ETAS v STEP
OEF testing strategies

J. Douglas Zechar*

Swiss Seismological Service

ETH Zurich

Department of Earth Sciences

University of Southern California

*In cahoots with Eberhard, Gerstenberger, Rhoades, Schorlemmer, & Werner

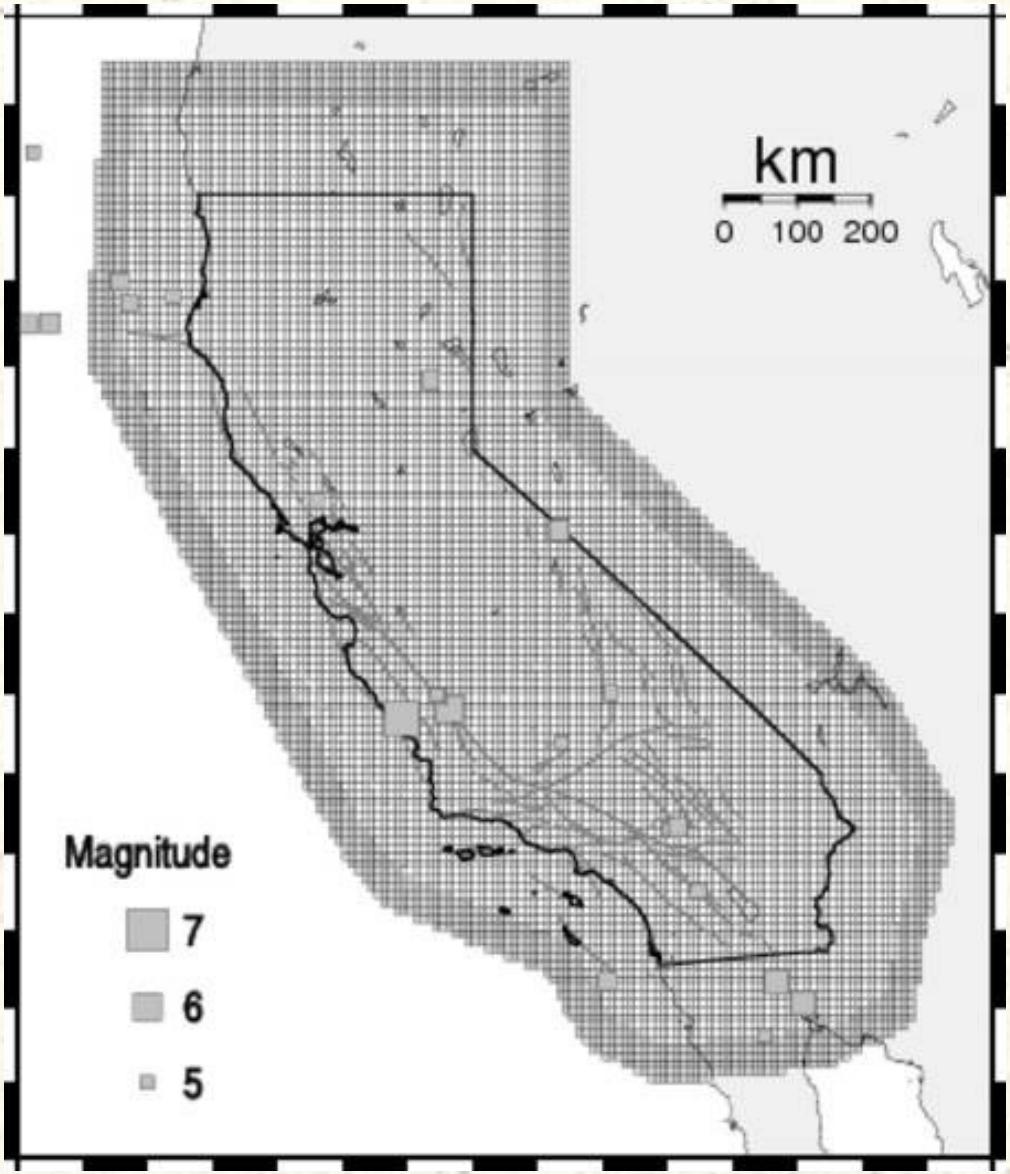
Daily forecast experiment in California

Target eqks with $M_{ANSS} \geq 3.95$ in California for $\Delta t = 1$ day

Lat/lon/mag bins are 0.1° x 0.1° x 0.1.

270 observed target eqks earthquakes between Sept. 2007 and August 2010.

2 participating models: ETAS (implemented by Zhuang) and STEP (implemented by Gerstenberger)



Zechar et al. in prep.

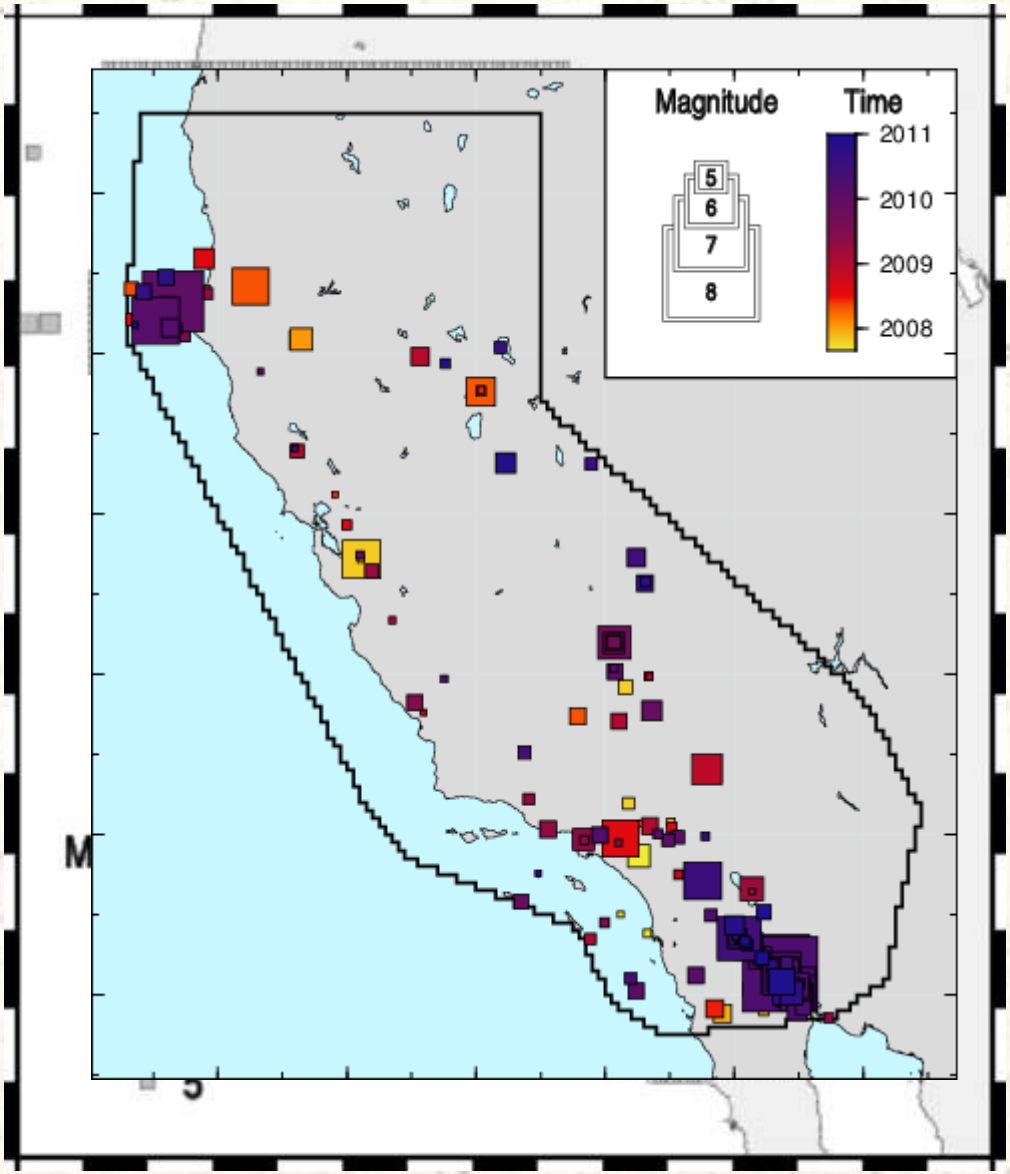
Daily forecast experiment in California

Target eqks with $M_{ANSS} \geq 3.95$ in California for $\Delta t = 1$ day

Lat/lon/mag bins are 0.1° x 0.1° x 0.1.

270 observed target eqks earthquakes between Sept. 2007 and August 2010.

