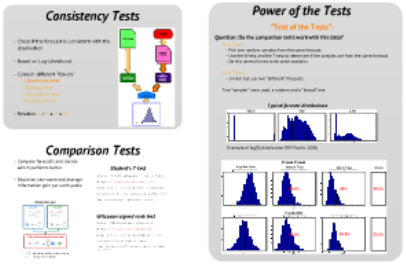
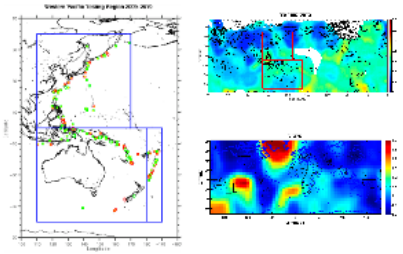
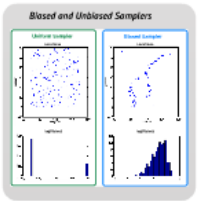
Details of the Test



Bonus Tracks



Bonus Tracks



Unfinished Business

CSEP West Pacific

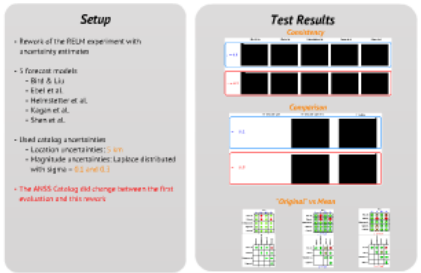
- 2 years of data are not enough, addition 1 year periods will be added
- 3 models are a start but not enough
- Magnitudes bins are missing and successor test regions like the global test region should consider them.

Accounting for Uncertainties

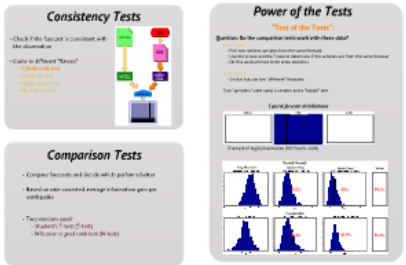
- The effect on the test outcome is dependent on the forecast and the location of the earthquakes
- Can help to evaluate the robustness of forecast for a certain test region and time period
- Uncertainty estimation is a valuable addition to the test procedure

West Pacific Testing Region

Uncertainties in RELM

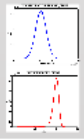


Testing of Forecasts



Uncertainty estimation

- Symmetric uncertainties in M0 - asymmetric uncertainties in Mw
- Because of Gutenberg-Richter Law there are more earthquakes below the magnitude threshold
- BUT: no higher probability for more earthquake in the test regions was observed



Motivation:

- Estimation of the uncertainties of the tests
- Measure the stability of the forecast models

Accounting for catalog uncertainties: Western Pacific and RELM

David Eberhard

SED ETH Zurich

Accounting for catalog uncertainties: Western Pacific and RELM

David Eberhard

SED ETH Zurich

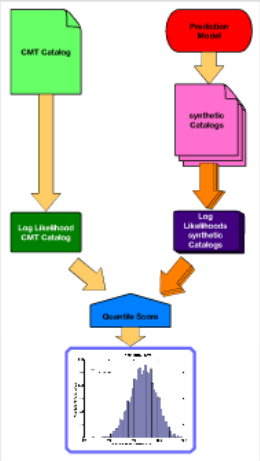
Motivation:

- Estimation of the uncertainties of the tests
- Measure the stability of the forecast models

Testing of Forecasts

Consistency Tests

- Check if the forecast is consistent with the observation
- Come in different "flavors"
  - L(ikelhood)-test
  - S(patial)-test
  - M(agnitude-test
  - N(umber)-test



Comparison Tests

- Compare forecasts and decide which performs better
- Based on rate-corrected average information gain per earthquake
- Two versions used:
  - Student's T-test (T-test)
  - Wilcoxon signed rank test (W-test)

Power of the Tests

"Test of the Tests":

Question: Do the comparsion tests work with these data?

Error Type I

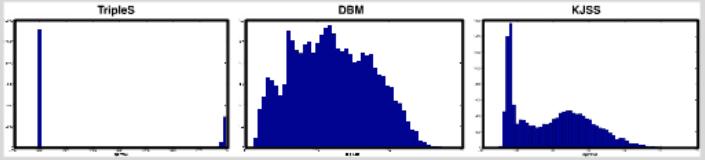
- Pick two random samples from the same forecast.
- Use the W-test and the T-test to determine if the samples are from the same forecast
- Do this several times to do some statistics

Error Type II

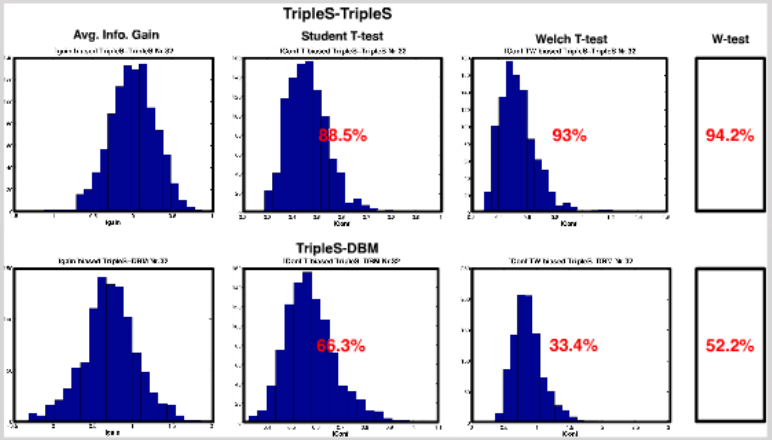
- Similar but use two "different" forecasts

Two "samplers" were used, a random and a "biased" one

Typical forecast distributions



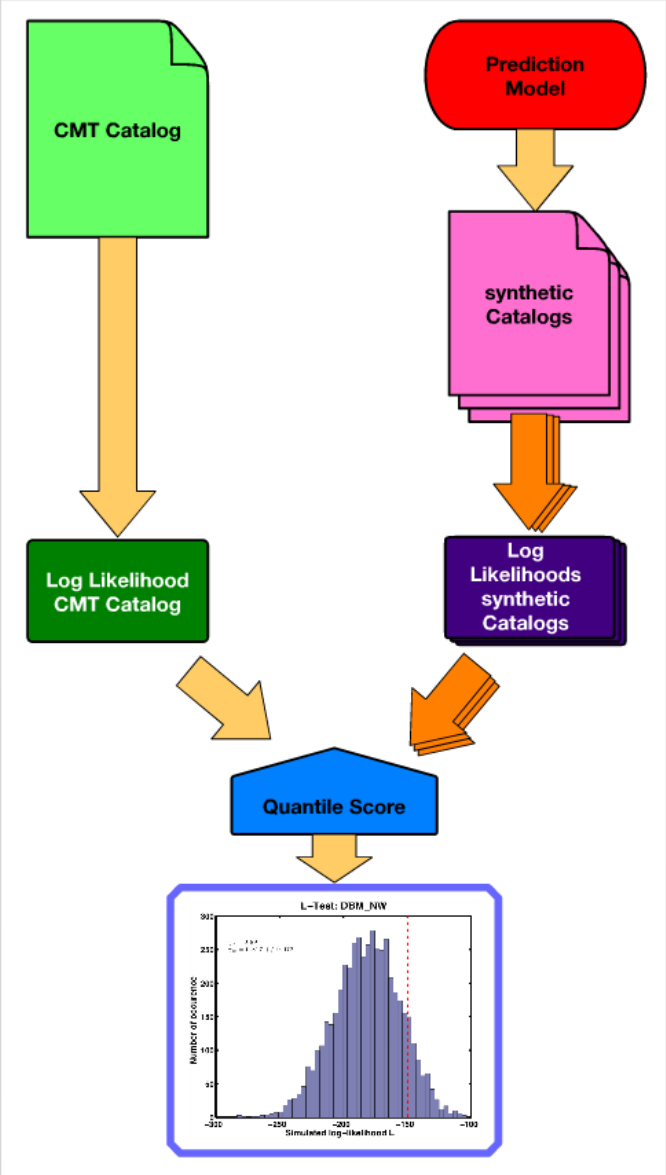
Example of log(A) distribution (NW Pacific 2009)



