SnT 2019
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

CTBTO PREPARATORY COMMISSION

Exploiting Bayesian inference priors to form synthetic waveform events or to validate events formed by automatic processing.

Noriyuki Kushida[*1] and Ronan Le Bras[*1]

[*1] Comprehensive Nuclear Test-Ban Treaty Organization, International Data Center

T3.5-P32

Introduction

Abstract

The Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization (**CTBTO**) has been developing and testing **NET-VISA**, a Bayesian automatic event detection and localization software package. In our testing at the CTBTO, NET-VISA shows better performance than its currently operating automatic association program. This implies that the priors of Bayesian inference built within the NET-VISA framework have a good ability to depict the relationship between sources of waveform events and detections on CTBTO's stations. At the same time, synthesizing a good event, or evaluating and validating events quantitatively are important topics to improve the performance of detecting waveform events, and therefore, detecting a nuclear explosion. In this context, we have been building methodologies to apply the NET-VISA priors onto this **synthetic event generation** and the **event evaluation**. In the presentation, the quality of synthetic events and the feasibility of validating events formed by the currently operating association program are discussed.

NET-VISA overview

NET-VISA is a waveform (seismic, hydro-acoustic and infrasound) event detector based on a generative probabilistic model. The basic idea of the generative model is that supposing probability distribution models which should explain the real world phenomena, and fitting (learning) the models with observed data. NET-VISA finds the hypothetical events which maximize the total probability of models which explain the observed detections at stations over the globe.

In the process of maximizing the total probability, NET-VISA repetitively creates hypothetical events and validates such events using the learned model. This process can be applied to the purposes of the present study, namely, building an event validator and a synthetic event generator.

Brief Summary

The developed event validator identifies false/true events over 90% of accuracy, and a synthetic event imbeded in real data was correctly reconstructed by NET-VISA

Event Validator

Objective

Global Association (GA), which is the conventional automatic waveform event detection software in CTBTO, is currently in operation. One of the weak points of such automatic waveform event detection software is that false events can be formed. At the same time, detailed analyses need to be performed on suspicious events manually by analysts. In line with the context, a method which evaluates the likeliness of a false/true event can reduce the analysts' burden.

Methodology

NET-VISA repetitively generates hypothetical events and evaluates the likelihood by scoring each event based on priors (learned model). This scoring process can be repurposed to evaluate the events formed by GA.

Results: True event: Positive score, False event: Negative score

DPRK events Scores (true events)

	REB (analysts' review)	SEL3 (GA)	SEL2 (GA)	SEL1 (GA)
	date	SCORE		
DPRK6	September 3, 2017	1073.8	528.0	528.0	329.8
DPRK5	September 9, 2016	929.4	461.3	462.1	286.3
DPRK4	January 6 , 2016	581.7	431.9	431.9	331.7
DPRK3	February 12, 2013	386.1	300.3	300.3	272.0
DPRK2	May 25, 2009	392.4	442.7	442.7	262.5
DPRK1	October 9, 2006	88.3	92.1	92.1	92.1

DPRK aftershock scores (true events)

SEL3 (GA) orid	original	modified rule *1
14801637	5.9	6.5
14885480	**-31.0**	-26.0 *2
14885852	9.8	10.2
14959072	2.4	2.4
15151076	**-3.0**	5.3
15165764	16.1	16.1
15165850	15.7	15.7
15398929	11.8	11.8
15722839	9.7	9.7
16768704	37.4	37.4

*1 Modified rule: GA events can contain physically unreasonable detections that turns the entire event identified false. Such detections have negative detection-wise score, and removed automatically.

False GA events formed around DPRK in 2018

- 47 out of 52 false events are scored negatively
  - 5 are wrongly labelled

Summery

	False GA events	DPRK events	Aftershocks (true events)
Correctly labelled	47 (90%)	6 (100%)	9 (90%)
Wrongly labelled	5 (10%)	0 (0%)	1 (10%)
Total	52 (100%)	6 (100%)	10 (100%)

Conclusion

- Over 90% of accuracy is achieved
  - Feasibility has been confirmed

*2 This event is located far from DPRK, and it can be considered that two different events are merged into one. Although an analyst formed an event based on this GA event, this should be labelled as false at this stage.

Synthetic Event Generator

Objective

In order to evaluate the performance of automatic waveform event detection software, including NET-VISA and GA, recording an event for which one knows the source characteristics accurately plays an important role, since we never can obtain the definitive list of natural earthquakes. In this regard, setting up man-made non-nuclear explosions would provide valuable data. However, obviously, this cannot be done frequently. On the other hand, building a synthetic event generator provides artificial data in a more convenient way. More importantly, synthetic event generator enable us to synthesize an event even in the area which we cannot intrude into.
Therefore, in the present study, we build a synthetic event generator using NET-VISA priors.

Methodology

The principal information that a synthetic generator should take in and provide is:
Input: location, depth, magnitude, and time of an event,
Output: detected arrival phases at each station.
In order to have NET-VISA form events, the following features are also necessary:
azimuth (and uncertainty), slowness (and *uncertainty*), arrival time (and uncertainty), amplitude of signal, and *noise level.*
In the list, features with under lines are estimated using NET-VISA priors or theoretical determined. Otherwise, approximated values are used.

Results

Reproducing natural earthquakes:

*3 Locations are computed using synthetic phases.
*4 Synth. Only + Common e.g. 8+22=30 in evid=16284708, is the number of synthetic phases actually generated. 36+22=58 phases are associated in REB.

	REB			sel3			Synthetic *3		Number of phases *4					
evid	lat	lon	mb	lat	lon	mb	lat	lon	Synth. only	REB only	Common	Synth. only	SEL3 only	Common
16284708	-54.21	-146.81	5.02	-54.22	-146.83	4.86	-54.23	-146.81	8	36	22	11	9	11
16874835	-21.08	-179.04	5.27	-21.04	-179.09	5.25	-21.19	-179.15	2	103	36	15	3	22
16716303	-8.31	116.63	4.71	-8.34	116.65	4.79	-8.3	116.62	0	28	23	7	13	17

Imbedding into operational data, and reconstruct with NET-VISA:

	time	lat	lon	depth	mb	num of phases
Input	1514765110	-54.209	-146.814	0	5.021139	32
Formed	1514765109	-54.256	-146.813	0	4.962301	33

One additional detection from the operational data is associated. Otherwise, the event is correctly reconstructed.

Disclaimer: The views expressed on this poster are those of the author and do not necessarily reflect the view of the CTBTO