Validating infrasound signal-parameter models using a global ground truth data set

Alexandra Nippress and David Green
AWE Blacknest, Reading, U.K. (alex@blacknest.gov.uk)

Introduction (i)

- At the International Data Centre (IDC) of the Comprehensive Nuclear-Test-Ban Treaty Organisation (CTBTO), a Bayesian monitoring system, NET-VISA (Arora et al., 2013) is being used to associate and locate seismic, hydroacoustic and infrasound events.

- For infrasound sources, this probabilistic model requires prior probability distributions for e.g., celerity and backazimuth. However, the probabilities (Fig.1) currently being used are based on the analyst reviewed Late Event Bulletin, which leads to priors that are likely biased by the models currently used at the IDC.

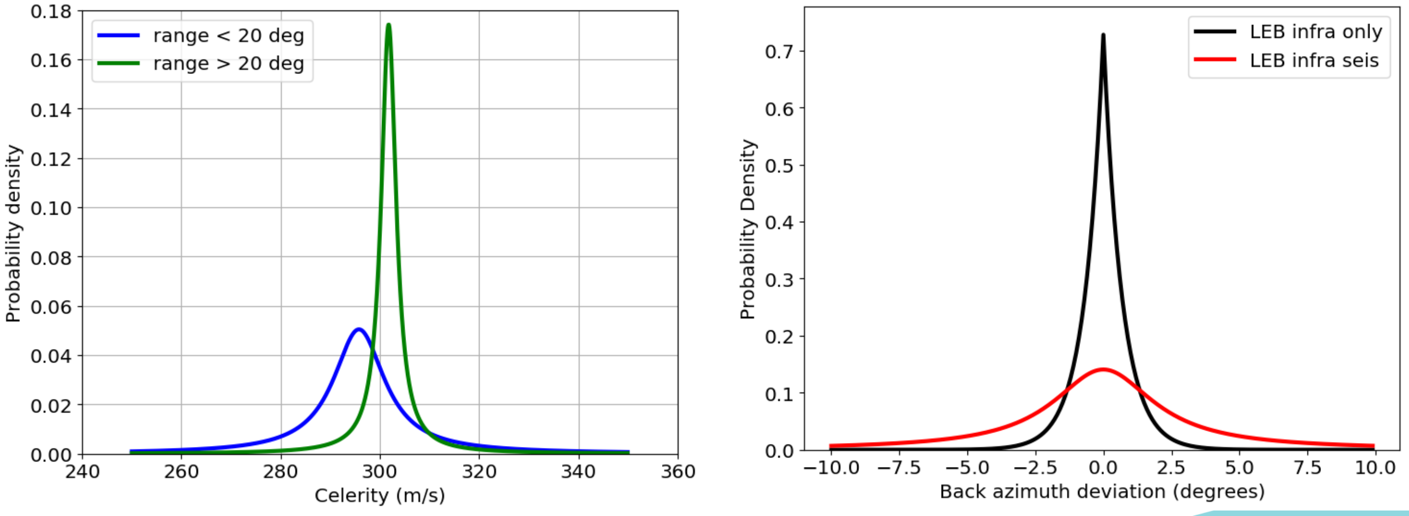


Figure 1. Prior probability distributions for celerity (left) and backazimuth (right) built using the 2012 LEB. (See Fig. 5, slide 7 for the IDC celerity-range model)

Introduction (ii)

- Using a global dataset of ground truth data (e.g., chemical explosions and known mine-blasts) (Fig. 2), we aim to develop infrasound parameter models that can be used to improve signal association algorithms such as NET-VISA by providing unbiased prior probability distributions for celerity, backazimuth and duration.

- Data sources include: the RDSS database[http://www.rdss.info - no longer available online] and events detailed in published literature, e.g., Buncefield explosion (Ceranna et al., 2009).

