VERIFICATION SCIENCE

Machine learning

for Comprehensive Nuclear-Test-Ban Treaty monitoring

BY STUART RUSSELL,
SHEILA VAIDYA
AND RONAN LE BRAS

The Comprehensive Nuclear-Test-Ban Treaty (CTBT) is gaining renewed attention in light of growing worldwide interest in mitigating the risks of nuclear weapons proliferation and testing. Since the Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) installed the first suite of sensors of the International Monitoring System (IMS) in the late 1990s, the IMS network has progressed steadily, providing valuable support for event diagnostics. This progress was highlighted at the International Scientific Studies (ISS) conference in Vienna in June 2009, where scientists and experts in the CTBT verification technologies met with policy makers to assess the current status of the CTBT's verification system.

In this article, we introduce a few concepts in machine learning and assess techniques that could provide both incremental and comprehensive value for event detection by increasing the accuracy of the final data product. The techniques could also be applied to refining on-site inspection (OSI) conclusions, and potentially reducing the cost of future network operations.

MACHINE LEARNING TECHNIQUES CAN HELP IMPROVE ACCURACY OF IDC'S FINAL OUTPUT

The IMS includes waveform physical sensor stations (seismic, hydroacoustic, and infrasound) connected by a worldwide communications network to a centralized processing system in the International Data Centre (IDC) in Vienna. The IDC operates continuously and in real time, performing station processing (analysis and reduction of raw seismic sensor data to detect and classify signal arrivals at each station) and network processing (association of signals from different stations that have come from the same event). Fully automated processing of the signals to produce a reliable catalogue of event reports is currently beyond the state-of-the-art, so the IDC analysts must post-process the output from the automated system to generate higher quality event bulletins for further distribution. Errors in automated processing include false detections and missed detections caused by station noise; incorrect classification of arrivals; and incorrect associations. Thus, opportunities exist at all levels of the IDC pipeline to apply techniques from machine learning to improve the accuracy of the final output.

We begin by explaining the basic ideas of machine learning, with special emphasis on data-driven and model-driven methods. We clarify how these methods may be applied to improve the performance of various parts of the IDC processing pipeline. Multiple teams at the ISS conference presented preliminary results that demonstrated improvements in phase classification as well as the rejection of spurious associations via some of these methods. Please see **www.ctbto.org/specials/the-international-scientific-studies-project-iss/** for more information.

The second section of the paper proposes a more radical revision of the IDC data processing approach using a model-driven Bayesian methodology[1]. This approach has several potential advantages, including globally optimal association sets, proper handling of non-detections as evidence, improved low-amplitude signal detection and noise rejection, continually self-calibrating sensor models, and optimal fusion of multiple sensor modalities.

We conclude that incorporating machine learning methods into the IDC framework could indeed improve the detection and localization of low-magnitude events, provide more confidence in the final output, and reduce the load on human analysts. The principal obstacles to rapid instantiation of machine learning methods within an operational context, however, are the availability of raw data for testing during algorithm development and the difficulty of evaluating and benchmarking the impact of local

[1] Bayes' theorem is a formal way of including prior knowledge in assessments of probability. It shows that evidence has a strong confirming effect if it was unlikely before being observed.

improvements on the overall system. We outline a programmatic construct for overcoming these hurdles by proposing to coordinate and drive data-related research and development initiatives through a virtual Data Exploitation Centre (vDEC), under the auspices of the CTBTO, for the evolution and evaluation of next generation data processing methods for CTBT verification.

BASIC CONCEPTS OF MACHINE LEARNING

The field of machine learning covers all computational methods for improving performance based on experience. The range of methods and settings is too vast to be sketched here in completeness, but there is a small set of key questions that must be answered when choosing a learning method:

- Which component of the overall system must be improved?
- How is that component represented – e.g. a weighted linear function, a complicated decision tree, or an impenetrable chunk of machine code?
- What existing data are relevant to that component?
- Do the data include the "right answers" – i.e. correct outputs for the component given the inputs?
- What knowledge is already available to constrain the design of the component?