2 participating models: ETAS (implemented by Zhuang) and STEP (implemented by Gerstenberger)



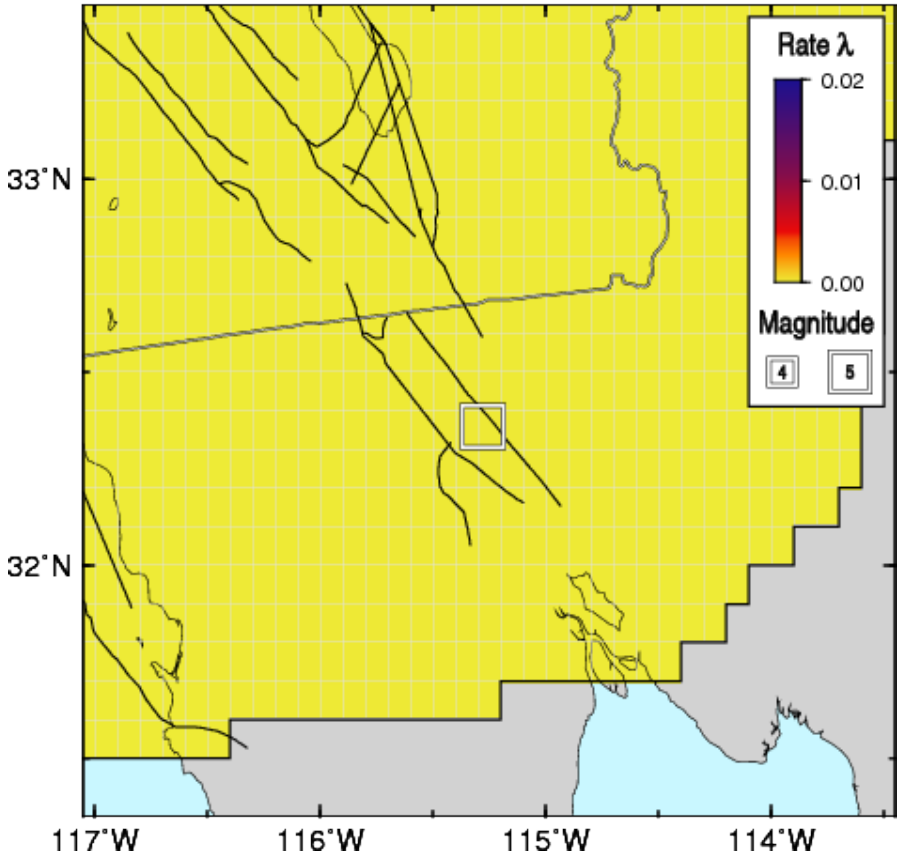
Zechar et al. in prep.

Testing:
Feb. 2008 swarm in Baja California

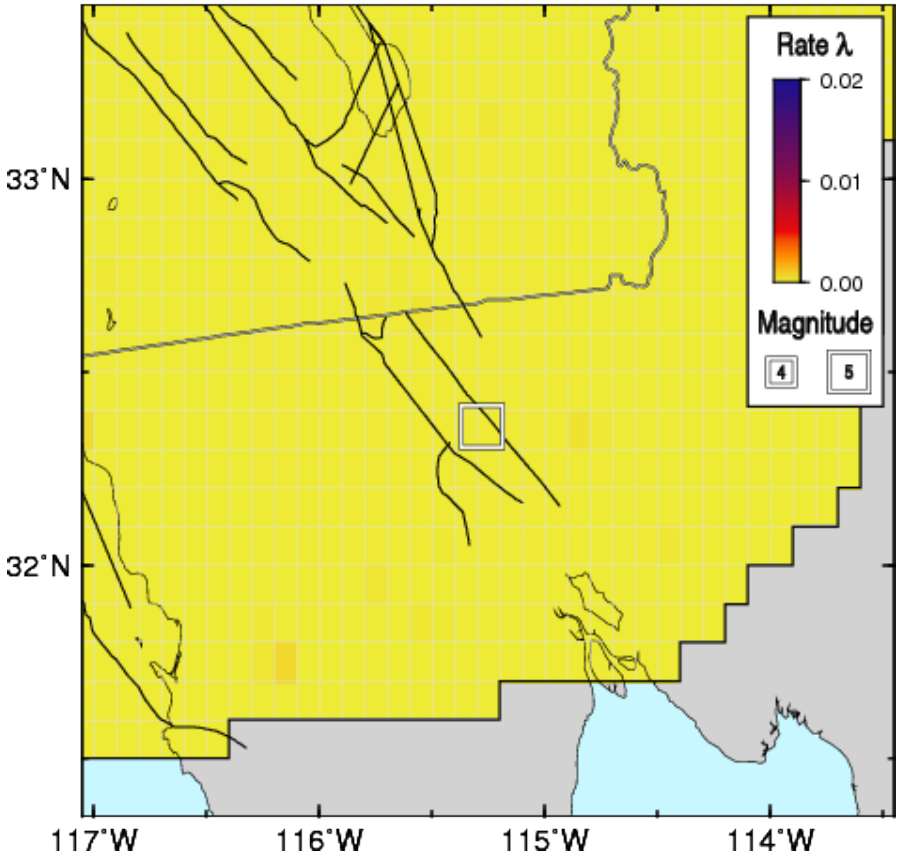
9 Feb. 2008

(background)

