Global Earthquake Activity Rate model 1 (GEAR1)

Peter Bird, Yan Y. Kagan, & David D. Jackson
UCLA

Corné Kreemer
UNR

Ross Stein
USGS & GEM Foundation

& Graeme Weatherill
GEM Foundation

for the SCEC CSEP/USGS/GEM Workshop, 6 September 2014,
Palm Springs, California

Goals:

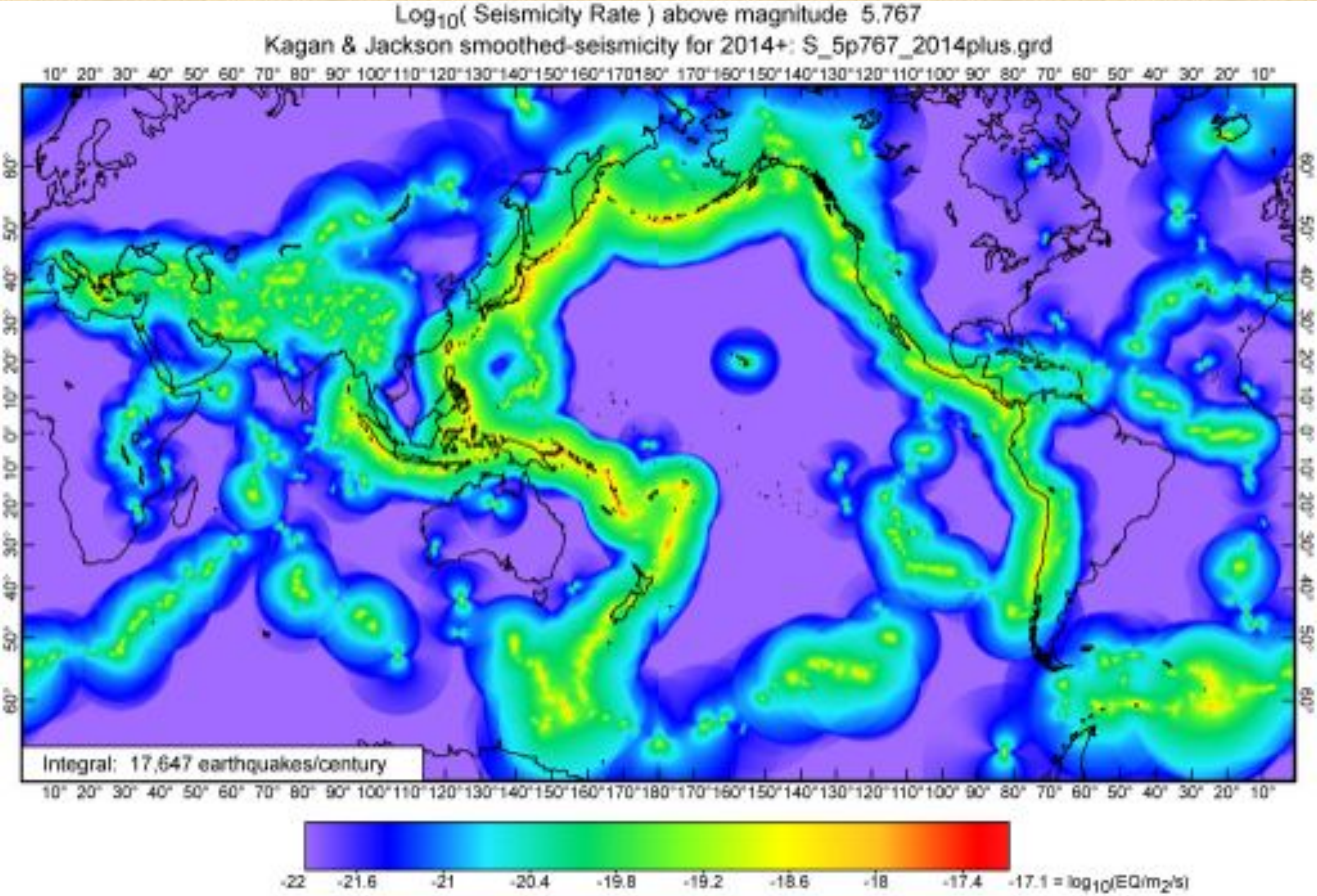
1) A global reference model should be based on global datasets that have been processed in uniform ways and that have been published.
   {GCMT (& ISC-GEM?) catalogs, and GSRM2.1 strain-rates (updating PB2002)}

2) A global reference model should blend the best features of competing seismicity models which have been previously proposed and tested.
   {Kagan & Jackson smoothed seismicity; Bird & Kreemer SHIFT model}

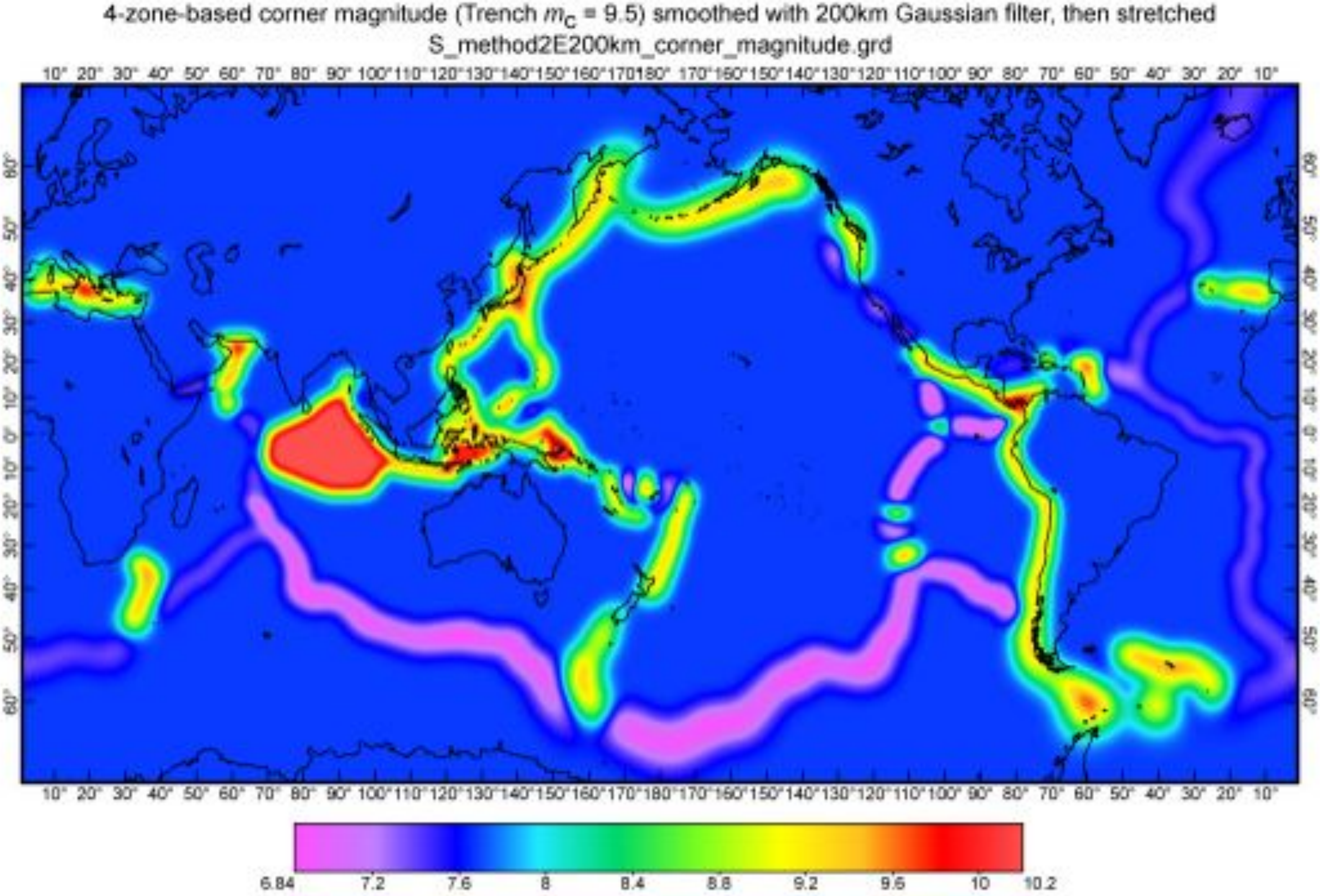
3) The calculation of the model should be algorithmic, using a published source code. In addition to providing transparency, this guarantees that updating can be straightforward. **{GEAR1_for_CSEP.f90}**

