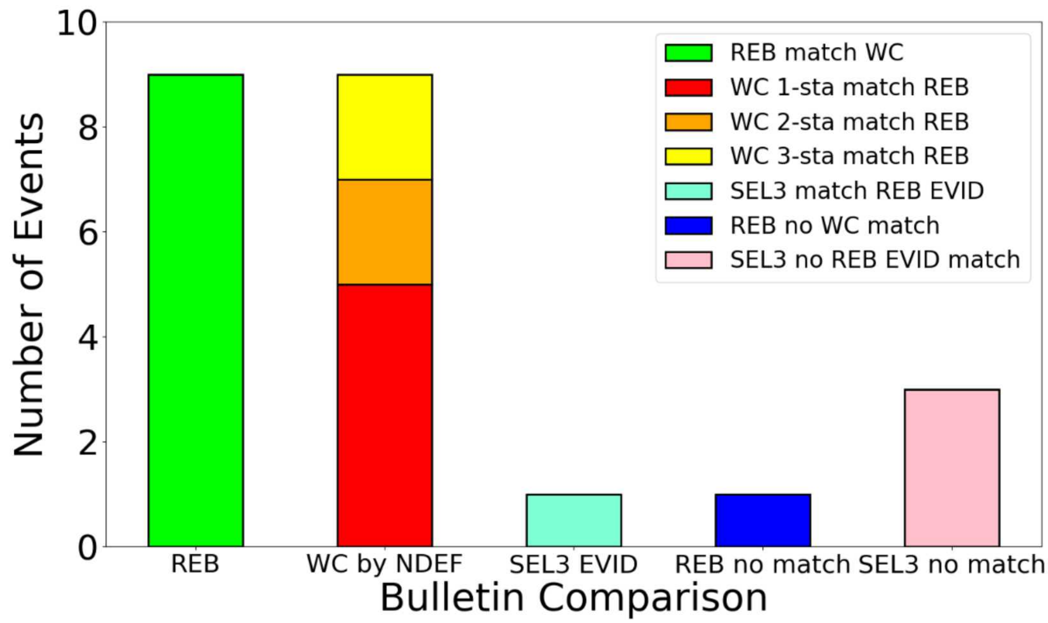

Figure 3-10. Number of Wyoming Week 2 events that are categorized in Figure 3-9 (same colors used on map). The first green column represents 9 REB events that matched waveform correlation (WC) events within tolerances of (1°, ±15s). The second column represents the WC events that matched the 9 REB events, color-coded by number of detecting stations (PDAR, NVAR, TXAR). The third column represents the 1 SEL3 event that was modified to become an REB event (same EVID). The blue column represents 1 REB event that was mistakenly excluded from processing and was not detected by waveform correlation. The pink column represents 3 SEL3 events that were discarded.

3.2.3. Scandinavia Week 1 Event Comparison

The SEL3 and REB events for Scandinavia Week 1 were mapped (Figure 3-11) with a color scheme to indicate what SEL3 events are traceable to REB events, which SEL3 events were discarded, and which REB events were manually built. The SEL3 event locations are scattered relative to the final REB locations. There were 37 SEL3 events during Week 1, but the analysts discarded all but 4 of those SEL3 events while manually building an additional 15 REB events.

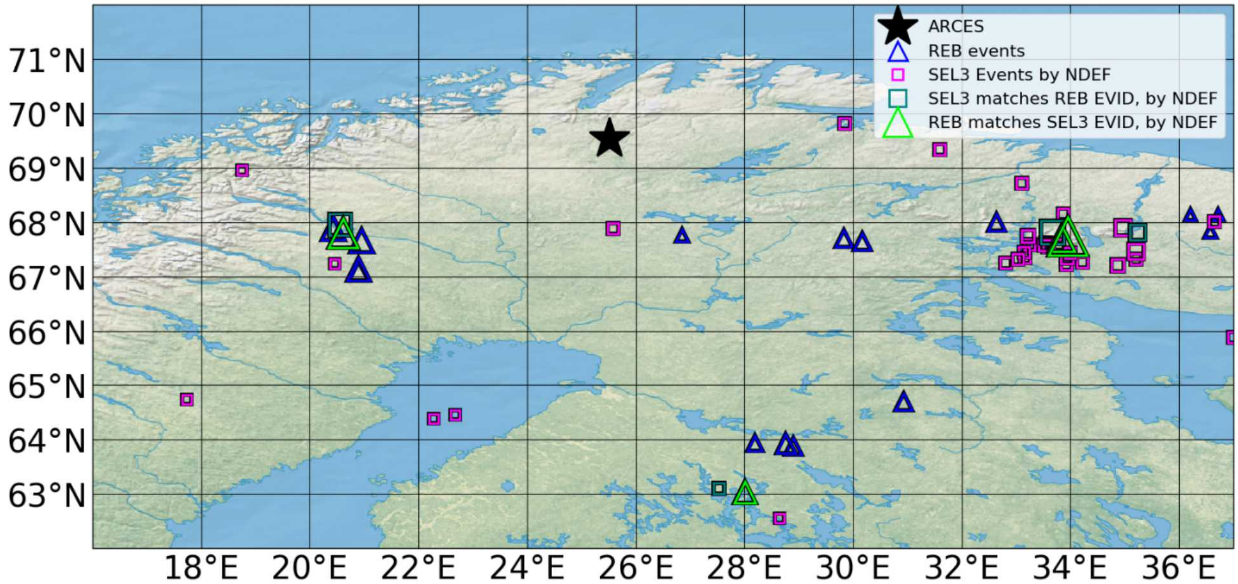

Figure 3-11. Map of SEL3 and REB events for Scandinavia Week 1, where SEL3 (square) and REB (triangle) event symbol size indicates the number of associated phases. The colors show pairs of 4 SEL3 (dark green) events that were modified to become REB (green) events (i.e., the EVIDs are the same), while 33 SEL3 (magenta) events were discarded and 15 REB (blue) events were manually built.

