ANALYST EVALUATION OF MODEL-BASED BAYESIAN SEISMIC MONITORING AT THE CTBTO

Nimar S. Arora[1], Jeffrey Given[2], Elena Tomuta[2], Stuart Russell[1,3], Spiro Spiliopoulos[2]

Bayesian Logic, Inc.[1], the Comprehensive Nuclear-Test-Ban Treaty Organization[2], and
the University of California at Berkeley[3]

Sponsored by the Comprehensive Nuclear-Test-Ban Treaty Organization

ABSTRACT

NET-VISA (NETwork processing Vertically Integrated Seismic Analysis) is a generative probabilistic model of global-scale seismology and an inference algorithm which deduces the seismic bulletin with the highest posterior probability given all the seismic detections (aka arrivals or triggers) and misdetections observed by a network of stations. It has been developed to potentially succeed the current automatic association tool – GA (Global Association). Early testing has already demonstrated that this probabilistic approach to treaty monitoring is able to find weak seismic events that were either mislocated or missed entirely by GA. In this research, we provide analyst verification of these early results.

Our general approach is to run GA and NET-VISA in parallel, on the same set of detections, all the way through to analyst evaluation and to compare the quality of the final products. Currently, the final GA bulletin SEL3 is the basis for the analyst-produced bulletin LEB. In parallel, we are now producing a new NET-VISA bulletin named VISA which is the basis for an analyst-produced bulletin LEB_VISA. We have compared the two LEB bulletins in terms of the number of events that they contain and the number that satisfy the REB (Reviewed Event Bulletin) criteria. We restrict our comparison to only the events that the analysts created from the underlying automatic bulletin, i.e., no manually added events were considered. This restriction together with the fact that the same analyst worked on any given time interval in either bulletins helps ensure that we are not measuring analyst skill.

Our experiments indicate that NET-VISA finds roughly 20% more events than GA. Of the events found by both NET-VISA and GA, NET-VISA associates on average two additional stations per event. With more associated stations and with a probabilistic model which accounts for misdetections and detection amplitudes, the locations determined by NET-VISA tend to be much closer to the final event locations after analyst review than those found by GA.

OBJECTIVES

The purpose of this study is to evaluate NET-VISA for routine operational use at the International Data Centre of the CTBTO. Towards this end, IDC analysts were asked to review an event list produced by NET-VISA in addition to the normal GA event list for each time slot allotted to them over a period of one week. We report here the subjective feedback of the analysts as well as the empirical analysis of the resulting analyst-reviewed bulletins.

This work builds on an earlier evaluation of NET-VISA (Le Bras et al., 2011) that compared the event in the NET-VISA event list directly with the events in the automatically-generated Third Standard Event List (SEL3) and the analyst-reviewed Late Event Bulletin (LEB).

RESEARCH ACCOMPLISHED

Background

NET-VISA (Arora, 2012, Arora and Russell, 2012) is designed to address the network processing problem in seismology, i.e. it deduces the set of seismic events given the set of detections on a network of stations. The general approach is to build a generative probabilistic model of global-scale seismology and to use this model to infer the set of seismic events which have highest posterior probability given the observed detections and misdetections.

Our approach differs in various aspects from the analysis currently in operations at the IDC, which processes the data in stages. In the IDC processing pipeline, detected signals from a single station are examined first to determine an initial phase label (e.g. Pn, P, Lg). Next, the detections from the entire network are clustered together to form events based on consistency in their arrival time, azimuth, slowness, amplitude, and phase label. During this stage, the phase labels may be modified according to limited transformation rules. The locations of these events are computed using the time, azimuth, and slowness of the associated detections and the assigned phase labels. Further attributes of the event like magnitude are then computed from the amplitude measurements. Finally, other global measures as well as other detection attributes like SNR are examined using simple heuristics to determine event quality, and then a decision is made to either keep the event or discard it.

In contrast, our approach incorporates all aspects of this pipeline into a single vertically integrated generative model. The inference in our approach uses all the attributes of the detections – time, azimuth, slowness, amplitude, signal-to-noise ratio—as well as all the misdetections to identify the events. The phase labels can be freely modified in order to associate detections with an event. Finally, the decision to keep or discard an event is based solely on the probability of the world where the event exists and the alternate world without the event.