ETAS



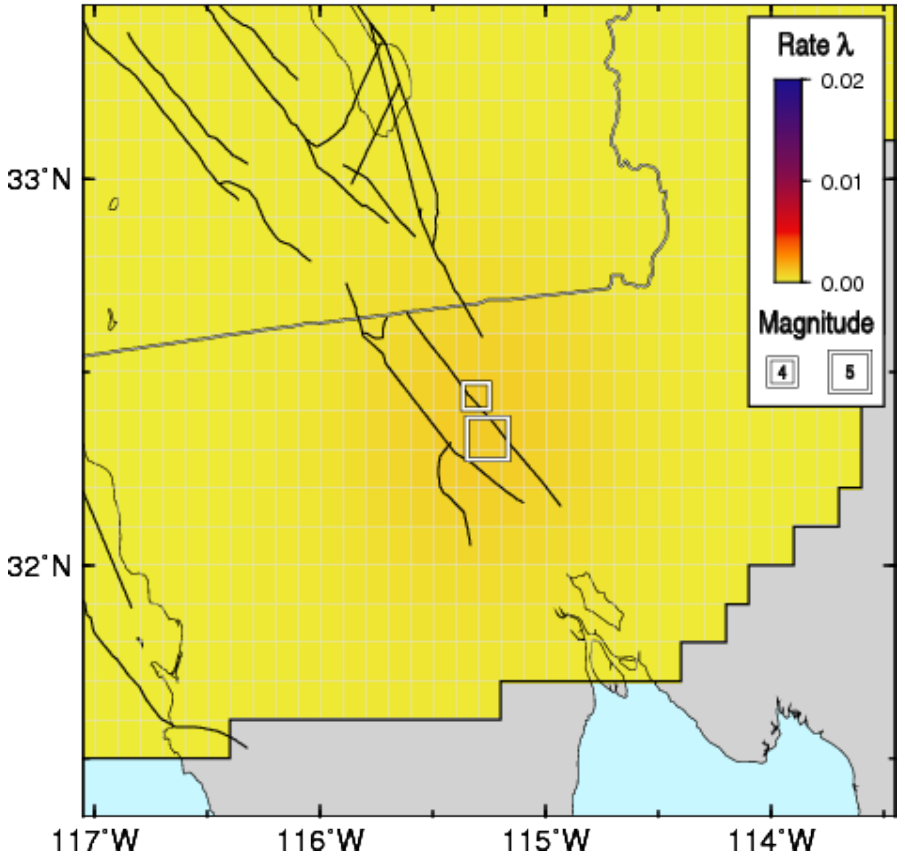
STEP



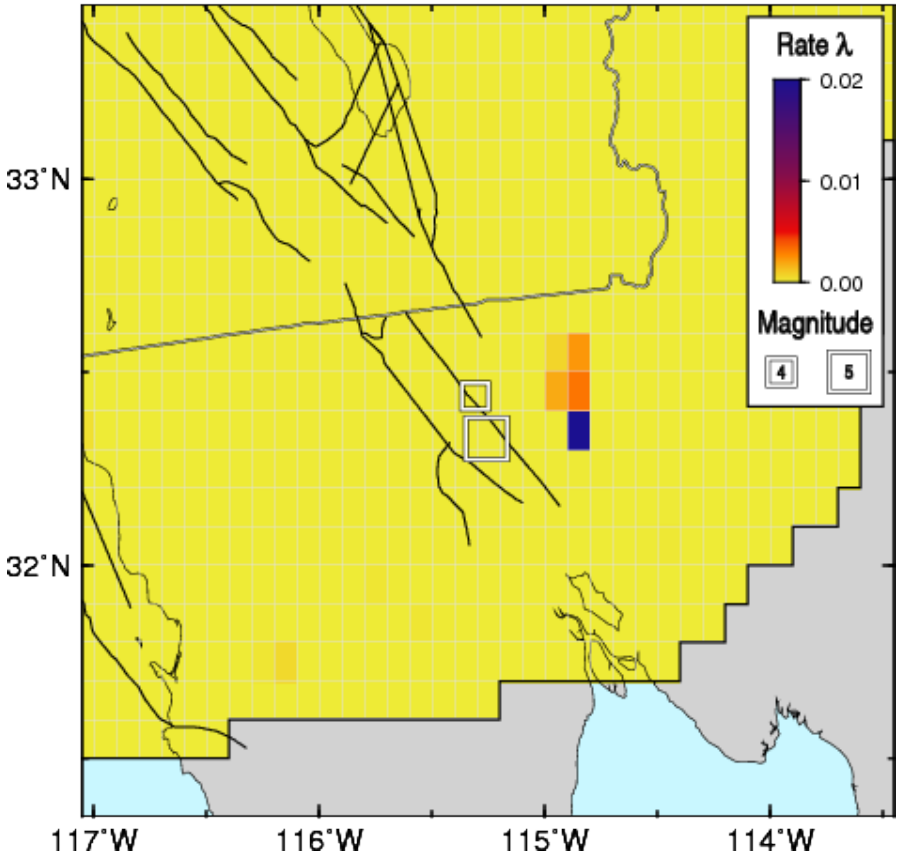
Testing:
Feb. 2008 swarm in Baja California

11 Feb. 2008

ETAS



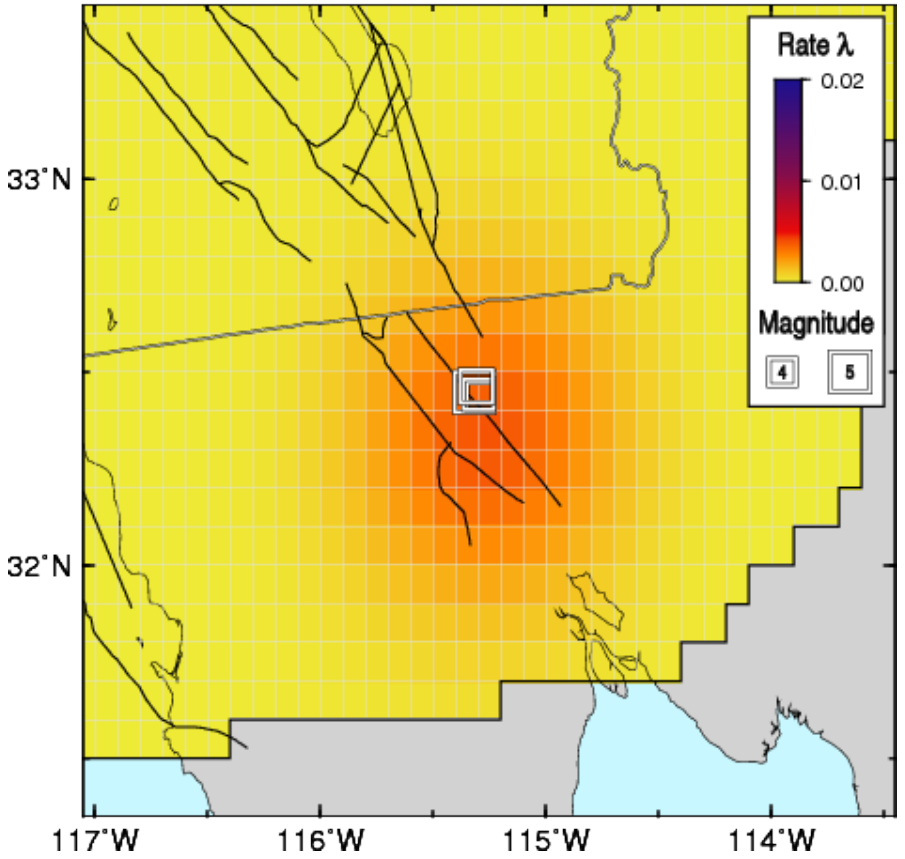
STEP



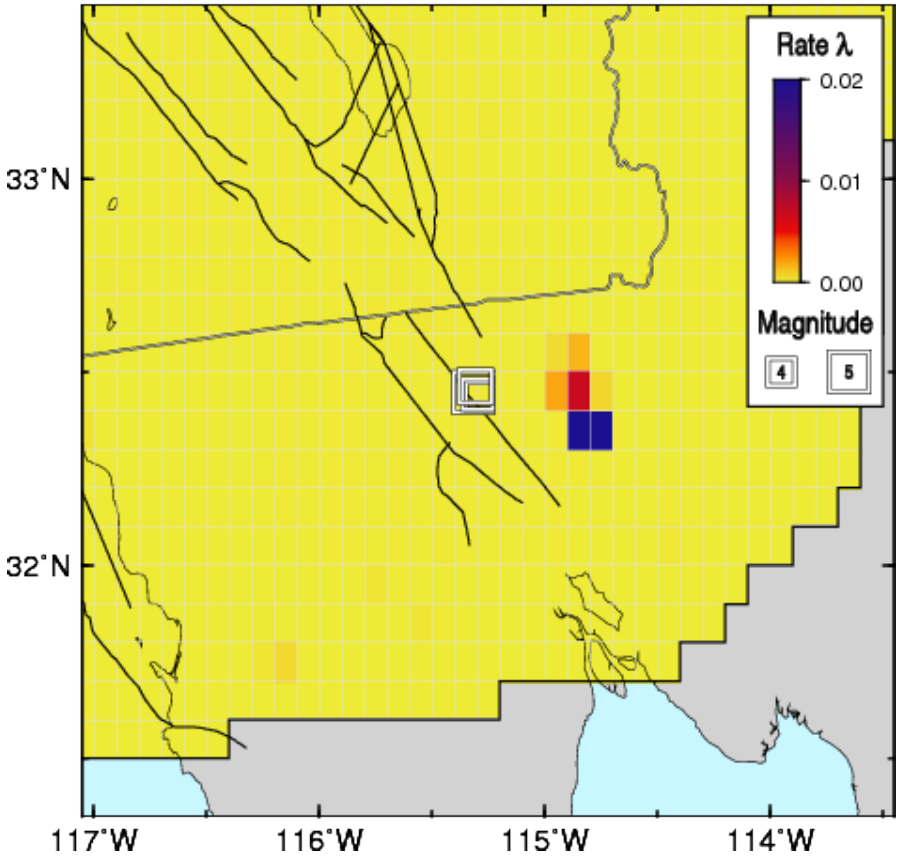
Testing:
Feb. 2008 swarm in Baja California

12 Feb. 2008

ETAS



STEP



Oops!