The bar graph (Figure 3-12) provides another view of the event data, where the colors match the categories on the map. There was a total of 19 REB events during Scandinavia Week 1. Only four REB events are traceable to SEL3 events by EVID. Thirty-three (33) SEL3 events were discarded by analysts and 15 REB events were manually built by analysts. Only 21% of REB events are traceable to the automatic pipeline for Scandinavia Week 1 in this study region.

The 4 paired SEL3-REB events are shown (Figure 3-13Figure 3-3) with connecting arrows to indicate the difference between the locations produced by the automatic pipeline and the final analyst-determined locations. This map provides perspective on the location inaccuracy even for the best SEL3 events that are retained by analysts. Most of the discarded 33 SEL3 events had few associated phases, which is illustrated by the size of the square symbol in the map (Figure 3-11).

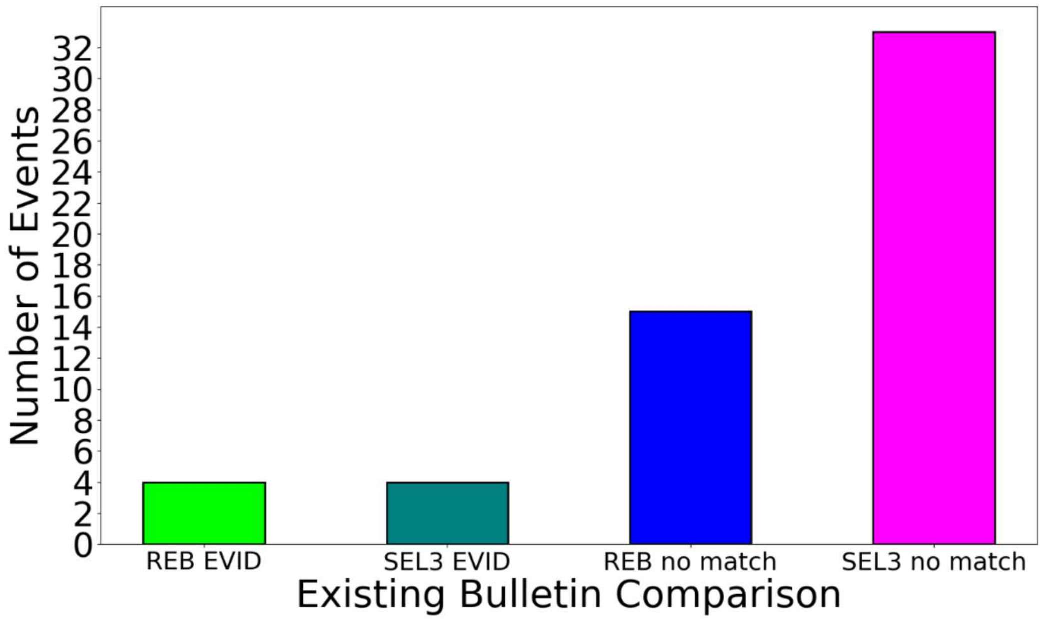

Figure 3-12. Number of Scandinavia Week 1 events that are categorized in Figure 3-11 (same colors used on map). The first two columns represent 4 SEL3 (dark green) events that were modified to become REB (green) events (same EVID). 15 REB (blue) events were manually built and 33 SEL3 (magenta) events were discarded by analysts.

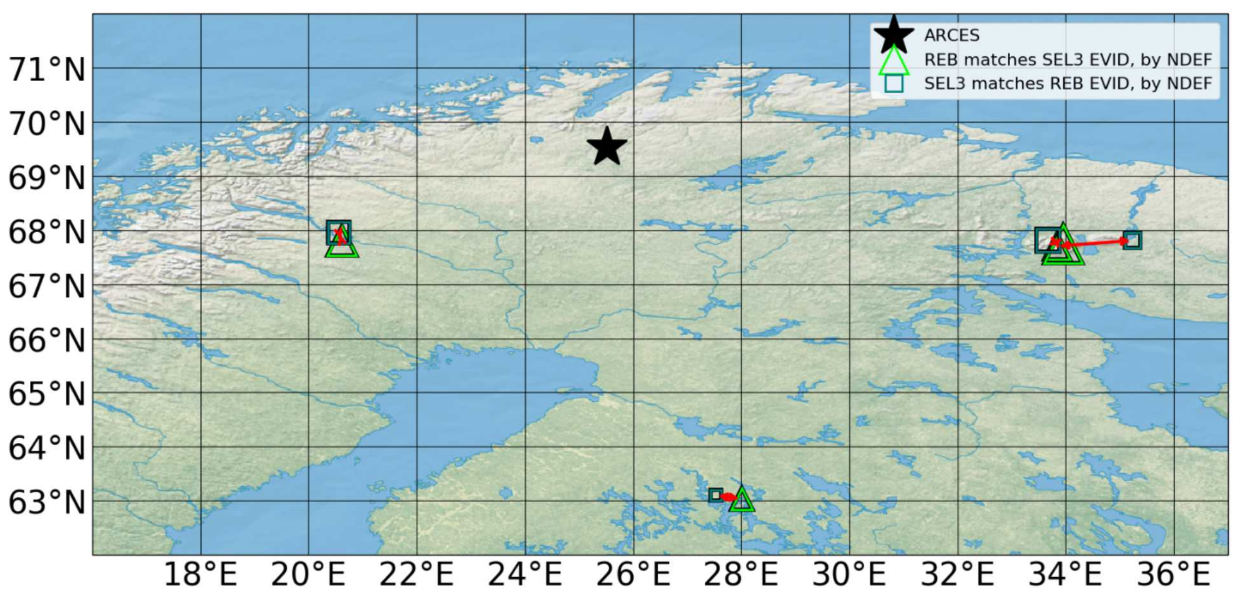

Figure 3-13. Map of 4 SEL3-REB event pairs for Scandinavia Week 1, with a red arrow to link each pair and show the analyst relocation of the SEL3 event to the final REB event location.

