SAND2011-2511C

PEDAL: A New Event Detection and Signal Association Algorithm

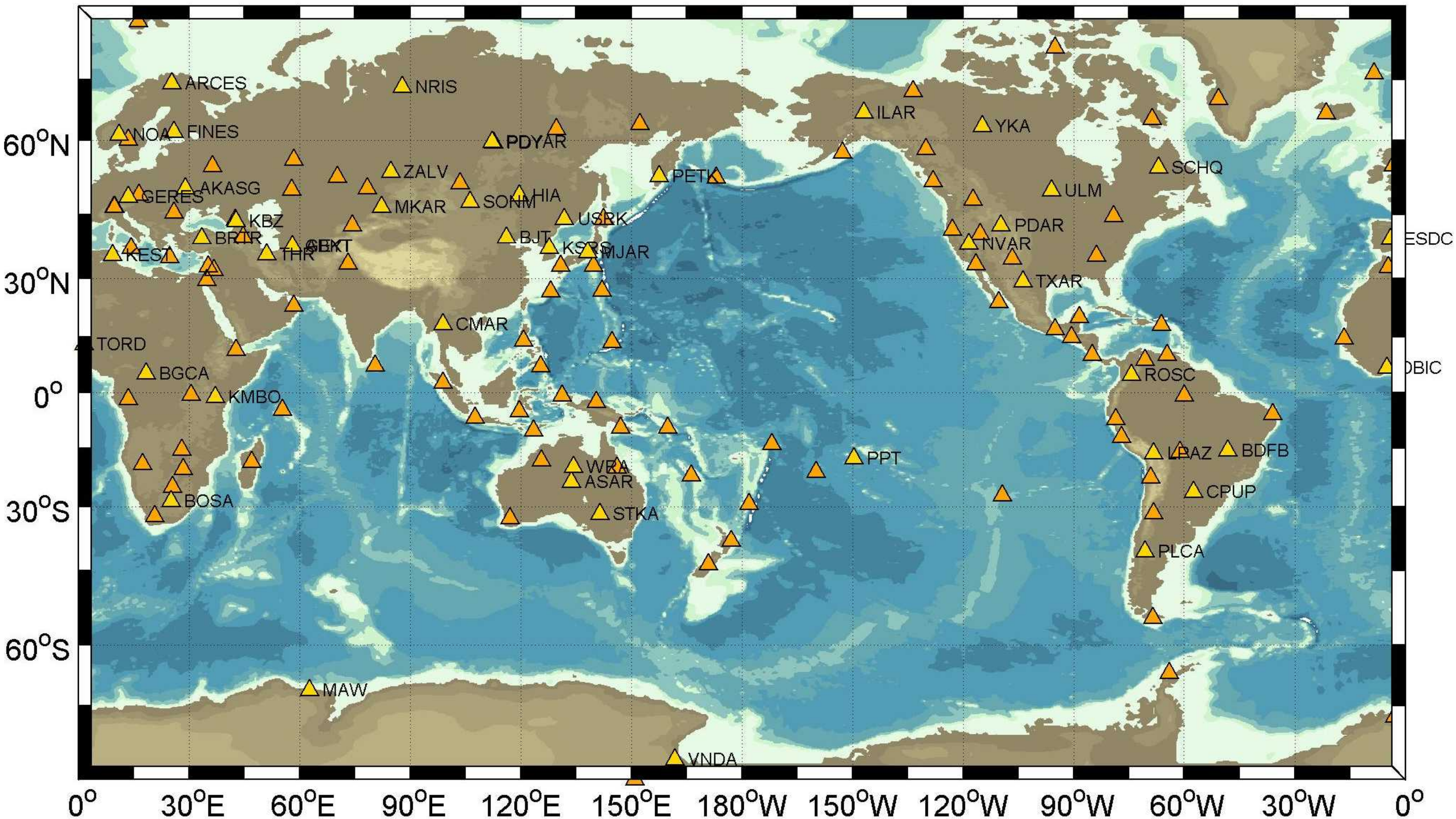
Sandy Ballard, Tim Draelos, Chris Young

Sandia National Laboratories

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

PEDAL

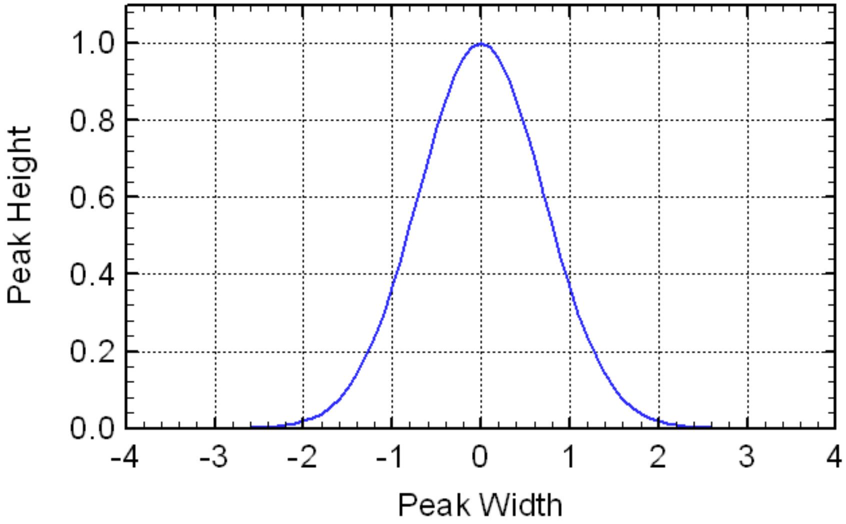
IMS Primary + Auxilliary Network



PEDAL

Consider a single arrival A_i, and a single 4D position ω, in space-time. The conditional 'probability' that A_i was generated by a seismic event, E_ω, which occurred at ω, is given by

$$P(T_i \mid E_\omega) = \exp\left[-\left(\frac{T_i - p_{T,i,\omega}}{\varepsilon_{T,i,\omega}}\right)^2\right]$$

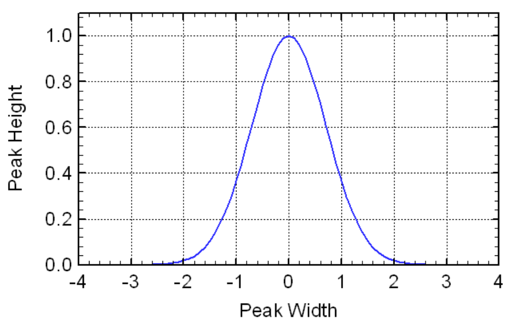


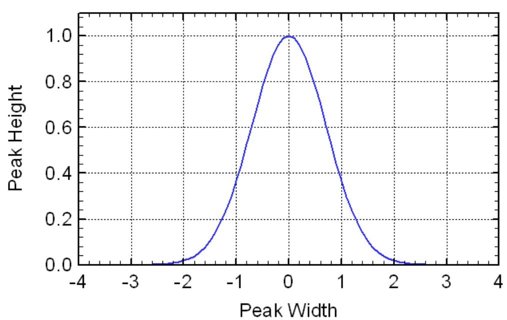
where T_i is the observed arrival time, $p_{T,i,\omega}$ is a prediction of the arrival time and $\varepsilon_{T,i,\omega}$ is a tolerance value, or estimate of the total uncertainty of the observation.

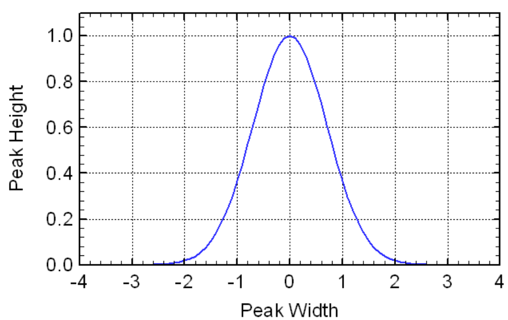
PEDAL

The conditional 'probability' of the observed arrival, A_i, given E_ω, is given by the joint 'probability' of all the components of arrival A_i

$$P(A_i \mid E_\omega) = \exp\left[-\left(\frac{T_i - p_{T,i,\omega}}{\varepsilon_{T,i,\omega}}\right)^2\right] \bullet \exp\left[-\left(\frac{az_i - p_{az,i,\omega}}{\varepsilon_{az,i,\omega}}\right)^2\right] \bullet \exp\left[-\left(\frac{sh_i - p_{sh,i,\omega}}{\varepsilon_{sh,i,\omega}}\right)^2\right]$$







Where $p_{T,i,\omega}, p_{az,i,\omega}$ and $p_{sh,i,\omega}$ are the expected arrival time, azimuth and horizontal slowness for an event at ω, and $\varepsilon_{T,i,\omega}$, $\varepsilon_{az,i,\omega}$, and $\varepsilon_{sh,i,\omega}$, are tolerance values.