The generative model consists of multiple subcomponents that are combined together by the independence and conditional independence assumptions in the model. The various subcomponents are:

- A distribution of possible event locations that is learned from the historical distribution of seismicity plus a uniform distribution to allow for explosions anywhere on the earth.
- The Gutenberg-Richter distribution of event magnitude.
- A station- and phase-specific detection probability that depends entirely on the event magnitude, depth, and distance to station. The probability model (logistic regression with local input features) is calibrated from historical data that record which events were detected by which stations.
- For each detection, the following predictive distributions over attributes of the detection:
    - A distribution for the measured arrival time given the predicted arrival time for the true phase from the event location to the station. NET-VISA currently uses the 1-D model iasp91 with station-specific corrections, i.e., the same model as the CTBTO. The uncertainty captures both the travel time uncertainty as well as the pick error. This is modeled as a Laplacian distribution.
    - A distribution for the measured amplitude given the event magnitude and distance – i.e., an attenuation model. The log amplitude is assumed to decay linearly with distance after an initial exponential drop off, with Gaussian error.
    - A multinomial distribution for the measured phase label given the true phase label, again estimated from historical data.

    - Distributions for the measured azimuth and slowness given the true azimuth and slowness for the event also follow a Laplacian distribution.

- Two types of false detections are modeled:
    - Independently generated false detections for each station are modeled by a time-homogeneous Poisson distribution. Each such detection has its amplitude drawn from an empirically estimated distribution (a mixture of Gaussians), its phase label from an empirically estimated multinomial distribution, and its azimuth and slowness from uniform distributions.
    - Coda false detections are randomly triggered by both true and false detections depending on the amplitude of the prior detection. The attributes of the coda detection are strongly correlated with the corresponding attributes of the prior detection.

The inference algorithm operates by a sequence of global and local moves which iteratively improve the deduced world, where a world is a complete specification of the set of events, their associated detections, as well as the independent and coda false detections. The initial world has no events, and all detections are marked as false. The following moves are then executed to improve the world:

- A *birth move* that adds a number of events. These events are created by inverting individual detections such that the proposed event is well supported by the other arrivals.
- An *improve detection move* that finds the best event and phase for each detection or marks it as false (independent or coda).
- An *improve event move* that changes the event location, time, and magnitude to the best possible given the currently associated set of detections and misdetections.
- Finally, a *death move* that kills unsupported events.

Analyst Evaluation

Procedure

The analyst evaluation of NET-VISA was designed to mimic the current procedures at the CTBTO for the review of the SEL3 bulletin, which is the final GA (Le Bras et al., 1994a,b) product. A new bulletin named VISA was produced by the NET-VISA software running 6 hours behind real time, similar to SEL3, and using the same automatic seismic detections as are available to SEL3. (Note that only seismic detections were used for the VISA bulletin in contrast to SEL3 which also includes infrasound and hydroacoustic detections.) The analysts were asked to review the VISA bulletin using the usual ARS (Analyst Review Station) software and to save the results in a new LEB_VISA bulletin. In the interest of time, the analysts were not asked to manually add any events to LEB_VISA, as is normally done for SEL3 using a tool called Scanner. In the opinion of the analysts, the manually added events are entirely due to the skill of the analysts and reflect nothing of the underlying automatic bulletin. It was thus decided to compare the LEB_VISA events with only the automatic events in the LEB bulletin.

In the normal procedure, the LEB bulletin that is produced by an associate analyst is reviewed by a lead analyst. To avoid this additional overhead for the LEB_VISA bulletin, only lead analysts were asked to review it. Finally, in order to rule out analyst skill from affecting the results, the analyst assigned to work on a time block for LEB_VISA was asked to work on the identical time block for LEB.

Empirical Analysis

A total of 27 hours of VISA events were reviewed by the analysts during the two week period from June 11 to June 25, 2012. In these 27 hours, the LEB bulletin had 101 automatic events and the LEB_VISA bulletin had 119 automatic events. The magnitude distribution of these automatic events is shown in Figure 1. The magnitude distribution shows that half of the additional events found by NET-VISA are weak regional events. The rest of the additional events are spread out evenly between low and high magnitudes.