For Scandinavia Week 1, the waveform correlation events are shown with the SEL3 and REB events (Figure 3-14). The REB events are matched to waveform correlation events within distance and time tolerances of (1°, ±15s). The waveform correlation events are adjacent to the REB events in comparison with the SEL3 events, indicating the potential improvement in initial location; the reader is also referred to the similar map (Figure 2-53) that does not include the SEL3 events for location comparison.

Three REB events were not matched by waveform correlation because there were no templates in this area (68°N, 37°E) and waveform correlation can only detect recurring events. Moreover, these three REB events are based on infrasound detections and were unlikely to be detected by waveform correlation with seismic stations even if there were template events for the location.

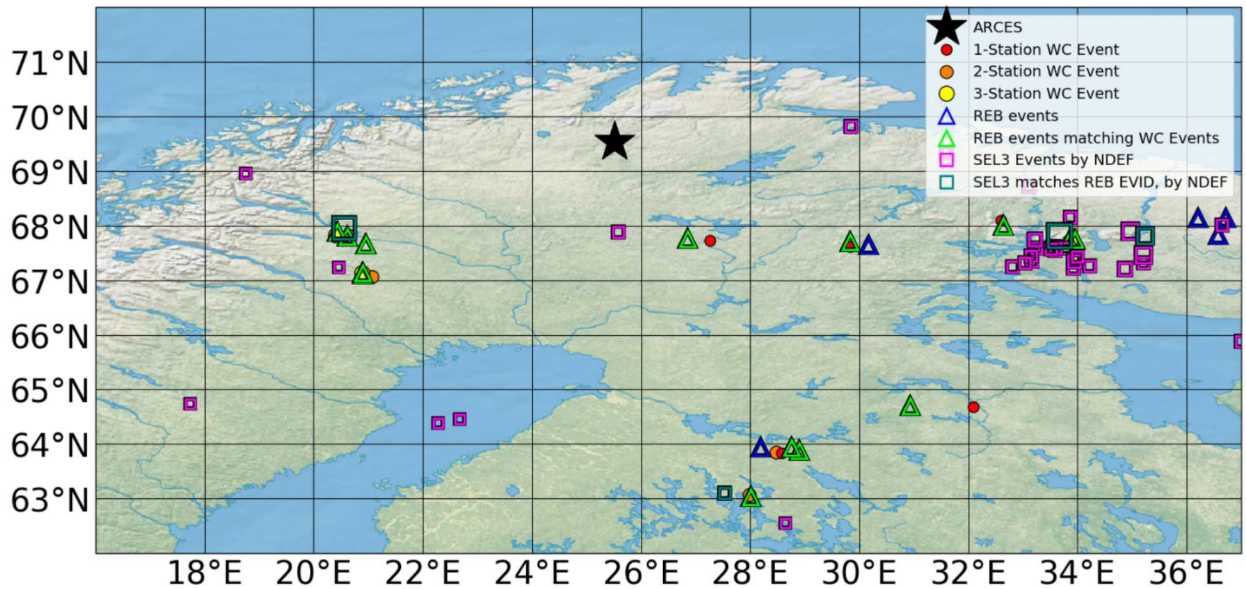

Figure 3-14. Map of SEL3, REB, and waveform correlation (WC) events for Scandinavia Week 1. The waveform correlation events are color-coded by number of detecting stations (ARCES, FINES, NOA). 14 REB (green triangle) events match WC events within tolerances of (1°, ±15s), while 5 REB (magenta triangle) events do not match within the tolerances. 4 SEL3 (dark green) events were modified to become an REB event (i.e., the EVIDs are the same), where the size of the square indicates the number of associated phases. 33 SEL3 (magenta) events were discarded by analysts.

The bar graph (Figure 3-15) provides another view of the event data by category counts, where the colors match the categories on the map. There was a total of 19 REB events in Scandinavia Week 1, and 14 of the REB events match waveform correlation events within a tolerance of (1°, ±15s), which is a tight enough tolerance to ensure that one and only one waveform correlation event matches each of the REB events. The numbers of matching waveform correlation events are further divided by number of detecting stations in the second column, showing that 9 of 14 REB events were detected by more than one station using waveform correlation for Scandinavia Week 1. In comparison, only four REB events were traceable to SEL3 events by EVID. Five REB events were not detected by waveform correlation, of which three were the infrasound detections. 33 SEL3 events were discarded by analysts.

Previously we pointed out that only 21% of REB events are traceable to the automatic pipeline for Week 1 in this study region. In contrast, 73% of REB events were detected by WC within the specified tolerances.