Consistency Tests

- Check if the forecast is consistent with the observation

- Come in different "flavors"
  - L(ikelhood)-test
  - S(patial)-test
  - M(agnitude-test
  - N(umber)-test



Comparison Tests

- Compare forecasts and decide which performs better
- Based on rate-corrected average information gain per earthquake
- Two versions used:
  - Student's T-test (T-test)
  - Wilcoxon signed rank test (W-test)

Power of the Tests

"Test of the Tests":

Question: Do the comparsion tests work with these data?

Error Type I

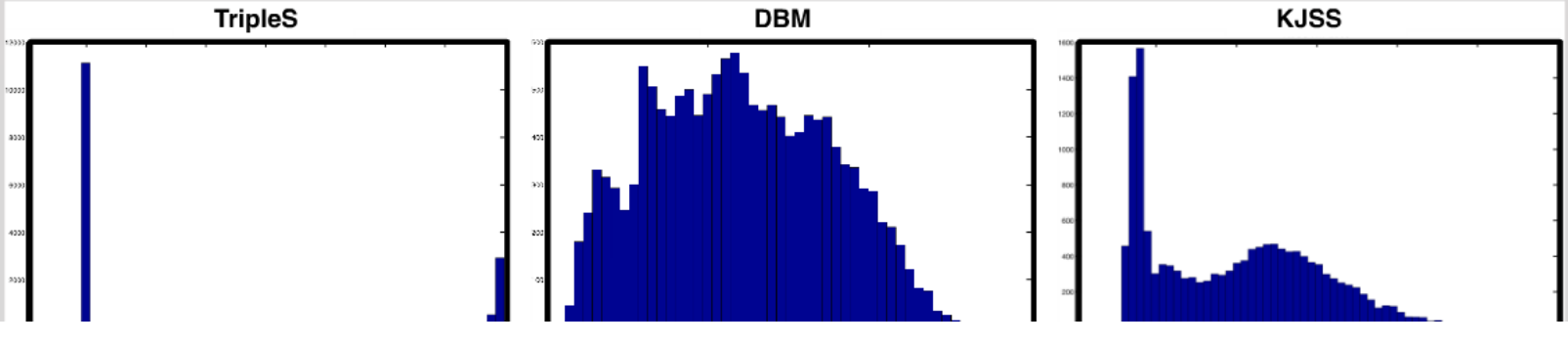
- Pick two random samples from the same forecast.
- Use the W-test and the T-test to determine if the samples are from the same forecast
- Do this several times to do some statistics

Error Type II

- Similar but use two "different" forecasts

Two "samplers" were used, a random and a "biased" one

Typical forecast distributions

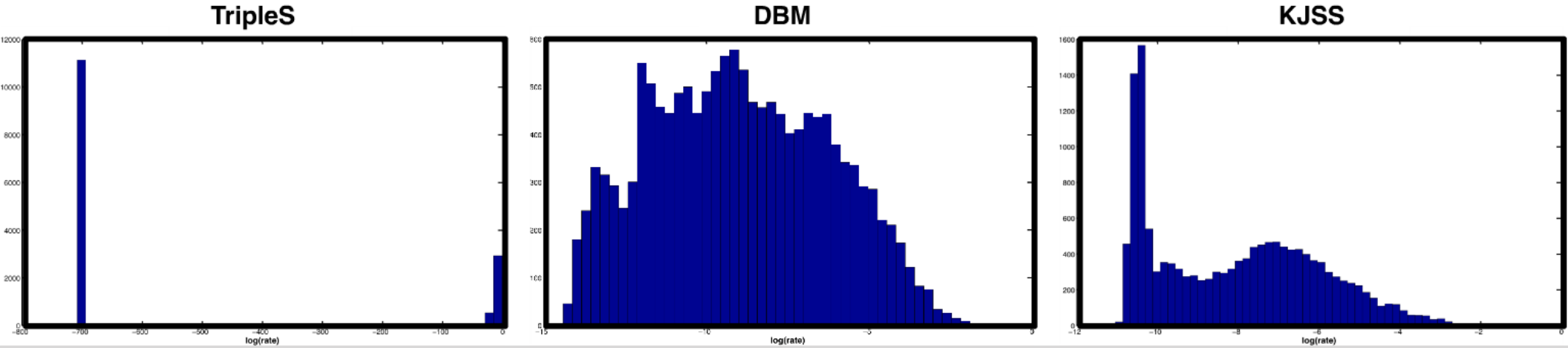


Error Type II

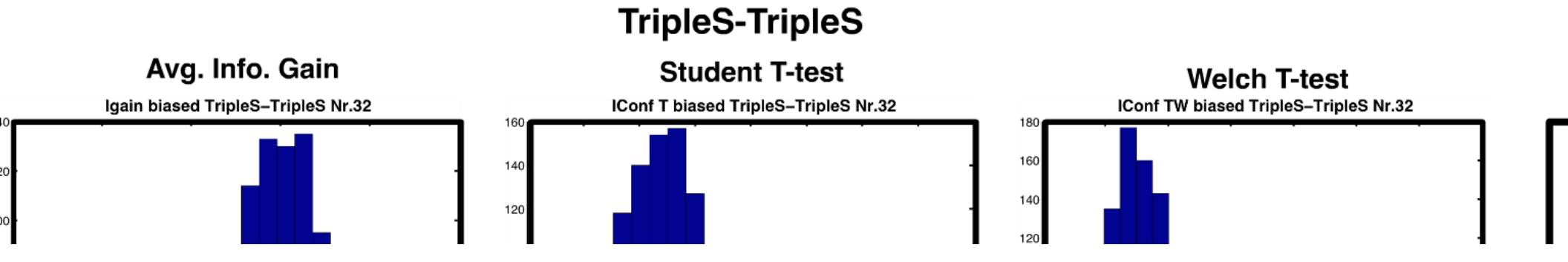
- Similar but use two "different" forecasts