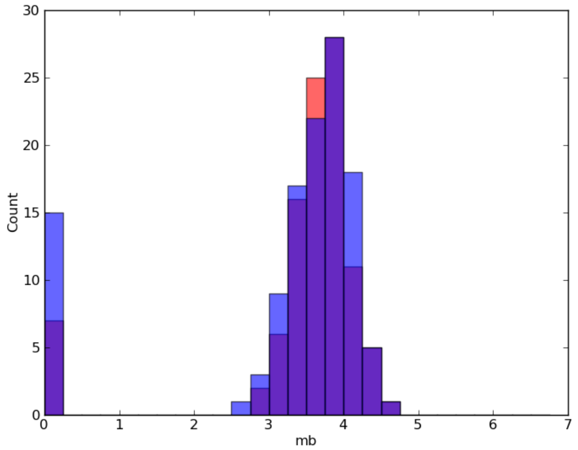

Figure 1. **The body-wave magnitude distribution of automatic LEB events, in red, and LEB_VISA events, in light blue. The overlapping counts are shown in dark blue. A magnitude of zero is reported for an event which has only regional phases, and body-wave magnitude cannot be computed.**

NET-VISA and GA had 92 common events. Of the nine GA events not in NET-VISA, five met the REB criteria. On the other hand, of the 27 NET-VISA events not in GA, 15 met the REB criteria. These 27 events included 25 that the analysts manually added in LEB (a total of 34 events were manually added in LEB). 2 NET-VISA events were missed even by the analysts (of these two one was a four-station regional event and the other a three-station magnitude 2.9 event, neither met the REB criteria).

For the common events, NET-VISA associated on average 2.2 more time-defining phases that were also retained by the analysts than GA. These additional phases translated into 1.8 additional stations, on average, with time-defining phases. The distribution for the number of additional phases associated by NET-VISA is shown in Figure 2 and the number of additional stations in Figure 3.

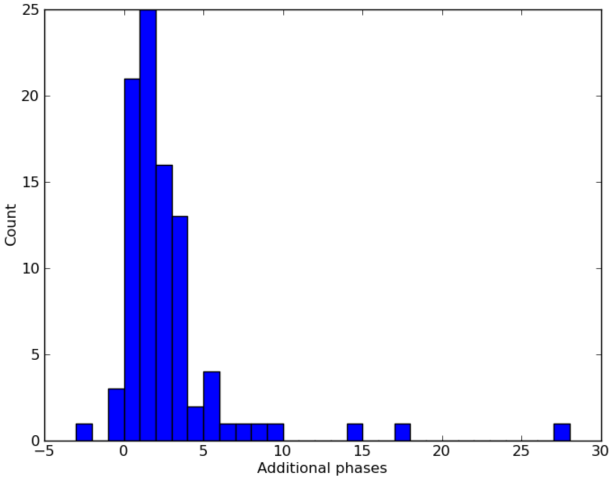

Figure 2. **The distribution of the number of additional time-defining phases found by NET-VISA versus GA.**

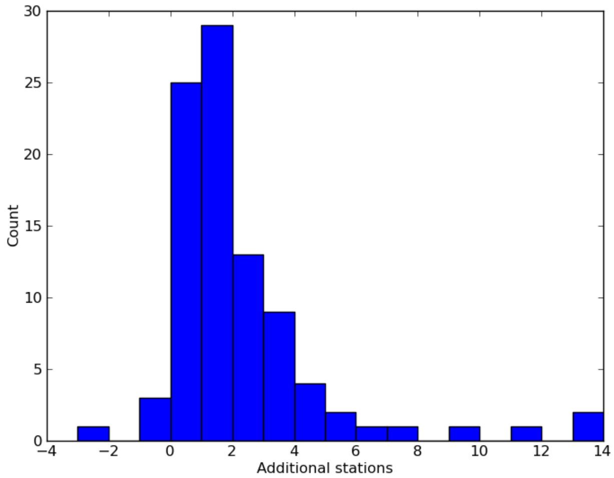

Figure 3. The distribution of the number of additional stations with time-defining phases associated by NET-VISA versus GA.