PEDAL

$$P(A_i \mid E_\omega) = \exp\left[-\left(\frac{T_i - p_{T,i,\omega}}{\varepsilon_{T,i,\omega}}\right)^2\right] \bullet \exp\left[-\left(\frac{az_i - p_{az,i,\omega}}{\varepsilon_{az,i,\omega}}\right)^2\right] \bullet \exp\left[-\left(\frac{sh_i - p_{sh,i,\omega}}{\varepsilon_{sh,i,\omega}}\right)^2\right]$$

Then we sum the conditional ‘probabilities’ for all arrivals

$$F_\omega = \sum_{i=1}^{NA} P(A_i \mid E_\omega)$$

This is a measure of the ‘fitness’ of E_ω . Because is it is a sum over the arrivals, it will grow if many arrivals have high conditional probability, i.e, ‘agree’ that E_ω occurred. But F_ω will not be diminished by arrivals that have low conditional probability.

PEDAL

Redefine ω to span only 3D space, independent of time. For all events E_ω, $(T_i - T_j)$ observed at two stations S_i and S_j is independent of event origin time. This is because $p_{T,i,\varpi} = T_o + p_{t,i,\omega}$ and $p_{T,j,\varpi} = T_o + p_{t,j,\omega}$ so $p_{T,i,\varpi} - p_{T,j,\varpi} = p_{t,i,\omega} - p_{t,j,\omega}$

$$P(A_i \cap A_j \mid E_\omega) = \exp\left[-\frac{\left(\left(T_i - T_j\right) - \left(p_{t,i,\omega} - p_{t,j,\omega}\right)\right)^2}{\varepsilon_{t,i,\omega}^2 + \varepsilon_{t,j,\omega}^2} - \left(\frac{az_i - p_{az,i,\omega}}{\varepsilon_{az,i,\omega}}\right)^2 - \left(\frac{az_j - p_{az,j,\omega}}{\varepsilon_{az,j,\omega}}\right)^2 - \left(\frac{sh_i - p_{sh,i,\omega}}{\varepsilon_{sh,i,\omega}}\right)^2 - \left(\frac{sh_j - p_{sh,j,\omega}}{\varepsilon_{sh,j,\omega}}\right)^2 \right]$$

where subscripts t refer to travel time and T to arrival time. T_o is origin time.

PEDAL

Now we sum the conditional probabilities of all pairs of arrivals

$$F_{\omega} = \sum_{i=1}^{NA-1} \sum_{j=i+1}^{NA} P(A_i \cap A_j \mid E_{\omega})$$

We search 3D space for points with high F_{ω} followed by a search of the time axis for an origin time, T_0, with high fitness.

$$F_{\omega,T_0} = \sum_{i=1}^{NA} \exp\left[-\left(\frac{(T_i - T_0) - p_{t,i,\omega}}{\varepsilon_{t,i,\omega}} \right)^2 - \left(\frac{az_i - p_{az,i,\omega}}{\varepsilon_{az,i,\omega}} \right)^2 - \left(\frac{sh_i - p_{sh,i,\omega}}{\varepsilon_{sh,i,\omega}} \right)^2 \right]$$

PEDAL

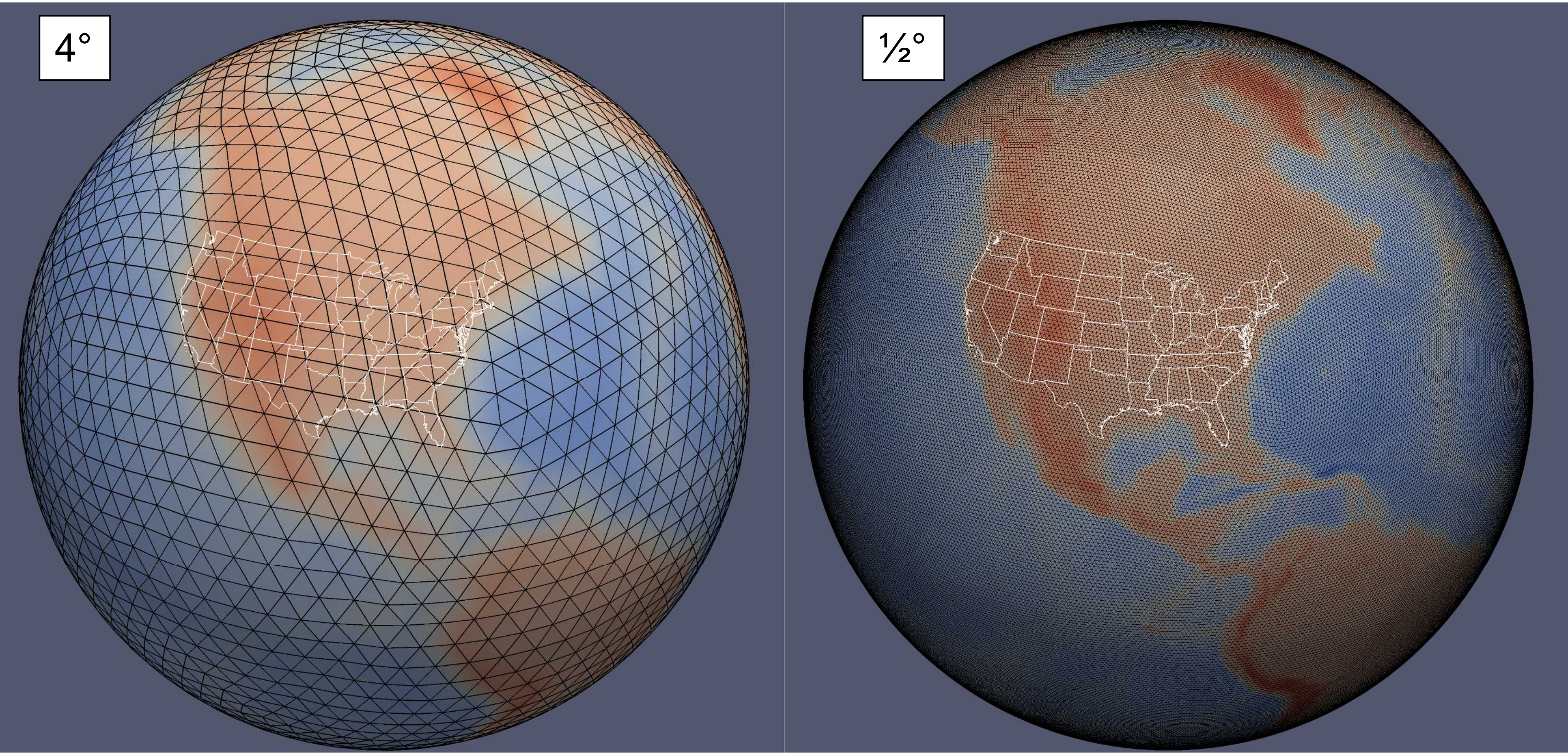
We will then empirically search for a threshold value of F_ω such that we can assert with 95% confidence that the set of arrivals $\boldsymbol{A}$, are consistent with an event having occurred at ω. This will be the conditional probability of the observations given the event, $P(\boldsymbol{A}|E_\omega)$.

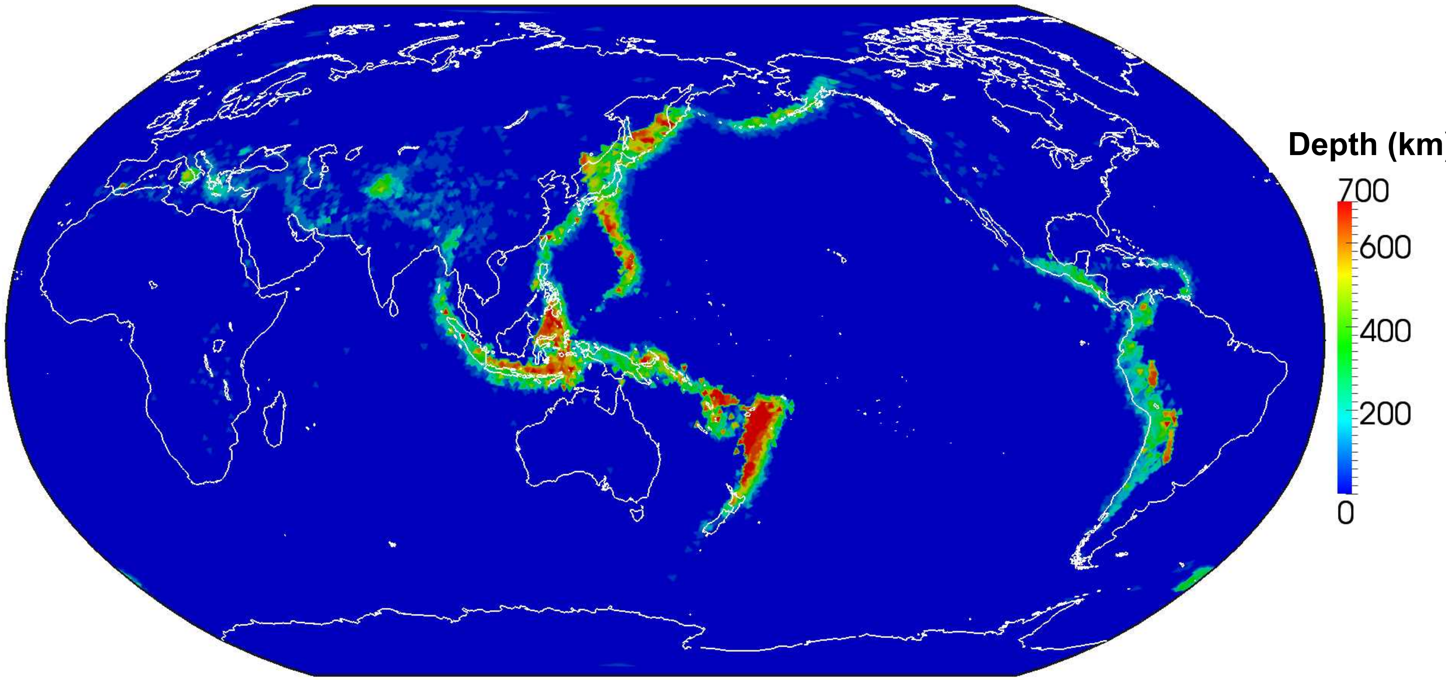
What we really want to know, however, is $P(E_\omega|\boldsymbol{A})$, the probability of an event having occurred at ω given the observed arrivals. To find this we apply Bayes Theorem:

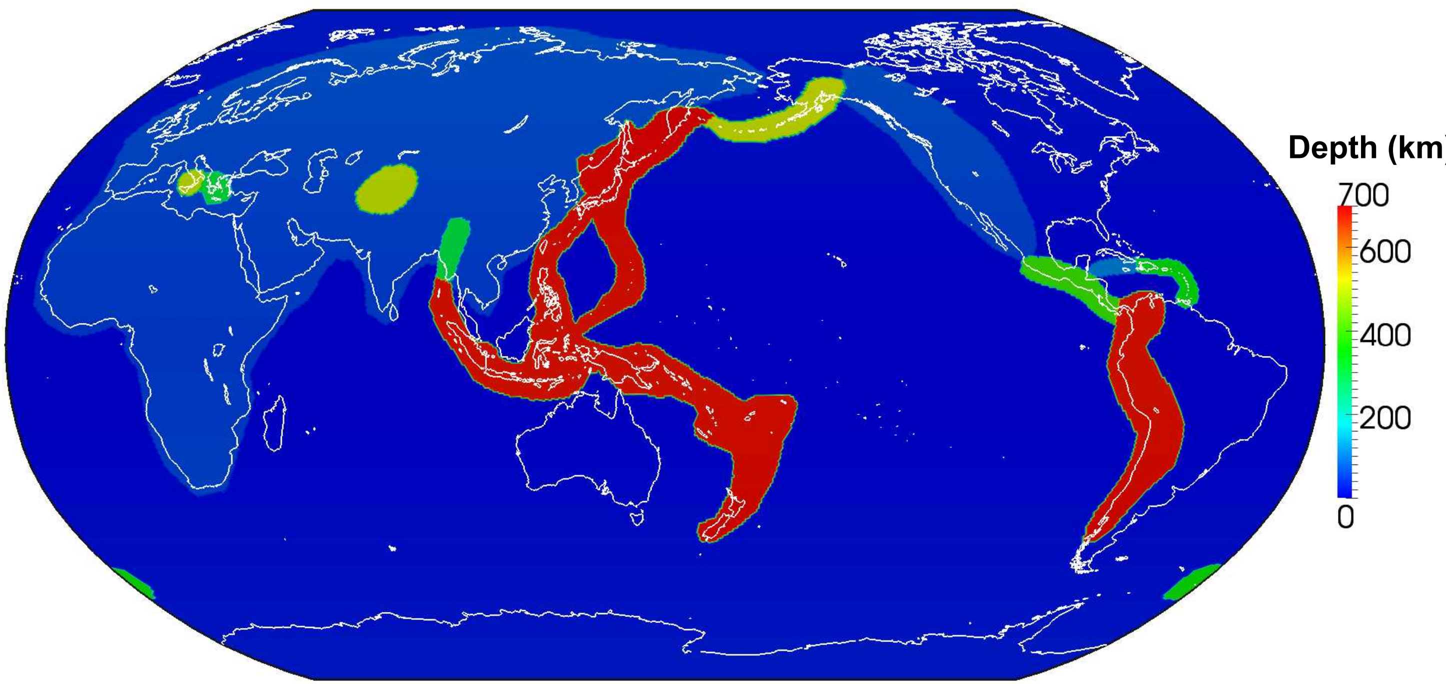
$$P(E_\omega \mid \boldsymbol{A}) = \frac{P(\boldsymbol{A} \mid E_\omega)P(E_\omega)}{P(\boldsymbol{A})}$$

$$= \frac{P(\boldsymbol{A} \mid E_\omega)P(E_\omega)}{P(\boldsymbol{A} \mid E_\omega)P(E_\omega) + P(\boldsymbol{A} \mid !E_\omega)P(!E_\omega)}$$

Grids




Depth (km)
700
600
400
200
0


Depth (km)
700
600
400
200
0

PEDAL

$$F_\omega = \sum_{i=1}^{NA-1} \sum_{j=i+1}^{NA} \exp\left[-\frac{\left(\left(T_i - T_j\right) - \left(p_{t,i,\omega} - p_{t,j,\omega}\right)\right)^2}{\varepsilon_{t,i,\omega}^2 + \varepsilon_{t,j,\omega}^2} - \left(\frac{az_i - p_{az,i,\omega}}{\varepsilon_{az,i,\omega}}\right)^2 - \left(\frac{az_j - p_{az,j,\omega}}{\varepsilon_{az,j,\omega}}\right)^2 - \left(\frac{sh_i - p_{sh,i,\omega}}{\varepsilon_{sh,i,\omega}}\right)^2 - \left(\frac{sh_j - p_{sh,j,\omega}}{\varepsilon_{sh,j,\omega}}\right)^2 \right]$$

Predictions and tolerance values at ~400K nodes for each of ~150 stations requires ~1.3 GB RAM.

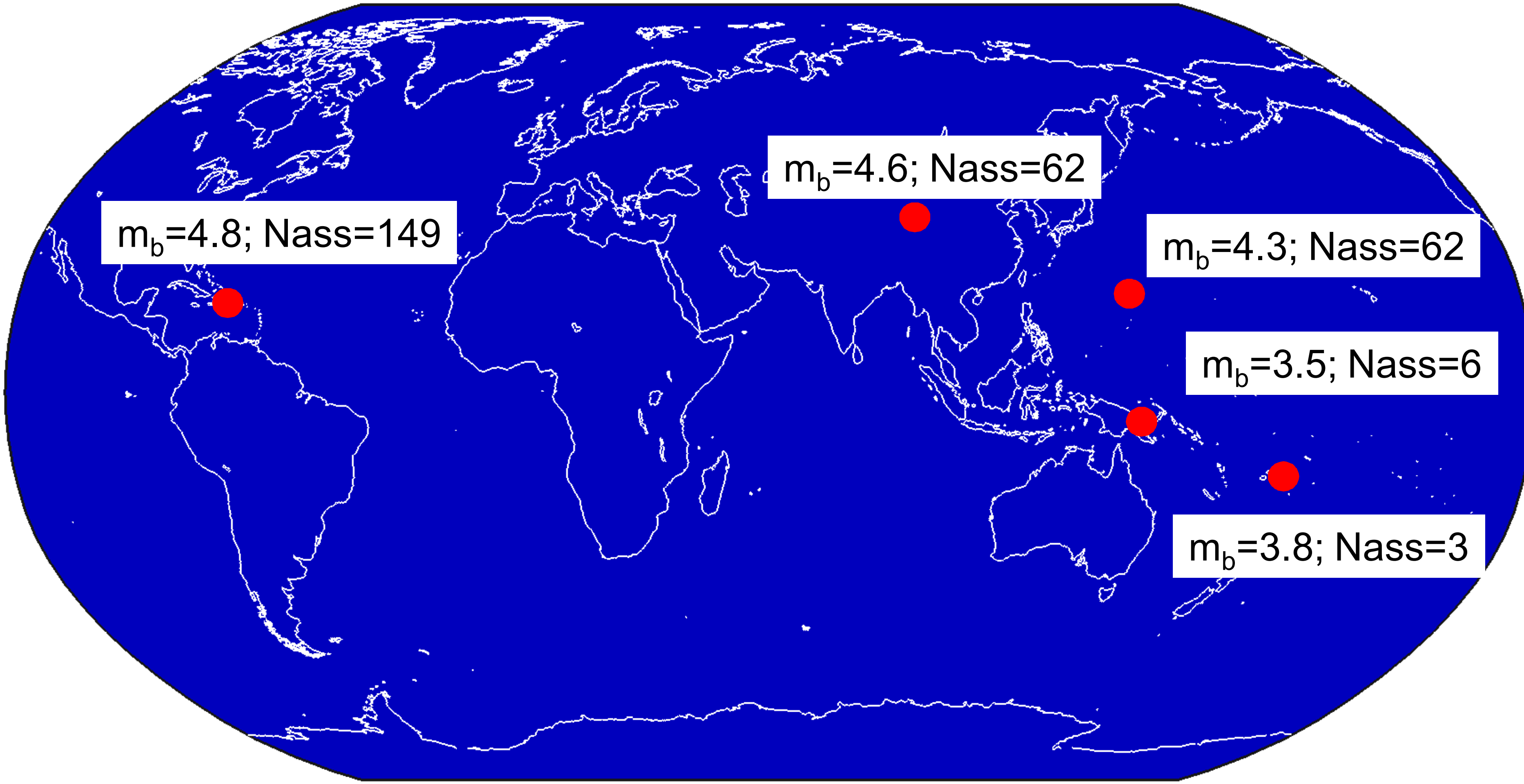
Equation is $O(n^2)$ where n is number of arrivals.