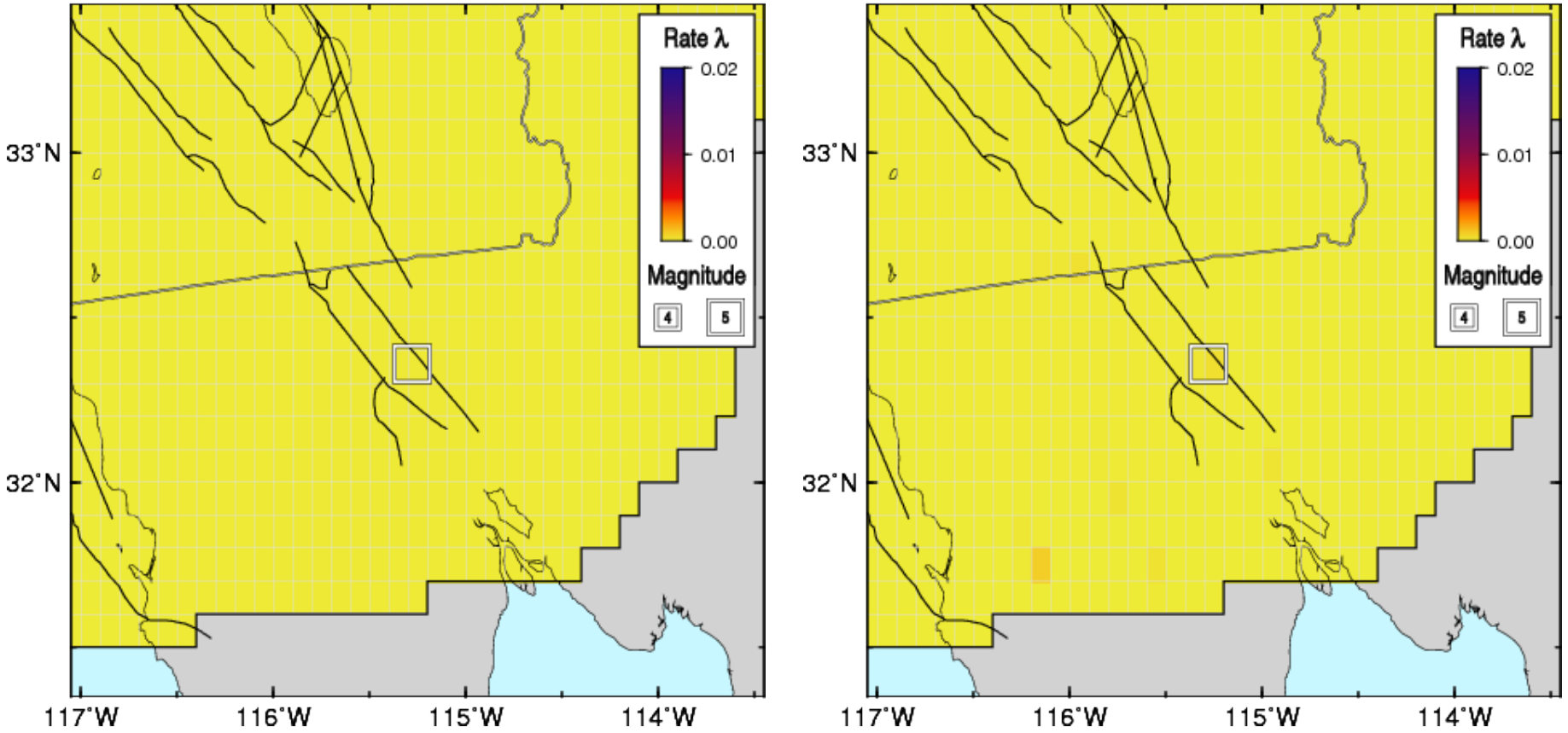
Testing:
Feb. 2008 swarm in Baja California

9 Feb. 2008

(background)

ETAS

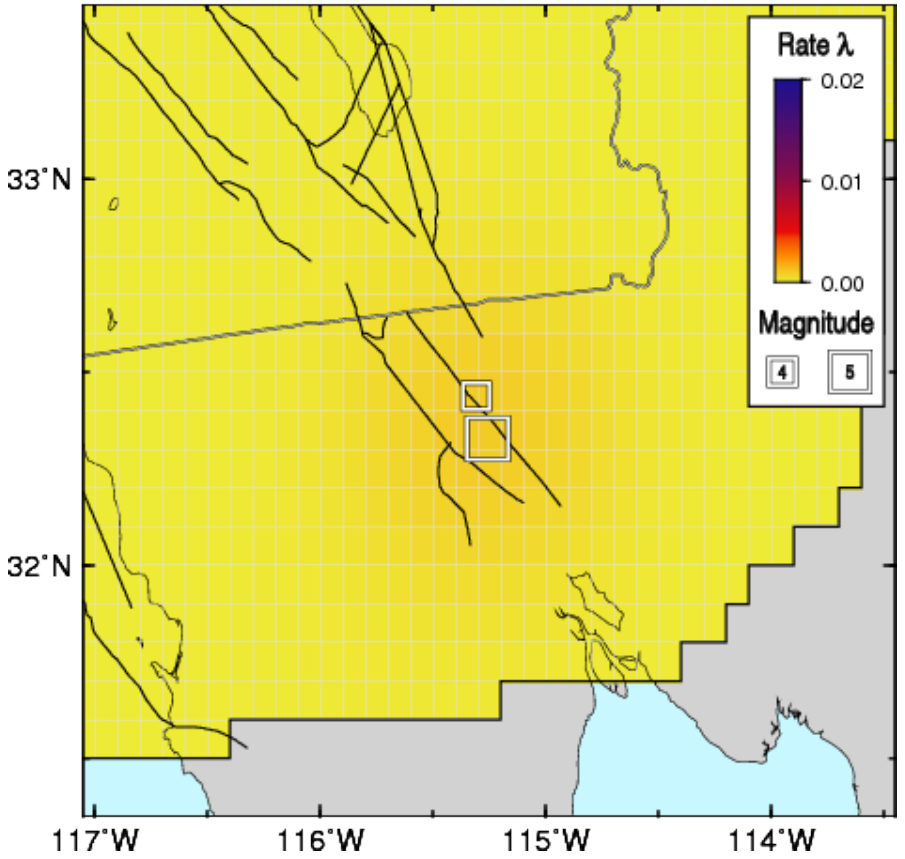
STEP, corrected



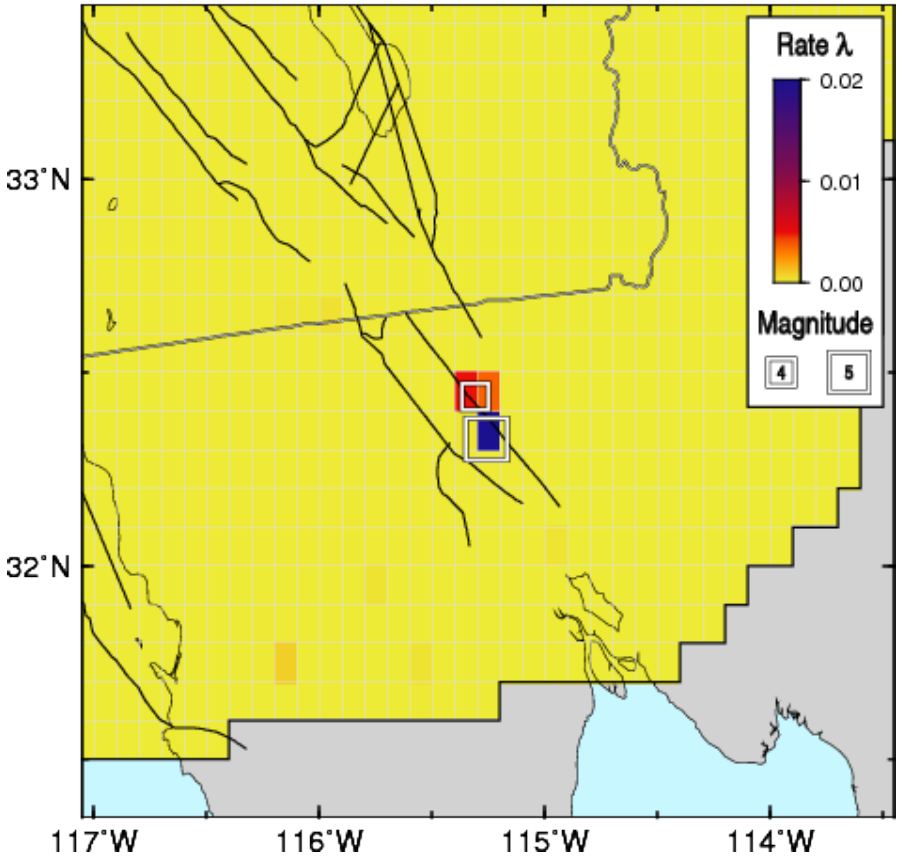
Testing:
Feb. 2008 swarm in Baja California