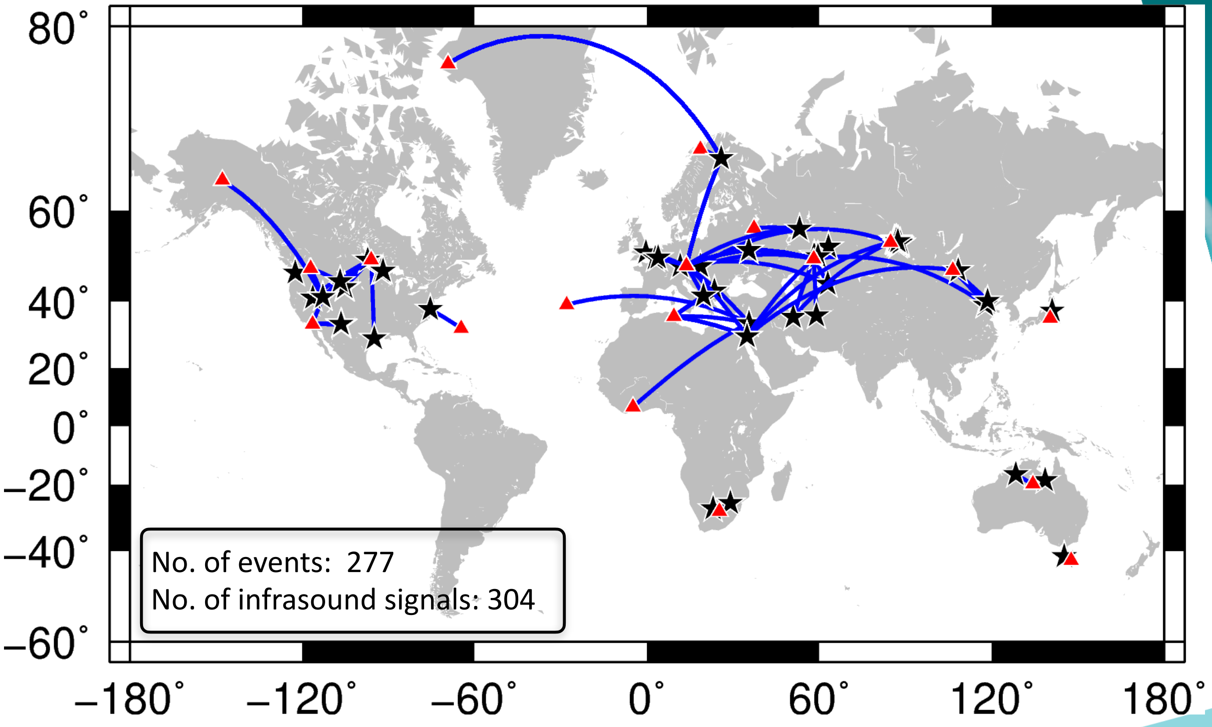


Figure 2. The International Monitoring System (IMS) of the CTBTO stations (triangles) and events (stars) processed. The blue lines indicate the station-event pairs and highlight the dominance of east-west paths.

Data processing using InfGEM (i)

- A software suite incorporating both interactive and automatic processing has been developed to allow consistent processing of all data: InfGEM (Infrasound Global Empirical Models).