Must evaluate exp() billions of times.

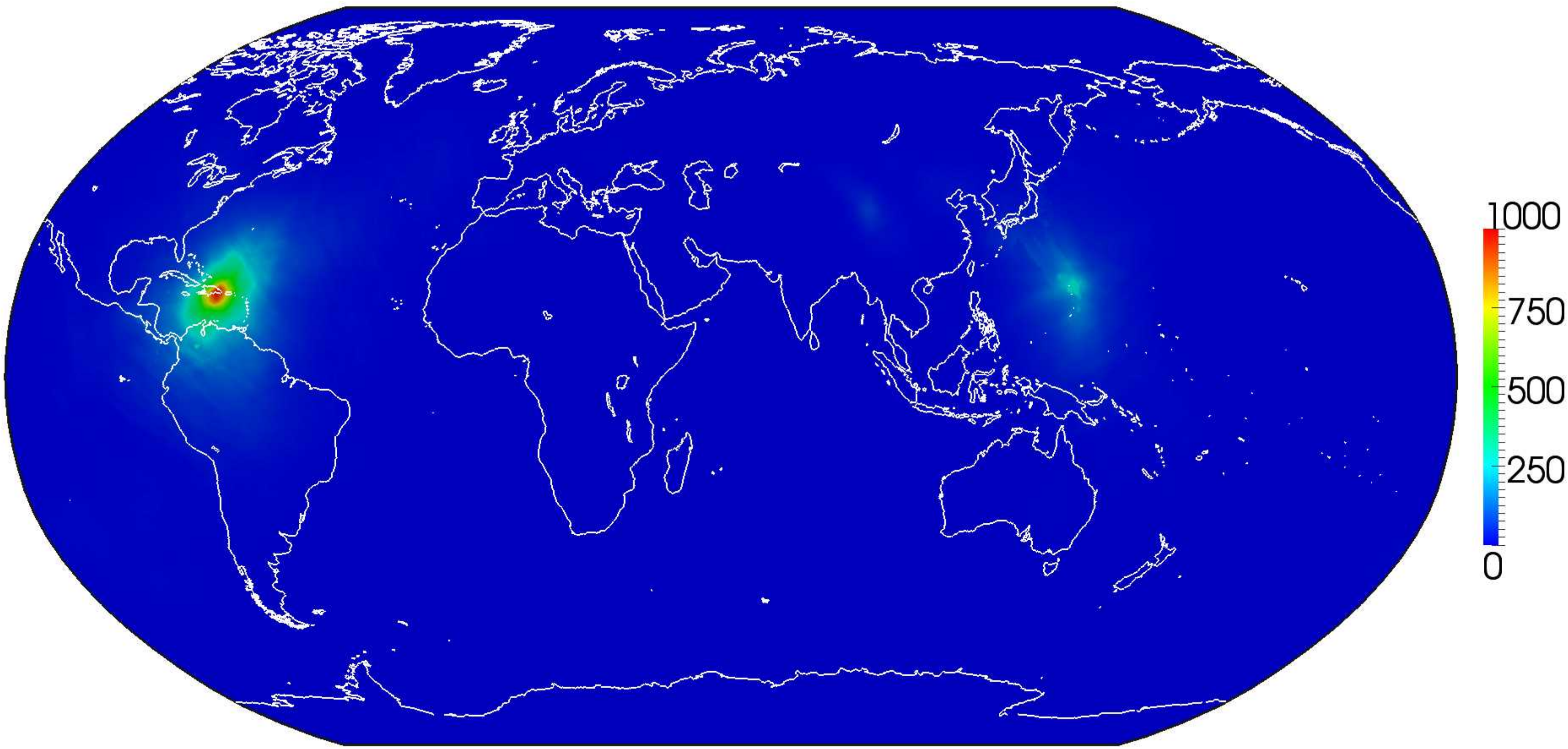
Parallelize on GPUs.

30 minute time window on 12/18/2008.

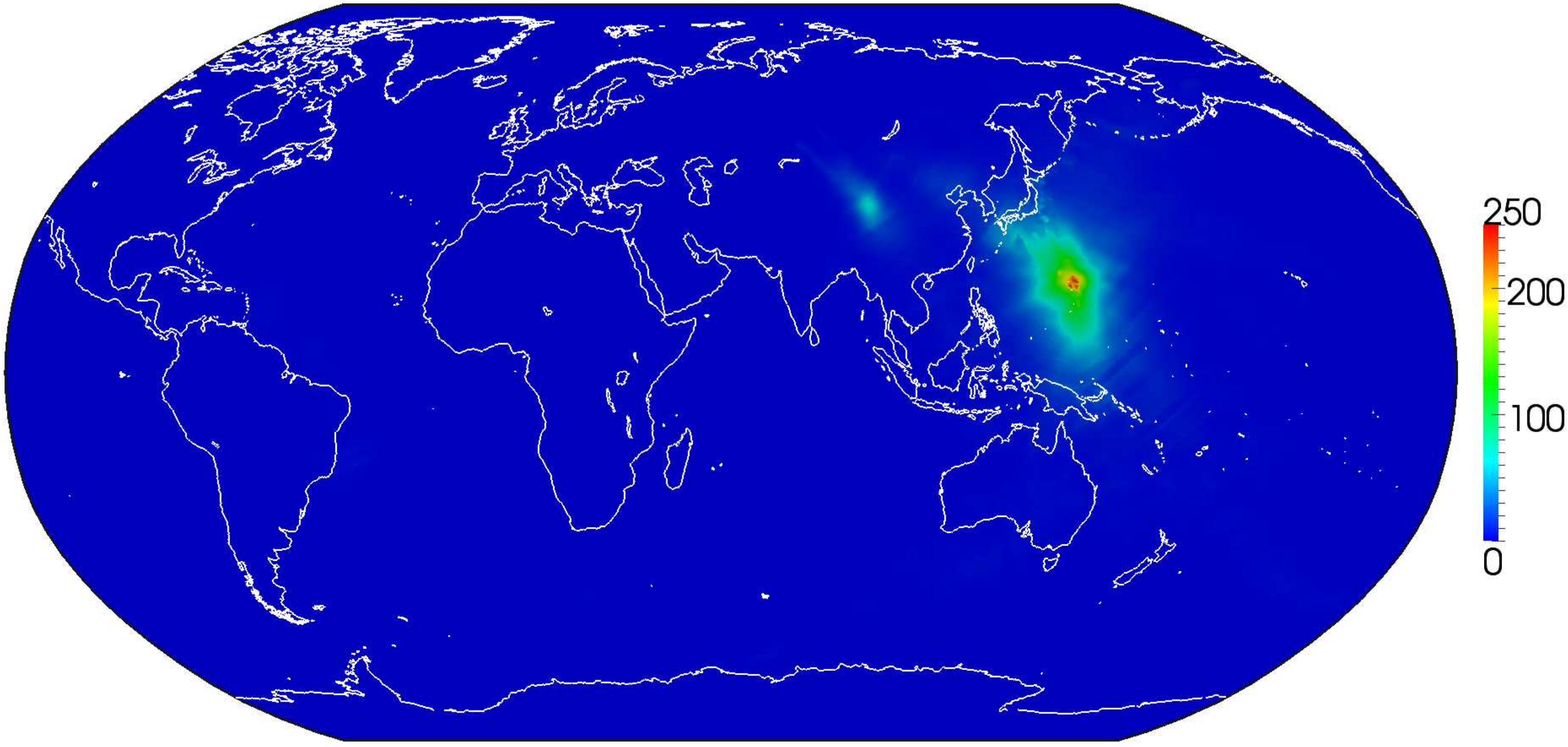
5 Known events; 560 automatically picked arrivals.