Two "samplers" were used, a random and a "biased" one

Typical forecast distributions



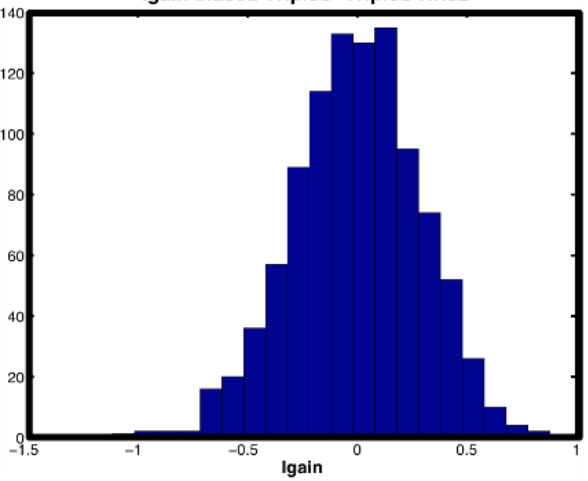
Example of log(A) distribution (NW Pacific 2009)



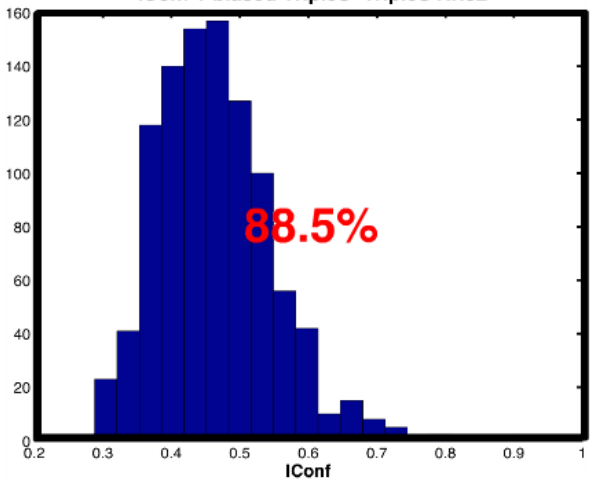
Example of log(A) distribution (NW Pacific 2009)