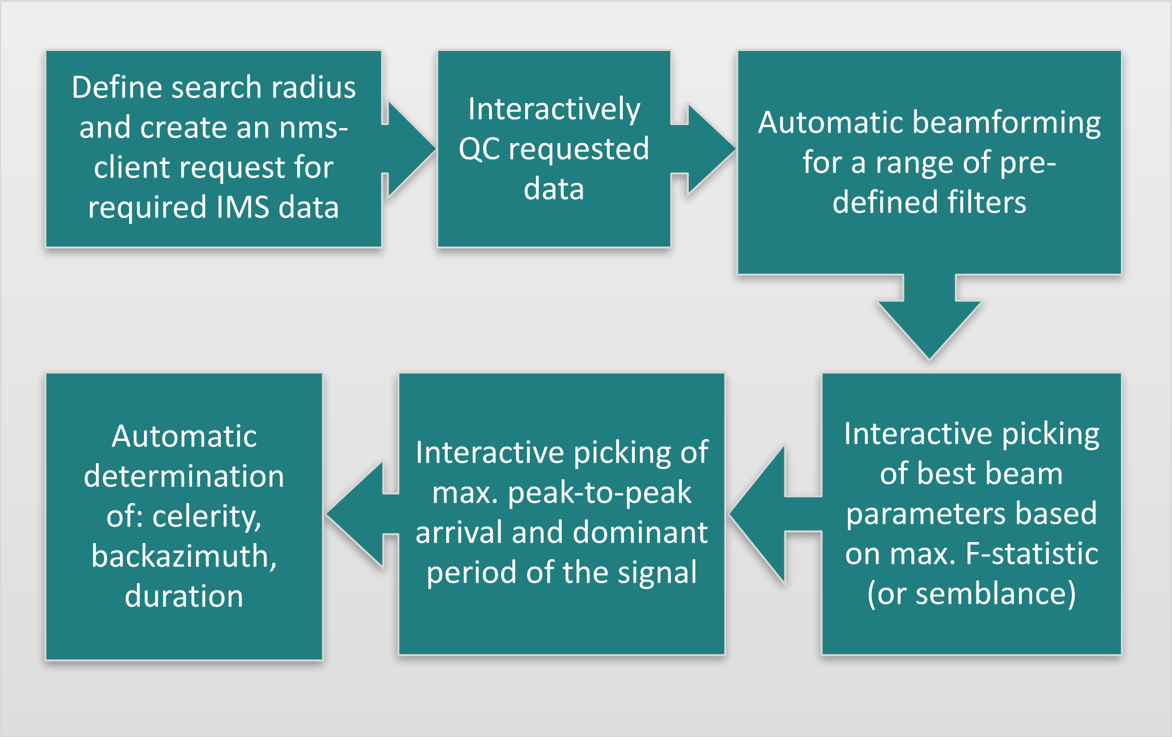


Figure 3. InfGEM processing steps.

Data processing using InfGEM (ii)

- Given an event, a search radius of 4000 km from the event is defined (with additional stations at greater ranges added where detections have been previously identified) and the data are processed as shown in Fig. 3 (previous slide).

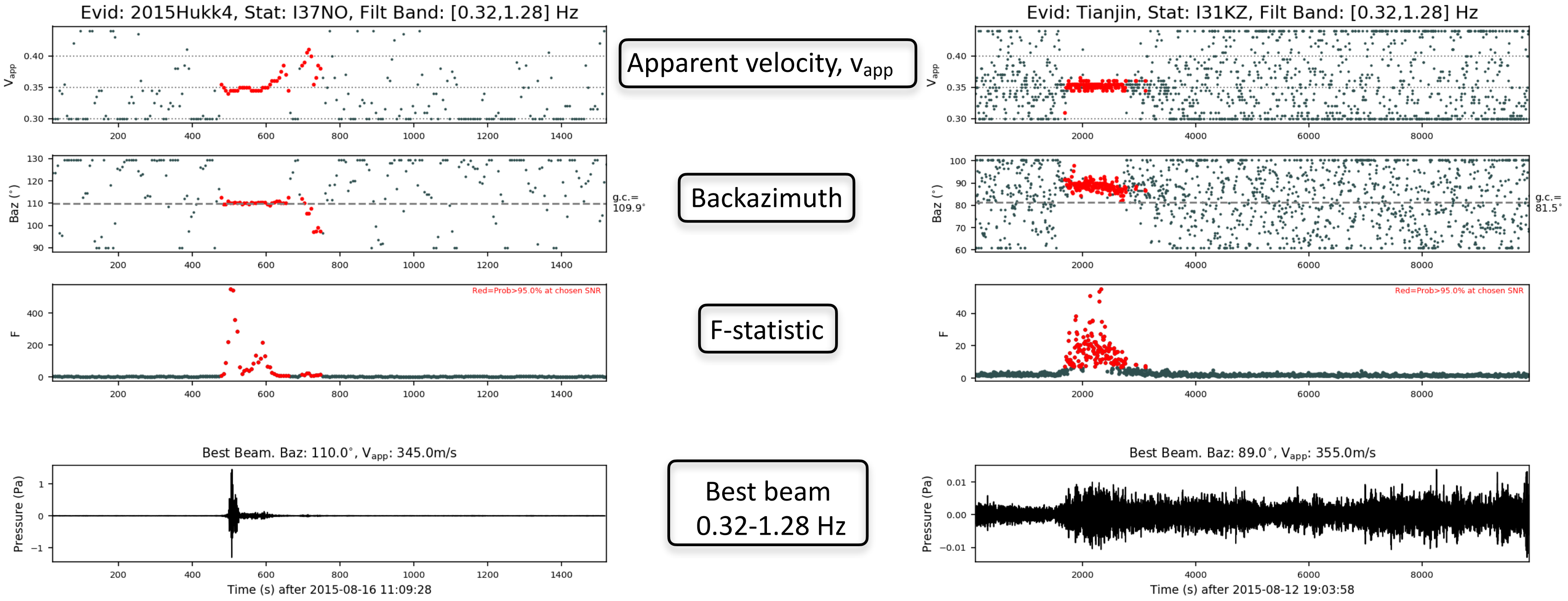
- Data are interactively checked for quality (QC'd) for e.g., data gaps, noisy channels, missing channels, etc.