11 Feb. 2008

ETAS



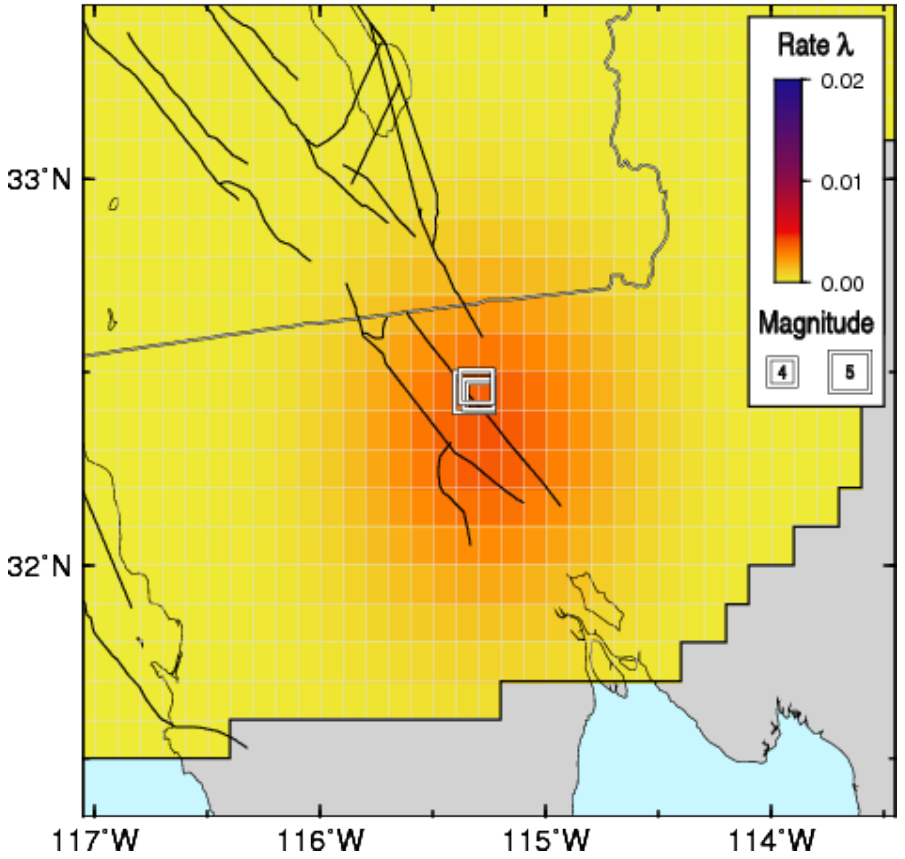
STEP, corrected



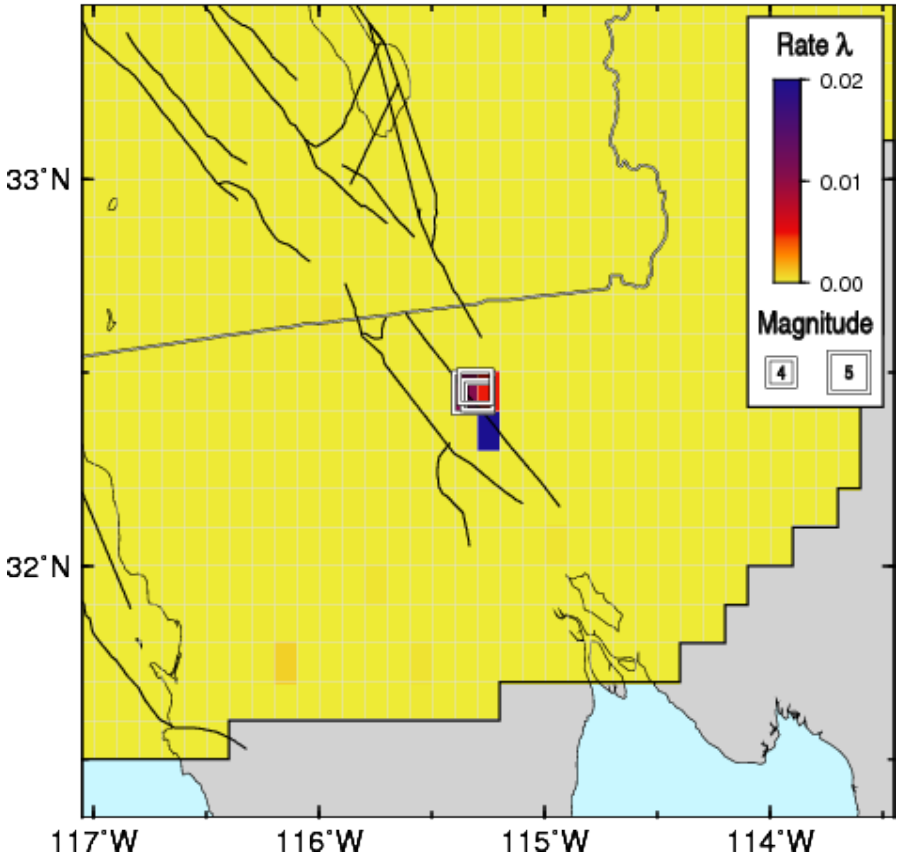
Testing:
Feb. 2008 swarm in Baja California

12 Feb. 2008

ETAS



STEP, corrected



Evaluation – consistency

The forecasts tell us the probability of any possible observation...

- How many earthquakes were expected, and how many were observed?
- What is the likelihood of the observation?

Evaluation – consistency

Number of earthquakes forecast

$$N_\Lambda = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} \lambda_{ijk}$$

Number of earthquakes observed

$$N = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} \omega_{ijk}$$

Log-likelihood of observation given forecast

$$L(\Omega|\Lambda) = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} \left(\omega_{ijk} \ln \lambda_{ijk} - \lambda_{ijk} - \ln(\omega_{ijk}!) \right)$$

For each experiment, we obtain one value for each of these measures; how can we tell if it's "good"?

Evaluation – consistency

To approximate the likelihood function of a forecast, produce many (1K-1M) stochastic catalogs consistent with the forecast, and compute the likelihood of each.

The percentile of the observed likelihood can then be thought of as a p-value.

We can do this for space-rate-magnitude forecast and each component (i.e., space, rate, and magnitude in isolation).

- Resulting tests are called L, M, N, and S.

Schorlemmer et al. 2007, 2010
Zechar et al. 2010
Werner et al. 2011

Evaluation – comparison

$$I_N(\boldsymbol{A},\boldsymbol{B}) = \frac{1}{N}\sum_{i=1}^{N}\left(\ln\frac{A_i}{B_i}\right) + \frac{N_{\boldsymbol{B}}-N_{\boldsymbol{A}}}{N}$$

I_N: rate-corrected average info. gain per eqk

$\boldsymbol{A}$, $\boldsymbol{B}$: forecasts to compare

N: number of target eqks

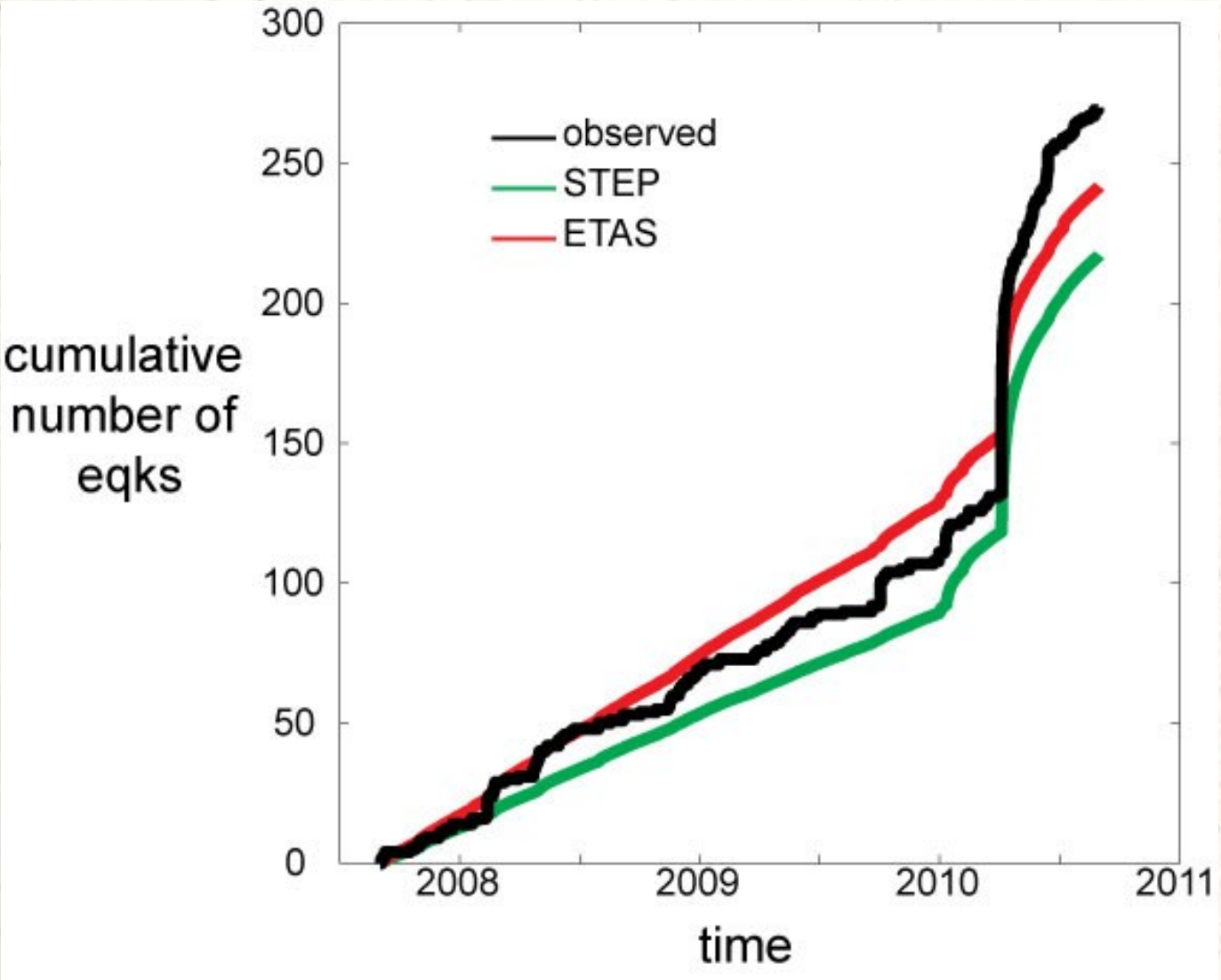
A_i: rate of forecast $\boldsymbol{A}$ in the bin of the i^{th} eqk

$N_{\boldsymbol{A}}$: total rate of forecast $\boldsymbol{A}$

Is I_N significantly different from zero? (Apply Student's paired t-test and Wilcoxon signed rank test to sample gains, resulting tests are called T and W.)