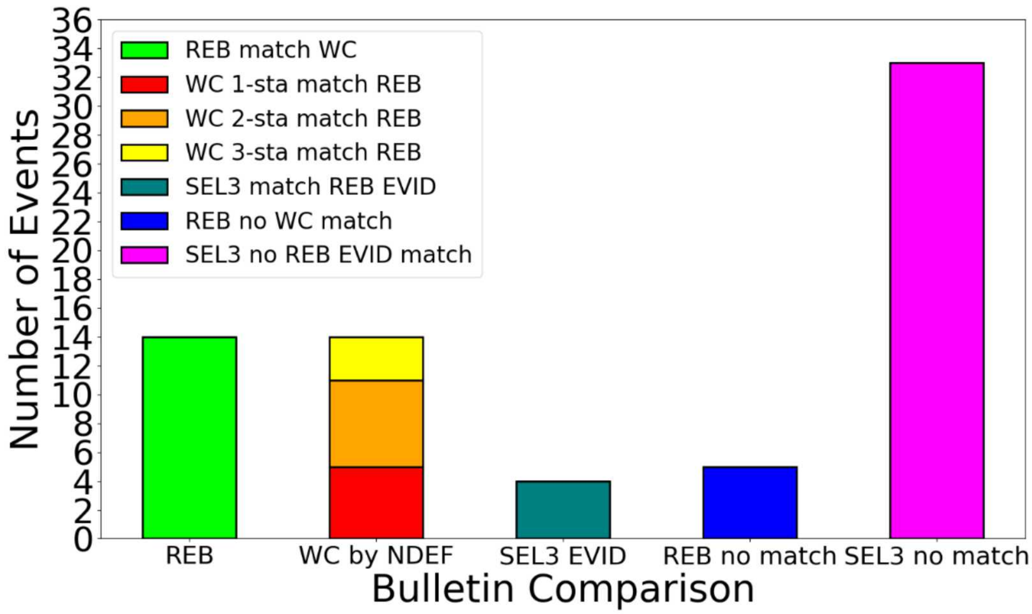

Figure 3-15. Number of Scandinavia Week 1 events that are categorized in Figure 3-14 (same colors used on map). The first green column represents 14 REB events that matched waveform correlation (WC) events within tolerances of (1°, ±15s). The second column represents the WC events that matched the 14 REB events, color-coded by number of detecting stations (ARCES, FINES, NOA). The third column represents the 4 SEL3 (dark green) events that were modified to become an REB event (same EVID). The blue column represents 5 REB events that were not detected by waveform correlation. The magenta column represents 33 SEL3 events that were discarded by analysts.

3.2.4. Scandinavia Week 2 Event Comparison

The SEL3 and REB events for Scandinavia Week 2 were mapped (Figure 3-11) with the same color scheme. Similarly to Week 1, there are a large cluster of SEL3 events near (67.5°N, 34°E) that were mostly discarded by the analysts. There were 22 SEL3 events during Week 2, but the analysts discarded all but 3 of those SEL3 events while manually building an additional 14 REB events.

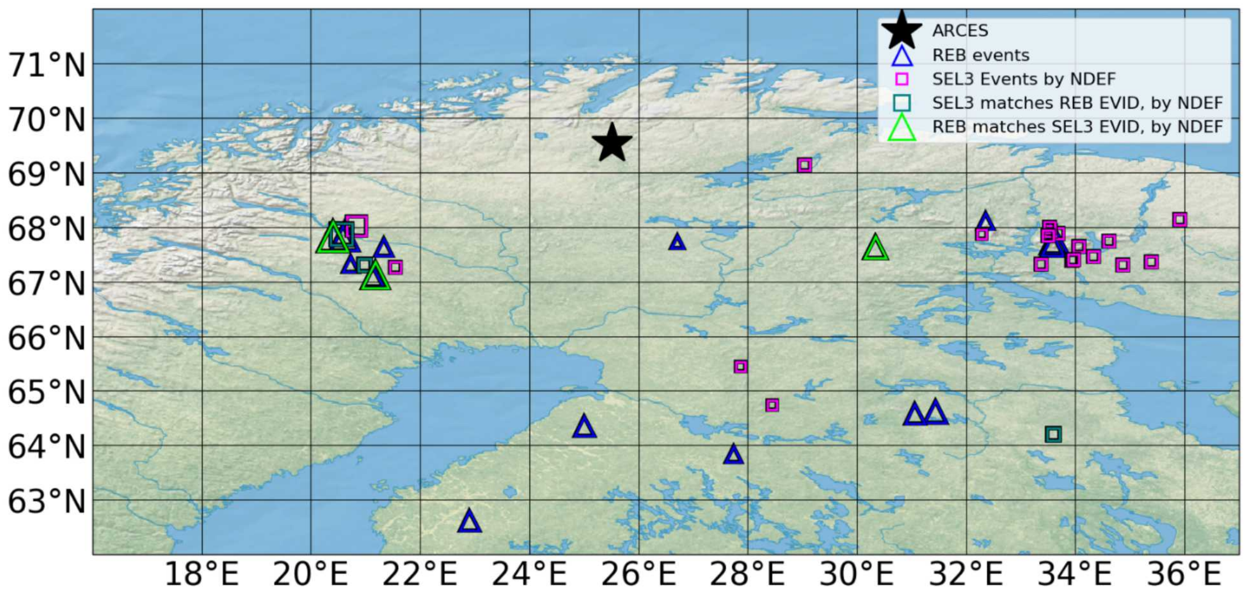

Figure 3-16. Map of SEL3 and REB events for Scandinavia Week 2, where SEL3 (square) and REB (triangle) event symbol size indicates the number of associated phases. The colors show pairs of 3 SEL3 (dark green) events that were modified to become REB (green) events (i.e., the EVIDs are the same), while 19 SEL3 (magenta) events were discarded and 14 REB (blue) events were manually built.

The bar graph (Figure 3-17) provides another view of the event data, where the colors match the categories on the map. There was a total of 17 REB events during Scandinavia Week 2. Only three REB events are traceable to SEL3 events by EVID. 19 SEL3 events were discarded by analysts and 14 REB events were manually built by analysts. Most of the discarded 33 SEL3 events had few associated phases, which can be seen by the size of the square symbol in the map. Only 18% of REB events are traceable to the automatic pipeline for Scandinavia Week 2 in this study region.

