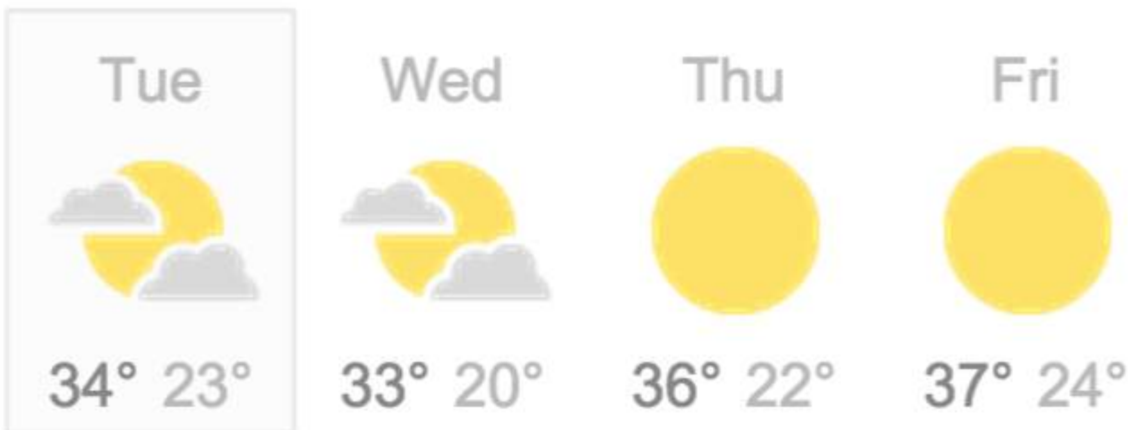
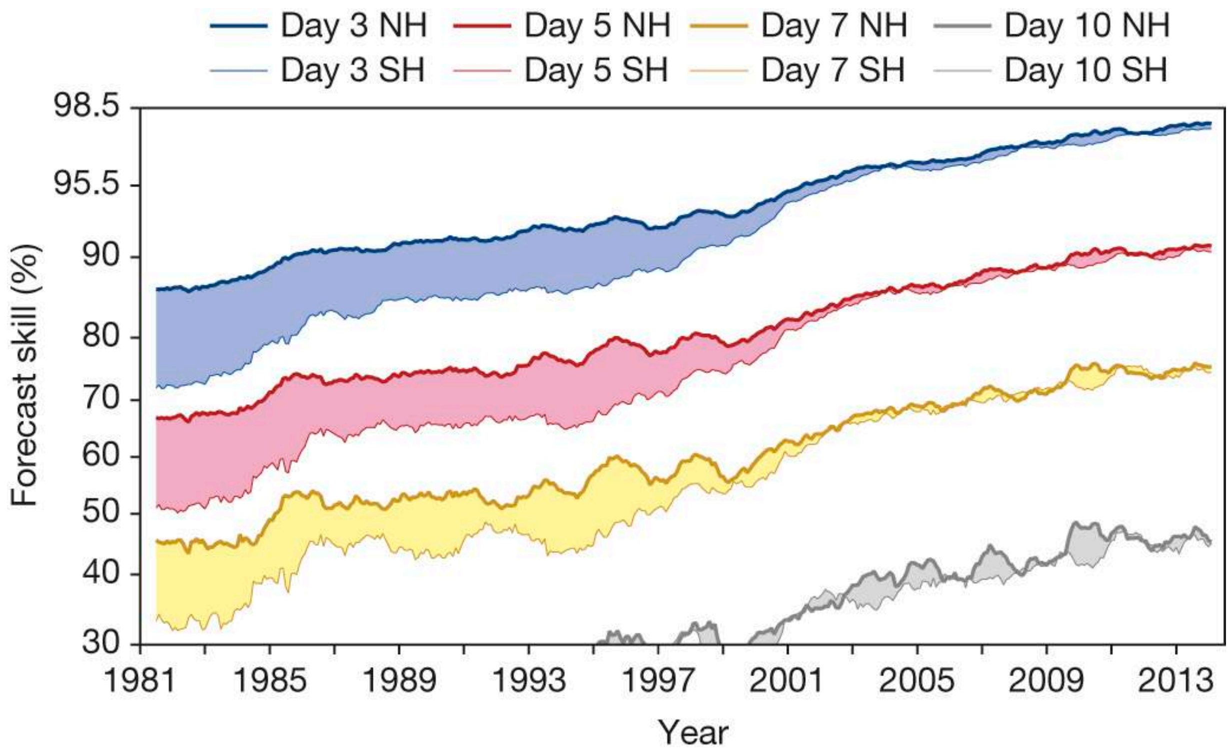

Tue
34° 23°
Wed
33° 20°
Thu
36° 22°
Fri
37° 24°

A measure of forecast skill at three-, five-, seven- and ten-day ranges, computed over the extra-tropical northern and southern hemispheres.



Correlation between the forecasts and the verifying analysis of the height of the 500 hPa level, expressed as the anomaly with respect to climatological height. Values greater than 60% indicate useful forecasts, while those greater than 80% represent a high degree of accuracy.

P Bauer *et al.* "The Quiet Revolution of Numerical Weather Prediction"
Nature **525**, 47-55 (2015) doi:10.1038/nature14956

nature

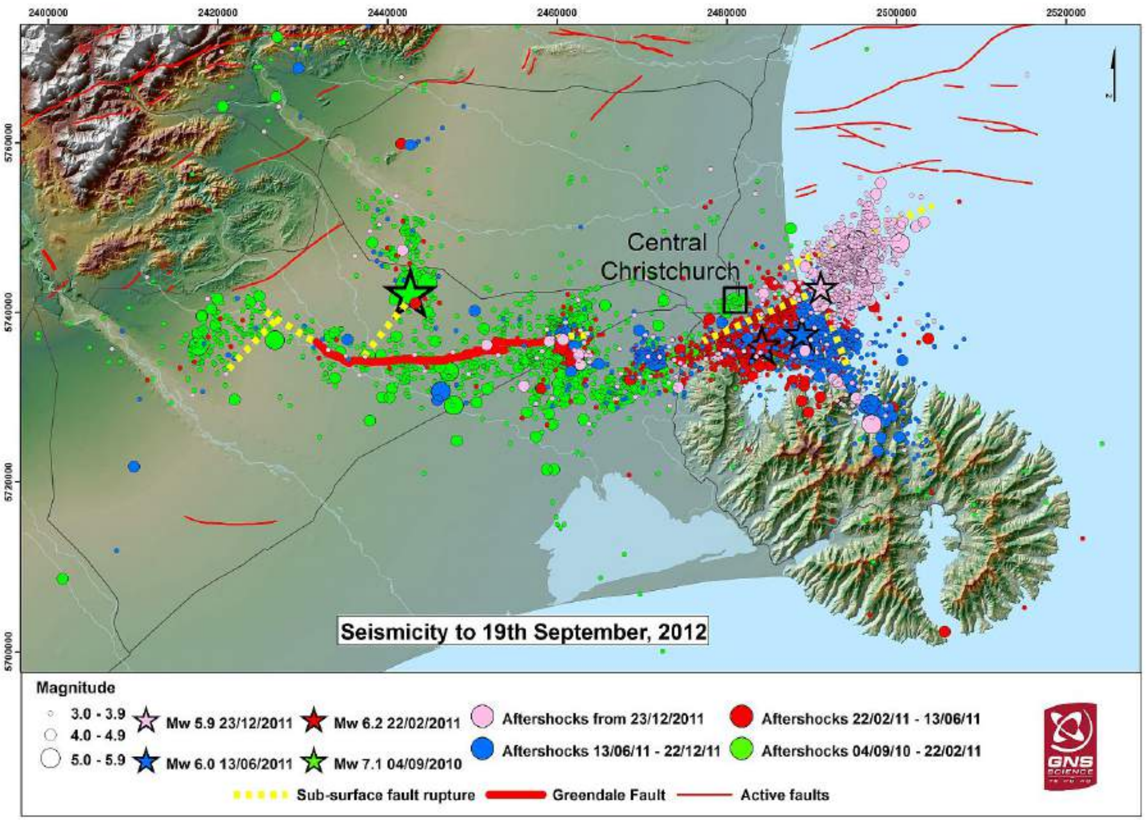
Physics-Based Earthquake Forecasting

Encouraging Results from a Retrospective CSEP Evaluation of Forecasting Models during the 2010 Canterbury Earthquake Sequence

Max Werner, M. Gerstenberger, M. Liukis, W. Marzocchi, D. Rhoades, M. Taroni, J. Zechar, C. Cattania, A. Christophersen, S. Hainzl, A. Helmstetter, A. Jimenez, S. Steacy & T. Jordan

University of Bristol, GNS Science, SCEC/USC, INGV, GFZ Potsdam, ISTerre Grenoble, University of Ulster, University of Adelaide

The Canterbury, NZ, Sequence



Complex:
- M7.1 Darfield (Sep ‘10)
- M6.2 Christchurch (Feb ‘11)
- M6.0 Christchurch (Jun ‘11)
- M5.9 Christchurch (Dec ‘11)

Devastating:
- Over 180 deaths
- $10-15 billion USD

Raised expected hazard:
- Time-dependence
- Gerstenberger et al. (2014), Earthquake Spectra

Wealth of data to study earthquake clustering and predictability.