- For each frequency band processed, the duration is automatically calculated, which allows the start and end celerities of the signal to be determined. For the time of the maximum peak-to-peak amplitude arrival, the celerity (referred to later as max. peak-to-peak celerity) and backazimuth deviation are recorded.
  - All results in this presentation focus on the [0.32, 1.28]Hz passband.

- Additional parameters such as noise amplitudes and SNR are also recorded.

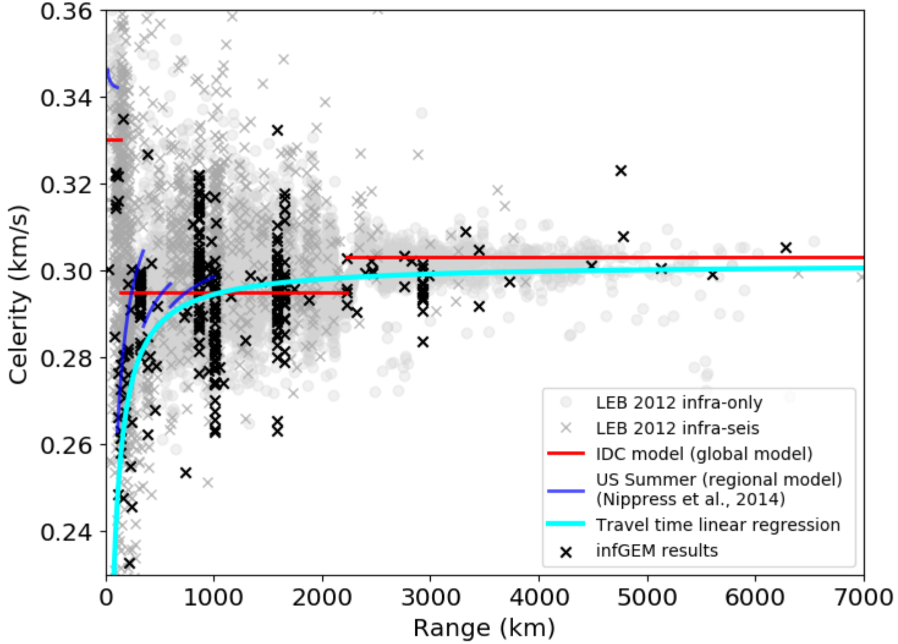
Example signal detections and processing

Figure 4. Left panel. Regional range event (Hukkakero, Finland) recorded at I37NO (Norway), range ~320 km. Stratospheric arrivals followed by thermospheric arrivals (the higher apparent velocity (v_{app}) signals), with a duration of 225 s, backazimuth deviation of 0.1° and max. peak-to-peak celerity 299 m/s. Right panel. Global range event (Tianjin, China) recorded at I31KZ (Kazakhstan), range ~4760 km. Stratospheric arrivals, with duration 1794 s, backazimuth deviation 7.5° and max. peak-to-peak celerity 323 m/s.