The 3 paired SEL3-REB events are mapped (Figure 3-18Figure 3-3) with connecting arrows to show the variation between the locations produced by the automatic pipeline and the final analyst-determined location. One of the SEL3 events had a very substantial relocation several degrees north.

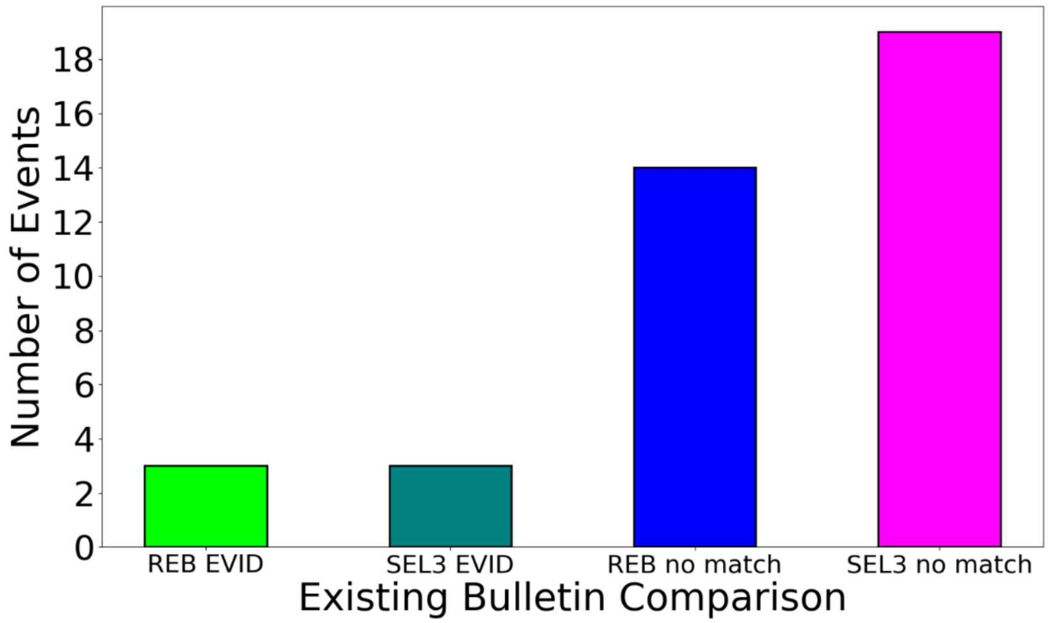

Figure 3-17. Number of Scandinavia Week 2 events that are categorized in Figure 3-16 (same colors used on map). The first two columns represent 3 SEL3 (dark green) events that were modified to become REB (green) events (same EVID). 14 REB (blue) events were manually built and 19 SEL3 (magenta) events were discarded by analysts.

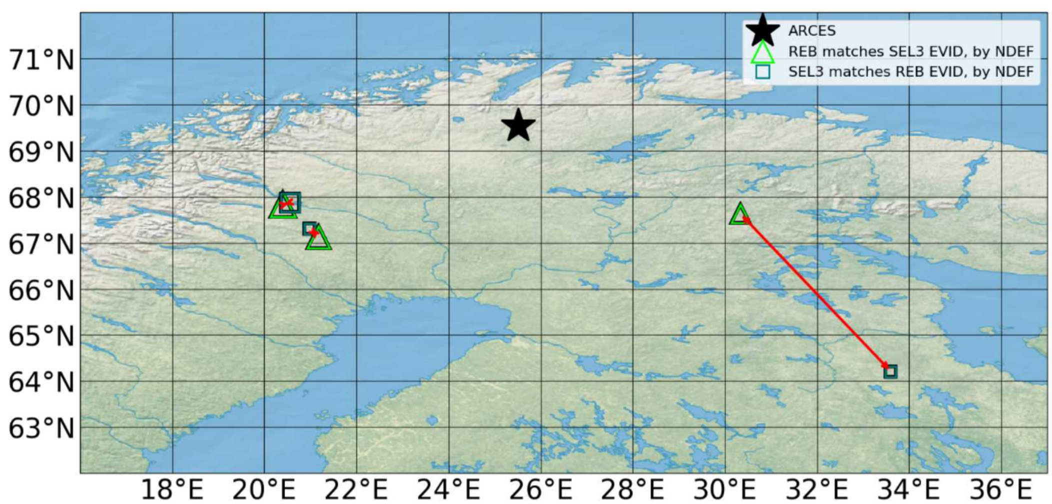

Figure 3-18. Map of 3 SEL3-REB event pairs for Scandinavia Week 2, with a red arrow to link each pair and show the analyst relocation of the SEL3 event to the final REB event location.

The waveform correlation events are mapped with the SEL3 and REB events (Figure 3-19) for Scandinavia Week 2. The REB events are matched to waveform correlation events within distance and time tolerances of (1°, ±15s). The waveform correlation events are adjacent to the REB events in comparison with the SEL3 events, indicating the potential improvement in initial location; the reader is also referred to the similar map (Figure 2-60) that does not include the SEL3 events for location comparison.

Three REB events were not matched by waveform correlation because an error in processing excluded the waveform data for that date. The remaining five REB events not found by waveform correlation include infrasound detections and had low amplitude seismic signals that were unlikely to be detected by waveform correlation.