Retrospective Canterbury Experiment

- Retrospective evaluation of the predictive skills of time-dependent earthquake forecasting models during the sequence

- Goals:
  - Improve our understanding of earthquake clustering
  - Evaluate newly developed physics-based and hybrid empirical/physical against empirical/statistical models
  - Characterize influence of real-time data on forecast quality
  - Help guide model development for Operational Earthquake Forecasting (OEF)

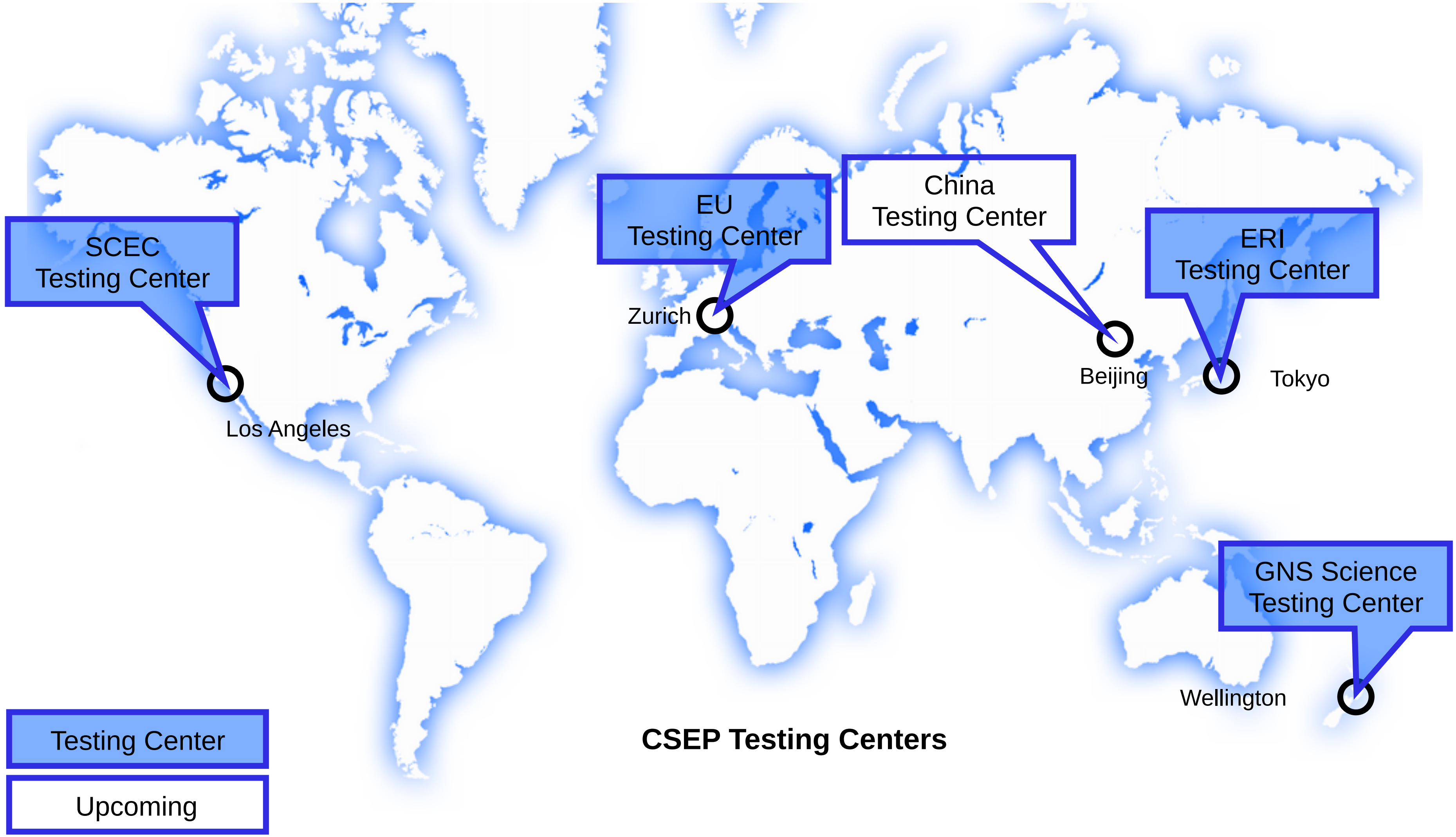
Collaboratory for the Study of Earthquake Predictability

- **CSEP's goal is an infrastructure for conducting earthquake predictability experiments and research. This entails:**
  - **Rigorous procedures for registering forecasting and prediction experiments**
  - **Reproducible evaluations of predictability hypotheses and forecasting models**
  - **Automated, blind, prospective testing in a standardized, controlled environment ("zero degrees of freedom" – the gold standard)**
  - **Select retrospective evaluations of models (an important milestone)**
  - **Community-endorsed standards for assessing forecasts & predictions**
  - **Experiments in a variety of tectonic environments**
- **Why?**
  - **Understand earthquake predictability, brick-by-brick**
  - **Guide model development and improve models**
  - **Reduce controversies surrounding earthquake prediction**
  - **Help government agencies in assessing the utility of earthquake forecasts and predictions in the context of risk reduction.**

Jordan (2006)

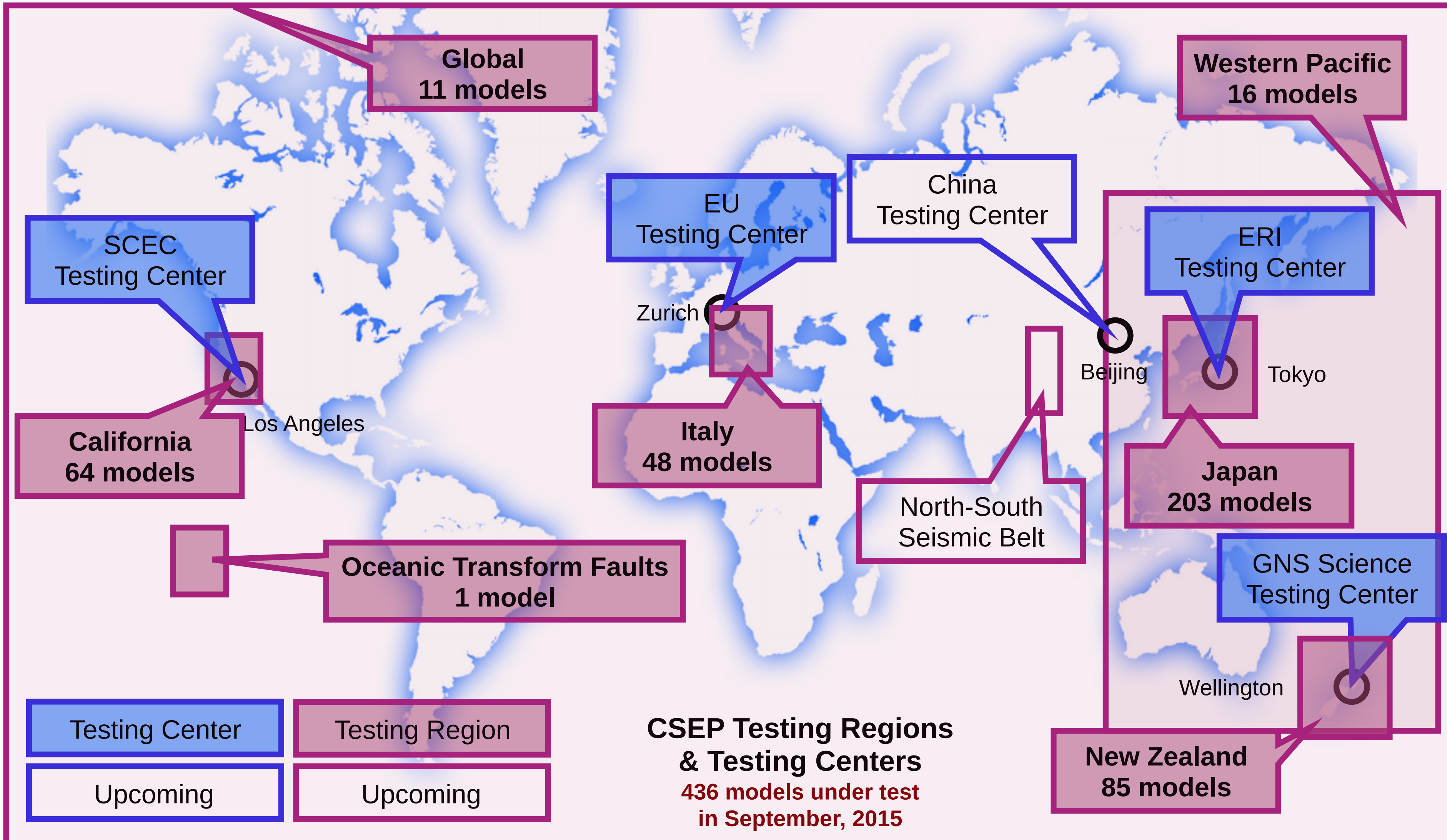
Collaboratory for the Study of Earthquake Predictability