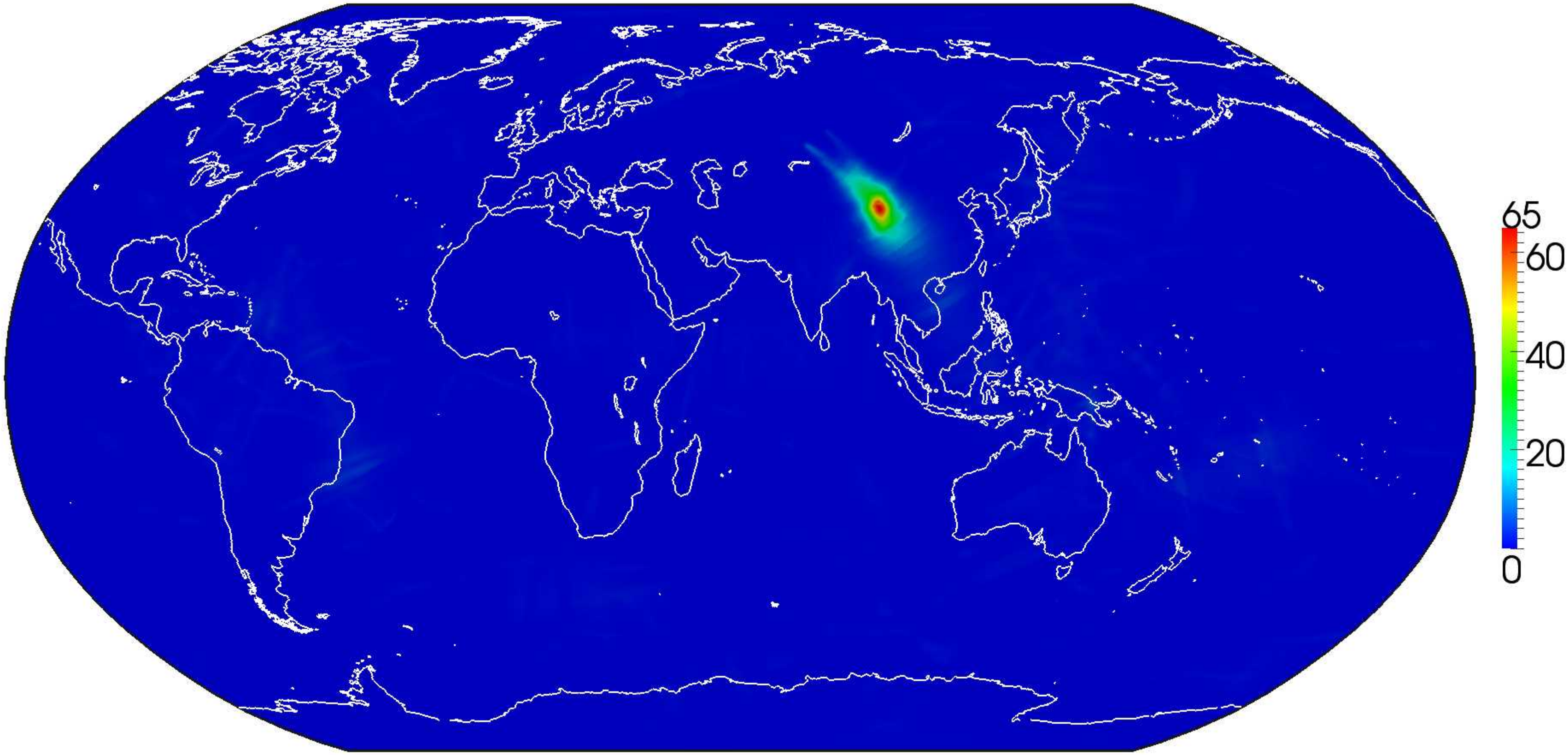
Magnitude 4.8 in Dominican Republic. 220 / 560 arrivals



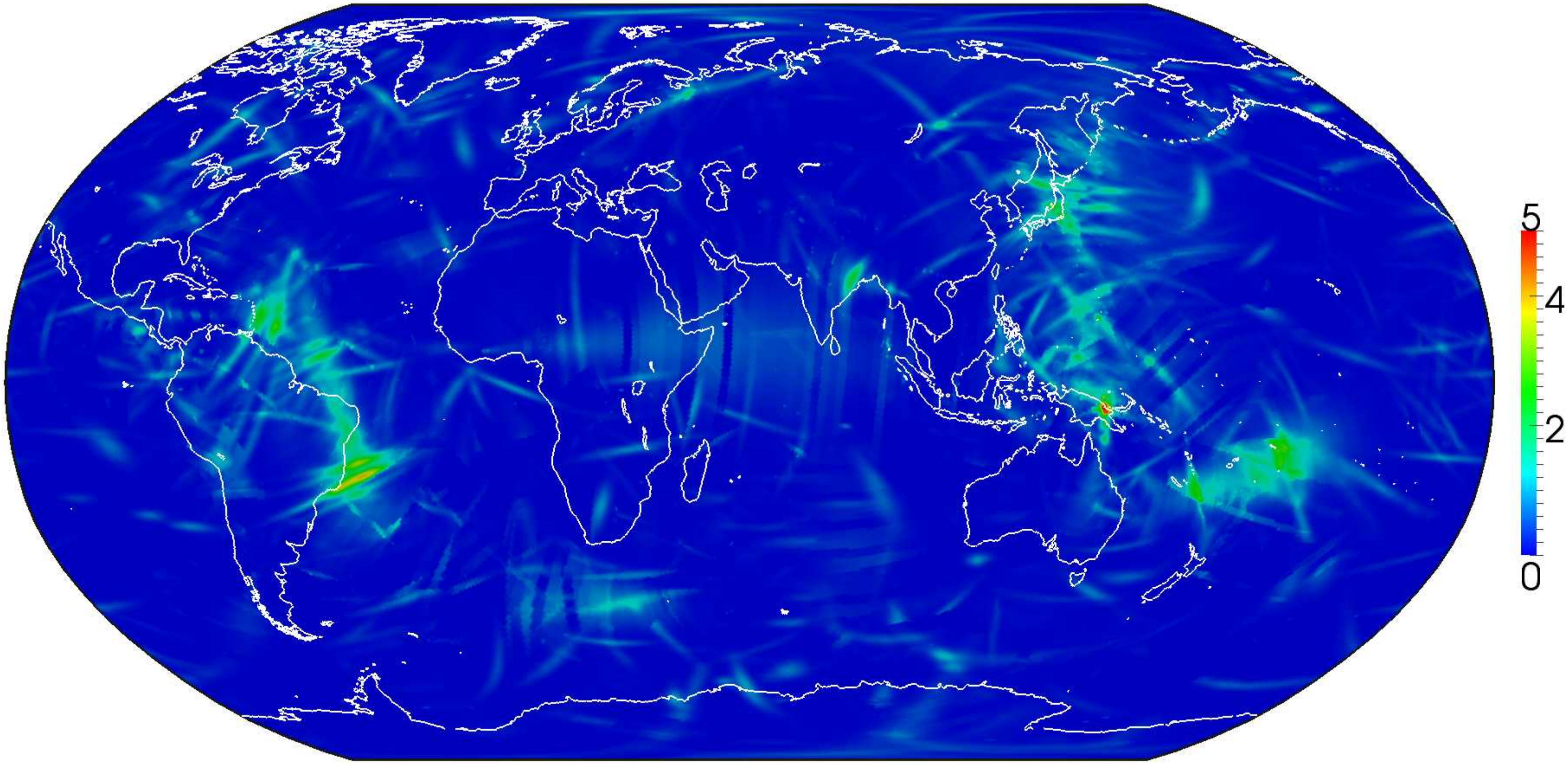
4 minutes later, magnitude 4.3 in Mariana Trench. 77 / 340 arrivals