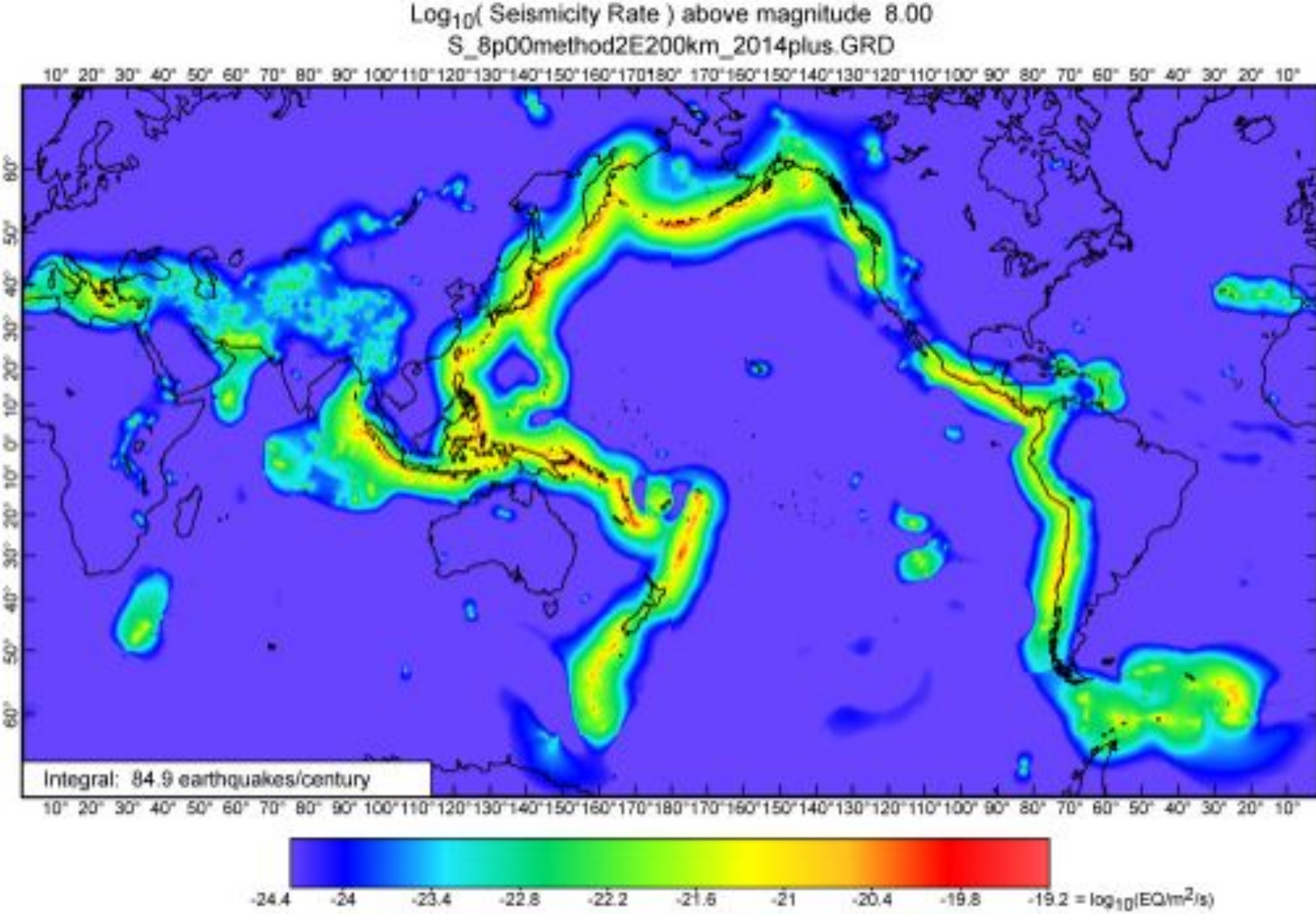
4) The model should be presented with sufficient specificity, and in a format, that makes it testable in a few years using established scoring and/or binary algorithms, and should not be significantly inferior in performance to any other forecast which meets the same criteria.
   {stationary total seismicity on a global grid of 0.1° cells; one depth bin -> 70 km; 31 magnitude bins}

Parent forecast “Seismicity” = optimally-smoothed GCMT seismicity from Kagan & Jackson:

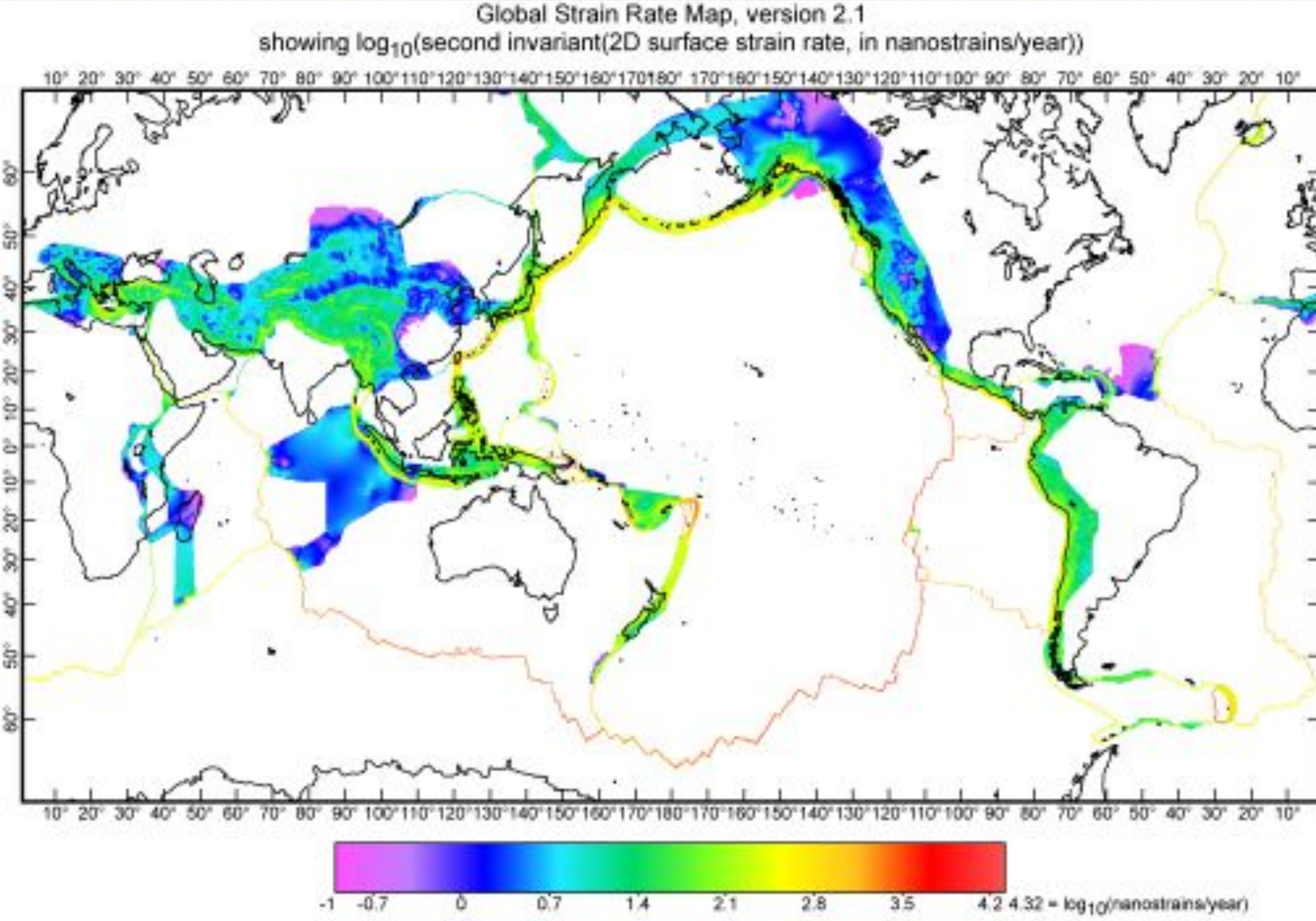


Extrapolation to higher threshold(s) uses corner magnitudes estimated for tectonic zones {*Kagan et al.*, 2010} and/or plate boundary classes {*Bird & Kagan*, 2004; *Kagan & Jackson*, 2013}:

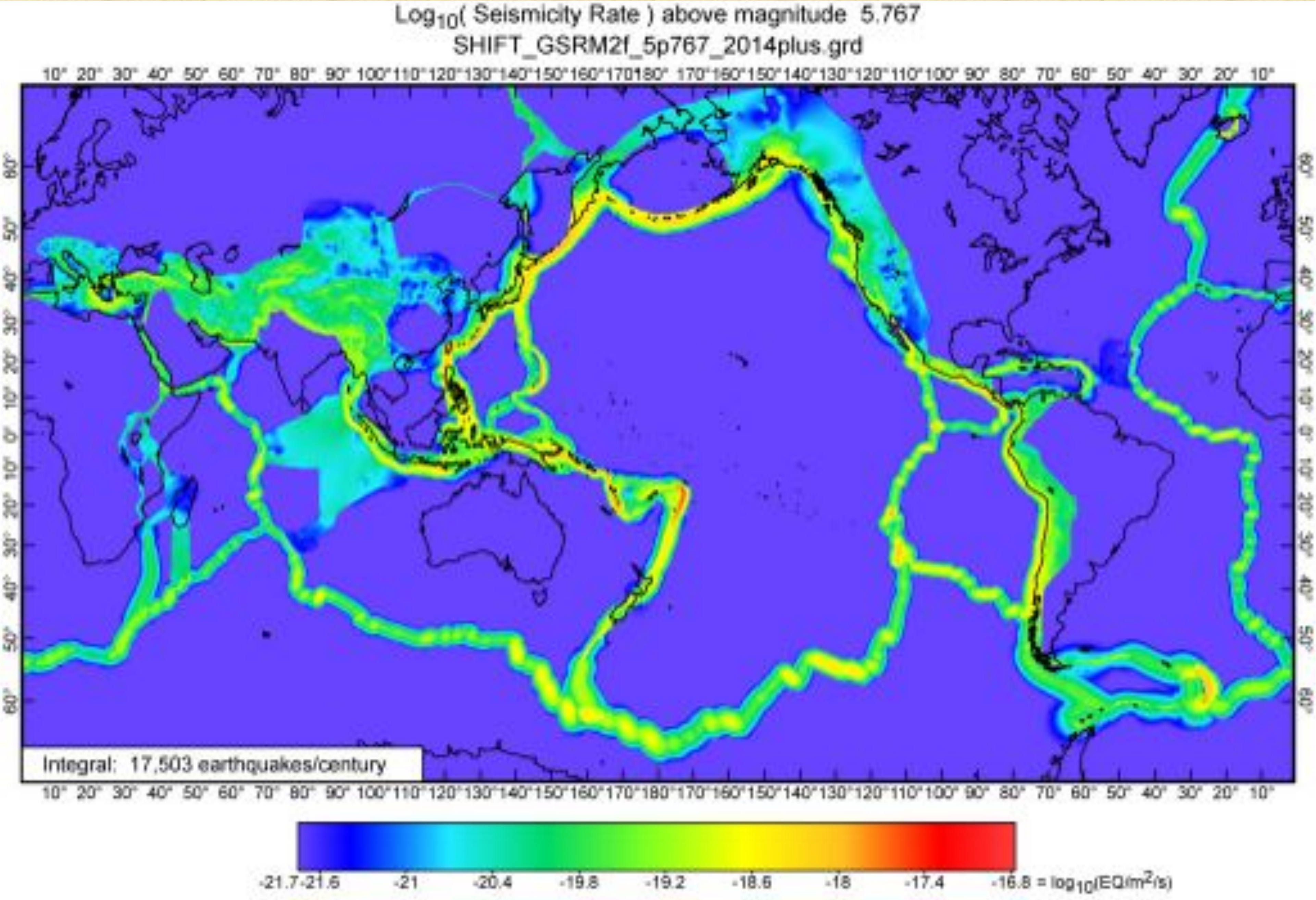



Log10(Seismicity Rate) above magnitude 8.00
S_8p00method2E200km_2014plus.GRD
Integral: 84.9 earthquakes/century
-24.4 -24 -23.4 -22.8 -22.2 -21.6 -21 -20.4 -19.8 -19.2 = log10(EQ/m2/s)

