



## **Big Data solution for CTBT monitoring: CEA-IDC joint global cross correlation project**

Dmitry Bobrov (1), Randy Bell (1), Nicolas Brachet (2), Pierre Gaillard (2), Ivan Kitov (3), and Mikhail Rozhkov (1)

(1) CTBTO, Austria (mikhail.rozhkov@ctbto.org), (2) CEA/DAM/DIF, F-91297 Arpajon France , (3) Institute of Geosphere Dynamics, Russian Academy of Sciences

Waveform cross-correlation when applied to historical datasets of seismic records provides dramatic improvements in detection, location, and magnitude estimation of natural and manmade seismic events. With correlation techniques, the amplitude threshold of signal detection can be reduced globally by a factor of 2 to 3 relative to currently standard beamforming and STA/LTA detector. The gain in sensitivity corresponds to a body wave magnitude reduction by 0.3 to 0.4 units and doubles the number of events meeting high quality requirements (e.g. detected by three and more seismic stations of the International Monitoring System (IMS)). This gain is crucial for seismic monitoring under the Comprehensive Nuclear-Test-Ban Treaty.

The International Data Centre (IDC) dataset includes more than 450,000 seismic events, tens of millions of raw detections and continuous seismic data from the primary IMS stations since 2000. This high-quality dataset is a natural candidate for an extensive cross correlation study and the basis of further enhancements in monitoring capabilities. Without this historical dataset recorded by the permanent IMS Seismic Network any improvements would not be feasible.

However, due to the mismatch between the volume of data and the performance of the standard Information Technology infrastructure, it becomes impossible to process all the data within tolerable elapsed time. To tackle this problem known as “BigData”, the CEA/DASE is part of the French project “DataScale”. One objective is to reanalyze 10 years of waveform data from the IMS network with the cross-correlation technique thanks to a dedicated High Performance Computer (HPC) infrastructure operated by the Centre de Calcul Recherche et Technologie (CCRT) at the CEA of Bruyères-le-Châtel.

Within 2 years we are planning to enhance detection and phase association algorithms (also using machine learning and automatic classification) and process about 30 terabytes of data provided by the IDC to update the world seismicity map. From the new events and those in the IDC Reviewed Event Bulletin, we will automatically create various sets of master event templates that will be used for the event location globally by the CTBTO and CEA.