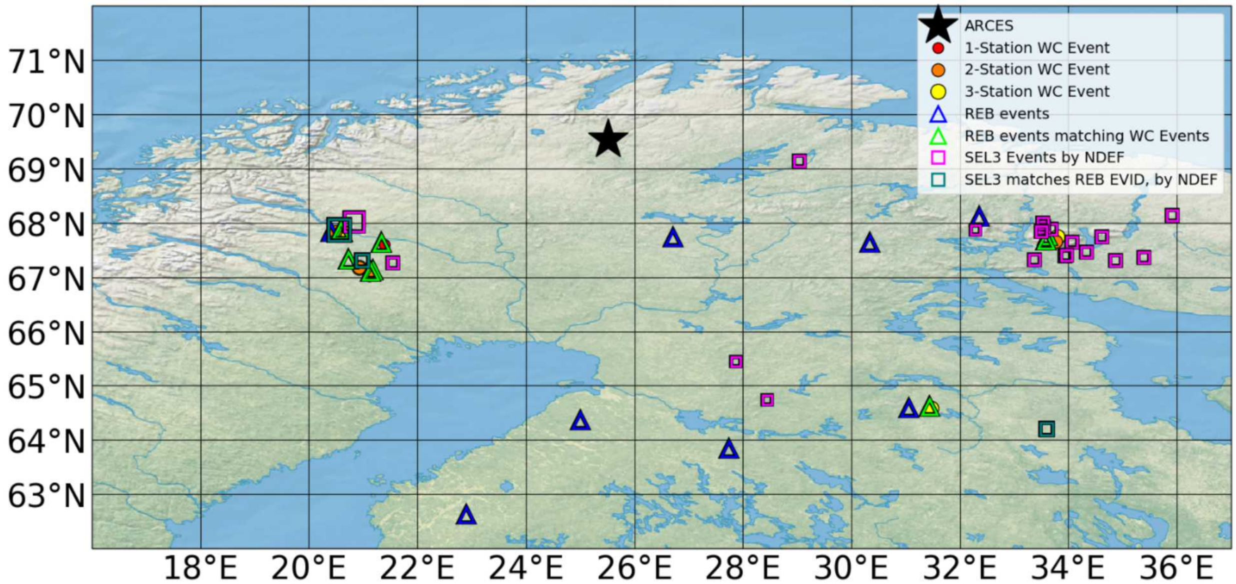

Figure 3-19. Map of SEL3, REB, and waveform correlation (WC) events for Scandinavia Week 2. The waveform correlation events are color-coded by number of detecting stations (ARCES, FINES, NOA). 9 REB (green triangle) events match WC events within tolerances of (1°, ±15s), while 8 REB (magenta triangle) events do not match within the tolerances. 3 SEL3 (dark green) events were modified to become an REB event (i.e., the EVIDs are the same), where the size of the square indicates the number of associated phases. 19 SEL3 (magenta) events were discarded by analysts.

The bar graph (Figure 3-20) provides another view of the event data by category counts, where the colors match the categories on the map. There was a total of 17 REB events in Scandinavia Week 2, and 9 of the REB events match waveform correlation events within a tolerance of (1°, ±15s), which is a tight enough tolerance to ensure that one and only one waveform correlation event matches each of the REB events. The numbers of matching waveform correlation events are further divided by number of detecting stations in the second column, showing that 8 of 9 REB events were detected by more than one station using waveform correlation for Scandinavia Week 2. In comparison, only three REB events were traceable to SEL3 events by EVID. Eight REB events were not detected by waveform correlation, of which five included infrasound detections. 19 SEL3 events were discarded by analysts.

Previously we pointed out that only 18% of REB events are traceable to the automatic pipeline for Week 2 in this study region. In contrast, 53% of REB events were detected by WC within the specified tolerances.

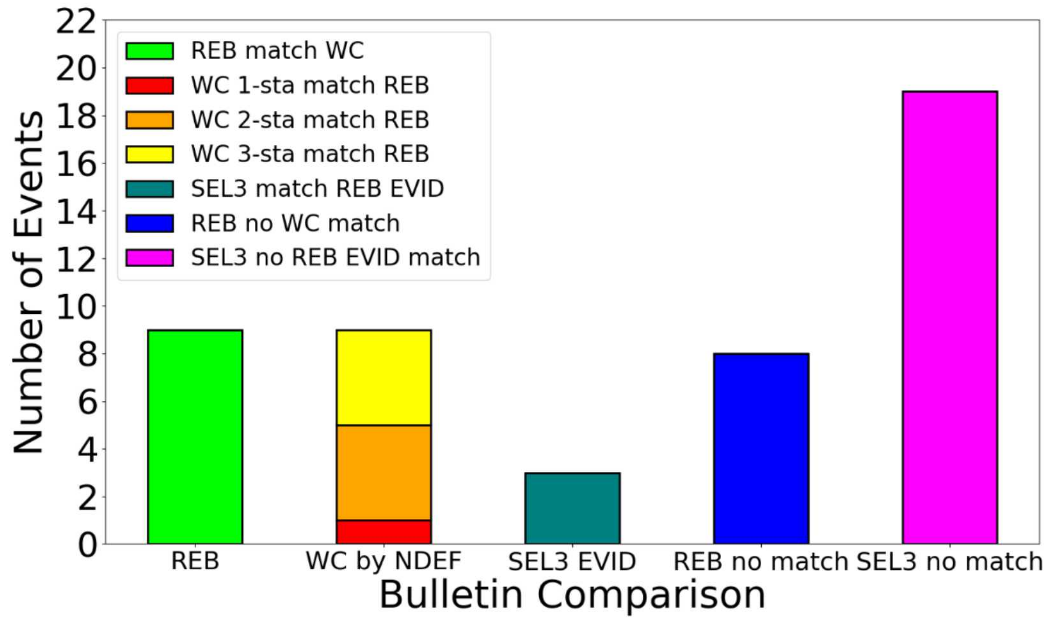

Figure 3-20. Number of Scandinavia Week 2 events that are categorized in Figure 3-19 (same colors used on map). The first green column represents 9 REB events that matched waveform correlation (WC) events within tolerances of (1°, ±15s). The second column represents the WC events that matched the 9 REB events, color-coded by number of detecting stations (ARCES, FINES, NOA). The third column represents the 3 SEL3 (dark green) events that were modified to become an REB event (same EVID). The blue column represents 8 REB events that were not detected by waveform correlation. The magenta column represents 19 SEL3 events that were discarded by analysts.