The locations of the GA events tended to move considerably more than NET-VISA locations after analyst review. Again restricting ourselves to the common events found by both GA and NET-VISA we found that GA locations changed on average 297 km versus 152 km for NET-VISA locations. The median location change for GA and NET-VISA was 82.5 km and 75.3 km respectively. The statistics of these location errors are displayed in Figure 4. The actual errors in location are displayed in Figure 5.

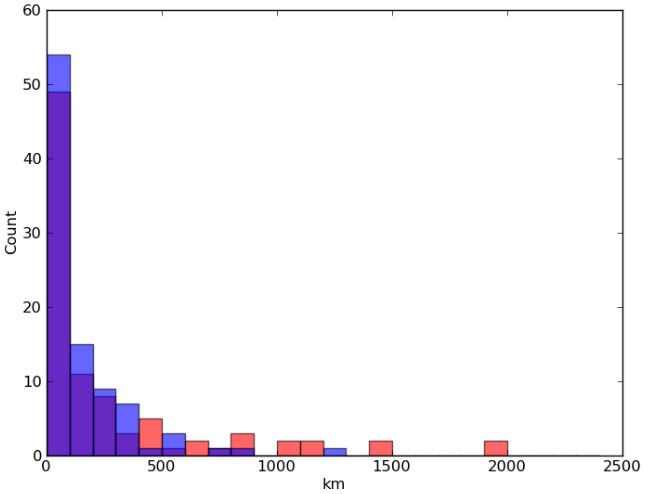

Figure 4. The distribution of the change in the GA location after analyst review (in red) versus the distribution for NET-VISA (in light blue, overlapping counts are shown in dark blue).

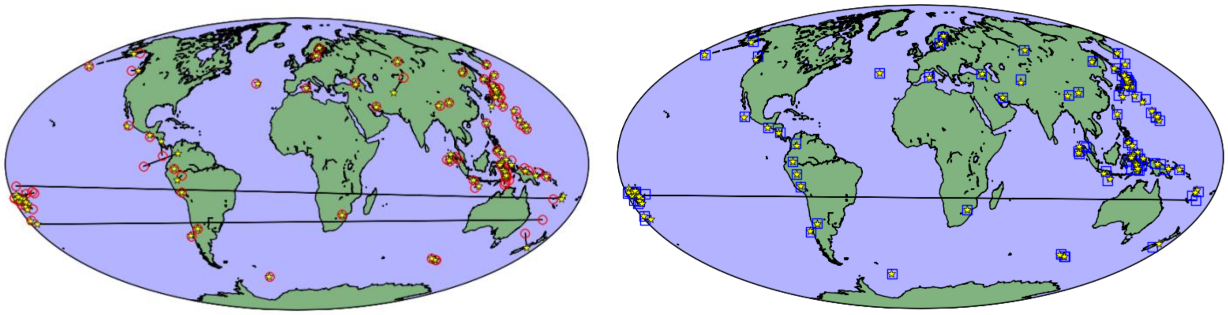

Figure 5. The changes in location of GA events (left figure, red circles) to the corresponding LEB events (yellow stars) and the changes in location of NET-VISA events (right figure, blue squares) to the corresponding LEB_VISA events (yellow stars). Straight lines (and not great circle lines) are drawn between corresponding events.

However, the improvement in NET-VISA locations is not entirely due to the additional phases and stations. In Figure 6, we plot the distribution in location errors for events with an identical number of stations with time-defining phases. For these events, the GA location changed on average 360.4 km and the VISA location by 226.4 km. The median location change for GA and NET-VISA was 157.9 km and 114.0 km respectively.