Cyber-infrastructure for blind, prospective and retrospective assessment of forecasting models in a variety of tectonic environments



Collaboratory for the Study of Earthquake Predictability

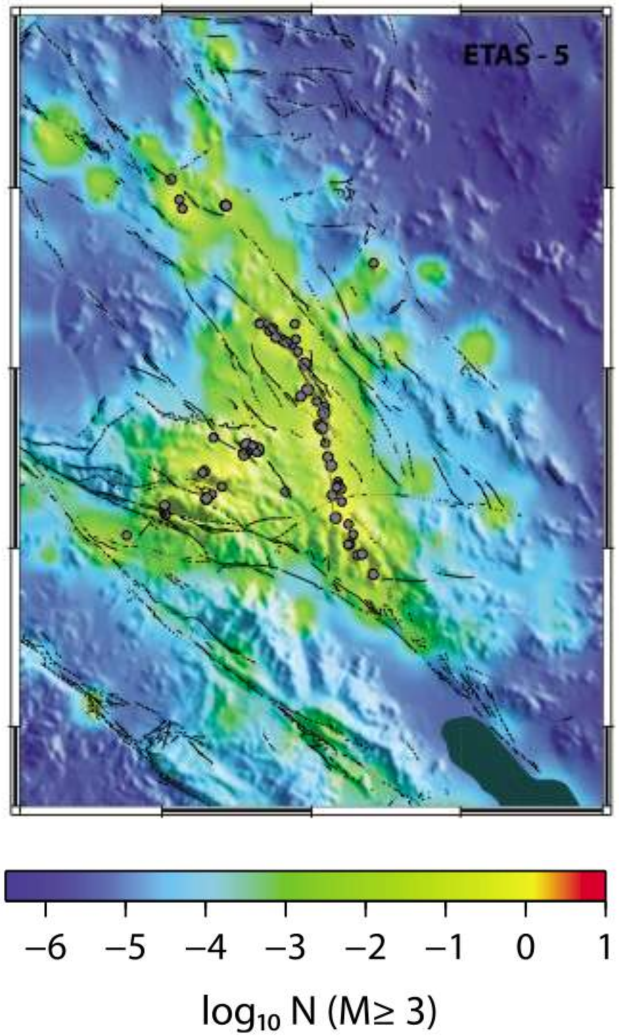
Cyber-infrastructure for blind, prospective and retrospective assessment of forecasting models in a variety of tectonic environments



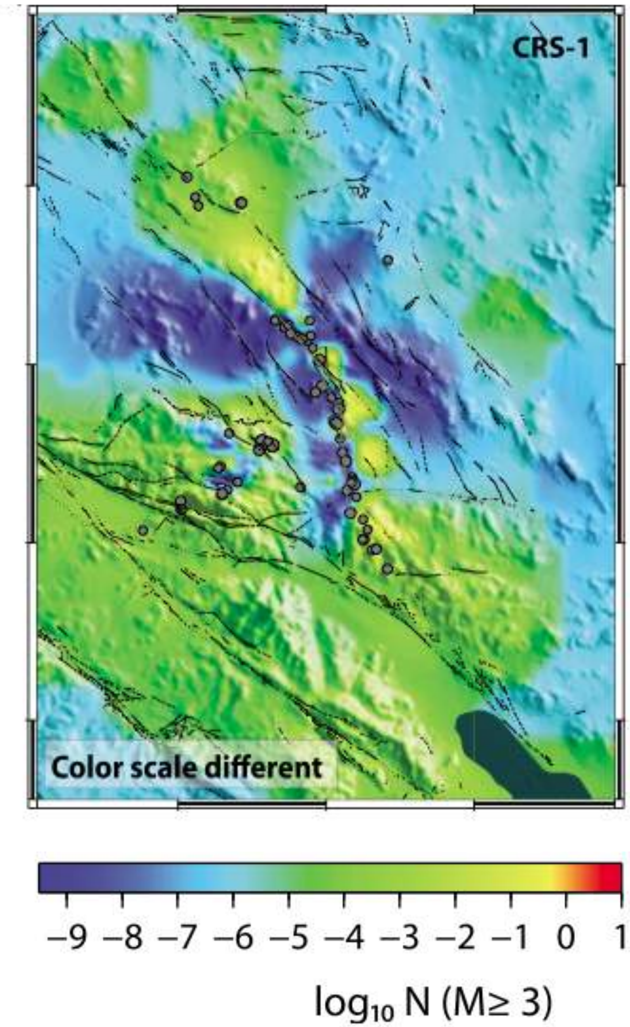
Retrospective Landers Experiment

CSEP evaluation of 1-day forecasts of 12 STEP, ETAS and Coulomb/rate-state models during 90 days after 1992 Landers earthquake [Woessner et al., 2011]

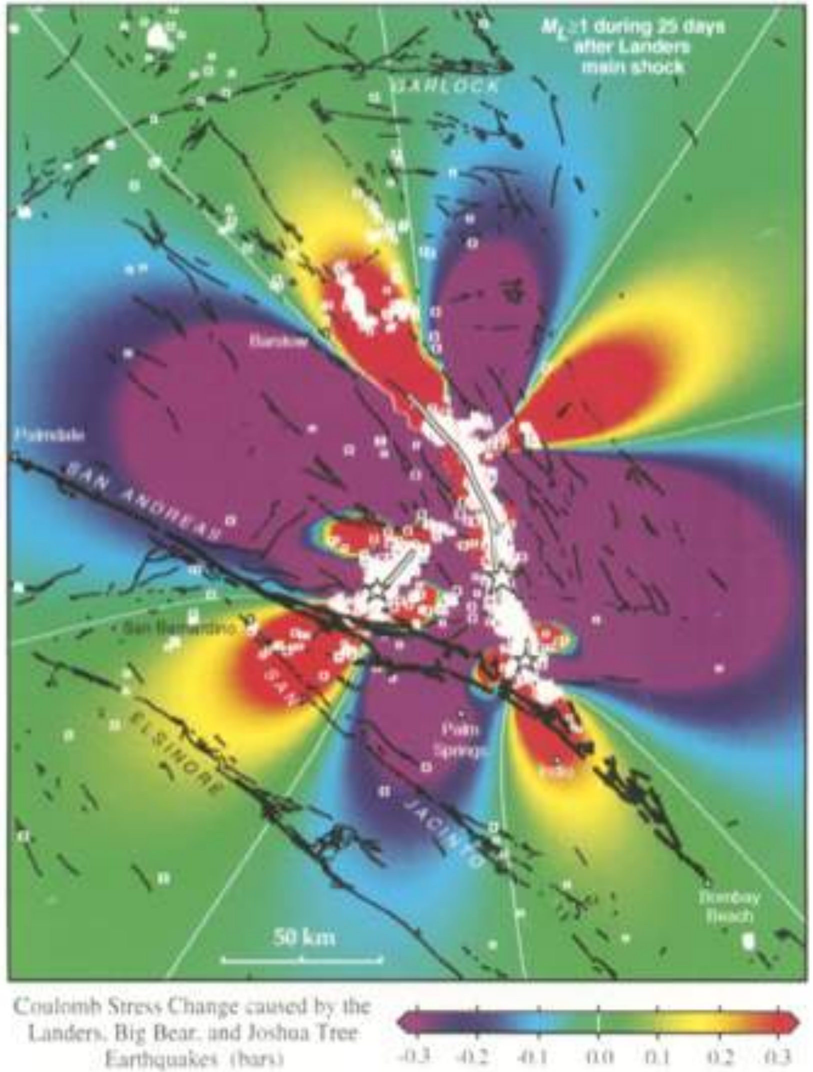
ETAS
(with early-aftershock smoothing)



Coulomb with rate-state friction
(mainshocks-only, parallel receivers)



the textbook example of static Coulomb stress triggering?