Results - Celerity (linear regression travel-time fit)

- Mean max. peak-to-peak celerity = 293 m/s
- Median max. peak-to-peak celerity = 295 m/s



Travel time residuals to a linear regression fit to the infGEM travel time results.

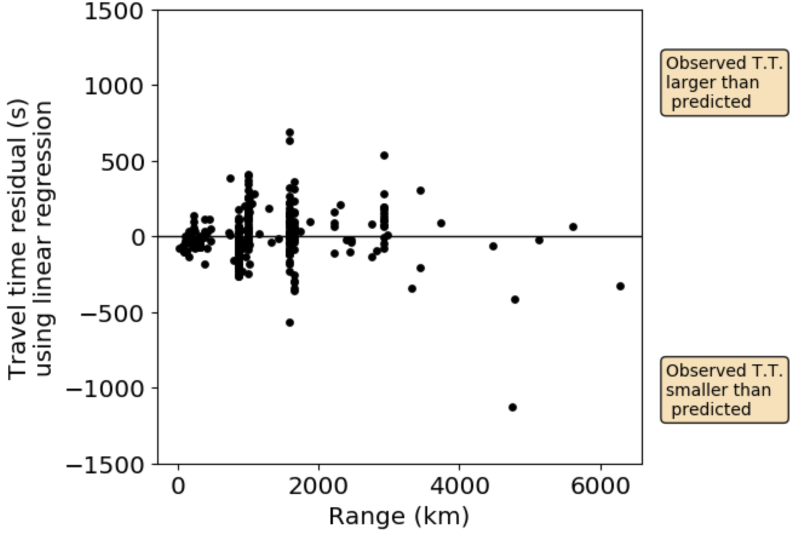
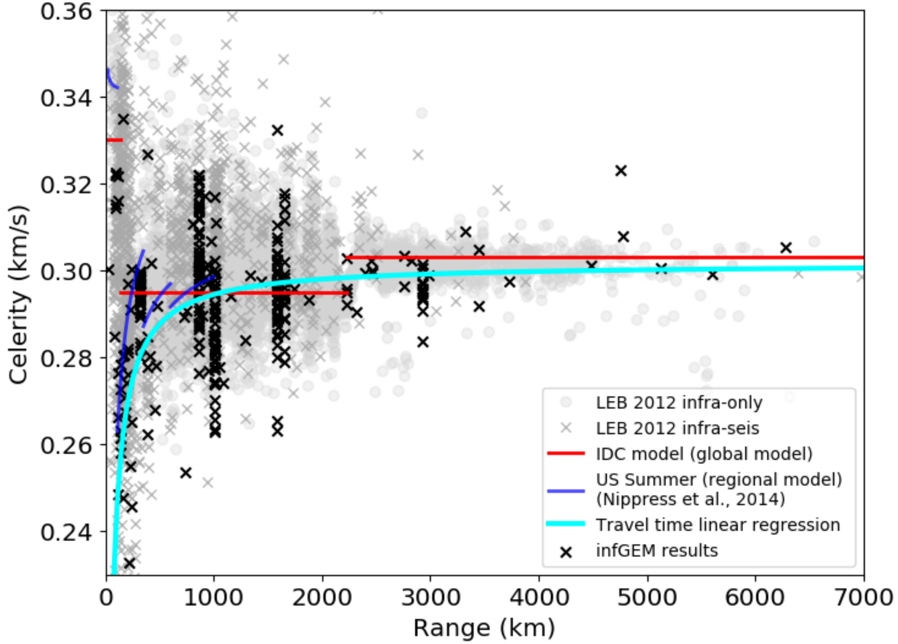


Figure 5. Celerity-range models and data (left panel). Travel time residuals between the infGEM results and the simple linear regression fit to the travel-times for the infGEM results (cyan on the left panel) (right panel).