3.3. Estimating Analyst Workload Reduction and Effectiveness

The discussion presented comparisons between events detected by the automated pipeline, waveform correlation results, and the REB in visual format through maps and bar graphs. In this section we will quantify the results for both regions and weeks of study through two calculations.

First, we estimate workload reduction for analyst-built events with Equation 3-1, which compares the number of REB events that were detected by waveform correlation but were not detected by the automated pipeline. This metric was chosen because analysts must find and build these events manually, and waveform correlation shows that it can assist by reducing the number of events built this way as well as providing a more accurate initial location for the events. The SEL3 events that are traceable to REB events because the EVIDs are identical is the second term in the denominator of Equation 3-1. The numerator is calculated from the number of REB events that are located within (1°, ±15s) distance and time tolerances of waveform correlation events, but then discounting the REB events that were already detected by the automatic pipeline as SEL3 events with the same EVID.

Equation 3-1: Estimated workload reduction for analyst-built events.

$$Workload\ Reduction = \frac{\#\ REB\ events\ matching\ WC\ not\ in\ SEL3}{\#\ REB\ events\ \ -\ \ \#\ SEL3\ events\ in\ REB} \times 100$$

The calculated values of estimated workload reduction for each region and week of study are reported in Table 3-1, which estimates the incremental improvement that waveform correlation can make to the existing automatic pipeline by finding events that analysts built manually. In this table, there is an alternate set of numbers in parentheses for Week 2 of each region that offsets the effect of the processing error to exclude the REB and SEL3 events that occurred on the last day of the week. This adjusts the calculations for a 6-day week.

Table 3-1. Estimated workload reduction for analyst-built mining blast events for each region and week of study. Week 2 of each region shows (adjustment) for a 6-day week.

Study Region	Period	# REB events	# SEL3 events in REB	# REB Events matching WC not in SEL3	Estimated Workload Reduction (%)
Wyoming	Week 1	13	5	5	62.5
	Week 2	10 (9)	1	8	88.9 (100)
Scandinavia	Week 1	19	4	10	66.7
	Week 2	17 (14)	3(2)	6(5)	42.9 (41.7)

The workload reduction metric is dependent upon the number of REB events that the analysts accepted from SEL3, calculating only the incremental improvement that waveform correlation can make to the existing pipeline. We also suggest an alternate metric, the effectiveness of waveform correlation for detecting mining blast events, that is independent of the performance of the existing automated pipeline (i.e., SEL3). Equation 3-2 calculates effectiveness as the percentage of all REB mining blast events that were detected by waveform correlation.

Equation 3-2: Effectiveness of waveform correlation to detect REB mining blast events.

$$Effectiveness = \frac{\#\ REB\ events\ matching\ WC}{\#\ REB\ events} x\, 100$$

The calculated values for effectiveness of waveform correlation to detect REB mining blast events for each region and week of study are reported in Table 3-2. There is an alternate set of numbers in parentheses for Week 2 of each region that offsets the effect of the processing error to exclude the REB and SEL3 events that occurred on the last day of the week. This adjusts the calculations for a 6-day week.

Table 3-2. Effectiveness of waveform correlation to detect REB mining blast events for each region and week of study. Week 2 of each region shows (adjustment) for a 6-day week.

Study Region	Period	# REB events	# REB Events matching WC	Estimated Effectiveness (%)
Wyoming	Week 1	13	10	76.9
	Week 2	10 (9)	9	90 (100)
Scandinavia	Week 1	19	14	73.7

Study Region	Period	# REB events	# REB Events matching WC	Estimated Effectiveness (%)
	Week 2	17 (14)	9(7)	52.9 (50)

3.4. Summary

The study results show that waveform correlation can detect many of the mining blast events that are in the REB. Furthermore, by assigning the location of the detection based on that of the template event(s), waveform correlation offers more accurate initial locations than the automated pipeline events (SEL3) for the regions and time periods studied. The results suggest that incorporating waveform correlation into the IDC pipeline may decrease analyst effort to build mining blasts if similar results can be obtained for other mining regions.

The waveform templates chosen for the stations in both regions were long in duration to offer a high time-bandwidth product. It has been shown in prior waveform correlation research studies that a high time-bandwidth product results in fewer detections because the complex waveform characteristics are difficult to match but that the detections are more reliable. For example, 1303 templates drawn from up to 10 years of history were included in the template libraries for Wyoming Week 1, yet only 5% of PDAR templates detected events; conversely 95% of the templates had no detections during that week. It is necessary to point out that a high time-bandwidth product requires a long duration template window. The template windows in this study varied between 50 to 160 s and when possible included multiple phase arrivals from the same event to increase waveform complexity. This approach will only work in mining regions when the mining blasts are separated by a time period exceeding the template window and are distinct in time without overlapping signals from other events. Some mines use blasting methods that may not work as well with long templates, so we suggest a broader survey of mining regions before broadly committing to the approach that worked in this study.