TripleS-TripleS

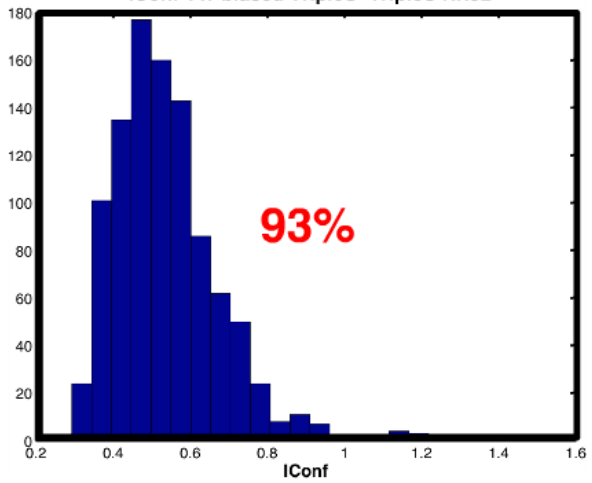
Avg. Info. Gain



Student T-test



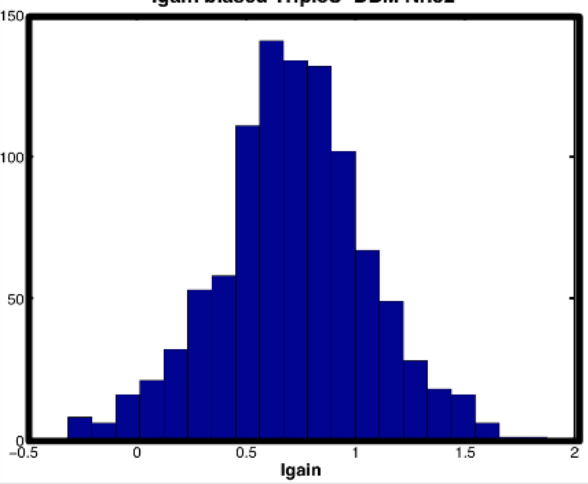
Welch T-test

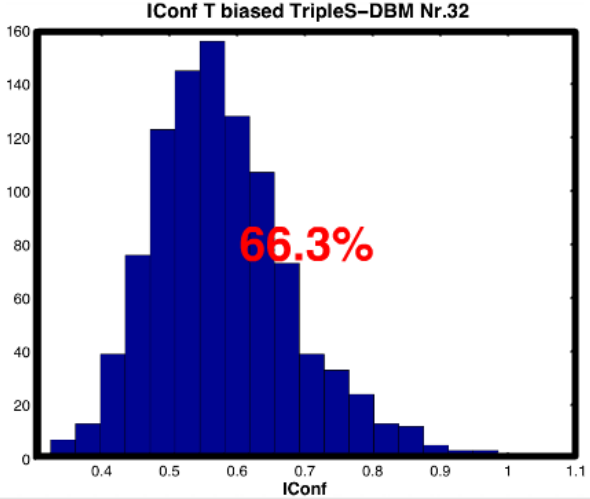


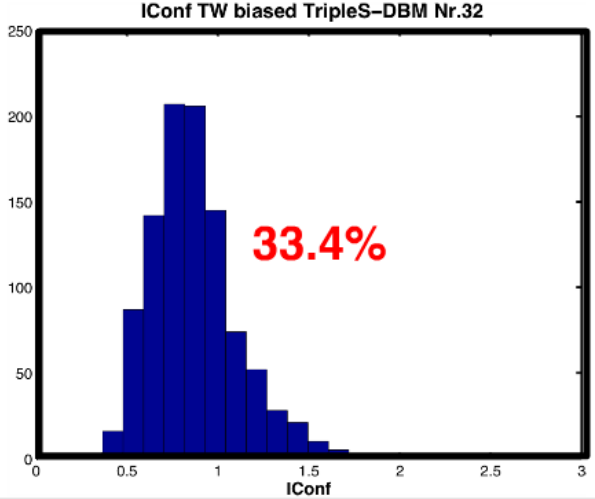
W-test

94.2%

TripleS-DBM

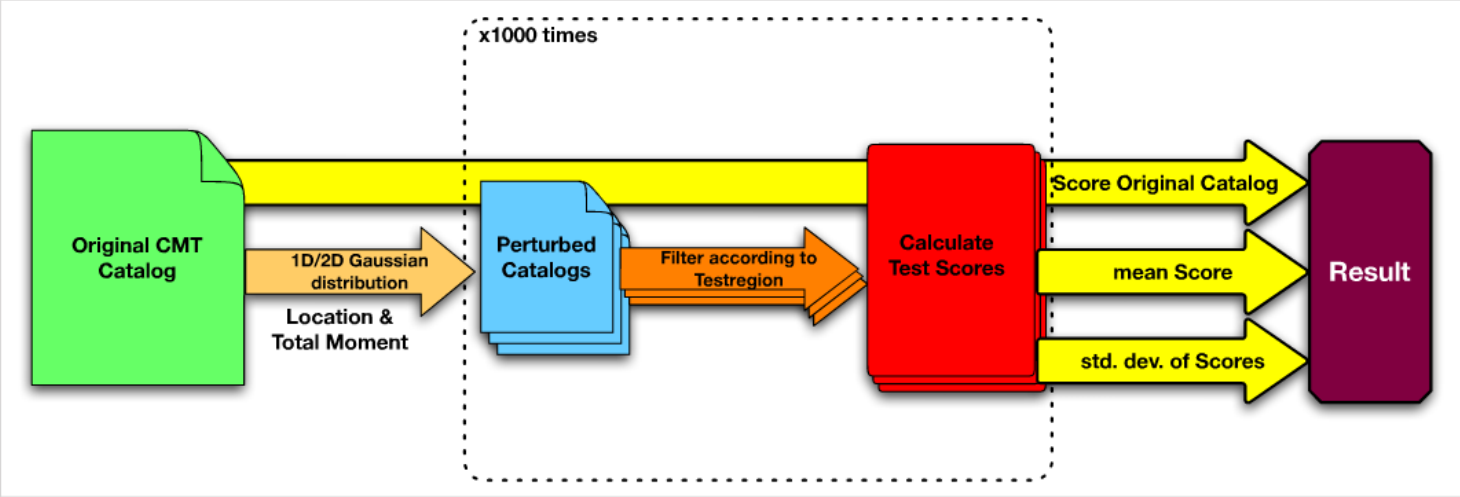




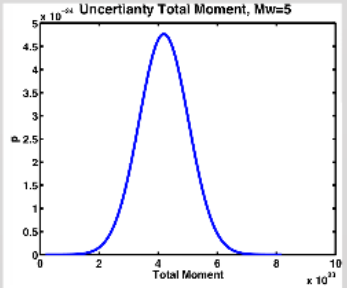


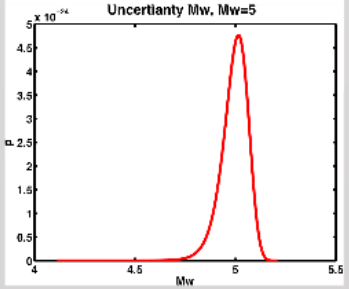
52.2%

Uncertainty estimation

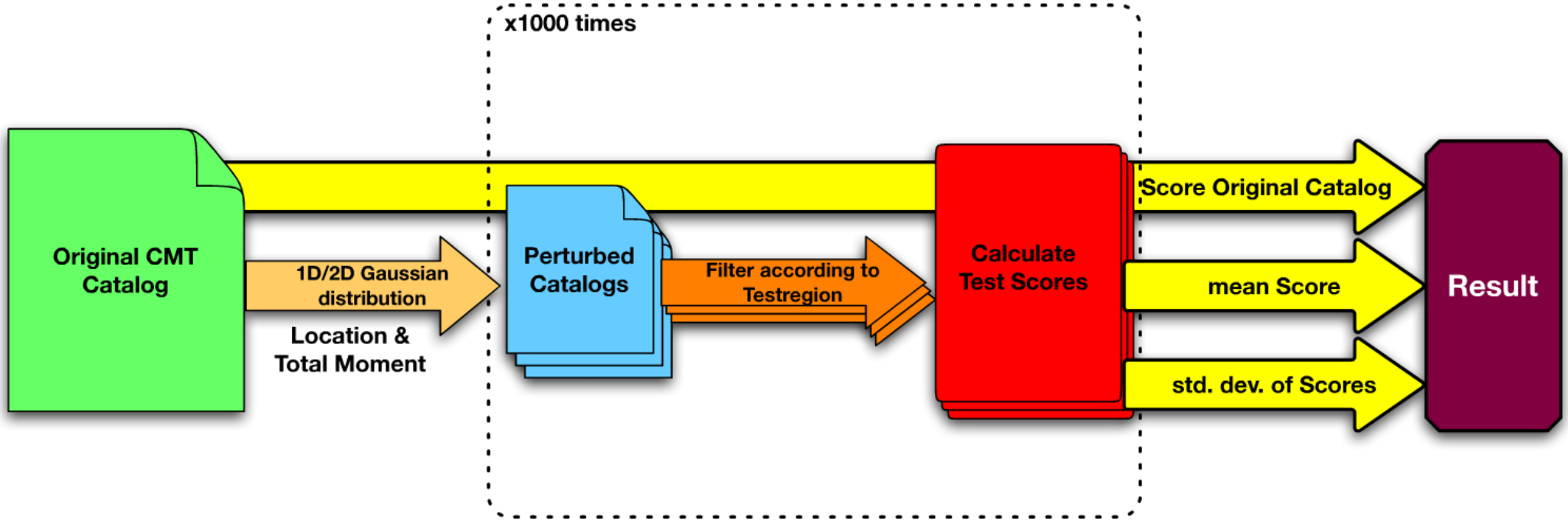


- Symmetric uncertainties in M0 - asymmetric uncertainties in Mw
- Because of Gutenberg-Richter Law there are more earthquakes below the magnitude threshold
- BUT, no higher probability for more earthquake in the test regions was observed