King and Cocco (2001)

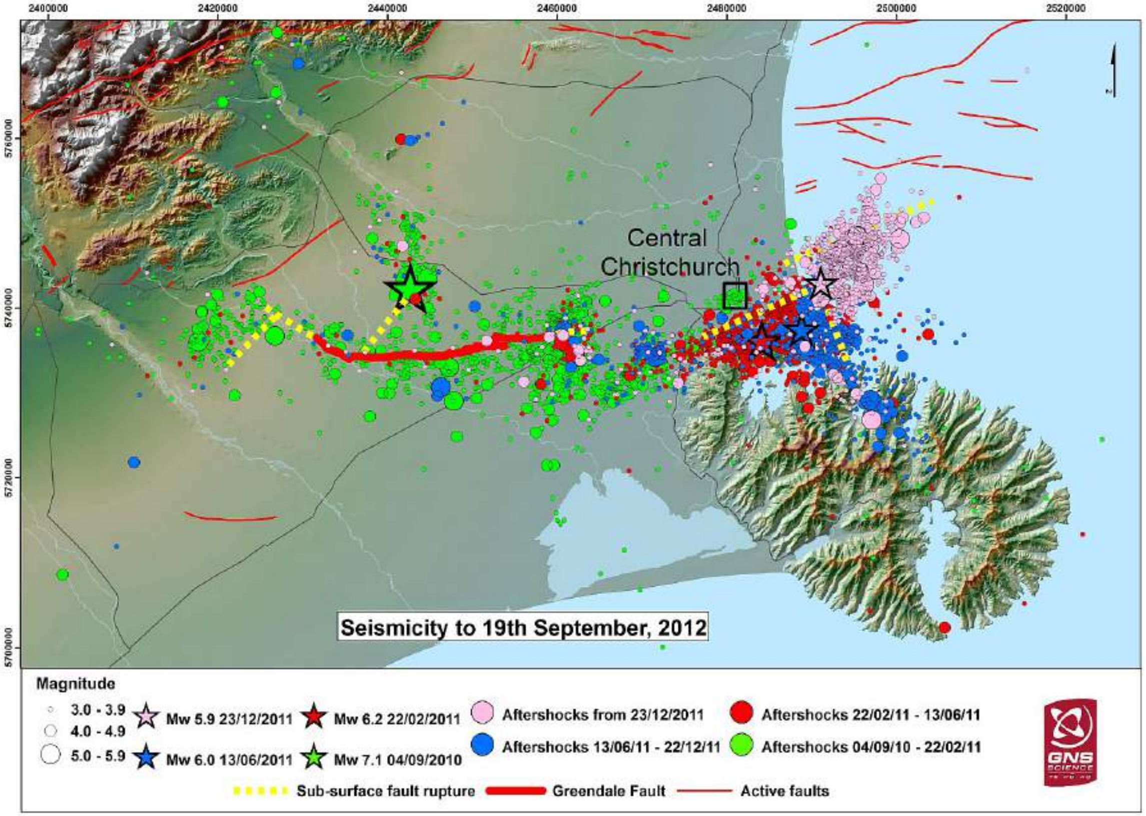
Results of the Retrospective Landers Experiment

Woessner et al. (2011)

Model	LL_S	Gain(S)	Rank
STEP-0	−5187.40	1.00	
STEP-1	−4099.87	3.02	8
ETAS-1	−3160.40	7.86	4
ETAS-2	−3012.83	9.14	3
ETAS-3	−3708.66	4.50	6
ETAS-4	−3308.43	6.76	5
ETAS-5	−2905.26	10.19	1
ETAS-6	−2907.27	10.17	2
CRS-1	-inf	0.00	11
CRS-2	−5351.49	0.85	10
CRS-3	−3932.49	3.58	7
CRS-4	−4298.86	2.47	9

- Both fixed-receiver (CRS-1) and optimally-oriented planes (CRS-2) Coulomb rate-state models predicted (too) severe stress shadows.
- CRS-3 accounted for uncertainty in the slip model, resulting in smoother and better forecasts.
- CRS-4 included poroelastic effects that did not improve forecasts.
- No CRS models provided better forecasts than the ETAS models.

The Canterbury, NZ, Sequence



Complex:

- M7.1 Darfield (Sep ‘10)
- M6.2 Christchurch (Feb ‘11)
- M6.0 Christchurch (Jun ‘11)
- M5.9 Christchurch (Dec ‘11)

Devastating:

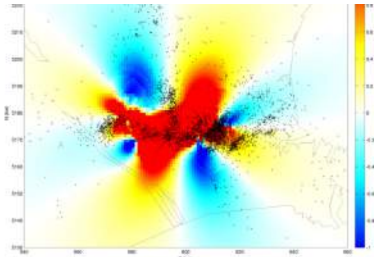
- Over 180 deaths
- $10-15 billion USD

Raised expected hazard:

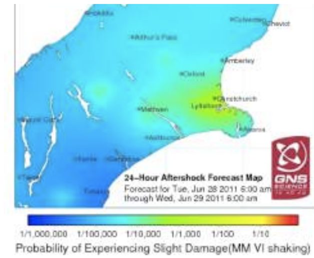
- Time-dependence
- Gerstenberger et al. (2014), Earthquake Spectra

Assess recently developed Coulomb/rate-state models & hybrid models.

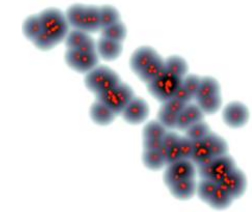
Canterbury Experiment: 15 Models



Coulomb-based



Statistical clustering



Smoothing

Type	Features	Names	Authors
Physical	Coulomb & Dieterich's (1994) rate-state friction	CRS0 (original + stress change uncertainties), CRS1 (+ receiver geometry uncertainties), CRS2 (+ available aftershock FMs) CRS3 (+ all aftershocks as sources) CRS4 (+ spatially-variable stressing rate)	Cattania et al.
Statistical	ETAS	ETAS 0 (epicenter) & ETAS 2 (fault geometry)	Hainzl et al.
	Space-time smoothing	K2 (kernels, with Gutenberg-Richter) K3 (kernels, with non-parametric magnitude)	Helmstetter & Werner
Hybrids	Spatial Coulomb & STEP	STEP-Coulomb	Steacy et al.
	Spatial Coulomb & ETAS	ETAS 1	Hainzl et al.
	ETAS & productivity	RETAS 0 (epicenter) & RETAS 2 (fault)	Hainzl et al.
	ETAS & productivity & Spatial Coulomb	RETAS 1	Hainzl et al.
Reference	Uniform Poisson	SUP	Rhoades

Caveat: *Retrospective*

- Not a fitting exercise!
  - Models (software) installed at CSEP
  - We requested that model parameters and model choices be made prior to forecasting experiment.
- But clearly not a zero-degrees-of-freedom environment …
  - Unconscious biases
  - Sanity checks
- An equal amount of bias between different models?
- All models are probably closer to the upper bound of their predictive skills.