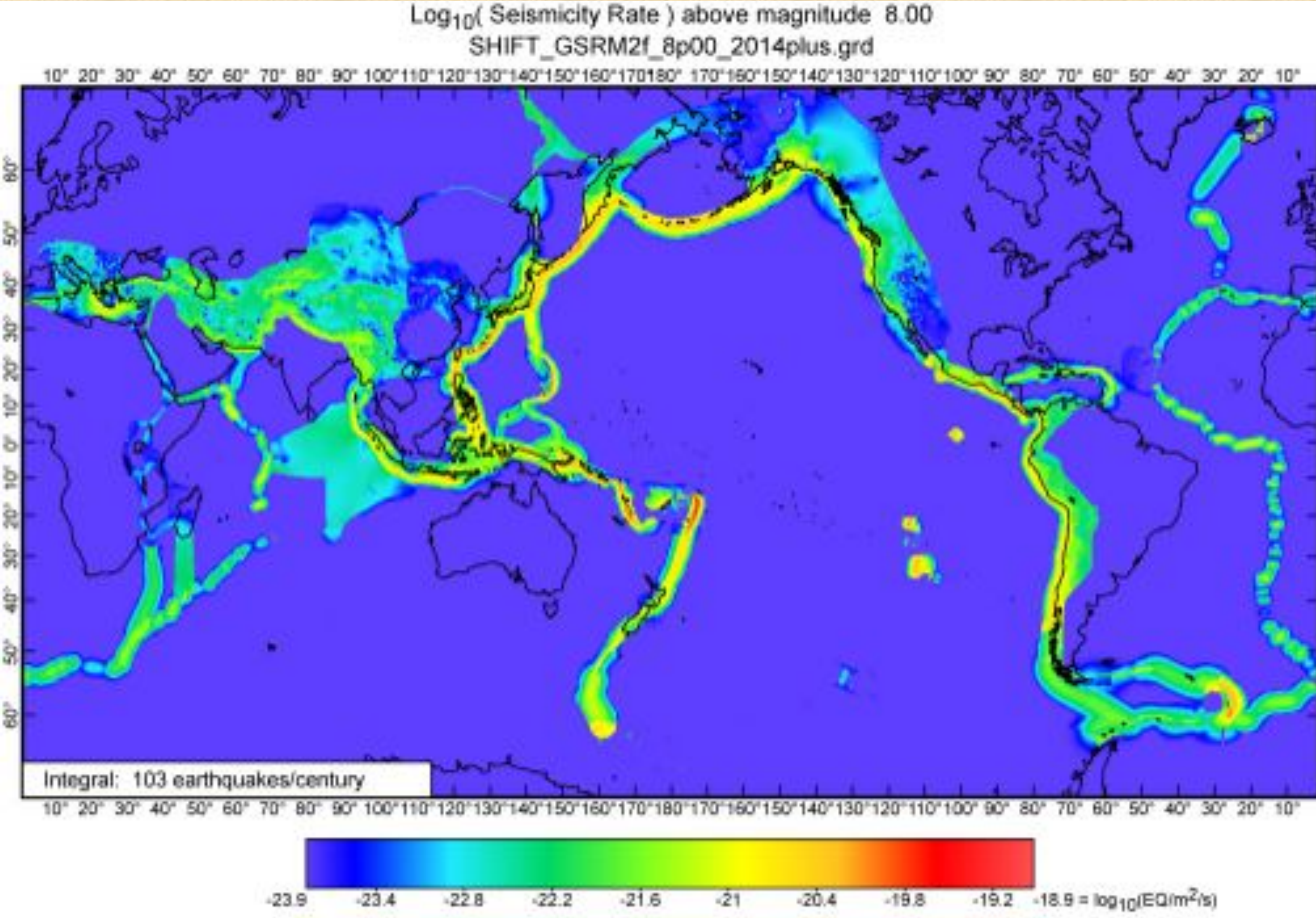
"Tectonics" parent forecast is based on GSRM2.1 {*Kreemer et al.*, 2014}, based on GPS & PT:



Conversion of strain-rates to seismicity is based on SHIFT hypotheses {*Bird et al.*, 2010} and seismicity parameters of plate-boundary classes {*Bird & Kagan,* 2004; *Bird et al.*, 2009}:



Extrapolation of "Tectonics" forecast to higher magnitudes is built-in {*Bird & Kreemer*, 2014?}:



HYBRID FORECASTS

One traditional kind of hybrid is a weighted linear-combination of S and T:

$$H_{ij} = \mathrm{N}(\sup((c\, S_{ij} + (1 - c)\, T_{ij}), b))$$

where c is to be determined. This kind of hybrid is appropriate if the two parent forecasts are regarded as expressing alternative measurements of the same underlying process, possibly with different error sources.

Another possible view might be that the two parent forecasts capture *independent* prerequisites for seismicity: there must be a continuing energy source for lithospheric deformation (some of which is elastic), *and* there must also be triggering by sudden stress changes (either static or dynamic) due to nearby earthquakes, to overcome slow rock-healing mechanisms and start a new earthquake rupture. In this view, it is more appropriate to *multiply* the S and T estimates:

$$H_{ij} = \mathrm{N}(\sup((S_{ij}^{d}\, T_{ij}^{(1-d)}), b))$$

where d is an exponent to be determined. This set of hybrids will be called "log-linear."

Finally, it is possible that both the Seismicity and Tectonics forecasts underestimate the true rates in different localities, and so taking the larger of the two in every cell might more successfully forecast future quakes. From this point of view, it may be natural to also consider an "envelope" hybrid which selects the greater of Seismicity or Tectonics:

$$H_{ij} = \mathrm{N}(\sup(S_{ij}, T_{ij}))$$

Information ***(I)*** scores
[*Kagan*, 2009, *GJI*]

- Scoring is applied to a map of the density of *conditional* probability that the next earthquake in the test area will be at a certain epicenter. Such maps always have an area-integral of 1, so the overall rate of the forecast is never a factor.
- Result is expressed as number of binary bits of information gain (over a spatially-uniform model) *per* {test, or virtual} *earthquake*, so the number of test earthquakes is never a factor.
- No simulation of virtual catalogs (from forecast) is needed.
- No changes are made to the test catalog (no declustering, and no random perturbations of: epicenter, depth, *m*).
- Finite-test effects cause only mild (and Gaussian) perturbations of the score from its ideal (long-term) value.

Information *(I)* scores

[*Kagan*, 2009, *GJI*]

$$I_1 = \frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{\lambda_i}{\xi}$$

where ***n*** is the number of test earthquakes,
λ_i is the conditional probability density of the cell in which the test earthquake epicenter occurred, and
ξ is the mean conditional probability density in the forecast region.
I like to call I_1 the "***success***" of the forecast.