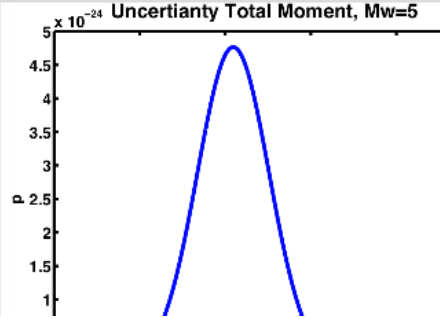




Uncertainty estimation



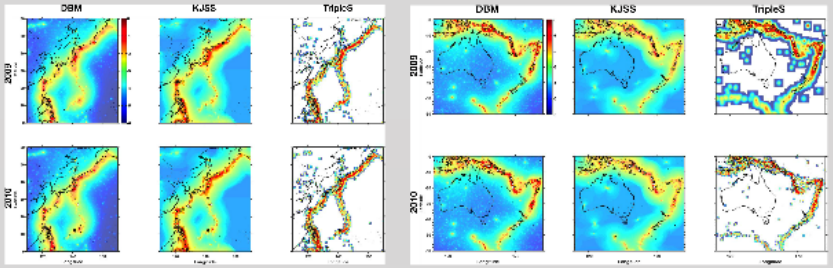
mmetric uncertainties in M0 - asymmetric
certainties in Mw



West Pacific Testing Region

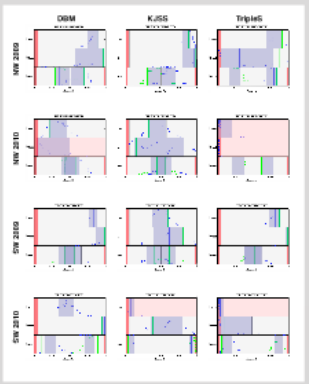
Setup

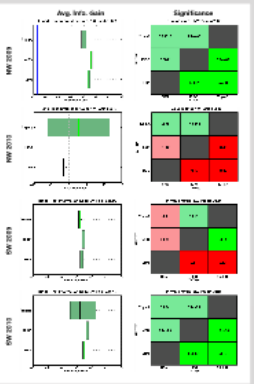
- Two sub-regions for testing: NW and SW Pacific
- Current 2 one-year periods are used (2009 and 2010)
- Poisson rate forecasts from 3 Models:
  - DBM: Double-Branching Model by Marzocchi and Lombardi
  - KJSS: Kagan and Jackson Smoothed Seismicity Model
  - TripleS: Simple Smoothed Seismicity Model by Zechar and Jordan
- Location uncertainties CMT: 30km
- M0 uncertainty CMT: 0.2*M0