Experiment Design

- Forecasting target:
  - Number of earthquakes
  - per 0.05 by 0.05 degree cell in testing region,
  - per 0.1 magnitude unit bin, above magnitude $M \geq 3.95$
  - per forecast horizon
- Several forecast horizons and evaluation periods
  - Start right *after* M7.1 Darfield earthquake
  - Forecast horizons: 1-yr, 1-mo, 1-day
  - Evaluation periods:
    - Non-overlapping, but …
    - *Re-generate* forecasts *right after* large quakes (Feb 11, June 11, Dec 11)
    - Forecasts with 1-yr horizon also evaluated on entire horizon
- Two "modes"
  - **Retrospective**, using best data available today as input
  - **Pseudo-prospective**, using preliminary data as input

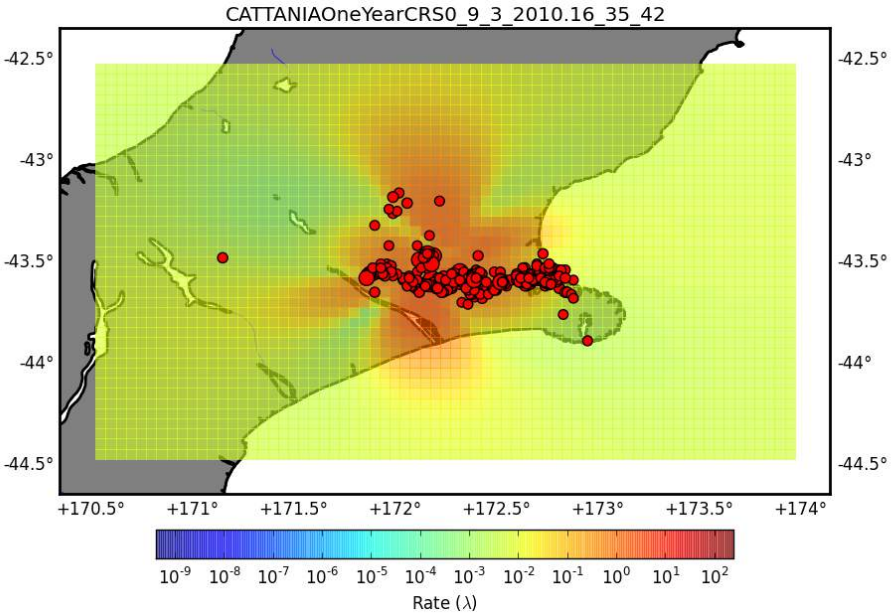
Data Sets

- Target data set to evaluate forecasts:
  - Best available earthquake catalog (394 eqks. M≥3.95)
  - Geonet 4 September 2010 – 1 March 2012

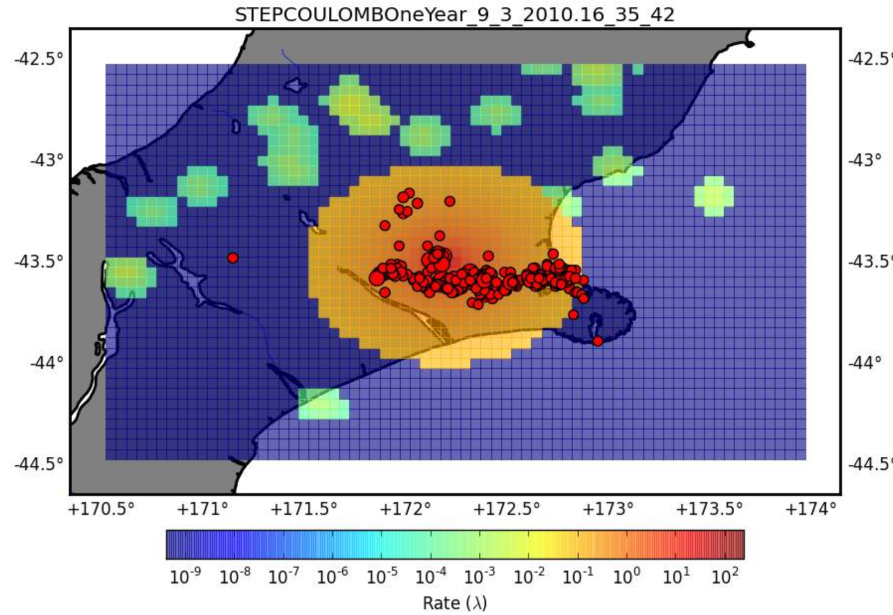
- Input/training data set for model input
  1. Mode 1 (retrospective)
     - Best available earthquake catalog from Geonet
     - Best available slip models (Beaven et al., 2012)
     - Best available focal mechanisms
  2. Mode 2 (pseudo-prospective)
     - Near-real-time earthquake catalog snapshots (captured by NZ CSEP Testing Center)
     - Preliminary slip models obtained within first 10 days (Holden et al., 2011 & personal comm.)

1-yr forecasts after M7.1 Darfield

Coulomb/rate-state model (CRS-0)
Cattania et al.



Hybrid STEP-Coulomb model
Steacy et al.



Testing period:

2010/09/03 – 2011/09/02

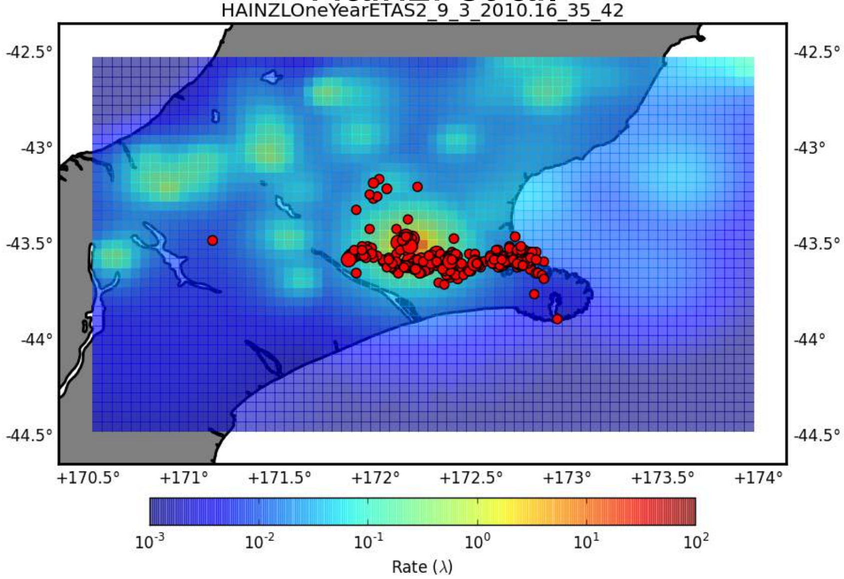
incl. M6.2 Feb and M6 June Christchurch eqks.

Target events:

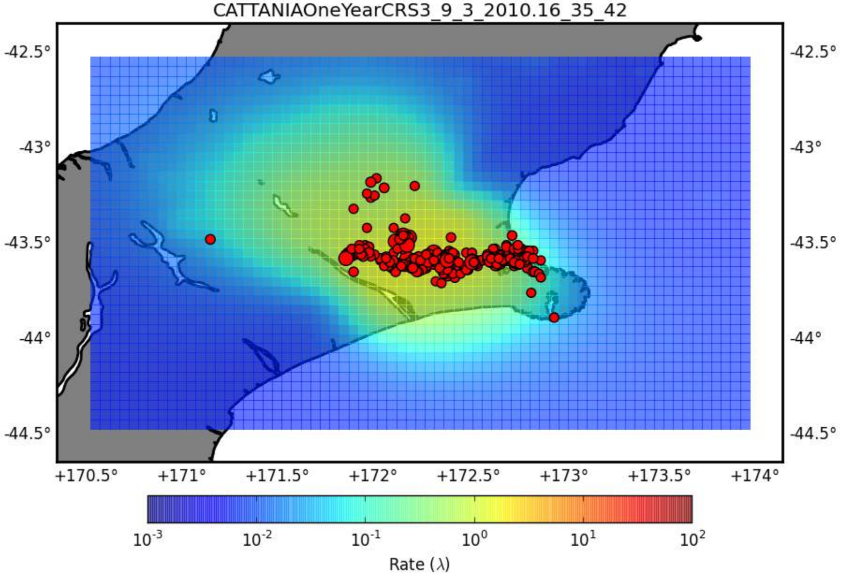
M ≥ 3.95 (328)

1-yr forecasts after M7.1 Darfield

ETAS-0
Hainzl et al.



Coulomb/rate-state model 3
Cattania et al.



Testing period:

2010/09/03 – 2011/09/02

incl. M6.2 Feb and M6 June Christchurch eqks.