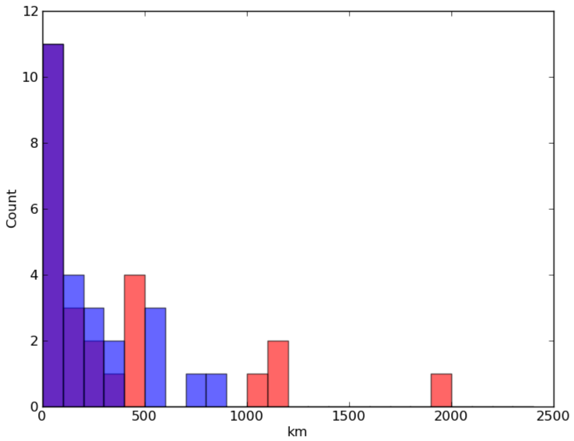

Figure 6. For events with an identical number of stations with time-defining phases, the distribution of the change in the GA location after analyst review (in red) versus the distribution for NET-VISA (in light blue, overlapping counts are shown in dark blue).

In Figure 7, we plot the distribution of depth errors for the common GA and NET-VISA events. The average of the absolute GA depth error was 70.3 km while NET-VISA had an error of 84.3 km. The median depth error was 0.0 km and 8.1 km for GA and NET-VISA respectively. This difference appears to be due to the fact that NET-VISA has no *depth prior*. Currently, the depth of each event in NET-VISA is assumed to be uniformly distributed. GA, on the other hand, has *depth cells* in regions of known deep seismicity.

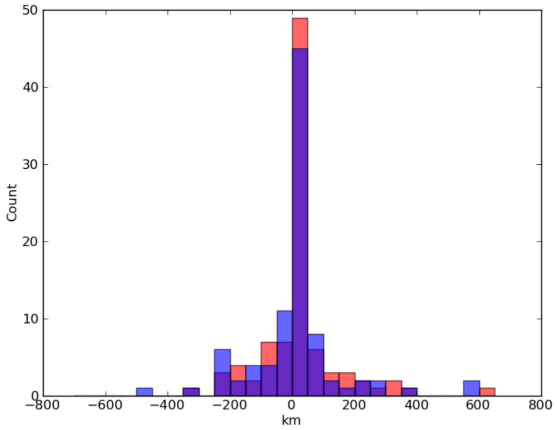

Figure 7. **The distribution of the errors in the GA depth (in red) versus the distribution for NET-VISA (in light blue, overlapping counts are shown in dark blue). A positive error represents a higher value for the automatic event versus the reviewed event.**

The number of false events found by GA and NET-VISA were comparable at 46 and 50 respectively. The magnitude distribution of these false events is displayed in Figure 8 and the locations of these events are displayed in Figure 9. The magnitude distribution shows that the NET-VISA false events are mostly in the 2 to 3 range (2 is the minimum magnitude that NET-VISA can report for an event, these are typically regional events). GA seems to have many more false events in the 3 to 4 range than NET-VISA.

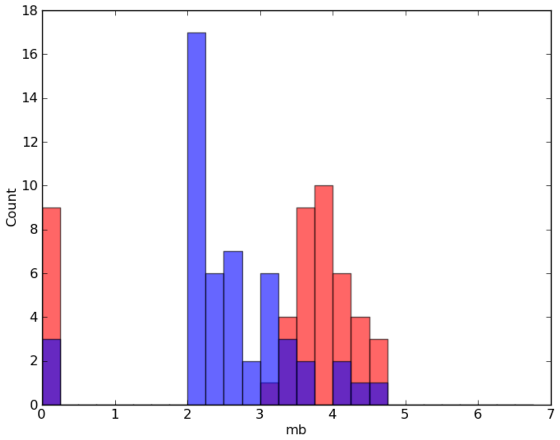

Figure 8. The distribution of the magnitudes of false events reported by GA (red) and NET-VISA (light blue). Overlapping counts are shown in dark blue.

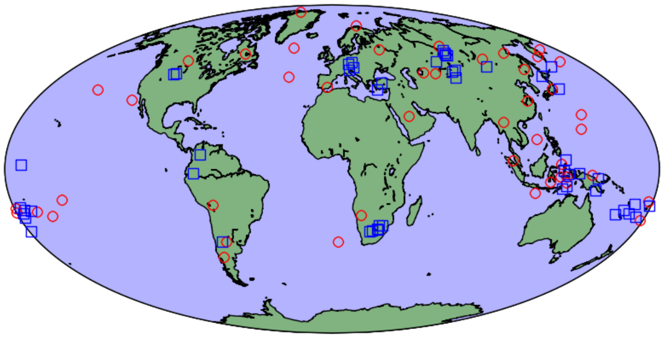

Figure 9. The locations of false events reported by GA (red circles) and NET-VISA (blue squares).