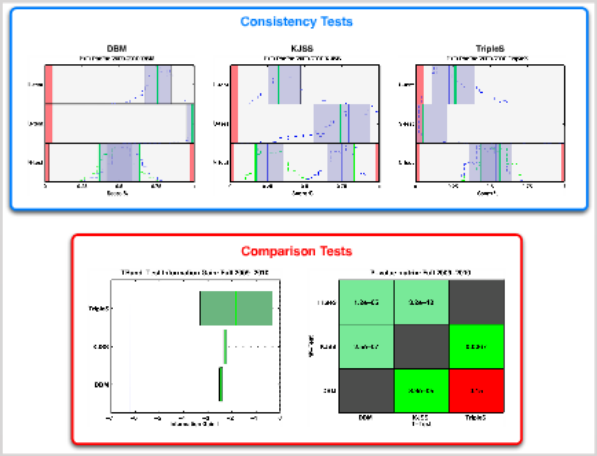
Test Results

"Single" Tests



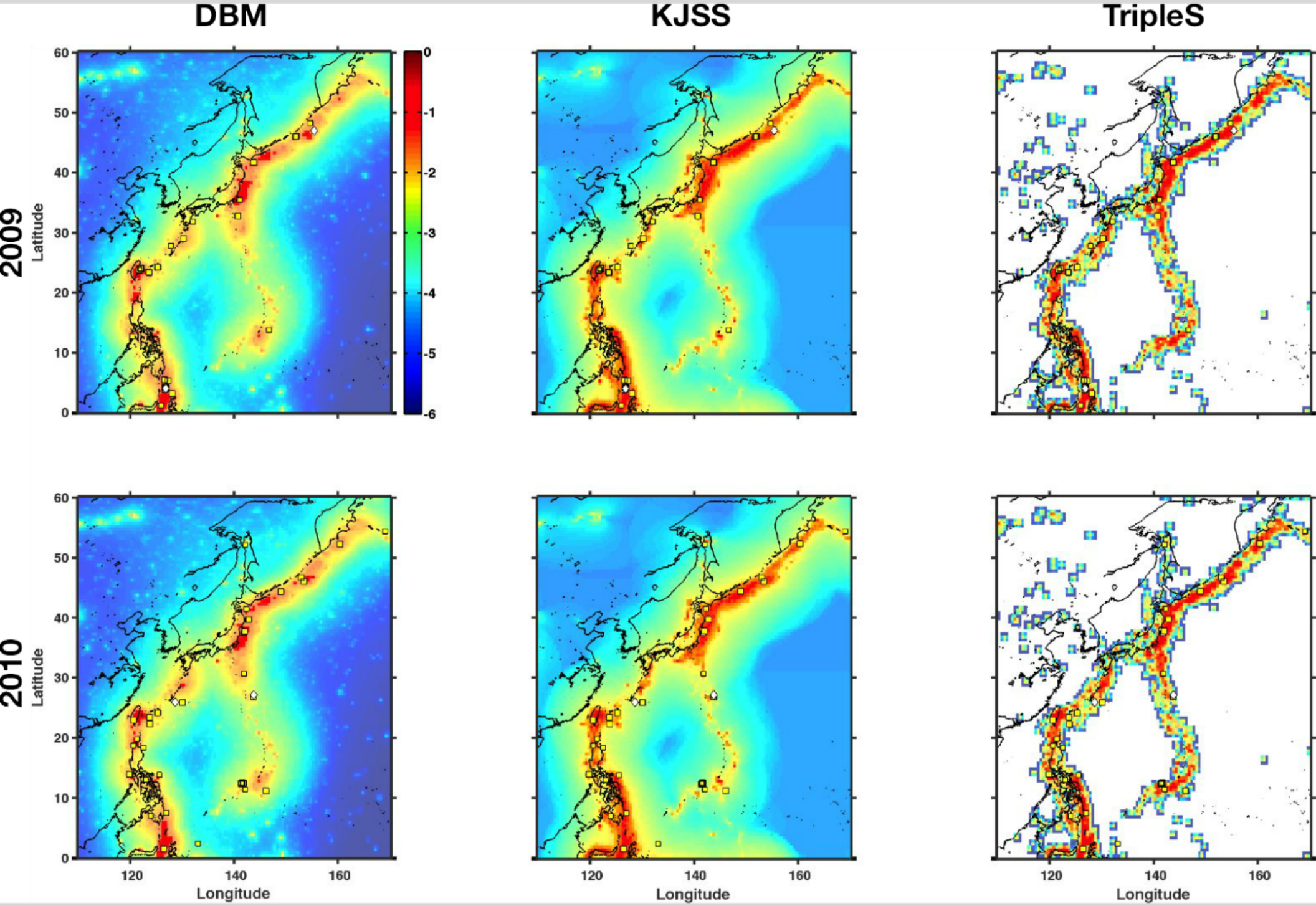


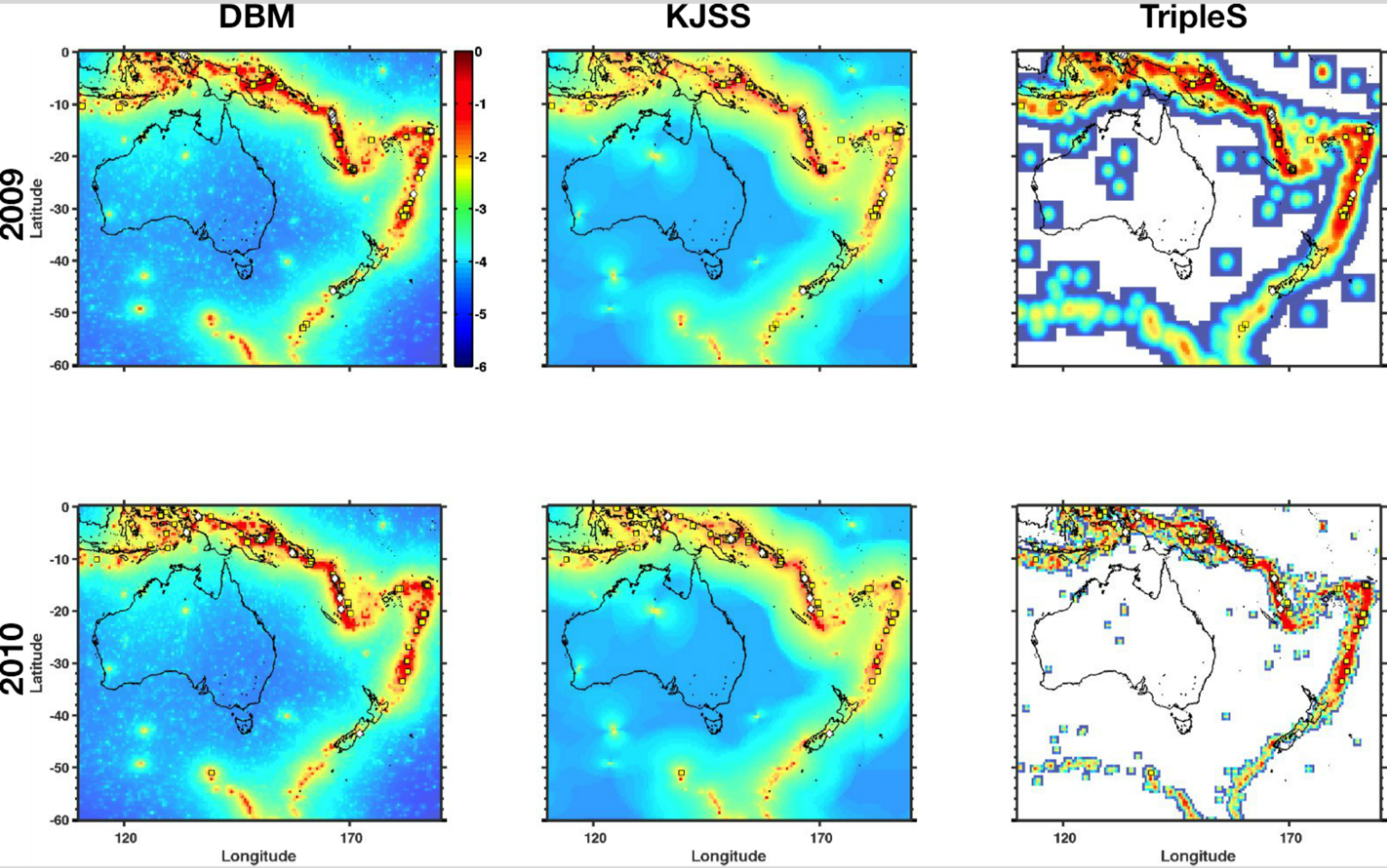
Test Summary



Setup

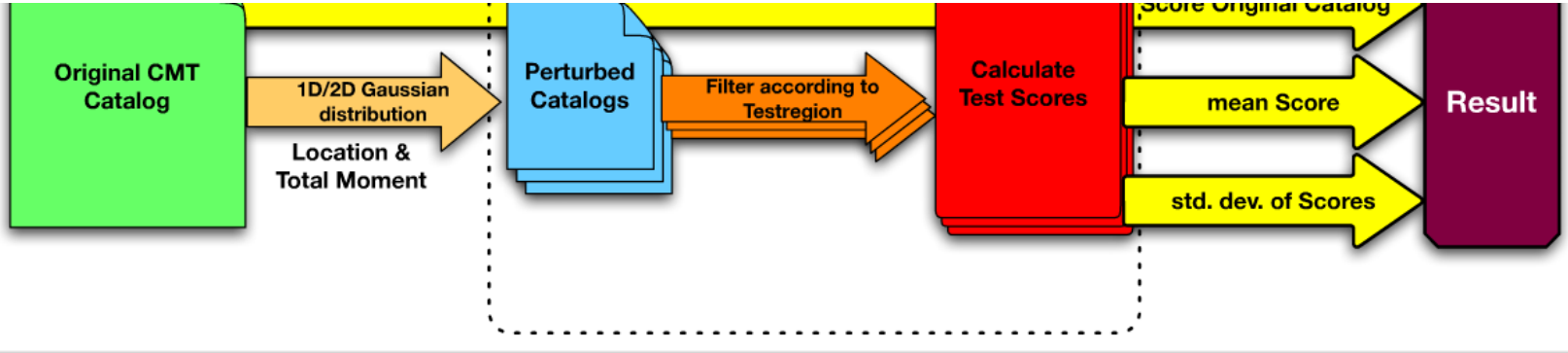
- Two sub-regions for testing: NW and SW Pacific
- Current 2 one-year periods are used (2009 and 2010)
- Poisson rate forecasts from 3 Models:
  - DBM: Double-Branching Model by Marzocchi and Lombardi
  - KJSS: Kagan and Jackson Smoothed Seismicity Model
  - TripleS: Simple Smoothed Seismicity Model by Zechar and Jordan
- Location uncertainties CMT: 30km
- M0 uncertainty CMT: 0.2*M0


DBM
KJSS
TripleS
2009
2010
Latitude
Longitude
0
10
20
30
40
50
60
120
140
160
0
-1
-2
-3
-4
-5
-6

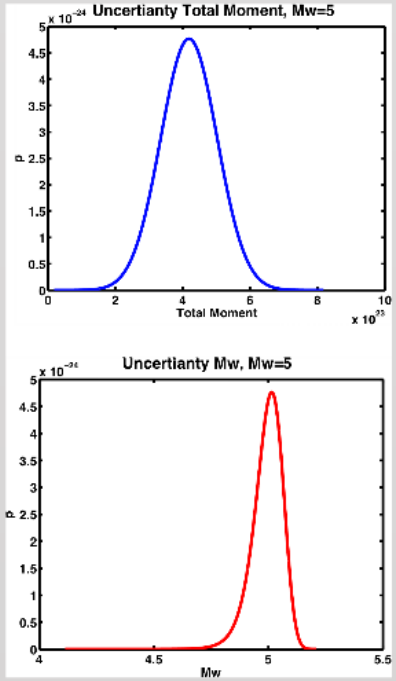

DBM
KJSS
TripleS
2009
2010
Latitude
Longitude
0
-10
-20
-30
-40
-50
-60
120
170
-1
-2
-3
-4
-5
-6