Information *(I)* scores

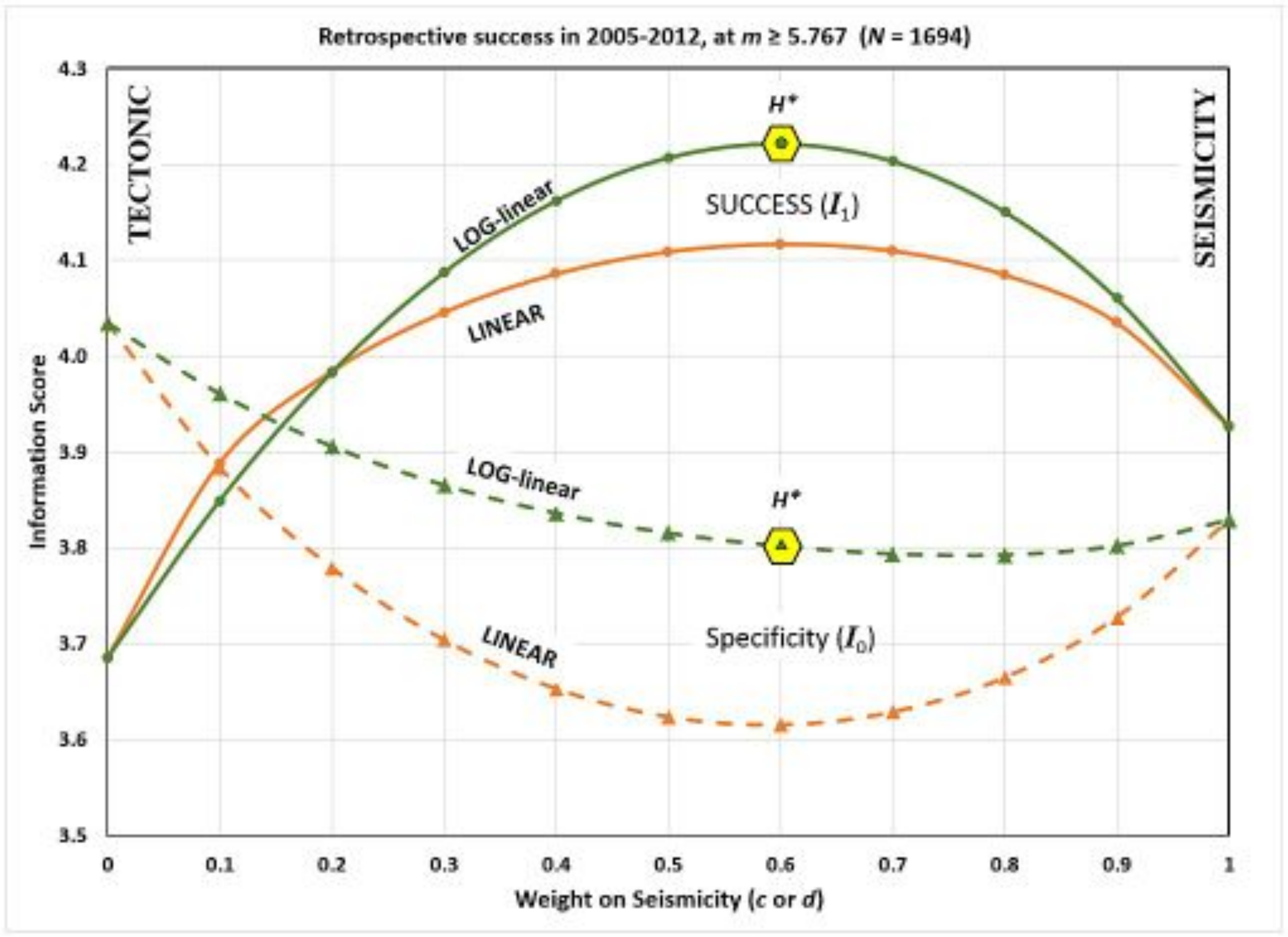
[*Kagan*, 2009, *GJI*]

$$I_0 = \sum_{i=1}^{N} \nu_i \log_2\left(\frac{\nu_i}{\tau_i}\right)$$

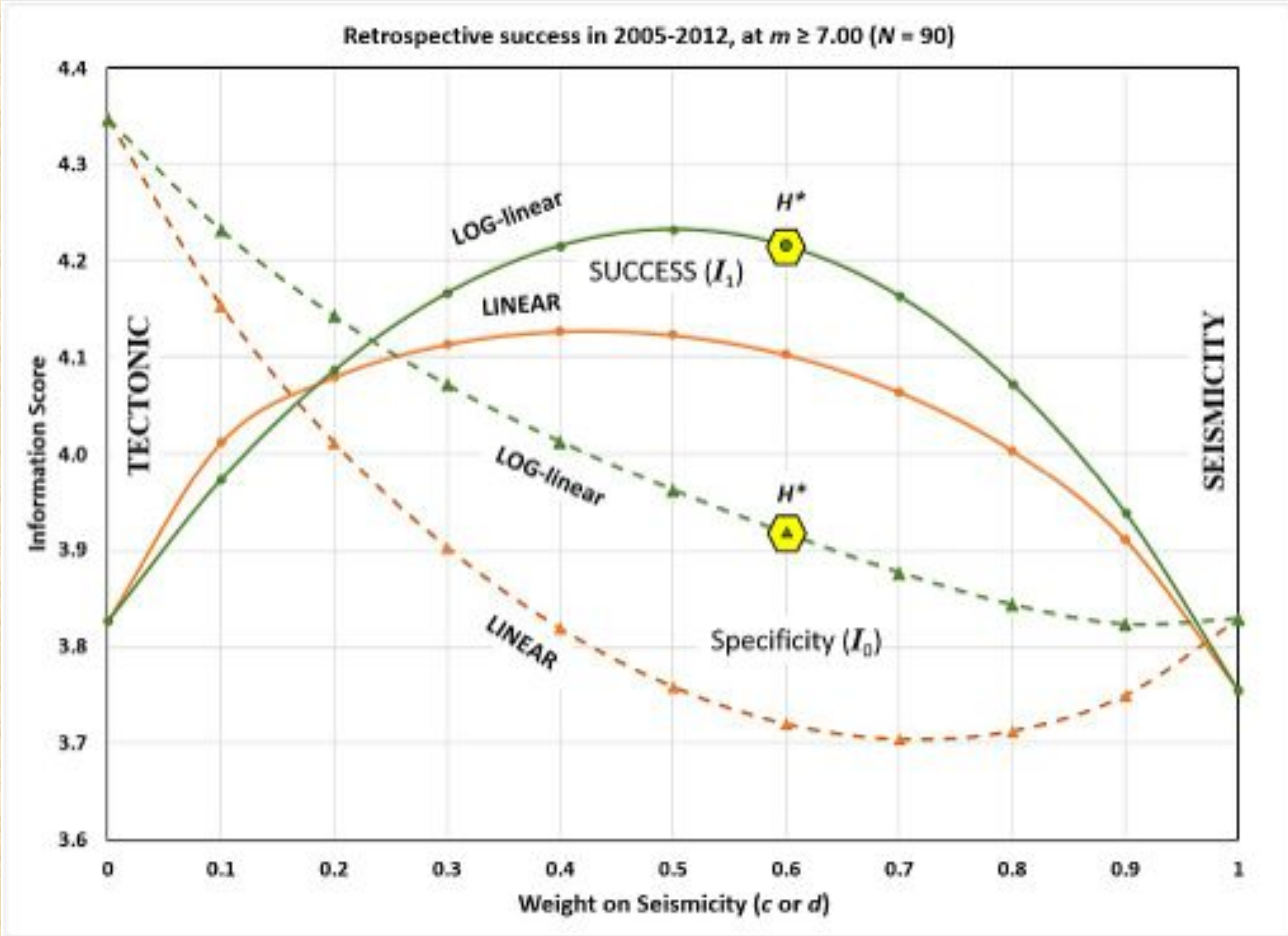
where $\boldsymbol{i} = 1, \ldots, N$ are the cells or gridpoints,
$\boldsymbol{\nu_i}$ are their normalized forecast rates (conditional probabilities),
and $\boldsymbol{\tau_i}$ are the normalized cell (or gridpoint) areas.

I like to call $\boldsymbol{I_0}$ the "***specificity***" of the forecast;
it gives the $\boldsymbol{I_1}$ score that would be found in a long
test if the forecast were exactly correct.