Results - Celerity (IDC model travel time comparison)

- Mean max. peak-to-peak celerity = 293 m/s
- Median max. peak-to-peak celerity = 295 m/s



Travel time residuals of the infGEM travel time results to the IDC model.

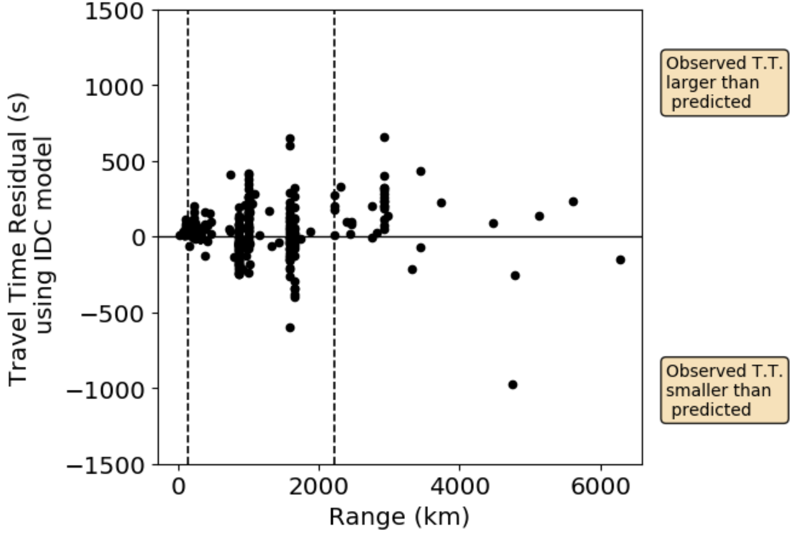


Figure 6. Celerity-range models and data (left panel). Travel time residuals between the infGEM results and the IDC model (red on the left panel) (right panel). The dashed black lines show the range at which the IDC model celerity changes.

Results - Celerity-range

- Across all ranges, the peak probability density is located within the 290-300 m/s celerity bin. This suggests that, for range vs. travel-time regression models, a single range-independent celerity model may be appropriate. This is in contrast to the current IDC global model that has different celerities at different ranges (Fig. 6, previous slide).

- The variation in the width of the celerity distribution as a function of range indicates that a range-dependent travel-time residual model should be considered.