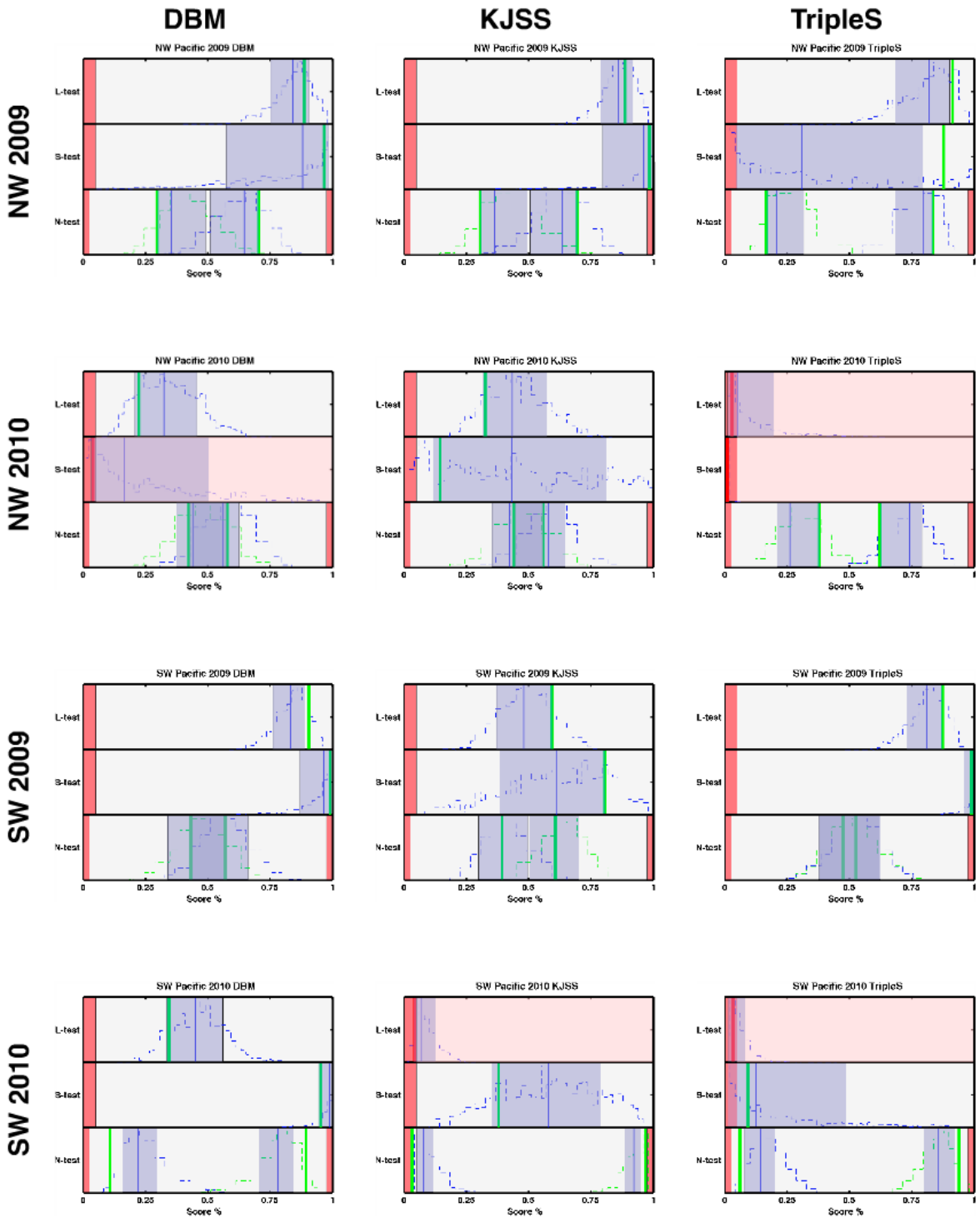
Setup

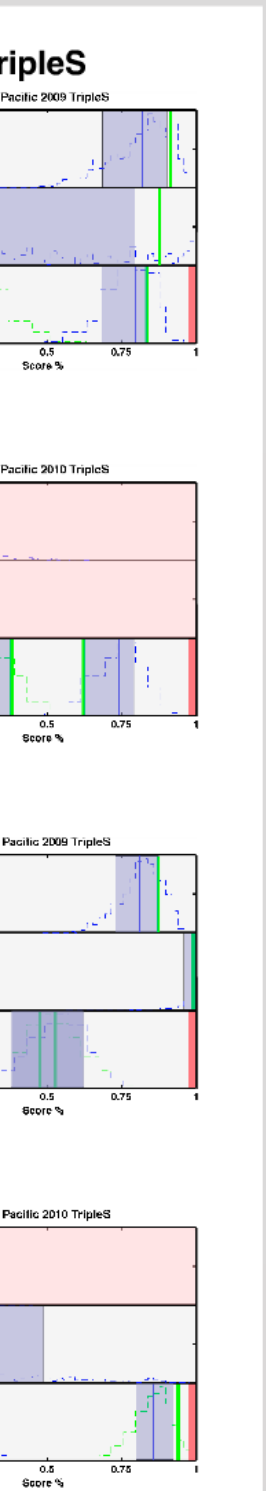
- Two sub-regions for testing: NW and SW Pacific
- Current 2 one-year periods are used (2009 and 2010)
- Poisson rate forecasts from 3 Models:
  - DBM: Double-Branching Model by Marzocchi and Lombardi
  - KJSS: Kagan and Jackson Smoothed Seismicity Model
  - TripleS: Simple Smoothed Seismicity Model by Zechar and Jordan
- Location uncertainties CMT: 30km
- M0 uncertainty CMT: 0.2*M0

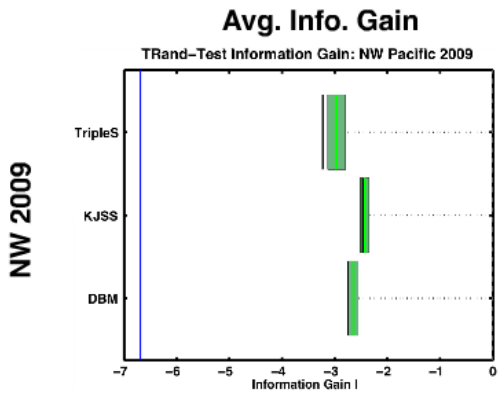


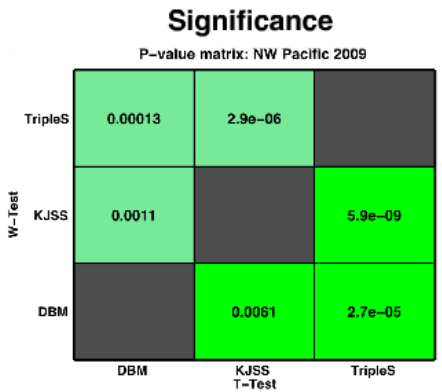
- Symmetric uncertainties in M0 - asymmetric uncertainties in Mw
- Because of Gutenberg-Richter Law there are more earthquakes below the magnitude threshold
- BUT, no higher probability for more earthquake in the test regions was observed

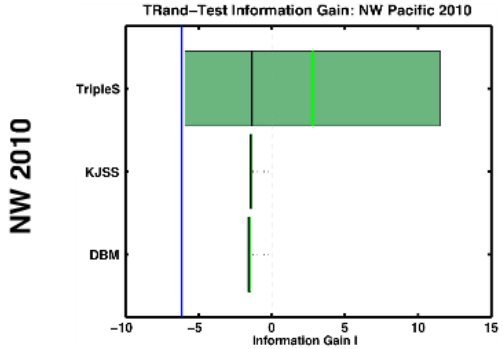


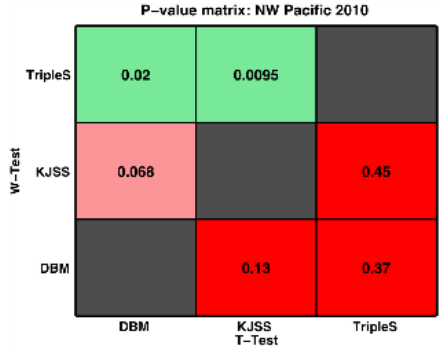

DBM
KJSS
TripleS
NW 2009
NW 2010
SW 2009
SW 2010
NW Pacific 2009 DBM
NW Pacific 2009 KJSS
NW Pacific 2009 TripleS
NW Pacific 2010 DBM
NW Pacific 2010 KJSS
NW Pacific 2010 TripleS
SW Pacific 2009 DBM
SW Pacific 2009 KJSS
SW Pacific 2009 TripleS
SW Pacific 2010 DBM
SW Pacific 2010 KJSS
SW Pacific 2010 TripleS
L-test
S-test
N-test
Score %
0
0.25
0.5
0.75
1