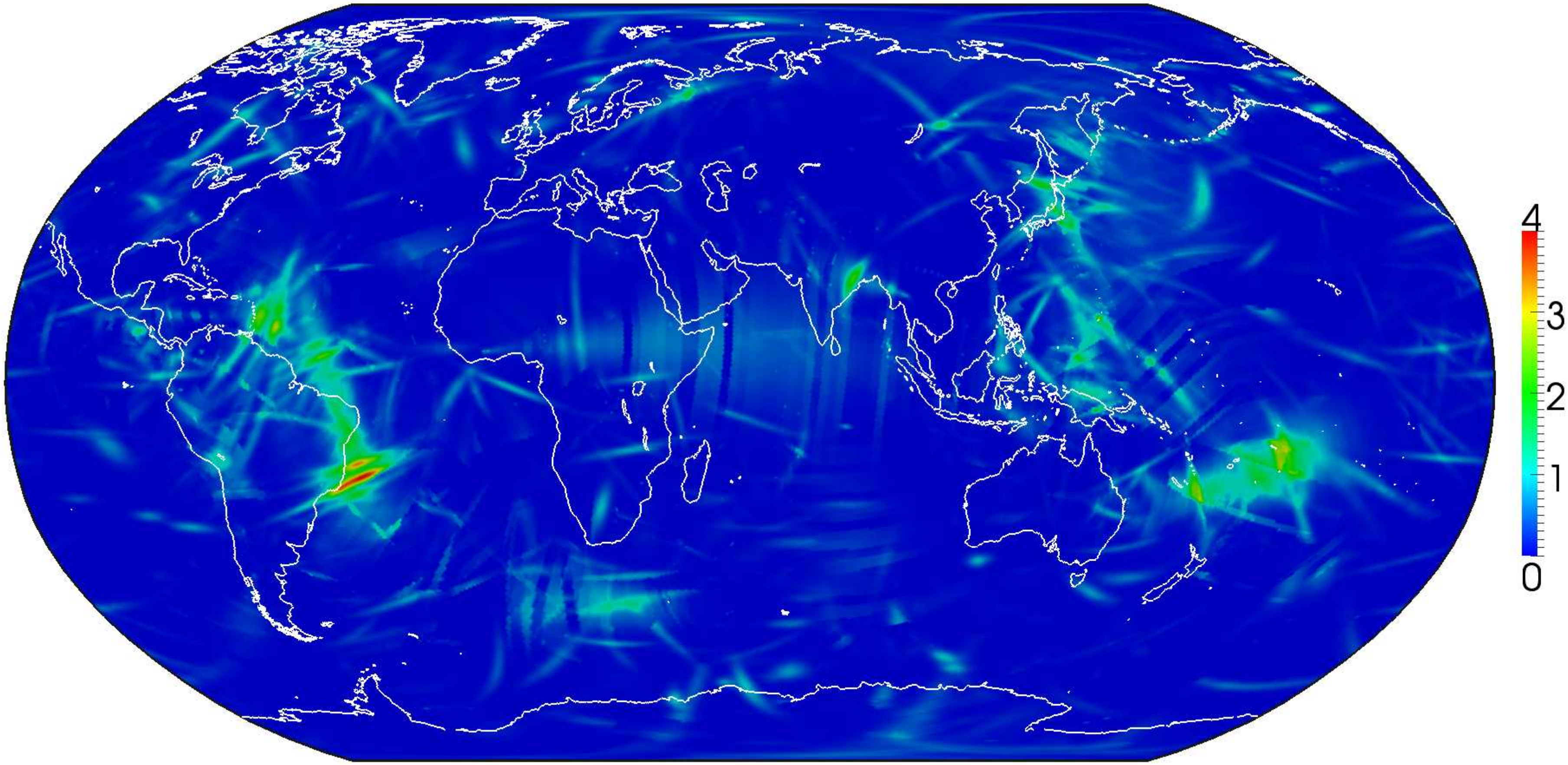
24 minutes, magnitude 4.6 in China. 21 of 263 arrivals



17 minutes, magnitude 3.5 near Papua New Guinea. 12 of 242 arrivals



Missed 3-station event near Fiji. Found spurious event near Brazil.



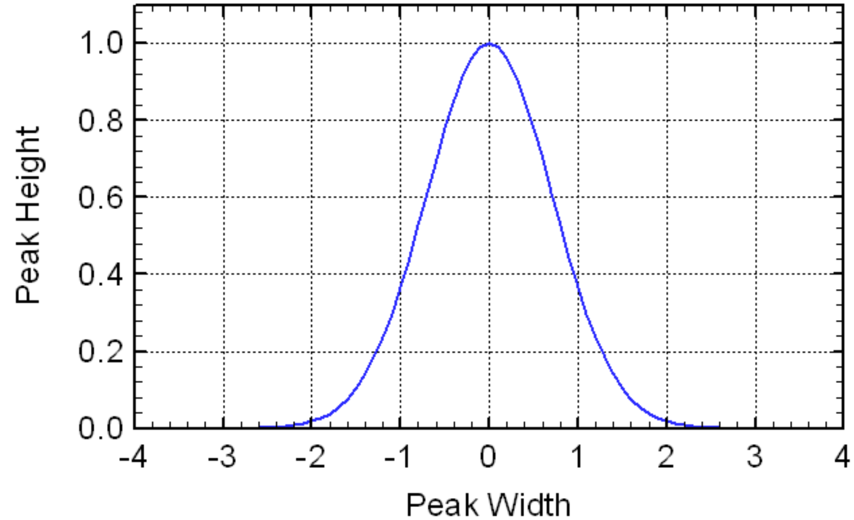
PEDAL

- We are developing a new signal association algorithm
- Exhaustive search of 4D parameter space for locations with maximum ‘fitness’ assuming all arrivals are P
- Association of arrivals to locations of maximum fitness
  - Empirical expected values and tolerances
  - Using list of secondary phases
- Computational demand addressed by use of GPUs
- Preliminary results are promising.
  - Need to test with much more data
- Need to assess:
  - Performance relative to IDC LEB
  - Ability to keep up with real time monitoring networks even during aftershock sequences.

PEDAL

Consider a single arrival A_i, and a single 4D position ω, in space-time. The conditional 'probability' that a seismic event, E_ω, occurred at ω, and generated arrival A_i with arrival time T_i, is given by

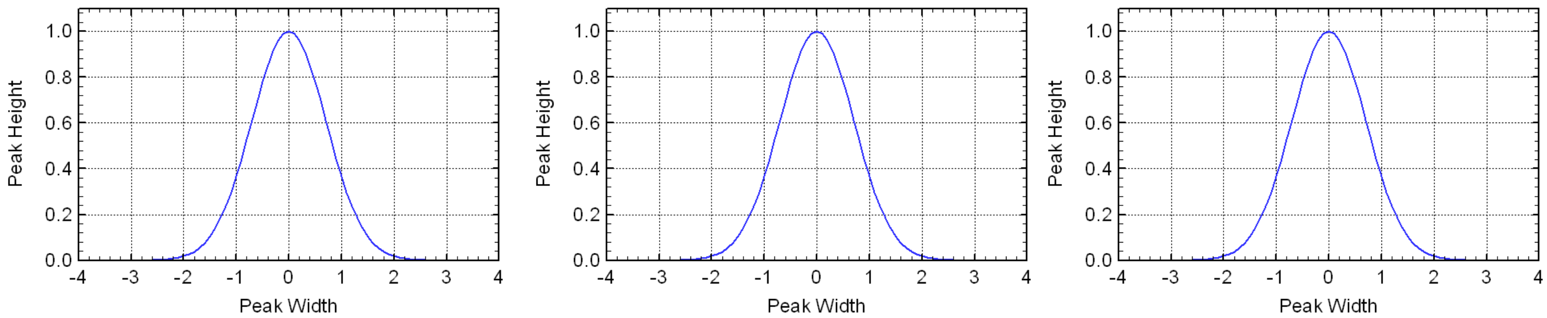
$$P(T_i \mid E_\omega) = \exp\left[-\left(\frac{T_i - o_\omega - p_{T,i,\omega}}{\varepsilon_{T,i,\omega}} \right)^2 \right]$$



where
o_ω is the origin time of event E_ω
$p_{T,i,\omega}$ is a prediction of the travel time and
$\varepsilon_{T,i,\omega}$ is a tolerance value.

PEDAL

$$P(A_i \mid E_\omega) = \exp\left[-\left(\frac{T_i - o_\omega - p_{T,i,\omega}}{\varepsilon_{T,i,\omega}}\right)^2\right] \bullet \exp\left[-\left(\frac{az_i - p_{az,i,\omega}}{\varepsilon_{az,i,\omega}}\right)^2\right] \bullet \exp\left[-\left(\frac{sh_i - p_{sh,i,\omega}}{\varepsilon_{sh,i,\omega}}\right)^2\right]$$

