



## Exploiting the skills of waveform data analysts in the quest for improved automatic processing

Robert G Pearce, Spilio Spiliopoulos, and staff Monitoring and Data Analysis Section  
CTBTO PrepCom, International Data Centre, Vienna, Austria (e-Mail: [spilio.spiliopoulos@ctbto.org](mailto:spilio.spiliopoulos@ctbto.org))

Over the past ten years analysts at the IDC have reviewed all signals detected automatically on IMS seismic, hydroacoustic and infrasound data which were associated to events in the final automatic event list (SEL3). They have retimed and/or changed the phase identification of many such detections, and almost 30% of events which appear in SEL3 have been discarded by analysts as invalid. Analysts have substantially modified a further 50% of SEL3 events. In addition, automatic processing misses potentially valid detections and events. Analysts must therefore build missed events from those automatically detected signals which have either not been associated to any event, or which have been wrongly associated, or which have been missed. Such events added by analysts constitute on average about 20% of events in the reviewed event list (LEB).

The large number of examples accumulated in the (automatic) SEL3 and in the (analyst-reviewed) LEB (more than 250,000 events and more than a million associated detections) invites data-driven searches for patterns in analyst actions that can drive improvements to automatic processing. However, such an approach would benefit from a formal understanding of what the analyst does correctly that the automatic processing does not.

When an analyst looks at a seismic signal, a wealth of experience, including that of previous examples, is instantly recalled and implicitly utilised. Some formal elements of this analyst experience which are not currently used effectively in automatic processing are the following:

1. Knowledge of signal characteristics specific to various commonly observed source-station paths
2. Knowledge of station combinations expected to detect signals from events in a given location or region
3. Expectation of similarities in signal characteristics observed at a range of stations for a given event
4. Ability to estimate signal onset time, azimuth and slowness more accurately than automatic processing
5. Ability to identify signal detections missed by automatic processing, and associate them to valid events
6. Rapid recognition of events which have an unrealistic combination of seismic, hydroacoustic or infrasound signals
7. Ability to identify and associate later-arriving seismic phases from large events
8. Ability to identify data artefacts and anomalies caused by station malfunction or noise sources which automatic processing has mistaken for real signals
9. Ability to correctly interpret complex combinations of signals.

For most of the above elements, negative evidence (i.e. the absence of a signal, for example at particular distances or at particular types of station) plays an important role.

Examples of invalid events from SEL3 will be shown together with co-located past examples which provide evidence that the invalidity could have been predicted automatically. Analyst thought processes used in discarding such events will be described. Examples of signal detections, either missed, mis-associated or detected but not associated in the (automatic) SEL3 will also be given, together with the evidence which has led to the analyst decision.