Rhoades et al. 2011

Evaluation – consistency
ETAS & STEP



STEP offers a better fit to individual numbers per day.

Overall, STEP significantly underestimates the total number of earthquakes.

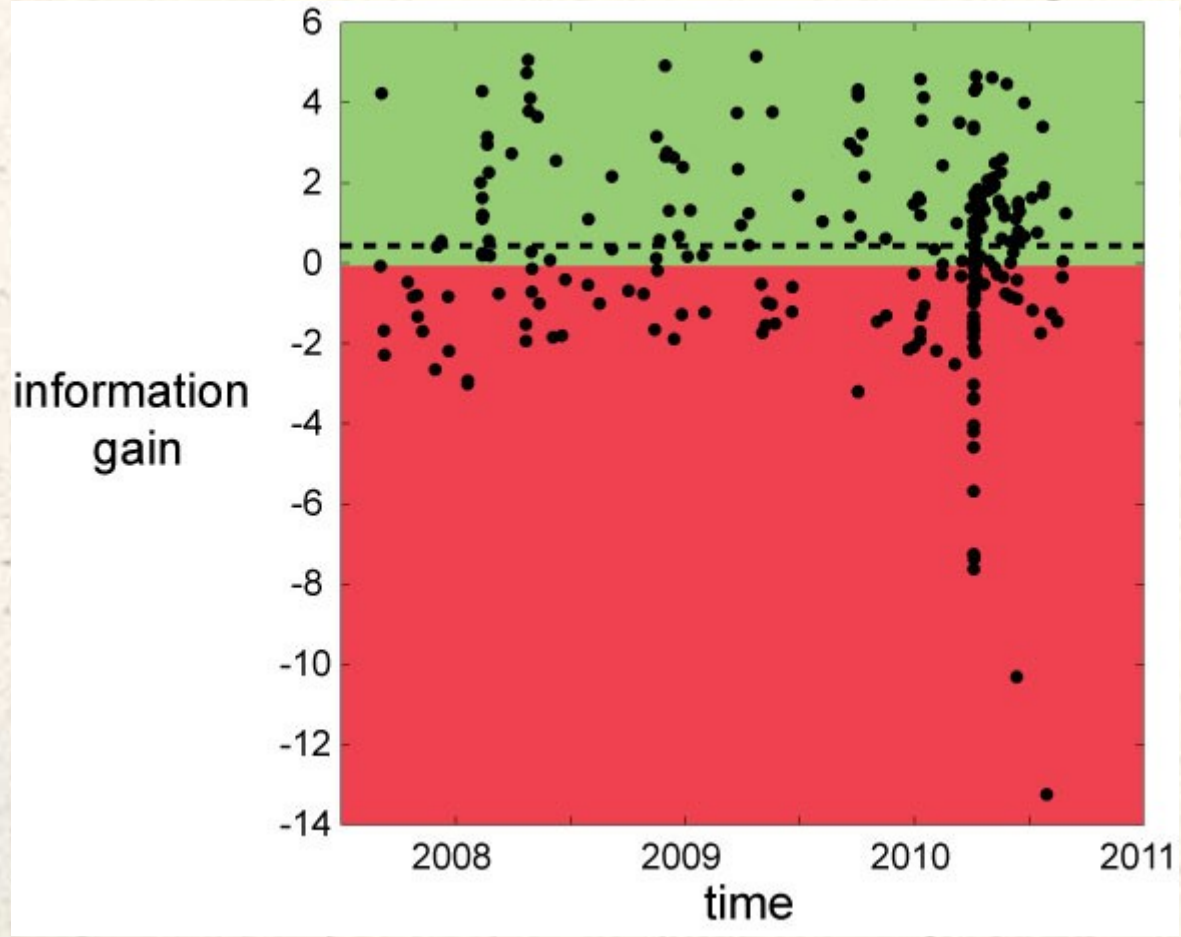
Should we “split” or should we “lump”?

Evaluation – consistency
ETAS & STEP

Evaluation Metric	Model	
	ETAS	STEP
Rate	18/1096 (1.000)	19/1096 (1.000)
Magnitude	5/161 (0.909)	2/161 (0.998)
Space	7/161 (0.699)	21/161 (0.000)
Space-Rate-Magnitude	47/1096 (0.877)	77/1096 (0.002)

- Table indicates frequency of inconsistency
- 95% confidence for each metric for each day

Evaluation – comparison ETAS & STEP



- I_N(STEP, ETAS) = 0.43
- With more than 99% confidence, we can say that STEP was better than ETAS.

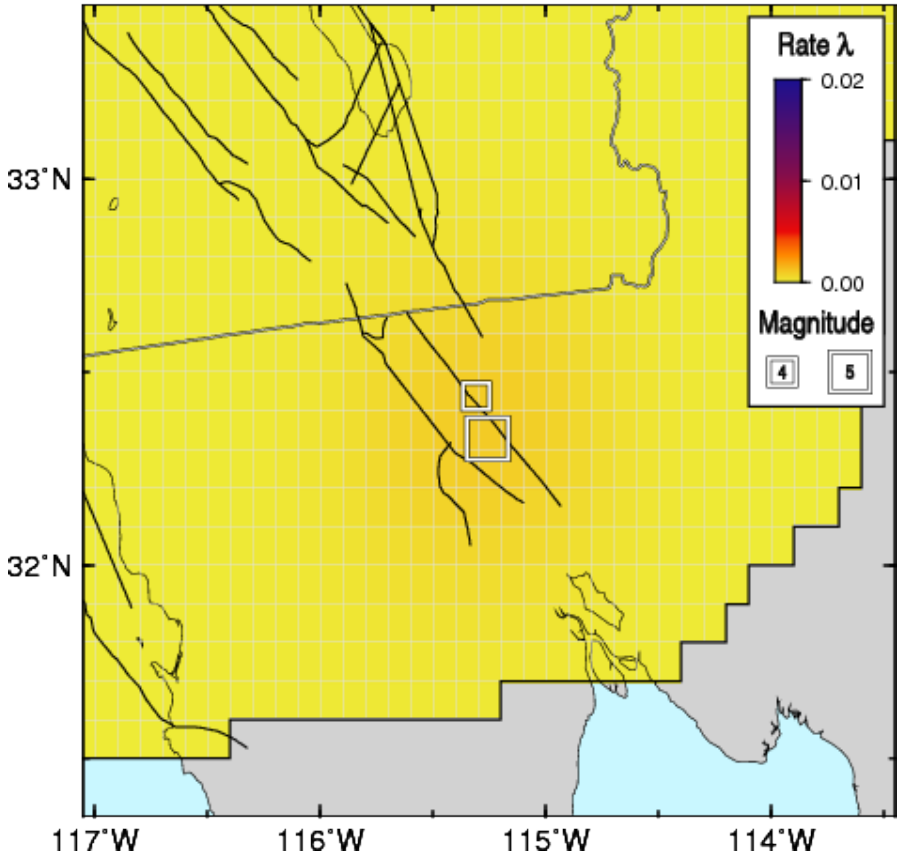
Evaluation – comparison ETAS & STEP

- I_N(STEP, ETAS) = 0.43
- With more than 99% confidence, we can say that STEP was better than ETAS.