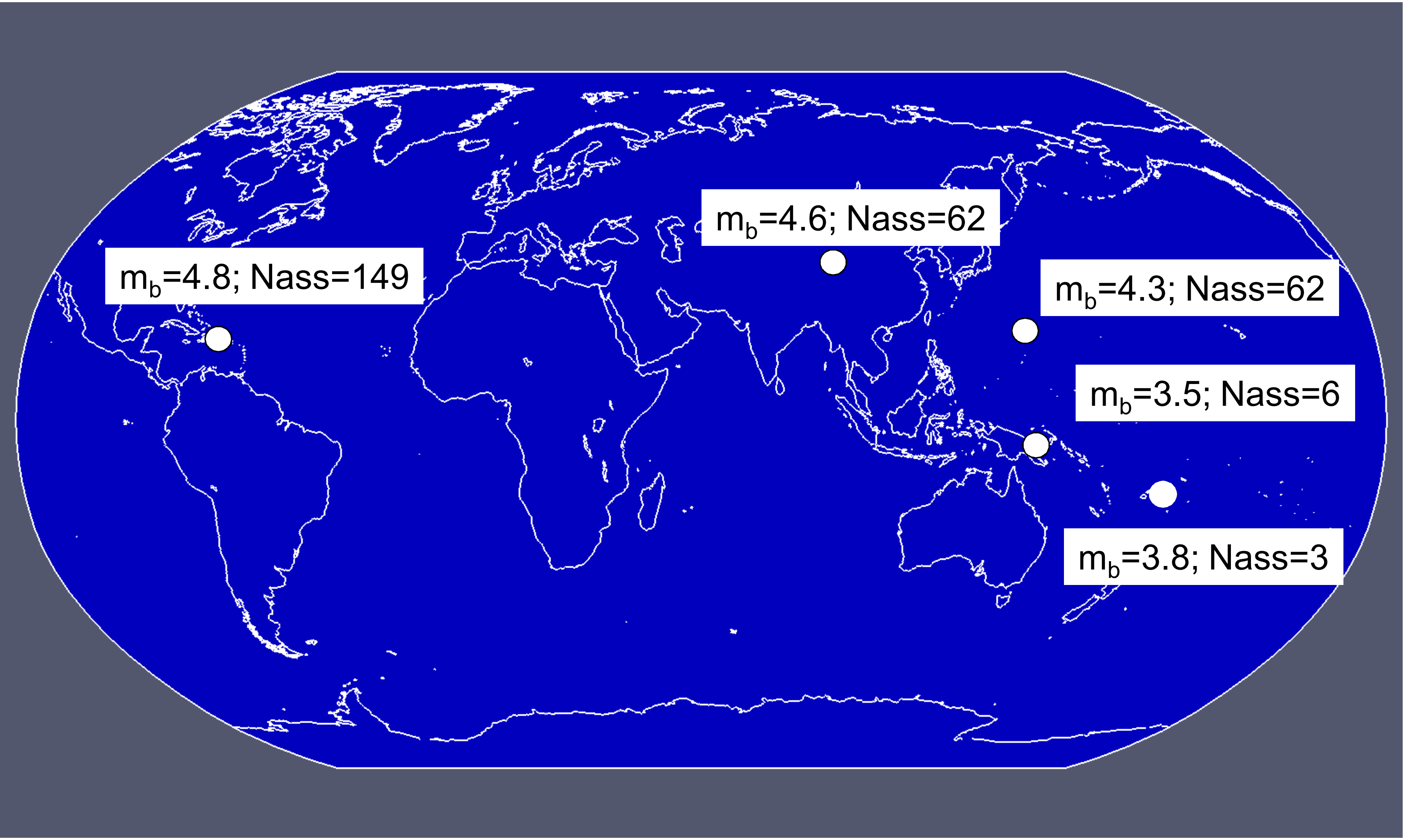


Fitness at point ω = Sum of Gaussians over all arrivals

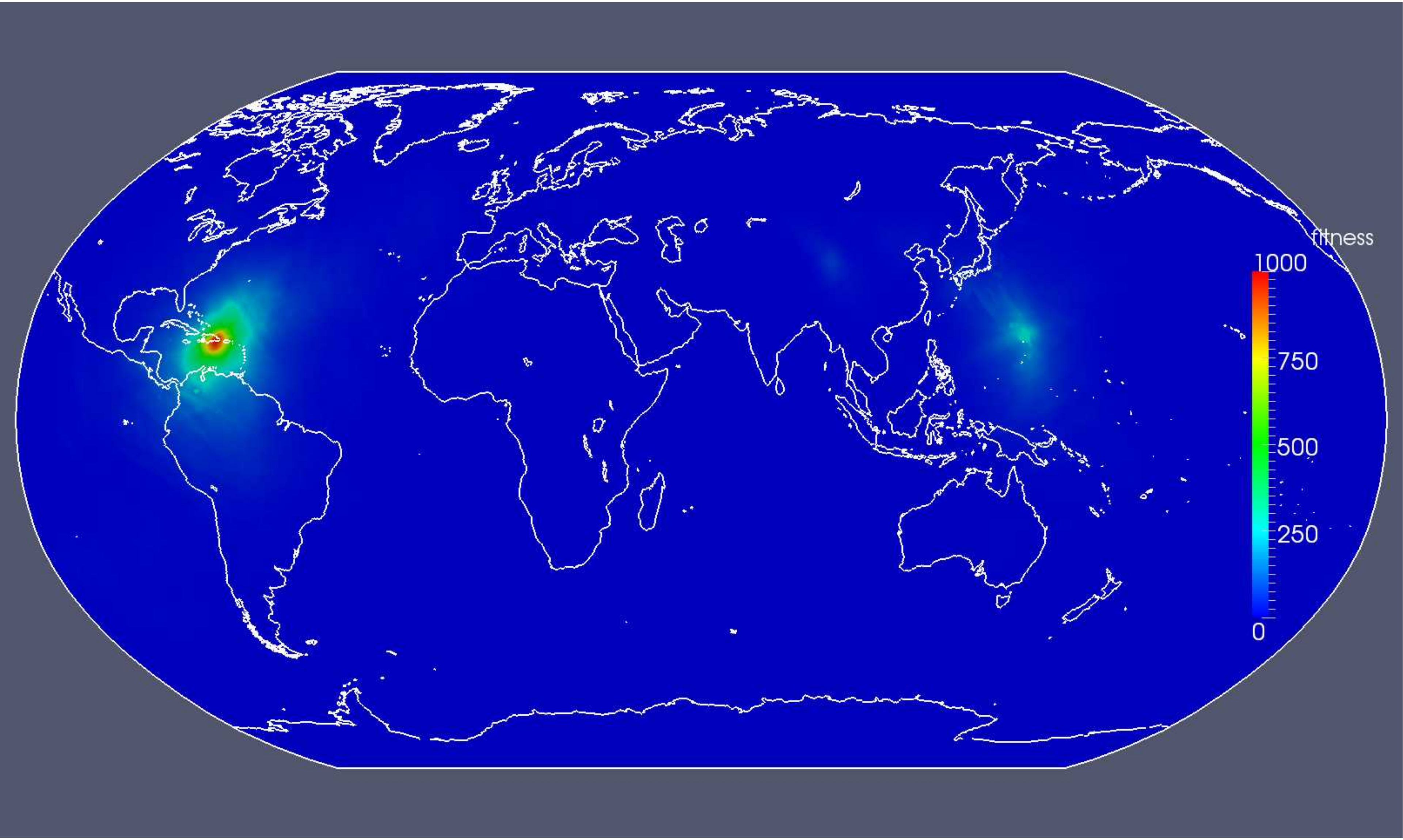
$$F_\omega = \sum_{i=1}^{NA} P(A_i \mid E_\omega)$$

30 minute time window on 12/18/2008.

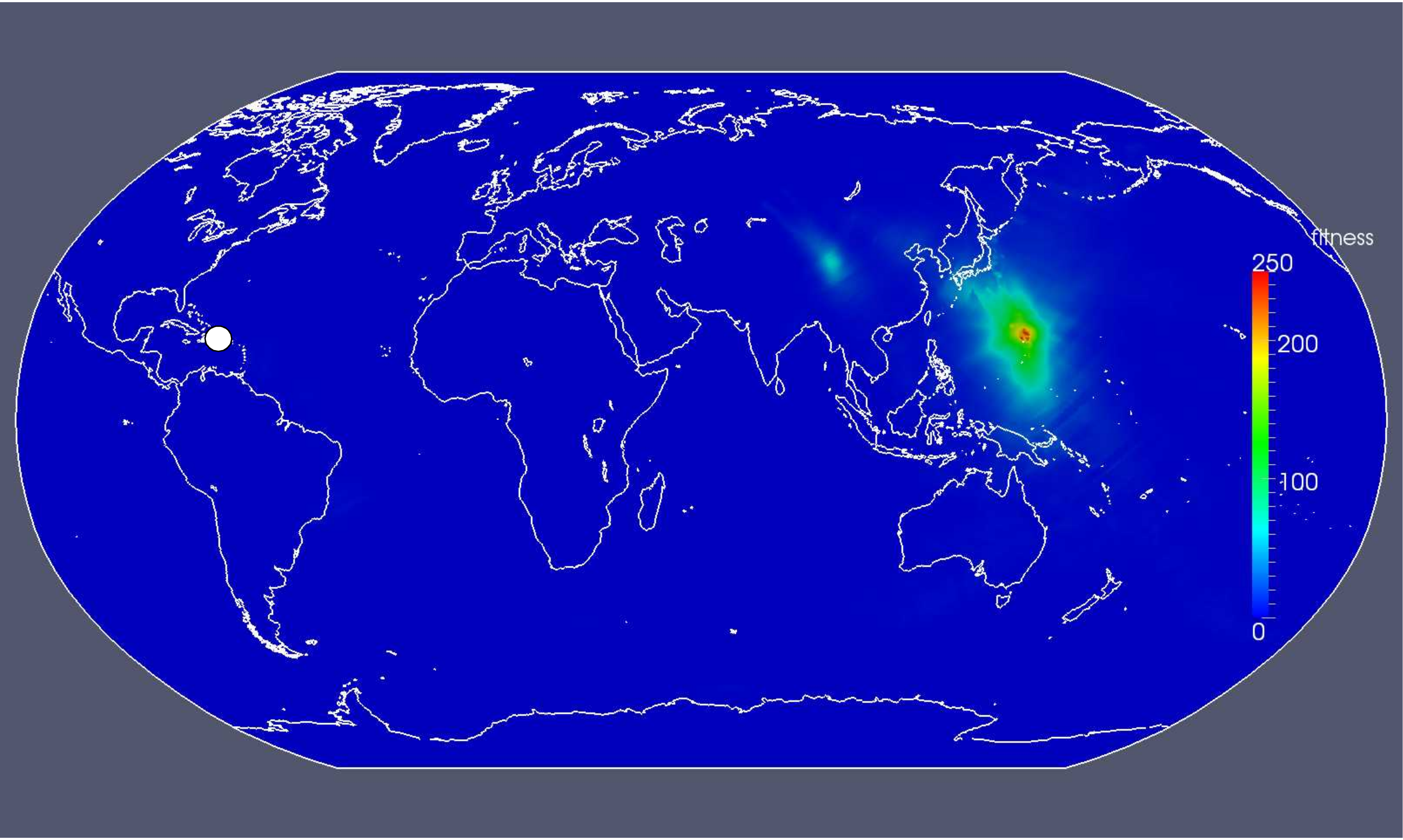
5 Known events; 560 automatically picked arrivals.