The preferred hybrid H* is identified using tests against GCMT, $m \geq 5.767$, 2005-2012:

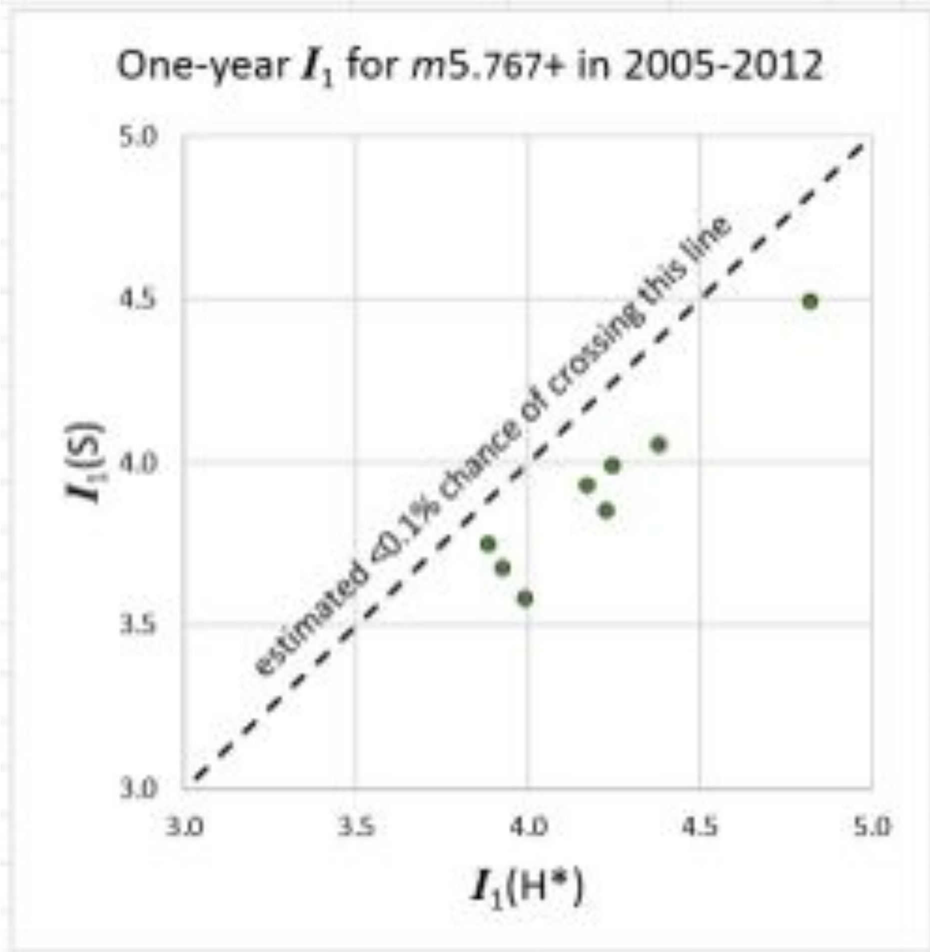


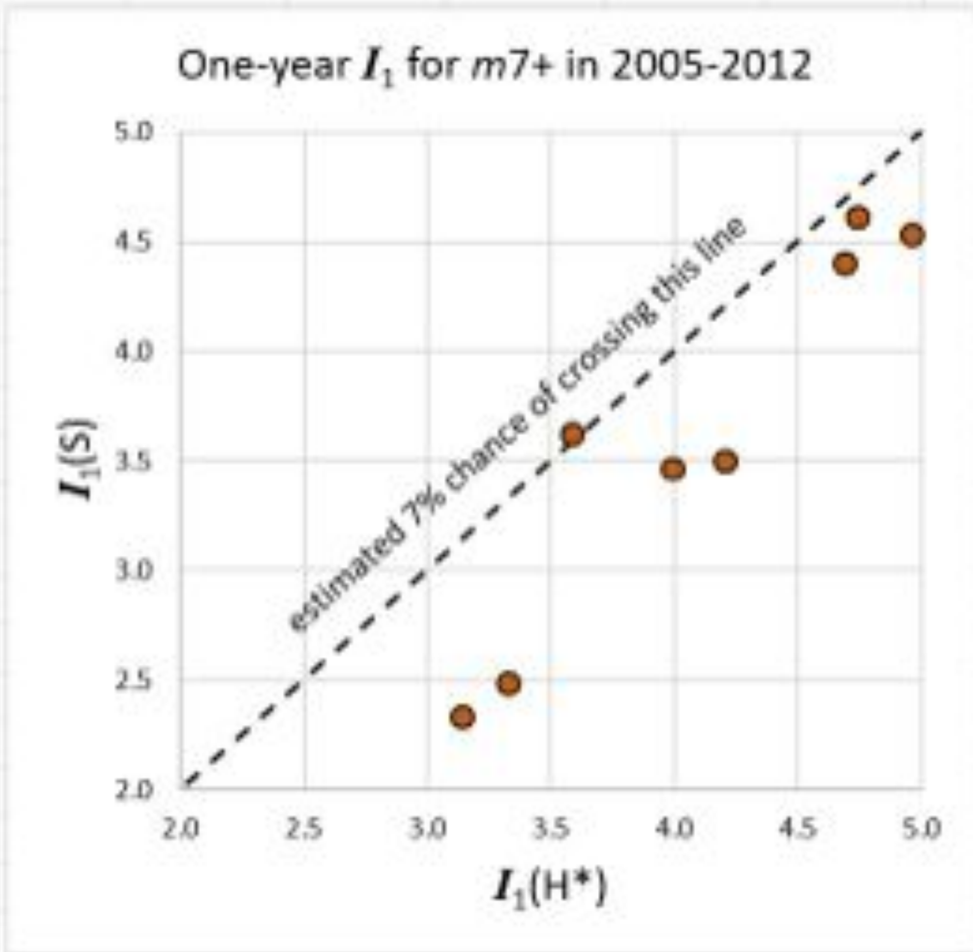
This preferred hybrid H* also outperforms both parents at $m \geq 7$ (GCMT, 2005-2012):



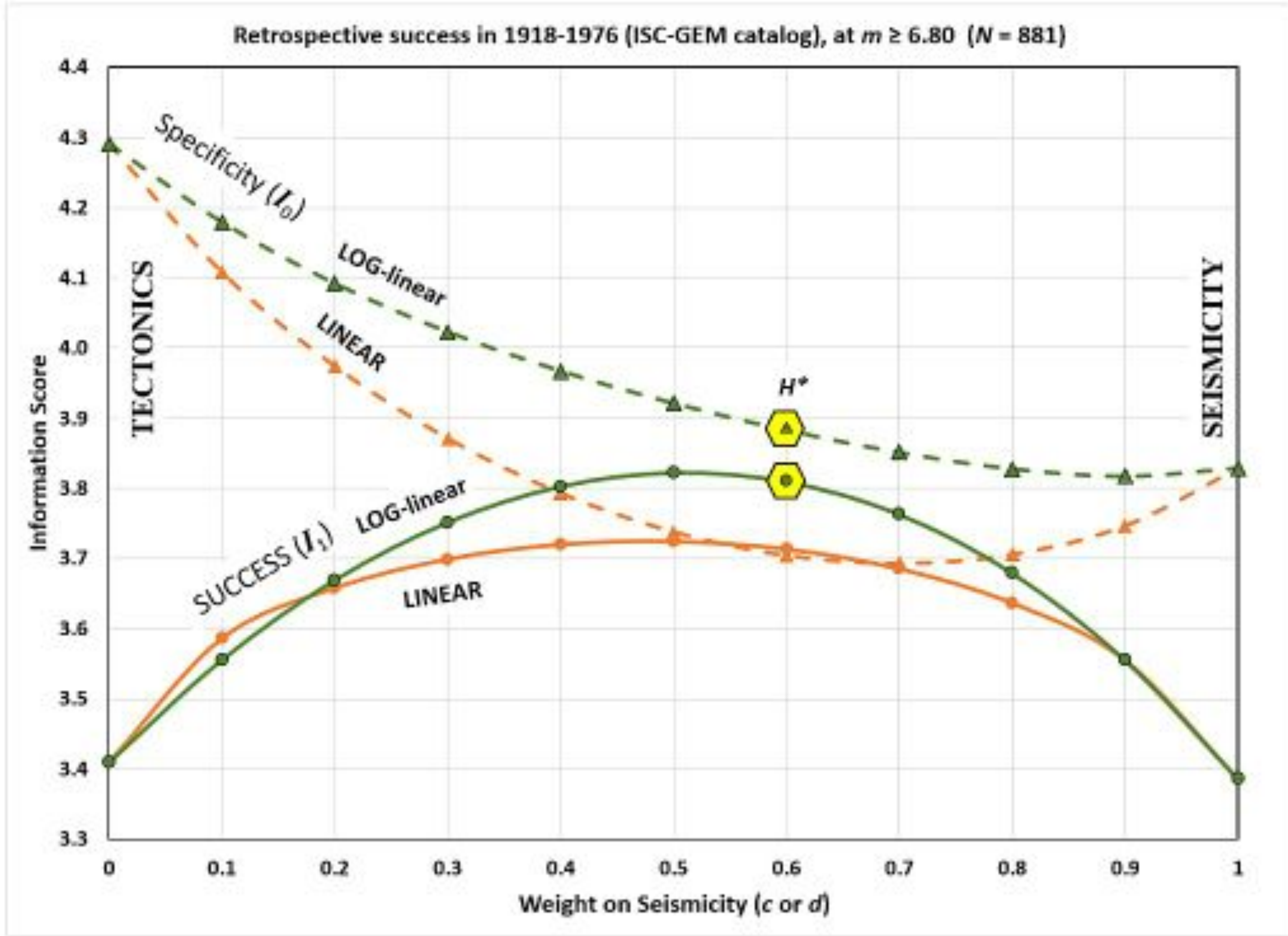
Sets of 1-year tests against GCMT show that varying I_1 scores of competing models are highly correlated, with correlation coefficients of 0.94 ($m \geq 7$) to 0.96 ($m \geq 5.767$).