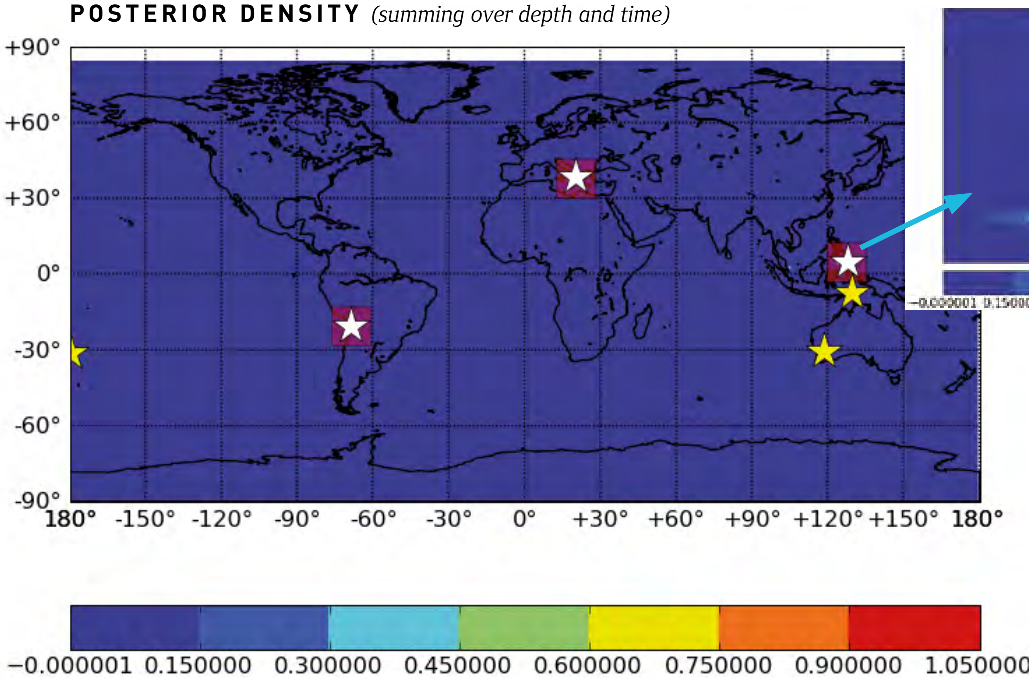
This article examines just two families of methods. The first, supervised data-driven learning, is appropriate for cases where data are plentiful and correct outputs are available, but little is known about the correct design of the component(s). The second, Bayesian model-driven learning, is effective in situations when significant prior knowledge is available; it does not require advance knowledge of the correct outputs for each component.

SUPERVISED DATA-DRIVEN LEARNING

The key idea here is many hundreds of years old: find a hypothesis that maximizes some combination of simplicity and explanation of the data. For example, suppose we want to classify detected seismic signals as **»P waves or S waves«** [see Glossary on p.35]. An unknown function f determines the true classification given the signal. In the supervised setting, we assume that we have a correctly labelled set of data – perhaps obtained from the final **»Reviewed Event Bulletin (REB)«** [see Glossary on p.35] or other authoritative source. The goal of learning is then to find a hypothesis h that is close to f in a precise sense: given a sufficient training set of examples, h should agree with f on the classification of almost all members of a previously unseen test set of unlabelled signals. The framework of machine learning provides guarantees on the possibility of finding such an h and predicts the amount of data necessary to find it.

This seemingly simple task encompasses a large range of activities, roughly characterized by the nature of the inputs, outputs, and the family of hypotheses considered. Supervised machine learning methods are readily applicable to IDC data sets for assisting the final diagnosis. Such methods were illustrated at the ISS conference last June. Several posters showed the value of incorporating off-the-shelf learning and classification methods to improve the accuracy of phase identification in station processing and to detect spurious events formed during network

POSTERIOR DENSITY *(summing over depth and time)*



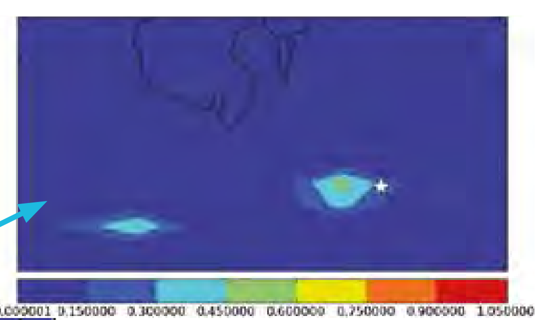


FIGURE 1:
Display showing preliminary results from the prototype NET-VISA system. White stars indicate true events, yellow stars indicate additional spurious events proposed in SEL3, and red squares show events proposed by our research prototype. The inset shows the posterior event distribution near the Sulawesi coast, Indonesia; the posterior is bimodal due to uncertainty in the association between events and detections.

processing. Examples of the benefits from data fusion were plentiful, and design concepts were presented for improving seismic database query processing, borrowing ideas from the Web-search environment. Underscoring the importance of machine learning, the Best Poster award at the conference went to a team that trained neural networks to identify false events in the **»SEL1 bulletin«** [see Glossary].