Target events:

M ≥ 3.95 (328)

Information Gains: T & W Tests

$$I_N(\mathrm{A,B}) = \frac{1}{N}\sum_{i=1}^{N}\left(\ln\frac{A_i}{B_i}\right) - \frac{N_A - N_B}{N}$$

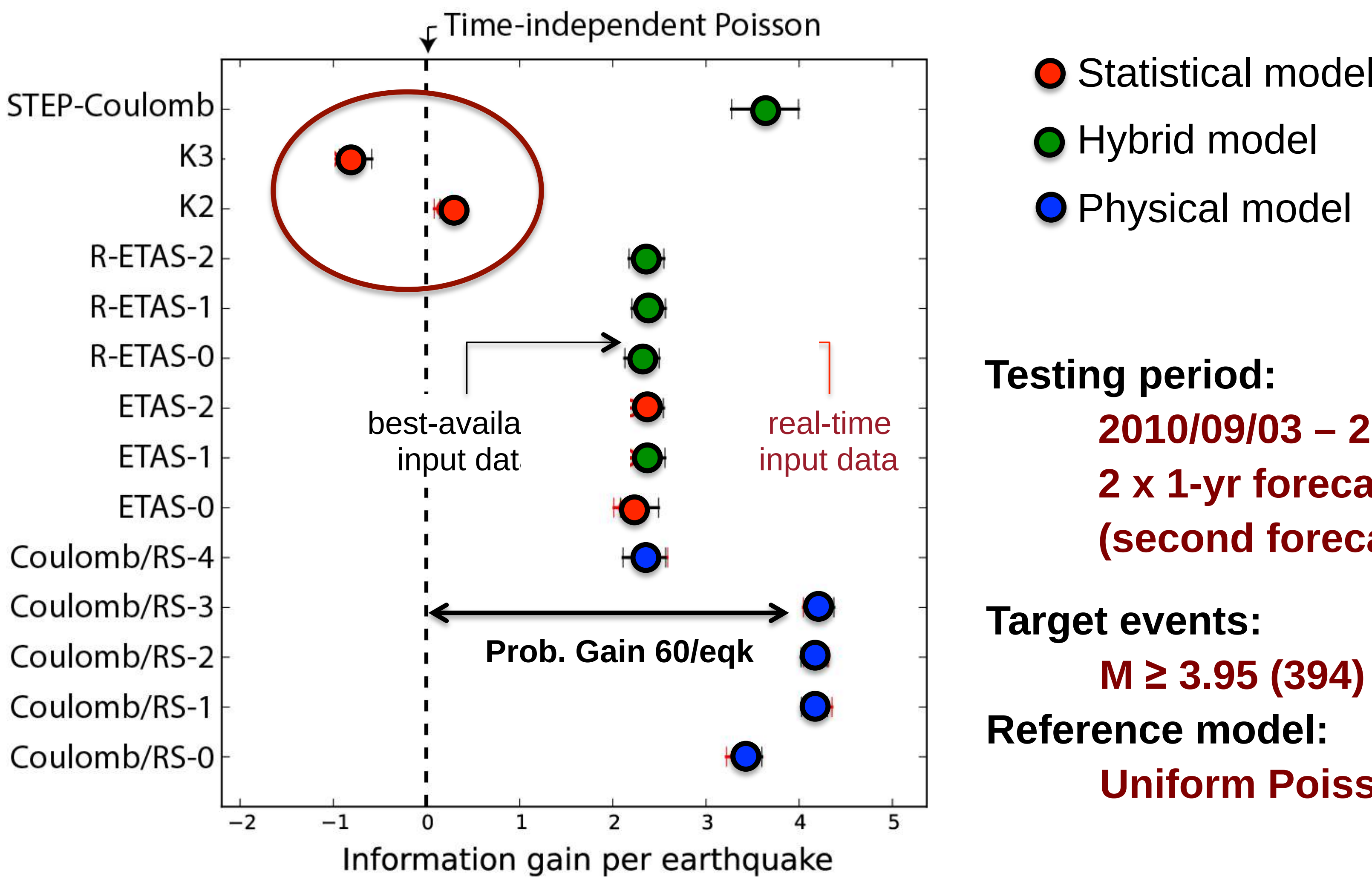
- A measure of predictive skill ("informativeness") of model A over model B

I_N:	rate-corrected information gain
N:	number of observed earthquakes
A, B:	forecasts to compare
A_i	rate of forecast A in bin of eqk i
N_A:	total rate of forecast A

Is I_N significantly different from 0?

- apply Student's paired t-test: T-test
- apply Wilcoxon signed rank test: W-test

Results: T/W tests 1-yr forecasts



Testing period:

2010/09/03 – 2012/02/29

2 x 1-yr forecasts

(second forecast scaled)

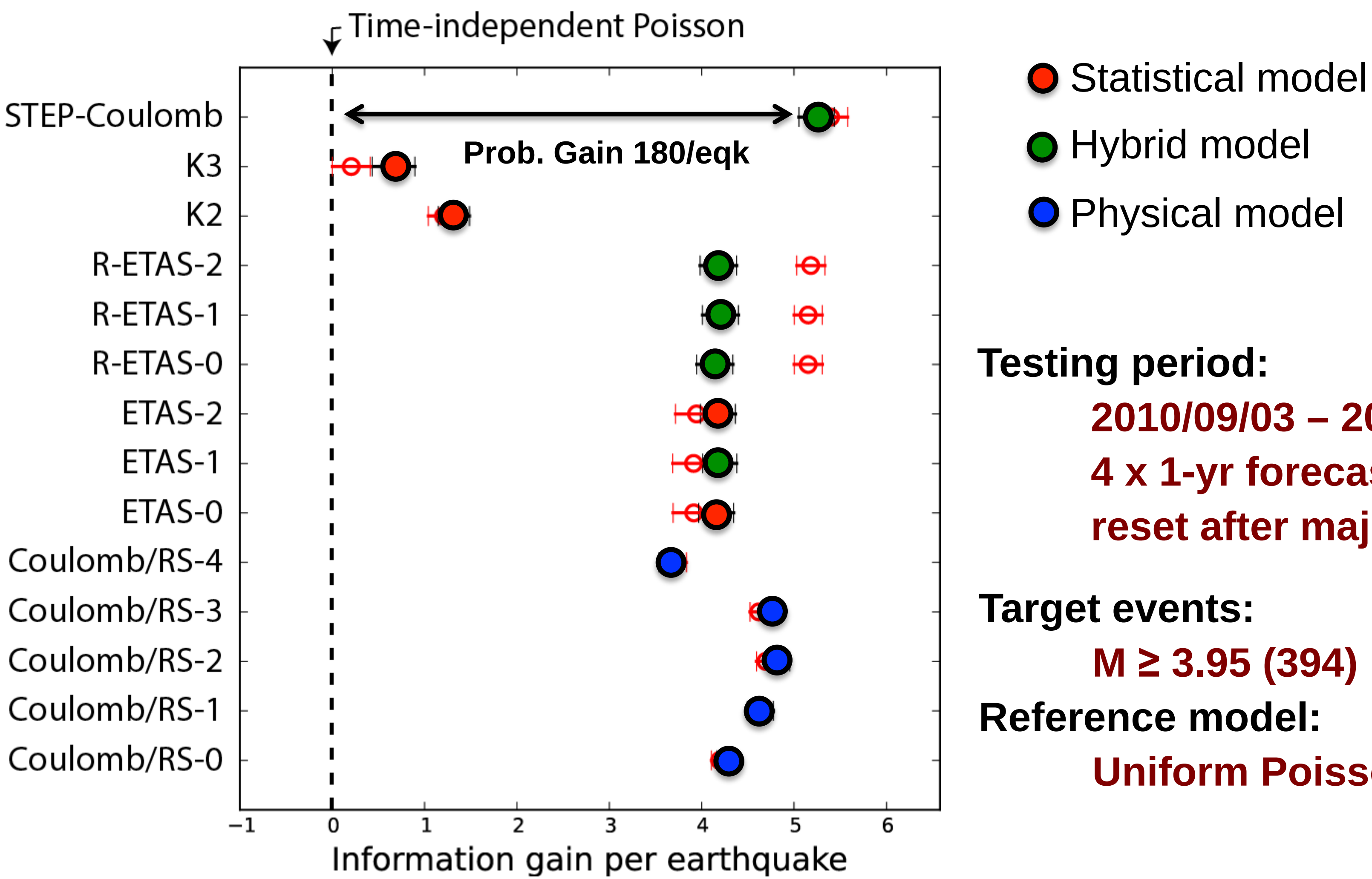
Target events:

M ≥ 3.95 (394)

Reference model:

Uniform Poisson

Results: T/W tests 1-yr forecasts



Testing period:

2010/09/03 – 2012/02/29

4 x 1-yr forecasts

reset after major eqks.