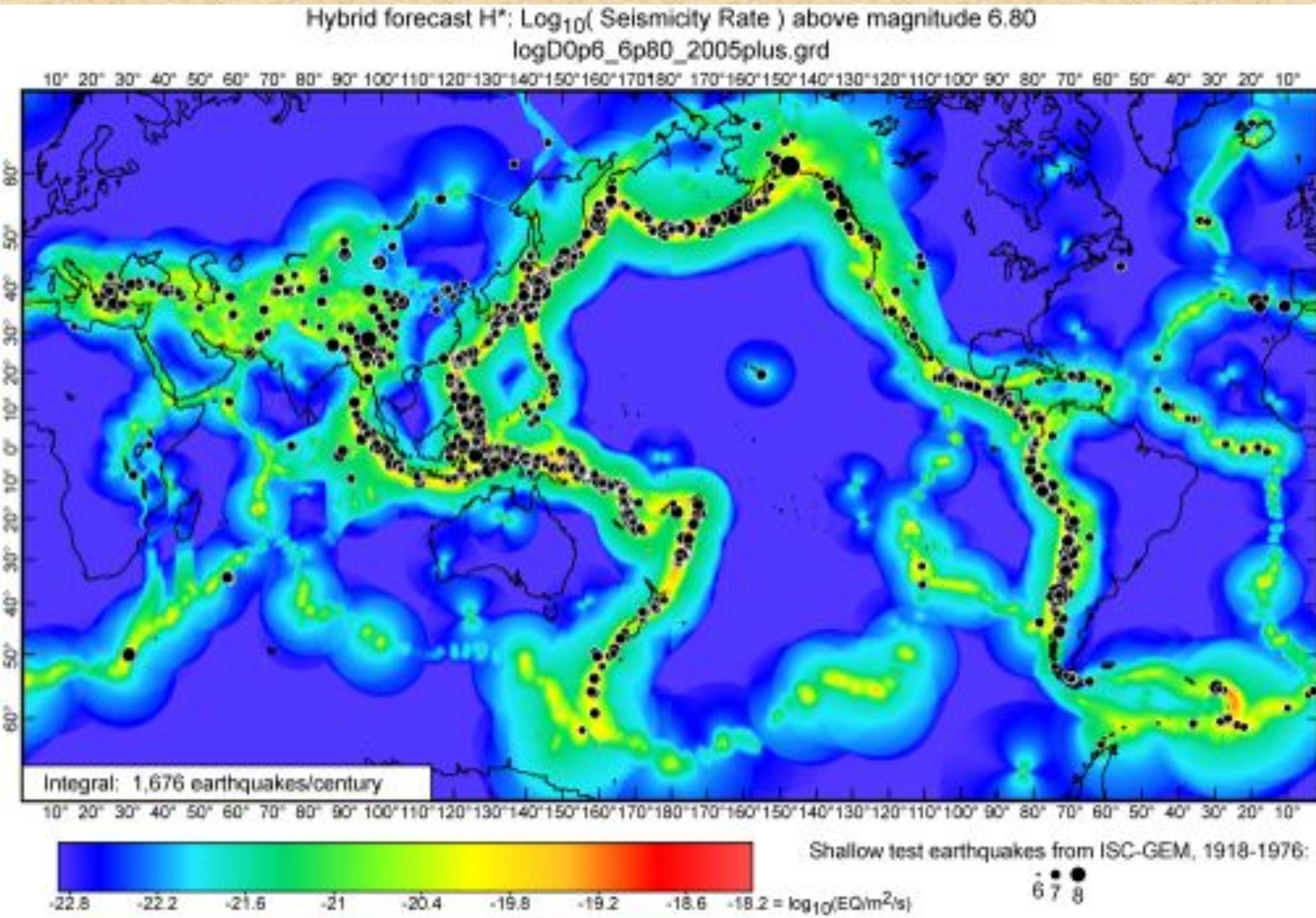
Therefore, we can show that the improvement we have identified, through hybridization, and using 8-year tests, has a significance of 9 standard deviations at $m \geq 5.767$, and of 4 standard deviations at $m \geq 7$.