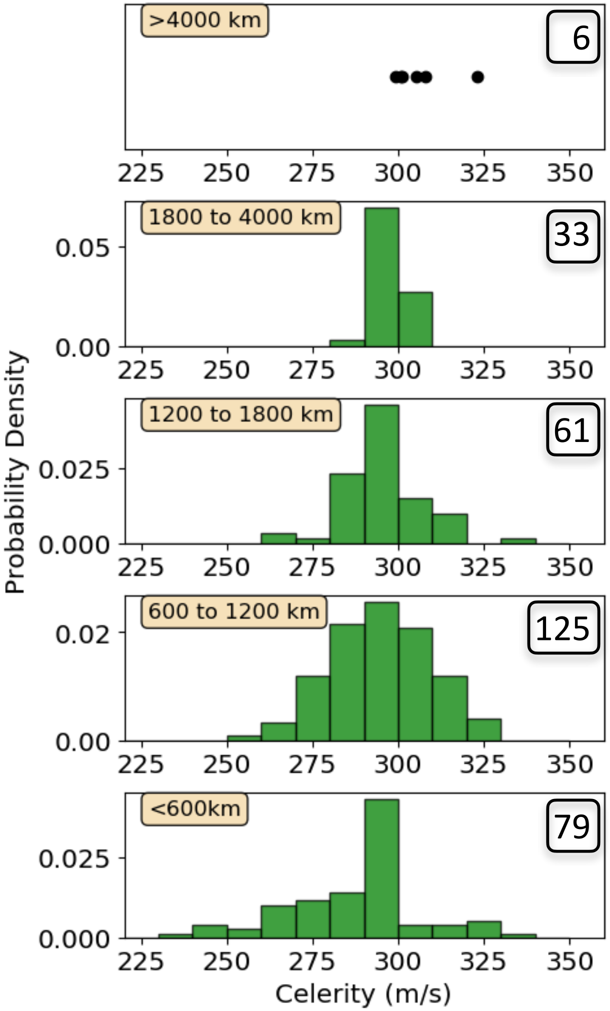


Figure 7. Histograms showing the variation in celerity distribution with range. The number in the top right shows the number of event detections in each range.

Results - Backazimuth

- Mean backazimuth deviation = -0.2°
- Median backazimuth deviation = 0.1°

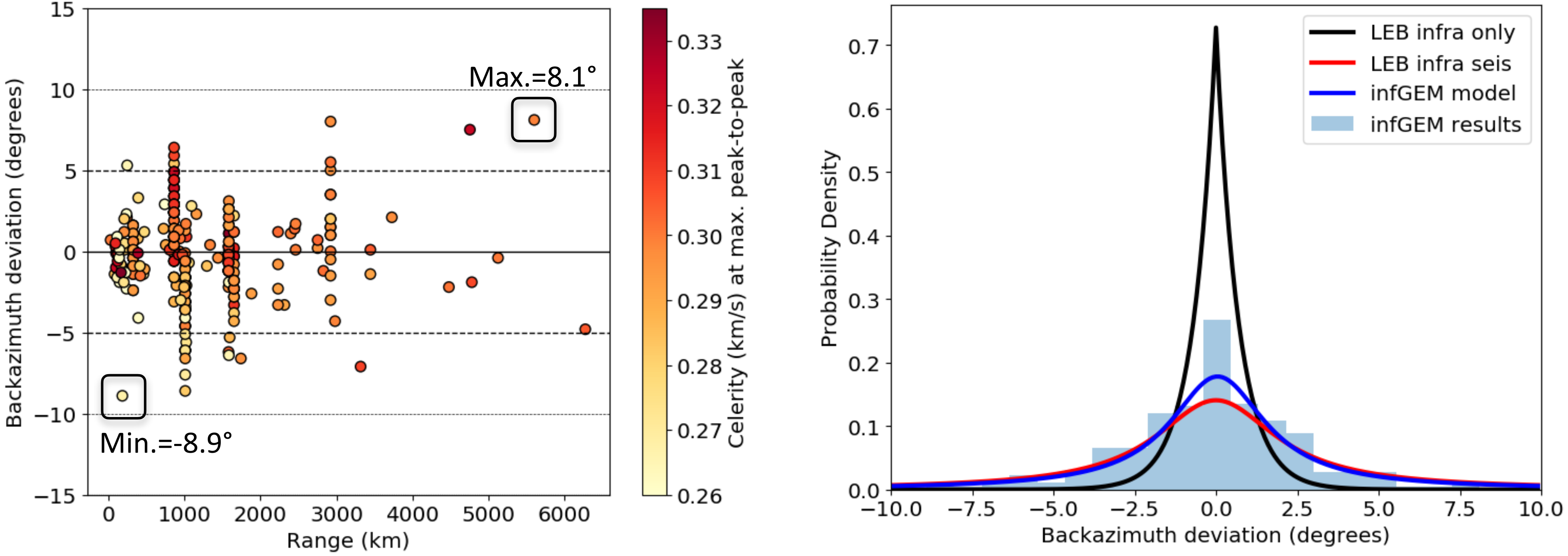


Figure 8. Backazimuth deviations do not show a clear range dependence or relationship with celerity (left panel). A Cauchy fit (blue line) to the infGEM results (histogram) is more similar to the NET-VISA prior for seismo-acoustic events (red line) and highlights the bias in the NET-VISA infrasound-only prior (black line) which is dominated by 2-station LEB events: 3920 vs. 1390 for >2 station events (right).