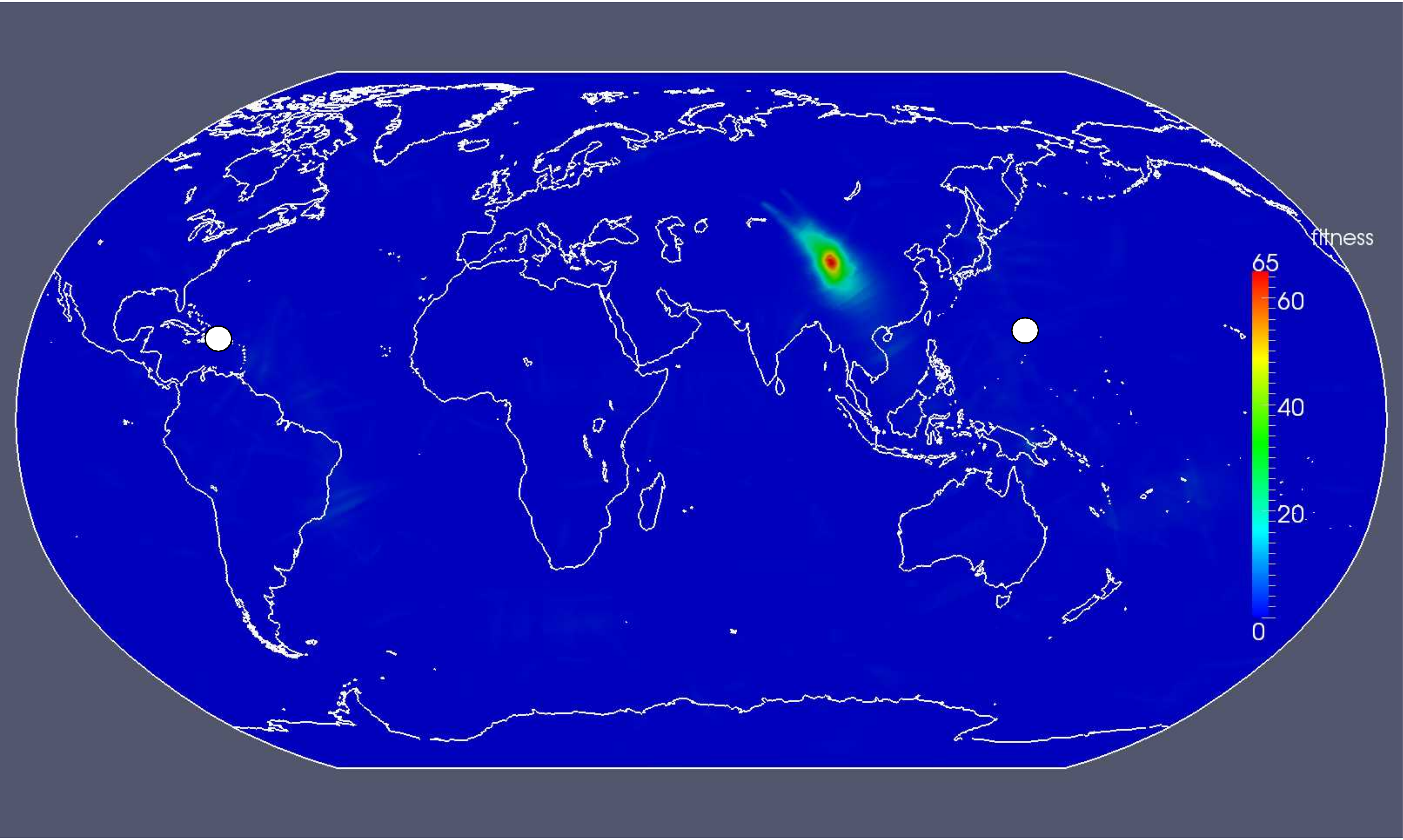
Magnitude 4.8 in Dominican Republic. 220 / 560 arrivals



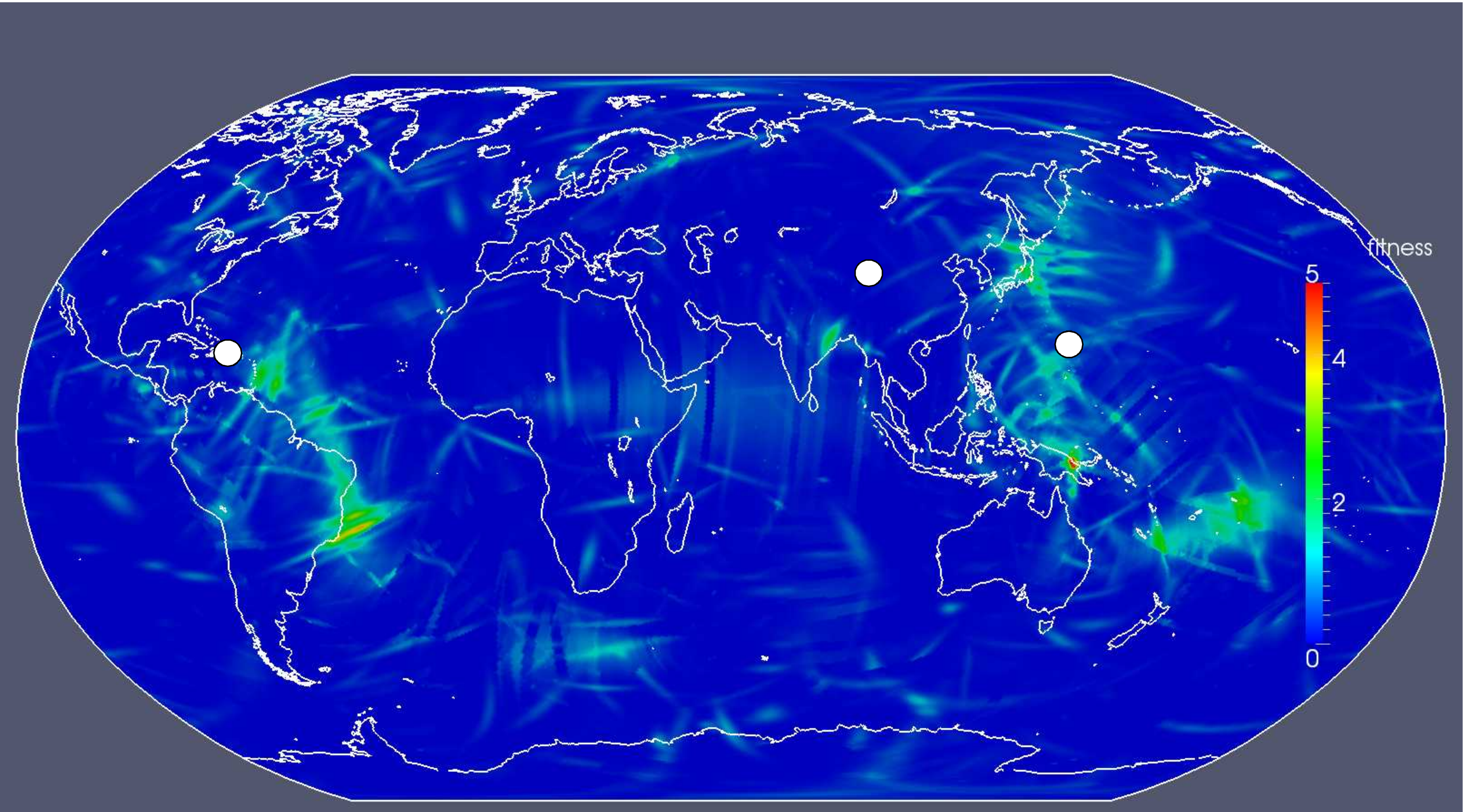
4 minutes later, magnitude 4.3 in Mariana Trench. 77 / 340 arrivals



24 minutes, magnitude 4.6 in China. 21 of 263 arrivals



17 minutes, magnitude 3.5 near Papua New Guinea. 12 of 242 arrivals



Missed 3-station event near Fiji. Found spurious event near Brazil.