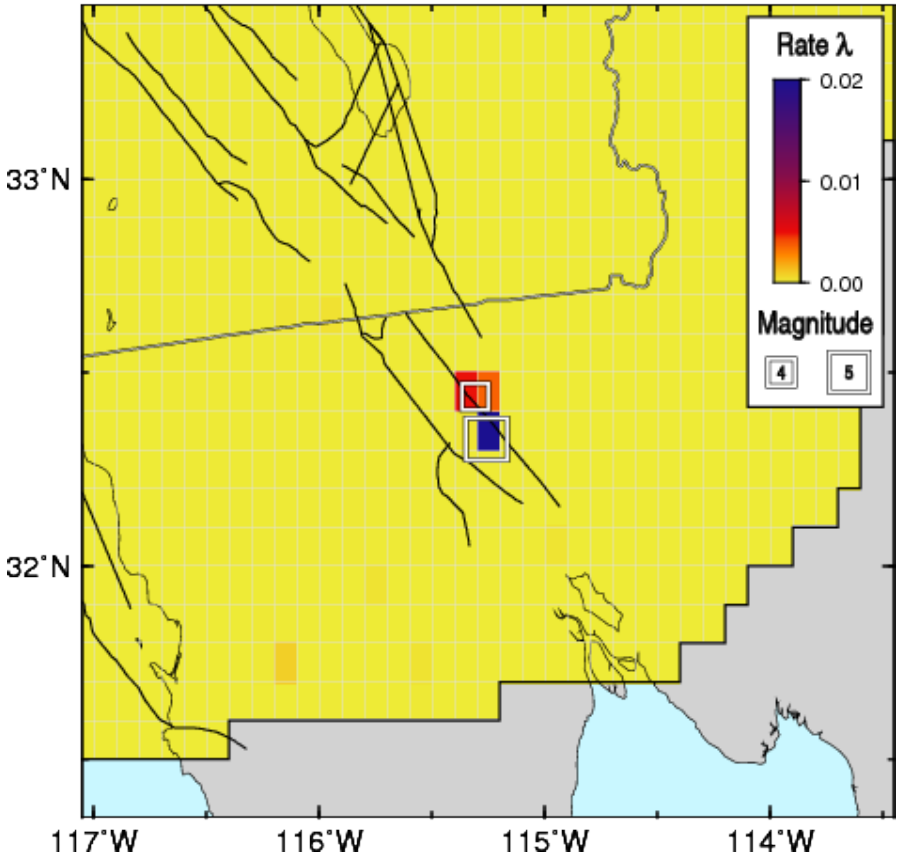
Testing:
Feb. 2008 swarm in Baja California

11 Feb. 2008

ETAS



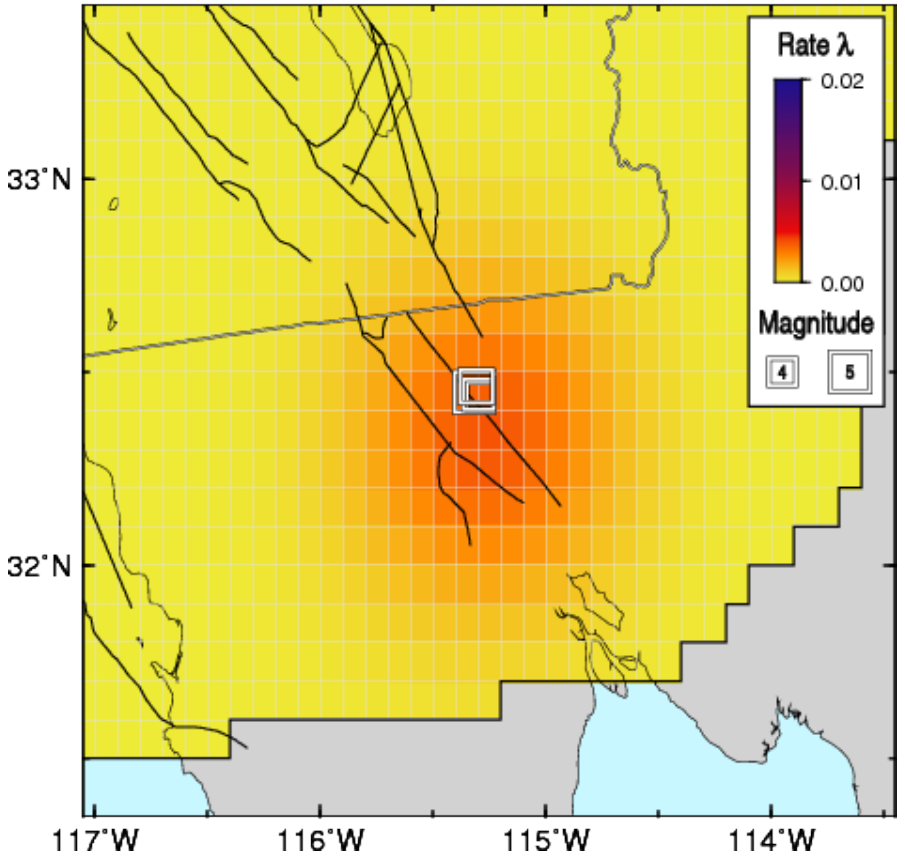
STEP, corrected



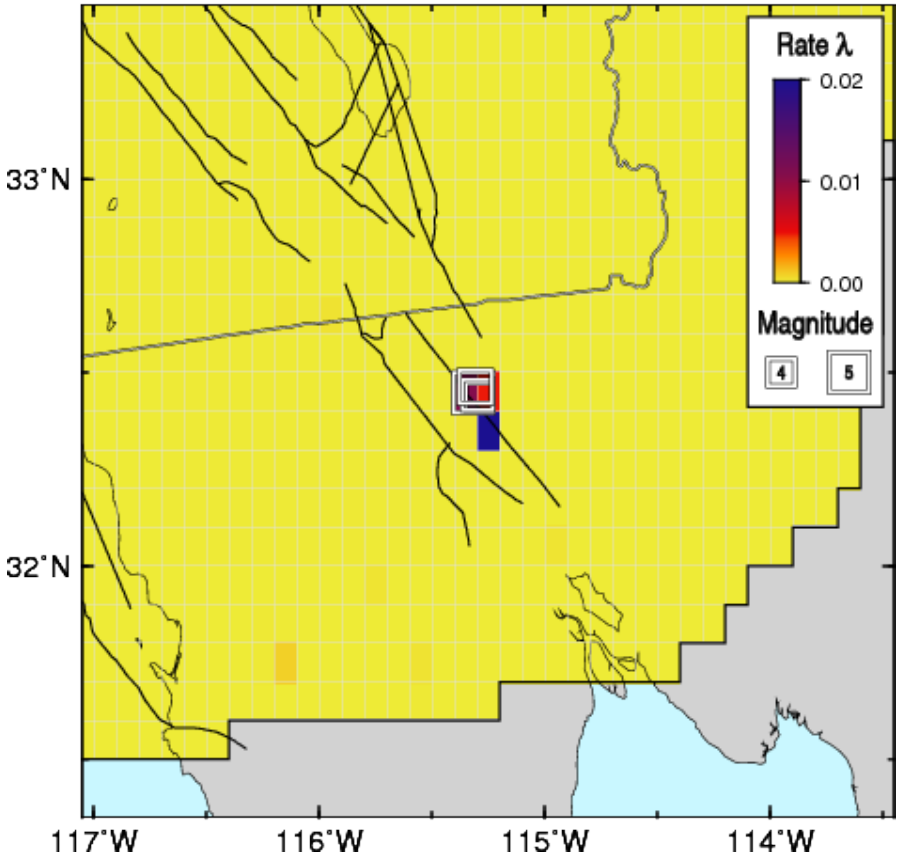
Testing:
Feb. 2008 swarm in Baja California