Subjective Feedback

The most common analyst feedback was regarding the usage of the so-called "depth phases." These are phases like Lg, which constrain events to be close to the surface, or pP which provide strong depth information. The analysts felt that because NET-VISA used these phases very early on in its processing, it severely biased the NET-VISA location and sometimes led to inferior or missed events.

The other frequent complaint was that NET-VISA seemed to mark all non-noise phase values of time, slowness, and azimuth as time-defining while analysts had specific policies regarding this. The analysts were thus forced to mark many of the attributes not time-defining before they could *locate* the event.

Some of the analysts were a bit puzzled by the *event definition criteria* used by NET-VISA. For example NET-VISA would sometimes create a very weak event with exactly one station, but with multiple phases – Pn, Sn, Lg. This was often because there was no data available in nearby stations to prove or disprove the existence of the event.

A more subtle issue was noticed by an analyst who rejected an S phase with slowness 7 for an event which already had a P phase with slowness 9 at the same station. The analyst pointed out that the S phase was normally expected to have a higher slowness than the P phase even with the usual measurement error.

Notwithstanding the above issues, the analyst had in general a very positive impression of NET-VISA. They noticed that many of the phases which they had to manually add were automatically associated. It was also immediately obvious to the analysts that NET-VISA found many more events which heretofore had to be built manually using Scanner. This earned NET-VISA the epithet "GA plus Scanner."

CONCLUSIONS AND RECOMMENDATIONS

The VISA bulletin is clearly an improvement over the current SEL3 bulletin as demonstrated by the number of events located (20% additional events), the location of the events, and the number of associated phases (2.2 additional phases and 1.8 additional stations). These improvements had previously been hypothesized by matching VISA and LEB events. However, with the analyst feedback there is now concrete evidence of this improvement.

The evaluation by the analysts has identified a number of potential improvements to NET-VISA.

- Depth phase travel times should be *tempered* to not influence the location very early on in the inference.
- Depth prior needs to be added to the model to better estimate the depth of events in regions of known deep seismicity.
- Phase attributes should be modeled as correlated to each other rather than independently generated.
- A careful analysis needs to be made of the small number of GA and Scanner events which NET-VISA is still missing to see if the birth proposer needs any improvement.

- A final round of post-processing needs to be done on the NET-VISA bulletin so that it has time-defining attributes marked as expected by analysts.

ACKNOWLEDGEMENTS

We are thankful to Dr. Lassina Zerbo, IDC Director, for allowing us to publish this research and for his continued support of the development of the NET-VISA software. The research in this paper could not have been possible without the diligence of the various lead analysts – Tajan, Ali Kasmi, Ezekiel Jonathan, Jane Gore, Marcela Villarroel and Kirill Sitnikov – who produced the LEB_VISA bulletin and provided detailed feedback.

REFERENCES

Arora, N. S., S. Russell (2012). A model of seismic coda arrivals to suppress spurious events, *European Geophysical Union* (EGU2012-6763*).*

Arora, N. S. (2012). Model-Based Bayesian Seismic Monitoring. University of California, Berkeley. Technical Report No. UCB/EECS-2012-125.

Le Bras, R., H. Swanger, T. Sereno, G. Beall, and R. Jenkins (1994a). Global association. Technical Report ADA304805, Science Applications International Corp, San Diego, CA.

Le Bras, R., H. Swanger, T. Sereno, G. Beall, R. Jenkins, and W. Nagy (1994b). Global association: Design document and user's manual. Technical Report SAIC-94/1142, Science Applications International Corp, San Diego, CA.

Le Bras, R., S. Russell, N. Arora, and V. Miljanovic. Machine Learning at the CTBTO (2011). Testing and Evaluation of the False Events Identification (FEI) and Vertically Integrated Seismic Association (VISA) Projects, in *Proceedings of the 2011 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, LA-UR-11-04823, Vol 1, pp. 313–321.

DISCLAIMER

The views expressed in this paper are those of the authors and do not necessarily reflect the views of the CTBTO Preparatory Commission.