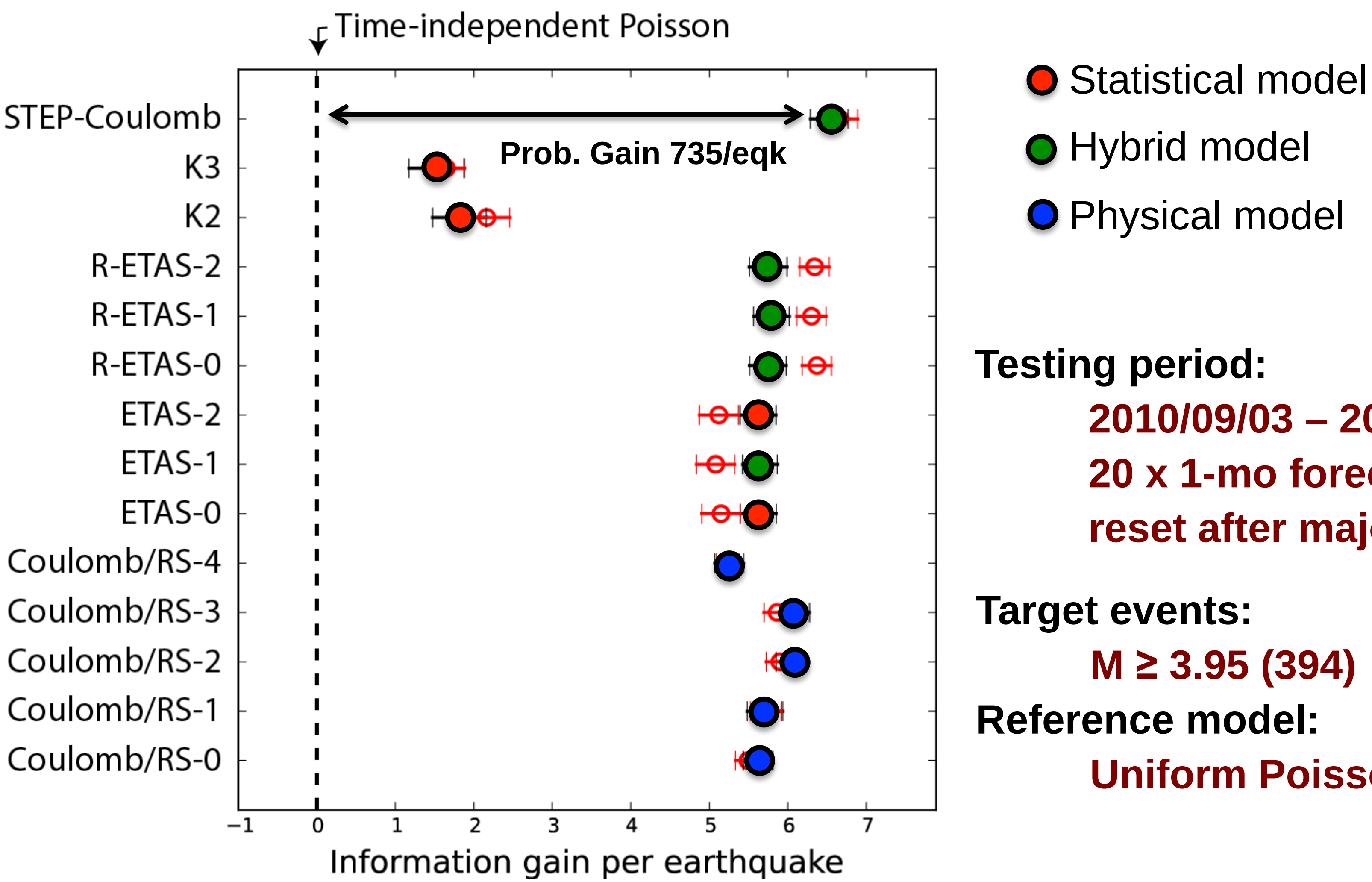
Target events:

M ≥ 3.95 (394)

Reference model:

Uniform Poisson

Results: T/W tests 1-mo forecasts



Testing period:

2010/09/03 – 2012/02/29

20 x 1-mo forecasts

reset after major eqks.

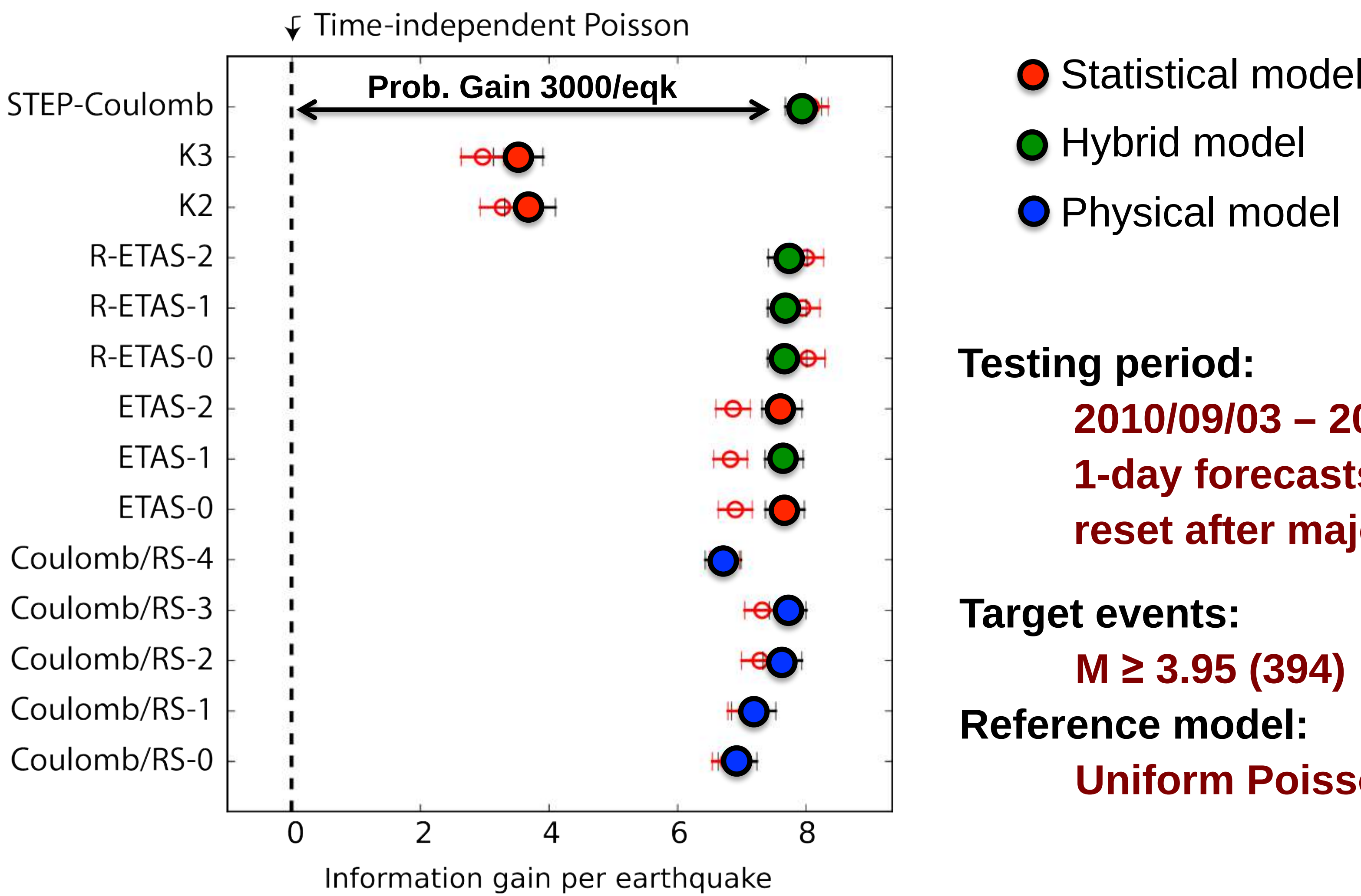
Target events:

M ≥ 3.95 (394)

Reference model:

Uniform Poisson

Results: T/W tests 1-day forecasts



Testing period:

2010/09/03 – 2012/02/29

1-day forecasts

reset after major eqks.