12 Feb. 2008

ETAS

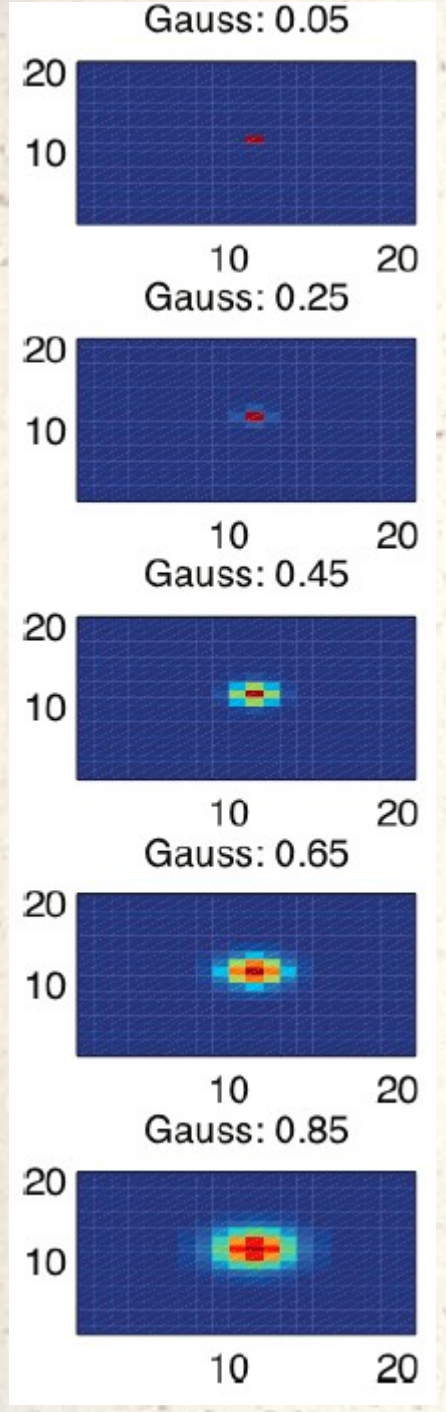


STEP, corrected

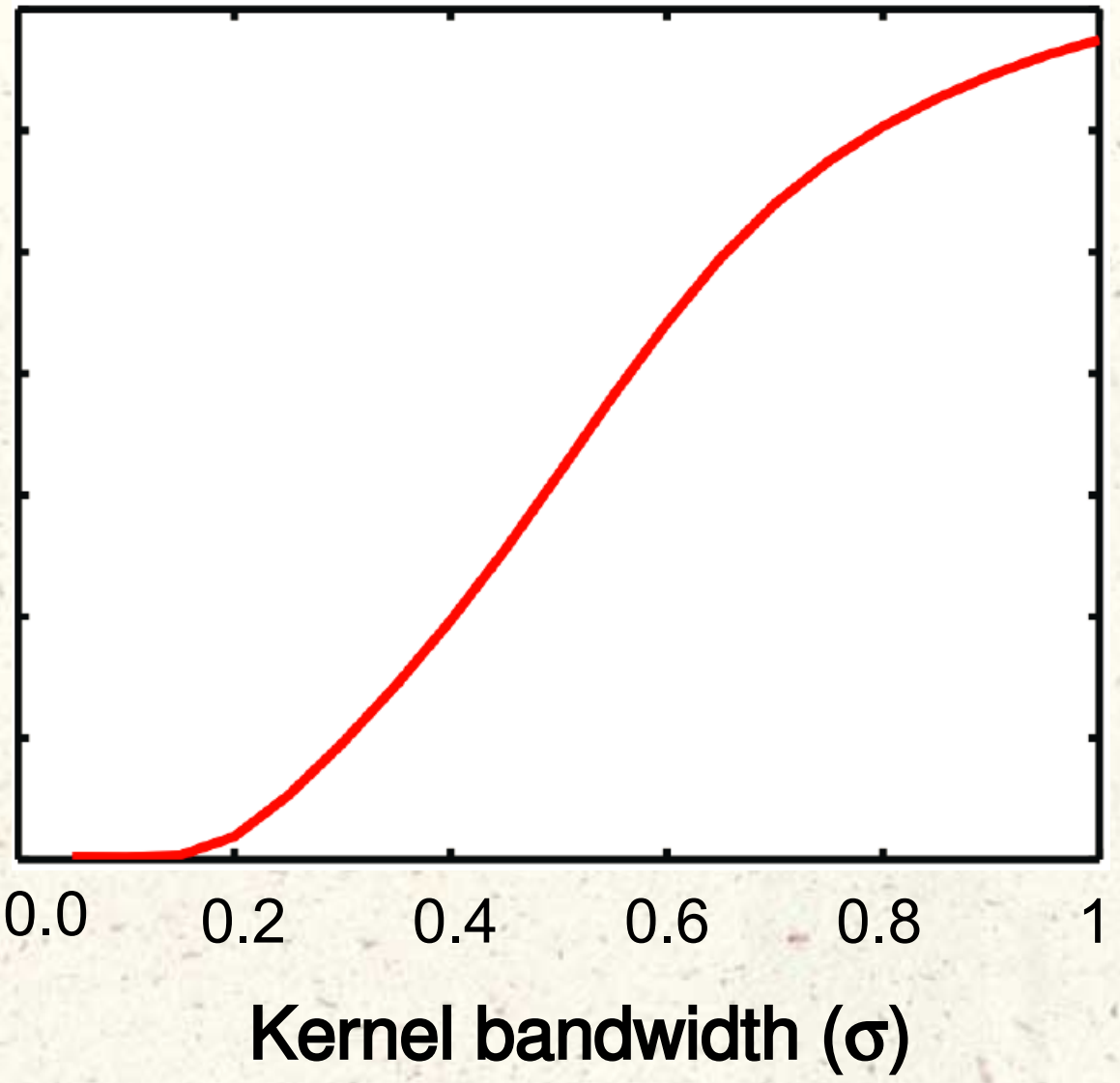


Is ETAS smoother than STEP?

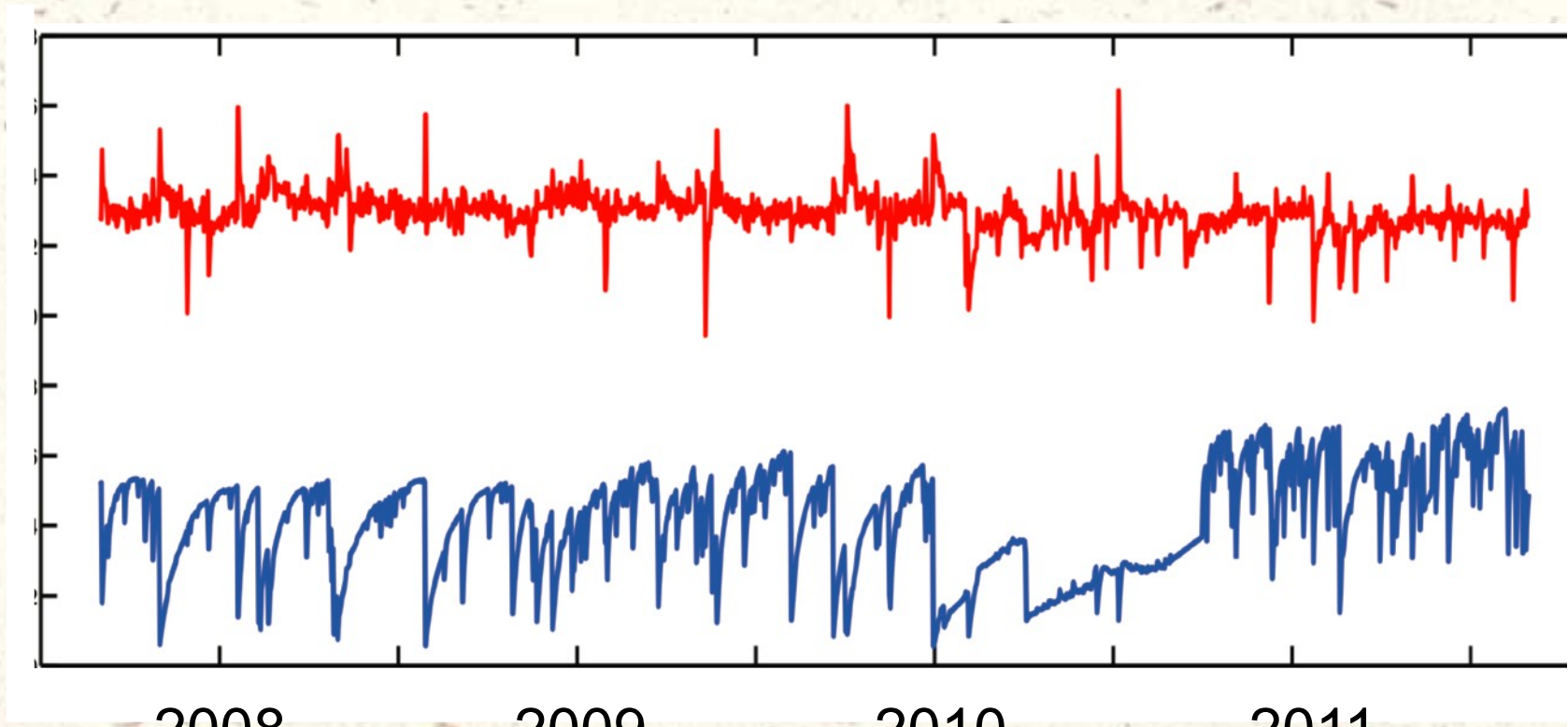
To characterize smoothness, compute 2D Fourier transform and subtract the mean value from the maximum...


Gauss: 0.05
20
10
10
20
Gauss: 0.25
20
10
10
20
Gauss: 0.45
20
10
10
20
Gauss: 0.65
20
10
10
20
Gauss: 0.85
20
10
10
20

smoothness


0.0
0.2
0.4
0.6
0.8
1
Kernel bandwidth (σ)

ETAS is much smoother than STEP



Smoothness is a double-edged sword

- STEP is “better” than ETAS because STEP is less smooth
- STEP fails S-test more frequently than ETAS because STEP is less smooth

Potential next steps

- See if complexities of STEP (sequence-specific parameterization, spatially-varying parameters) provide strong advantage
  - How often does STEP go into these modes?
- Re-analyze with updated data
- More splitting vs. lumping
- Bayesian analysis
- Construct ensemble forecasts