The emphasis in this study was to reproduce the REB bulletin, not to increase the number of mining blast detections. Even so, waveform correlation detected an order of magnitude more events for all regions and weeks of study than appeared in the REB. The detections may be legitimate events yet could increase analyst workload if such a sensitive detection mechanism is incorporated in the automatic pipeline. The time-reverse method of template cross-correlation threshold setting works well for very high time-bandwidth templates, and when a template is created to properly window an arrival the detections are credible. One obvious means to reduce the number of unwanted waveform correlation detections is to increase the thresholds determined by the time-reverse method by a standard amount (e.g., increase them by 5%). However, we found that some detections that matched REB events were close to the time-reverse threshold values; thus, we believe that a more complex approach is needed. We recommend further investigation of the following issues before integrating waveform correlation into an operational pipeline:

1. **Template selection methods**. Choosing template waveforms that have a high time-bandwidth product and properly windowed arrivals with sufficiently high signal-to-noise ratio is crucial to the performance of a waveform correlation system. A large and diverse template library is important to gain the most information from the detections such as accurate origin time and location, but this quickly becomes unwieldy for human review. Algorithms for template review, selection and windowing are needed to set up a global collection of template libraries. In addition, we recommend the ability to disable templates

in an operational pipeline if undesirable detection trends appear; for example, a template with a high-amplitude spike can generate thousands of false detections.

2. **Selecting practical waveform correlation detections**. Waveform correlation is a sensitive detection technique and has been proven to detect signals that are an order of magnitude smaller (and more) than the template waveform. Waveform correlation will not reduce the overall workload of the analyst if significant time must be spent rejecting additional small events. The approach of including waveform correlation detections as input to the pipeline associator, currently under investigation by the IDC, may prevent flooding the analyst with small events. However, we recommend continued research to develop methods to select practical waveform correlation detections for operational pipelines that can be adjusted according to the concept of operations and operational goals of the monitoring agency.
3. **Determination of analyst workload reduction**. Further study is required to develop an estimate of analyst workload reduction in the presence of a waveform correlation extension to the operational pipeline. Our results presume that events not detected by SEL3 that are later manually built by analysts are time-consuming, and that a waveform correlation event with assigned accurate initial location will save the analysts time. The IDC initiated the project that this study is part of, and we expect that the IDC's final report will provide an evaluation of results that includes a method for determining a realistic estimate of analyst workload reduction for true and false waveform correlation detections. Such a realistic estimate was deemed to be beyond the scope of the analysis for this report.

The conclusion of this report, based on processing data for two non-adjacent weeks from two regions that are far from each other and very different, is that waveform correlation shows promise as a technique for reducing human analyst workload for recurring mining blasts. The addition of valid waveform correlation events to the automatic pipeline would lead to fewer events that have to be manually built by the analysts. Moreover, the accuracies of locations assigned by waveform correlation is better than those of SEL3 events, which could save analyst time spent in correcting associations and relocating. We recommend that further research focuses on the issues enumerated above to develop pragmatic tools that are required to adopt waveform correlation into an automatic pipeline for recurring mining blast events.

REFERENCES

[1] Slinkard, M., S. Heck, D. Schaff, N. Bonal, D. Daily, C. Young, and P. Richards (2016). Detection of the Wenchuan Aftershock Sequence Using Waveform Correlation with a Composite Regional Network, *Bull. Seismol. Soc. Am.* **106**, 1371-1379. doi: 10.1785/0120150333.

[2] Slinkard, M. E., D. B. Carr, and C. J. Young (2013). Applying waveform correlation to three aftershock sequences, *Bull. Seismol. Soc. Am.* **103**, 675-693. doi: 10.1785/0120120058.

[3] Sundermier, A., R. Tibi, and C. J. Young (2019). Applying Waveform Correlation to Aftershock Sequences Using a Global Sparse Network, *Technical Report SAND2019-10184*.

[4] Slinkard, M., D. Schaff, N. Mikhailova, S. Heck, C. Young, and P. G. Richards (2014). Multistation validation of waveform correlation techniques as applied to broad regional monitoring, *Bull. Seismol. Soc. Am.* **104**, 2768-2781. doi: 10.1785/0120140140.

[5] Anderson, J., W. E. Farrell, K. Garcia, J. Given, H. Swanger (1990). Center for Seismic Studies Version 3 Database: Schema Reference Manual. *Technical Report C90-01*. http://jkmacc-lanl.github.io/pisces/data/Anderson1990.pdf

[6] Ganter, T., A. Sundermier, and S. Ballard (2018). Alternate Null Hypothesis Correlation: A New Approach to Automatic Seismic Event Detection, *Bull. Seismol. Soc. Am.* **108**, 3528-3547. doi: 10.1785/0120180074.

[7] Tibi, R., C. Young, A. Gonzales, S. Ballard, and A. Encarnacao (2017). Rapid and robust cross-correlation-based seismic signal identification using an approximate nearest neighbor method, *Bull. Seismol. Soc. Am.* **107**, 1954–1968. doi: 10.1785/0120170011.

[8] Schaff, D. P., and P. G. Richards (2004). Lg-wave cross correlation and double difference location: application to the 1999 Xiuyan, China, sequence, *Bull. Seismol. Soc. Am.* **94**, 867-879. doi: 10.1785/0120030136.

[9] Schaff, D. P., P. G. Richards, M. Slinkard, S. Heck, and C. Young (2018). Lg-Wave Cross Correlation and Epicentral Double-Difference Location in and near China, *Bull. Seismol. Soc. Am.* **108**, 1326-1345. doi: 10.1785/0120170137.

[10]Sundermier, A., M. Slinkard, J. Perry, D. P. Schaff, C. J. Young, P. G. Richard (2017). Automated Techniques for Waveform Correlation Applied to Regional Monitoring of Eastern Asia. Poster presented at: *CTBT: Science and Technology 2017 Conference (SnT2017)*, June 26-30, Vienna, Austria.

[11]Arora, N. S., S. Russel, and E. Sudderth (2013). NET-VISA: Network processing vertically integrated seismic analysis, *Bull. Seismol. Soc. Am.*, **103**, no. 2A, 709-729. doi: 10.1785/0120120107.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Technical Library	01177	libref@sandia.gov

This page left blank

This page left blank