However, none of these supervised learning methods, as currently conceived, are likely to overcome the fundamental limitations of bottom-up, localized processing of signals and detections. Seismic data analysis on a global scale cannot be decomposed into independent local decisions about detections and associations; the ambiguities inherent in the data are best resolved by a comprehensive analysis of the kind offered by integrated probabilistic inference methods.

Moreover, such methods can easily integrate the best Earth models as well as detailed models of sensor artifacts and failures, and missing data. Such an approach is discussed in the following sections.

BAYESIAN MODEL-DRIVEN LEARNING

When there is substantial prior knowledge available – for example, that of seismic phases and signal propagation – this knowledge can improve prediction accuracy and reduce the amount of data needed for learning. Bayesian methods are well-suited to this context.

In general, Bayesian inference yields a posterior probability distribution over a set of hypotheses given some evidence. In the CTBT setting, a hypothesis might be a collection of seismic events (natural or man-made) occurring over space and time; the evidence is provided by the sensor data. The inference process is based on a model with two components: The **»prior probability distribution«** [see Glossary] over hypotheses; for the CTBT problem, this would include the natural seismicity distribution on Earth. The conditional probability distribution for the evidence given each possible hypothesis; in our case, this part of the model describes how signals propagate through the Earth and how they are detected by sensors, as well as the ways in which noise signals arise. In seismology, this is often called the forward model.

Bayes' rule simply multiplies these two components together to give the posterior probability distribution over the set of hypotheses, given the available evidence. Because there are infinitely many possible hypotheses (each a set of seismic events), the calculations involved are nontrivial and require efficient inversion of the forward model.

As a side effect of the inference process, the Bayesian approach generates information that can be used to continuously adapt the model to better explain the data. This adaptation requires no **»ground truth«** [see Glossary] (unlike supervised learning methods) and hence provides a technical foundation for continuous self-calibration and sensor diagnostics.

VERTICALLY INTEGRATED SEISMIC ANALYSIS

While the current IDC data analysis pipeline is functioning effectively, we believe that its overall serial nature imposes unnecessary limitations on system performance that can be largely overcome by a vertically integrated probabilistic approach. Recent advances in modelling capabilities and in general-purpose inference algorithms such as »Markov Chain Monte Carlo (MCMC)« [see Glossary] suggest that it is in fact possible to address problems as complex as seismo-acoustic event detection via a completely integrated, model-based probabilistic system derived from first principles. A research prototype system (Network Vertically Integrated Seismic Analysis – NET-VISA – for CTBT verification) is currently under development with the goals of testing it within the IDC domain.

Once data samples – currently, just the IDC arrival detections but eventually the full waveforms – are supplied to the system, MCMC probabilistically infers a posterior distribution over possible event locations, times, and magnitudes. In essence, MCMC efficiently samples over hypothetical worlds to obtain estimates that converge to the true posterior given the evidence. The fact that MCMC computes posterior probabilities – the best possible answers given the data – takes the algorithm itself off the table; to get better answers, one must either improve the model or add more sensors.

One important benefit of the vertically integrated approach is that signals need not be analyzed at each station in isolation. Suppose that a hypothetical event has been formed from detections at three other stations, such that the event's location, time, and magnitude imply an arrival at a fourth station in a given time window. If a signal is present – even well below the usual signal-to-noise-ratio threshold – it can be picked out and associated with the event. On the other hand, if no signal is present, the event may be disconfirmed by the (absence of) evidence. The smaller the window, the more pronounced this effect will be. Thus, a strong, and thus far unexploited, interaction exists between the accuracy of the travel time model and the ability to pick out signals from noise at a particular station.

The NET-VISA research prototype has been tested on a small two-hour segment of parametric data from the IDC (i.e. above-threshold P-wave detections, rather than raw waveforms). The segment includes three events that generated three or more arrivals, and the prototype recovers all three perfectly. In comparison, the IDC **»SEL3 bulletin«** [see Glossary] includes three additional events which are not well supported by the evidence (see FIG. 1). On a more comprehensive test with a week's worth of data comprising nearly thousand events, NET-VISA showed significant gains in detection sensitivity compared to SEL3, particularly at lower magnitudes.