Results - Duration

- Duration shows a general increase with range, but no clear relationship with SNR, consistent with previous results (Green & Nippress, 2019).
- Future work will consider along-path variability of the effective sound-speed and its relationship to the duration of the signal.

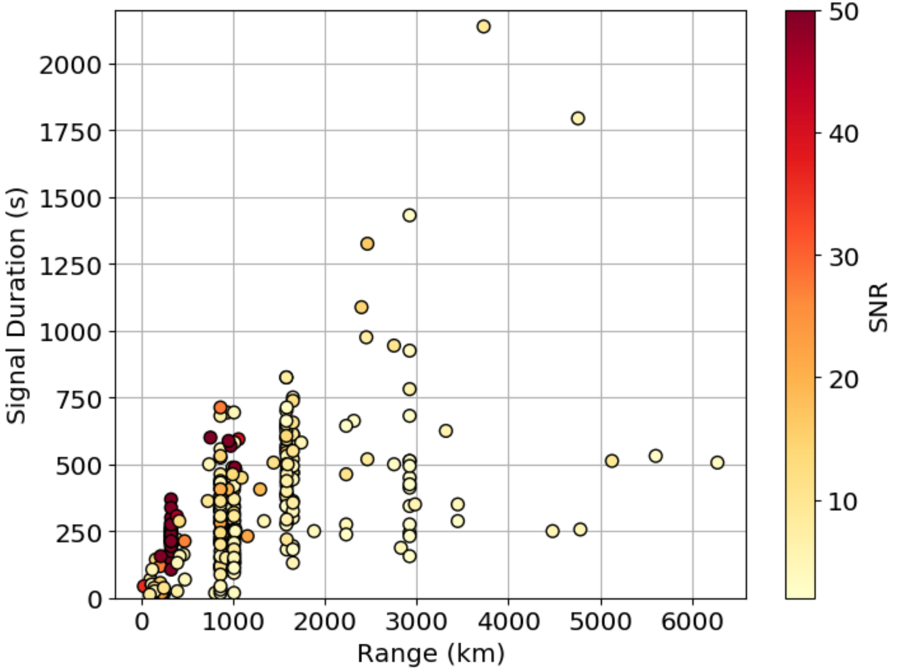


Figure 9. Duration with range results, colour scale from signal-to-noise ratio (SNR).

Summary

- InfGEM results for ground truth data suggest current celerity-range models are not capturing the full data variability at close ranges and are adding unwarranted complexity at long ranges. Travel time residual models need to account for the changes in celerity distribution with range.
- Backazimuth deviation results using ground-truth data are more consistent with the NET-VISA static prior built from seismo-acoustic not infrasound-only events - suggests the seismo-acoustic prior should be used for processing.
- Duration results consistent with previous work (Green & Nippress, 2019).
- Dataset is still lacking event detections at ranges >3000 km, need to account for this when developing travel time residual models.
- The current ground-truth dataset is dominated by events and detections in the northern hemisphere.

References

Arora, N. S. et al., (2013), NET-VISA: Network processing Vertically Integrated Seismic Analysis, *BSSA*, **103** (2A), doi:10.1785/0120120107.

Ceranna, L. et al., (2009), The Buncefield explosion: A benchmark for infrasound analysis across Central Europe, *GJI*, **177** (2), 491-508.

Green, D. N. and Nippress, A., (2019), Infrasound signal duration: the effects of propagation distance and waveguide structure, *GJI,* **216** (3), 1974-1988.

Nippress, A. et al., (2014), Generating regional infrasound celerity-range models using ground-truth information and the implications for event location, *GJI,* **197** (2), 1154-1165.