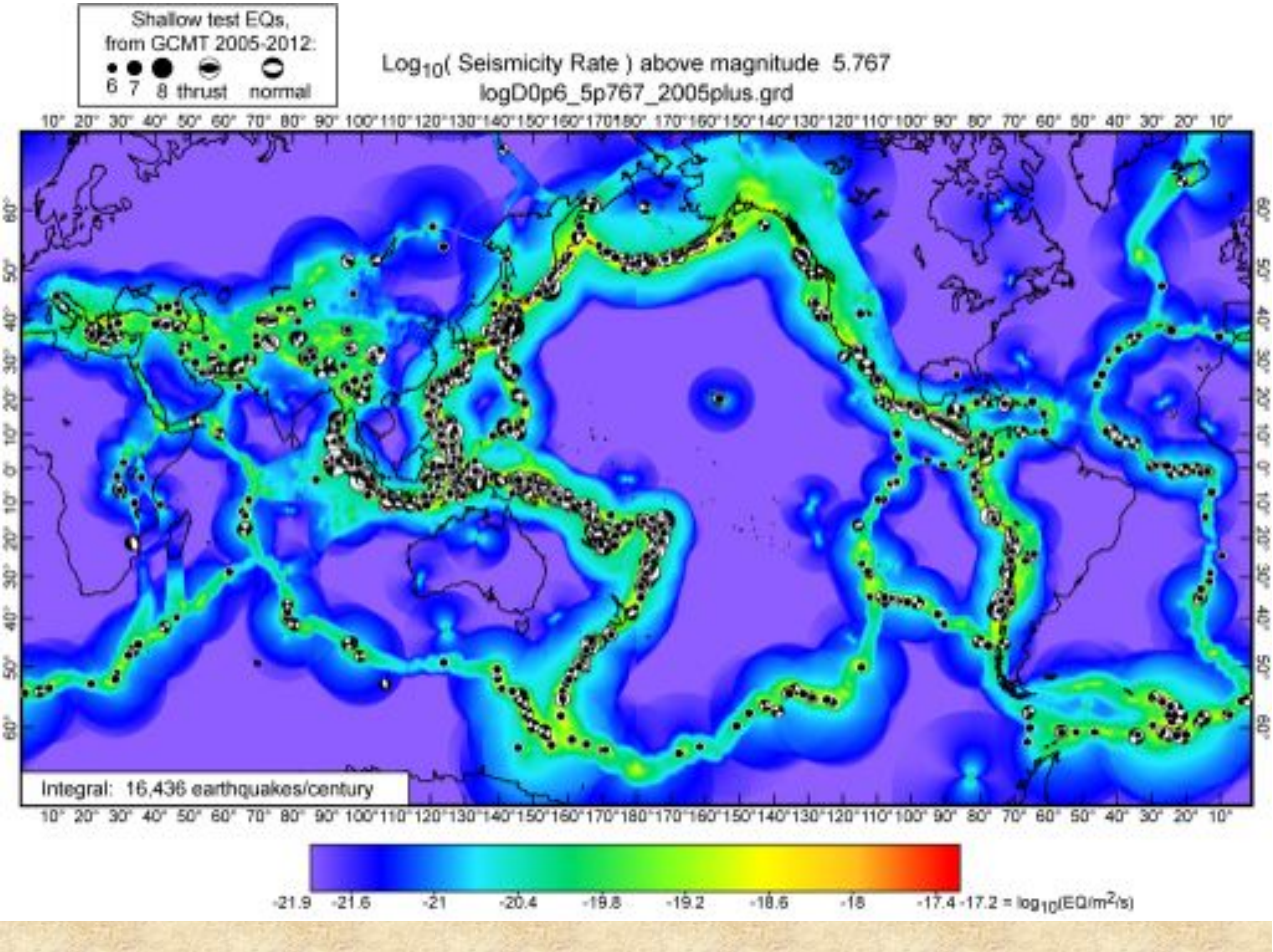


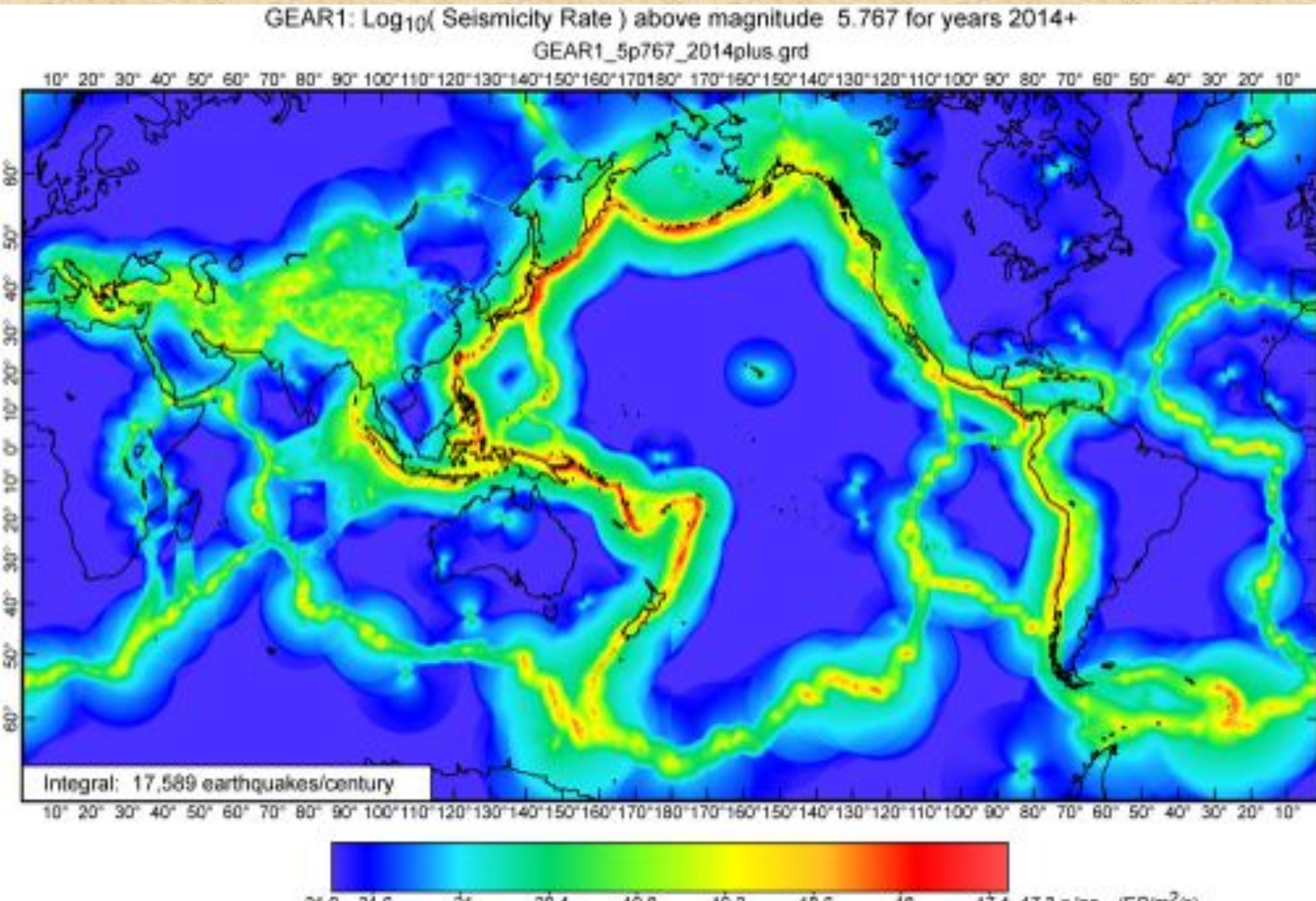
Excitingly, preferred hybrid H* also outperforms both parents in retrospective tests against the complete part of the new ISC-GEM catalog ($m \geq 6.80$ in 1918-1976).





Hybrid forecast H*: Log10(Seismicity Rate) above magnitude 6.80
logD0p6_6p80_2005plus.grd
Integral: 1,676 earthquakes/century
-22.8 -22.2 -21.6 -21 -20.4 -19.8 -19.2 -18.6 -18.2 = log10(EQ/m2/s)
Shallow test earthquakes from ISC-GEM, 1918-1976:
6 7 8


Shallow test EQs,
from GCMT 2005-2012:
6 7 8 thrust normal
Log10(Seismicity Rate) above magnitude 5.767
logD0p6_5p767_2005plus.grd
Integral: 16,436 earthquakes/century
-21.9 -21.6 -21 -20.4 -19.8 -19.2 -18.6 -18 -17.4 -17.2 = log10(EQ/m2/s)


GEAR1: Log10(Seismicity Rate) above magnitude 5.767 for years 2014+
GEAR1_5p767_2014plus.grd
Integral: 17,589 earthquakes/century
-21.9 -21.6 -21 -20.4 -19.8 -19.2 -18.6 -18 -17.4 -17.3 = log10(EQ/m2/s)