VIRTUAL DATA EXPLOITATION CENTRE

Based upon the information above, we believe that the CTBTO could benefit greatly from a strategic thrust focused on improving techniques for processing IMS and on-site inspection (OSI) data sets, taking into consideration the state-of-the-art in machine learning, the advances in data structures and query techniques, and the shaping of sensor data for more accurate exploitation and inference. The long-term goal of such an effort should be to assist the CTBTO analyst in making more robust and expedient decisions, aided by a historical perspective, in the face of rapidly growing multi-sensory information and the importance of more accurate and timely event characterization. To facilitate such an endeavour, a valuable next step will be the creation of a virtual Data Exploitation Centre (vDEC) hosted by the IDC, which will connect international experts (academic, government, and commercial) in different disciplines with the IDC/OSI framework, to assess, develop and implement upgrades to the current data processing infrastructure for event detection and localization. vDEC's charter will be to advance the state-of-the-art in data processing in coordination with the operational arm of the IDC so as to provide a smooth transition from research into the production environment.

THE WAY FORWARD

We have summarized applications of machine learning to CTBT verification, including near-term improvements to components of the current IDC pipeline, as well as a more substantial architectural overhaul based on vertically integrated probabilistic models that connect underlying seismic events to measured signals. Such models could improve seismic phase classification, identify spurious associations through global optimization, characterize station drift/noise, use the absence of detections to disconfirm hypotheses, perform time-localized »sub-threshold« signal detections, combine multiple inputs, and cumulatively, lower the threshold for event detection and localization. Taken a step further, continuous sensor self-calibration could lead to better sensor design and layout and potentially mitigate the cost of future network operations.

To coordinate and guide machine learning and data exploitation methods development in support of Treaty verification, we recommend a focus centre (vDEC) under the CTBTO umbrella, which will leverage multidisciplinary expertise to incubate, test and evolve next generation data solutions for IDC/OSI missions.

GLOSSARY

GROUND TRUTH:
Seismoacoustic sources whose location, depth and origin time, (together with their uncertainties), are known to high precision, either from non-seismic evidence, or using exceptionally good coverage of seismometers close to the event.

MARKOV CHAIN MONTE CARLO (MCMC):
is a technique for generating random samples from a specified probability distribution by simulating a Markov chain. A Markov chain is a mathematical model for a probabilistic system whereby the next state of the system depends only on the current state and not on the previous states.

P (PRIMARY) WAVES AND S (SHEAR OR SECONDARY) WAVES:
P waves are compressional and analogous to a sound wave in air or water. They can pass through any kind of material. S waves move perpendicular to the direction of the waves' propagation and can only exist in the solid Earth.

PRIOR PROBABILITY DISTRIBUTION:
*reflects the probabilities one assigns to a set of hypotheses before seeing any evidence; the *posterior* probability distribution reflects the revisions to the prior in the light of specific evidence.*

REVIEWED EVENT BULLETIN (REB):
A bulletin listing events and signal measurements at each station that detected an event, derived from waveform data that have been reviewed by a human analyst.

STANDARD EVENT LIST (SEL):
A bulletin listing events based on the processing of waveform data. The first Standard Event List, SEL1, includes seismic and hydroacoustic data. Based on SEL1, additional seismic data may be requested from auxiliary seismic stations. Results listed in SEL2 also include the processing of auxiliary seismic and infrasound data. The third list, SEL3, adds processing of data arriving late from all monitoring stations.

BIOGRAPHICAL NOTES

STUART RUSSELL
joined the University of California, Berkeley, USA, in 1986, where he is the Chair of Electrical Engineering and Computer Sciences. Dr. Russell is a winner of the Computers and Thought Award, the principal research award in artificial intelligence, and co-authored the standard textbook in the field of artificial intelligence.

SHEILA VAIDYA
is Deputy Programme Director of Non-proliferation at the Global Security Principal Directorate, Lawrence Livermore National Laboratory, California, USA. During her career, Dr. Vaidya has built programmes in data exploitation, remote sensing, high performance embedded computing, integrated circuit manufacturing, quantum electronics, and semiconductor materials, devices and circuits.

RONAN LE BRAS
joined the CTBTO's International Data Centre (IDC) in 2001 and is now Head of the Software Integration Unit. Dr. Le Bras has contributed key items to the IDC system and managed projects and teams in nuclear monitoring for the past 16 years.