Target events:

M ≥ 3.95 (394)

Reference model:

Uniform Poisson

Possible Interpretations of the Performance of Coulomb/RS models

University of BRISTOL

- **Random luck?**
  - **Earthquake sequences span wide range of patterns**
  - **If random fluctuations, even odds for empirical & Coulomb models (not bad!).**
  - **Collect further data points!**
- **Improved (over-) fitting?**
  - **Philosophy: forecasts can only use/fit data up to the time when forecasts are issued.**
  - **Numbers of parameters b/w models *about* equal (a handful).**
  - **Port models to prospective CSEP experiments and test with zero degrees of freedom!**
- **Improved model components/choices?**
  - **All CRS models propagate uncertainties in slip to stress & seismicity rates.**
  - **CRS1 additionally uses uncertainties in receiver fault geometry.**
  - **CRS2/3 both use aftershocks as stress sources in addition to the large M6 mainshocks**
- **Does Coulomb/rate-state mechanism apply more to Canterbury?**
  - **1992 Landers earthquake is a textbook example of the Coulomb stress change hypothesis.**
  - **But CSEP comparison showed poor performance of predictive skill.**
  - **Lots of additional triggering mechanisms reported for Landers. Less so for Canterbury?**

University of BRISTOL

Implications for OEF

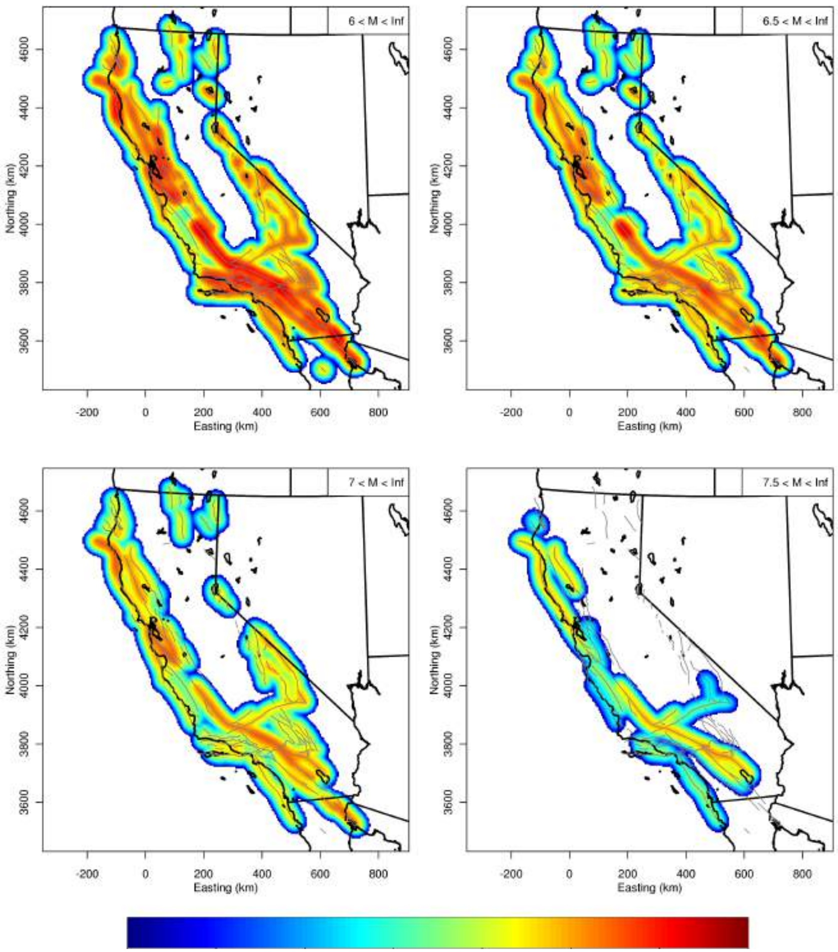
- Time-dependent OEF candidate models can reach probability gains per earthquake of up to 3,000 per eqk over a time-independent model. Absolute probabilities (per week, per month) remain low, but may still be useful.

- The effect of real-time (incomplete) data as model input for OEF candidate models is model-dependent. Real-time data *tends* to degrade forecast quality.

- In my view, these CSEP results are as unbiased as retrospective experiments come, and they suggest that recently developed physics-based and hybrid time-dependent forecasting models could contribute meaningfully to skillful OEF systems.

- We need more retrospective experiments (to accumulate results fast) and prospective experiments (to make sure we're not fooled by randomness and retrospective bias) of physics-based forecast models.

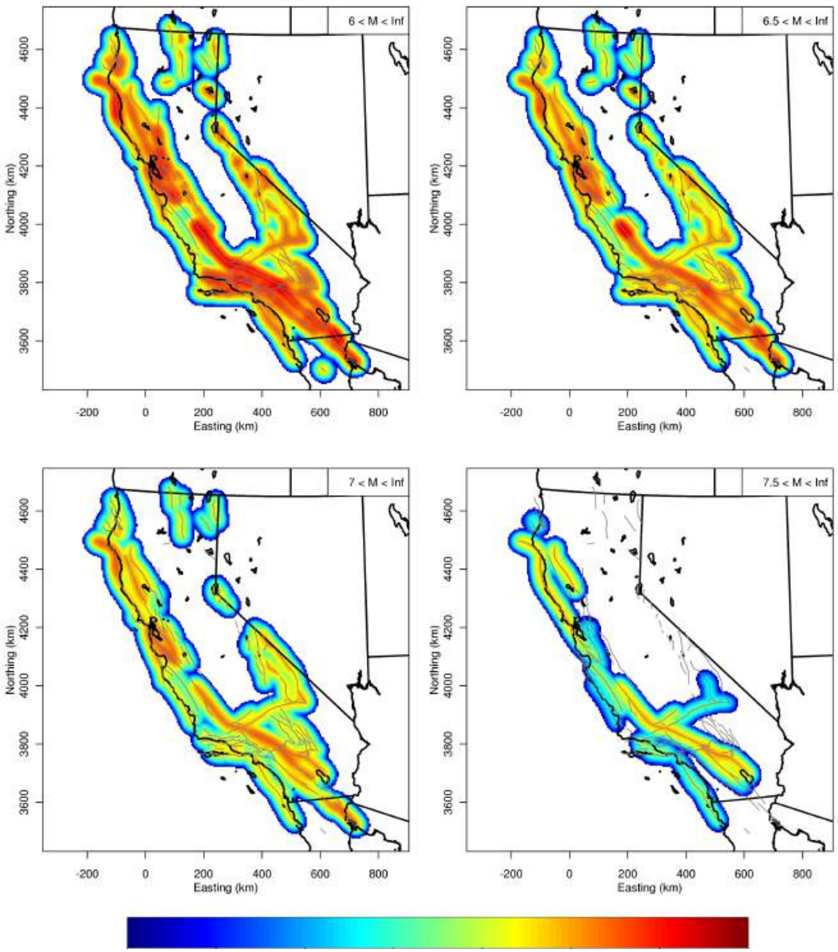
Earthquake Rupture Simulators



Tullis et al., SRL special issue, 2012

- Opportunities
  - Building physical understanding
  - Identifying predictions that could be sought for in data
  - Greater (anticipated) predictive skill than empirical models
  - Contributing to the SCEC system-science master model
  - Physics-based PSHA

Earthquake Rupture Simulators



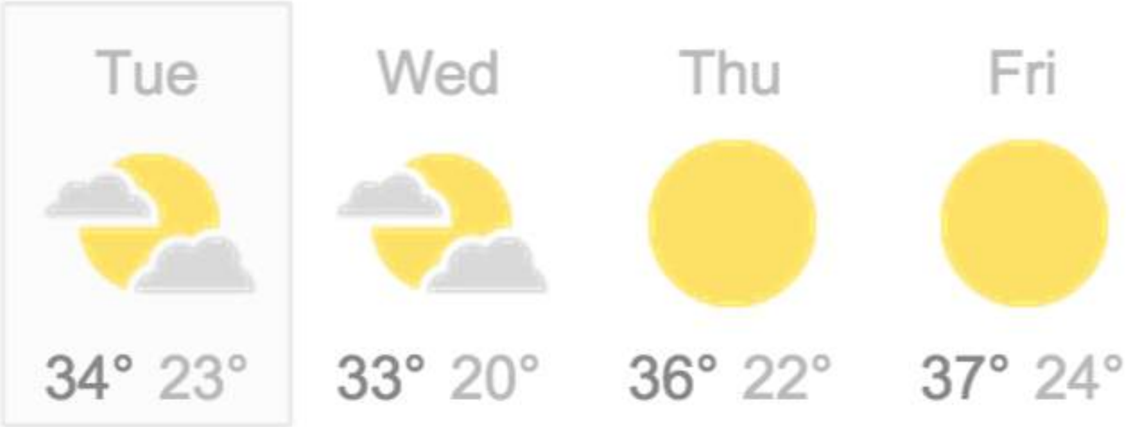
Tullis et al., SRL special issue, 2012

- Challenges
  - Approximating the (inferred) physics adequately
  - Understanding the importance of off-fault dynamics on on-fault earthquakes
  - Calibrating simulators to data
  - What is the *theoretical* predictive skill?
  - Testability of simulator forecasts

Conclusions

- Some physics-based earthquake forecasting models have reached an important milestone:
  - Results from the retrospective Canterbury experiment show that recently developed models can improve on forecasts of empirical/statistical models.
- The next important milestone will be to demonstrate that potential in prospective forecasting experiments.

- Physics-based earthquake rupture simulators could help us identify new questions to ask of data and new targets for improving predictive skill.


Tue
Wed
Thu
Fri
34° 23°
33° 20°
36° 22°
37° 24°