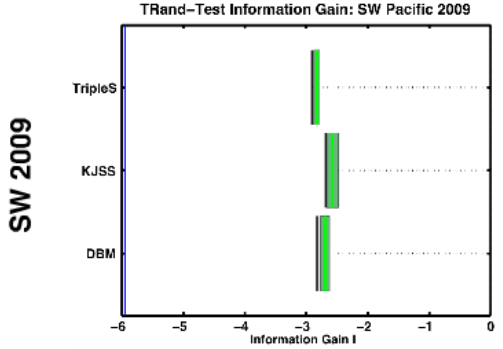

ripleS
Pacific 2009 TripleS
0.5
0.75
1
Score %
Pacific 2010 TripleS
0.5
0.75
1
Score %
Pacific 2009 TripleS
0.5
0.75
1
Score %
Pacific 2010 TripleS
0.5
0.75
1
Score %

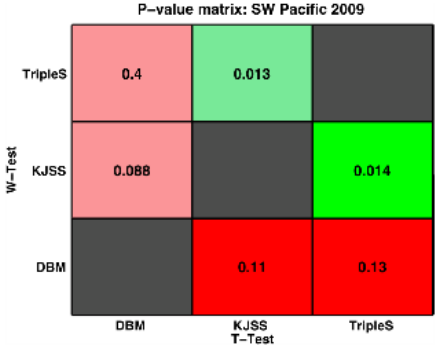

Avg. Info. Gain
NW 2009
TRand−Test Information Gain: NW Pacific 2009
TripleS
KJSS
DBM
−7
−6
−5
−4
−3
−2
−1
0
Information Gain I

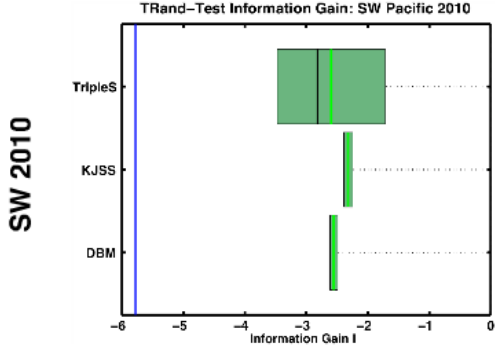

Significance
P−value matrix: NW Pacific 2009
TripleS
KJSS
DBM
W−Test
0.00013
2.9e−06
0.0011
5.9e−09
0.0061
2.7e−05
DBM
KJSS
TripleS
T−Test

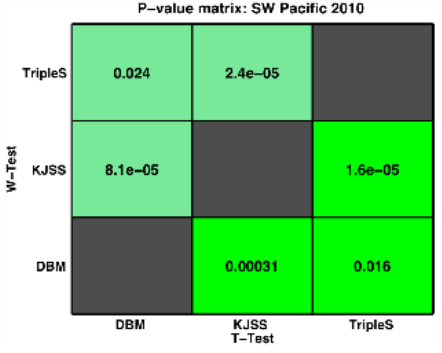

NW 2010
TRand−Test Information Gain: NW Pacific 2010
TripleS
KJSS
DBM
−10
−5
0
5
10
15
Information Gain I


P−value matrix: NW Pacific 2010
TripleS
KJSS
DBM
W−Test
0.02
0.0095
0.068
0.45
0.13
0.37
DBM
KJSS
TripleS
T−Test

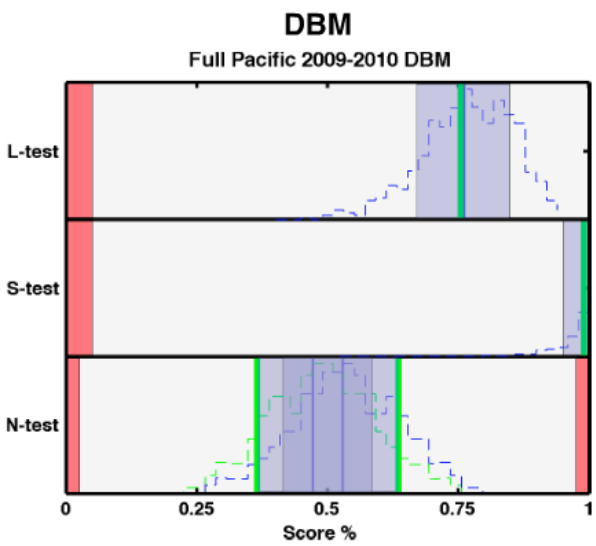

SW 2009
TRand−Test Information Gain: SW Pacific 2009
TripleS
KJSS
DBM
−6
−5
−4
−3
−2
−1
0
Information Gain I

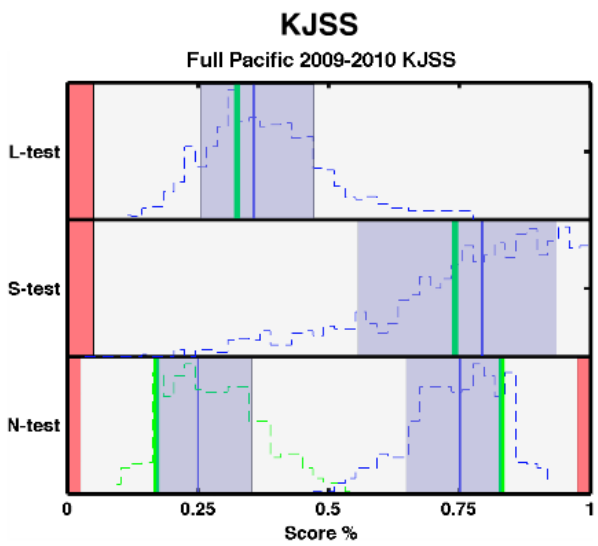

P−value matrix: SW Pacific 2009
TripleS
KJSS
DBM
W−Test
0.4
0.013
0.088
0.014
0.11
0.13
DBM
KJSS
TripleS
T−Test

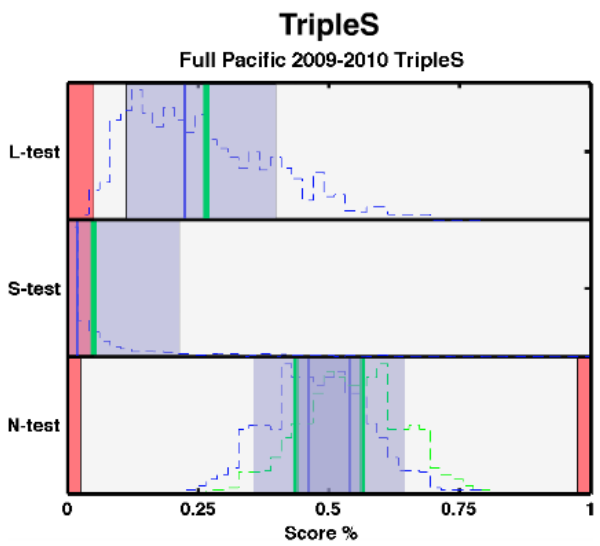

SW 2010
TRand−Test Information Gain: SW Pacific 2010
TripleS
KJSS
DBM
−6
−5
−4
−3
−2
−1
0
Information Gain I


P−value matrix: SW Pacific 2010
TripleS
KJSS
DBM
W−Test
0.024
2.4e−05
8.1e−05
1.6e−05
0.00031
0.016
DBM
KJSS
TripleS
T−Test

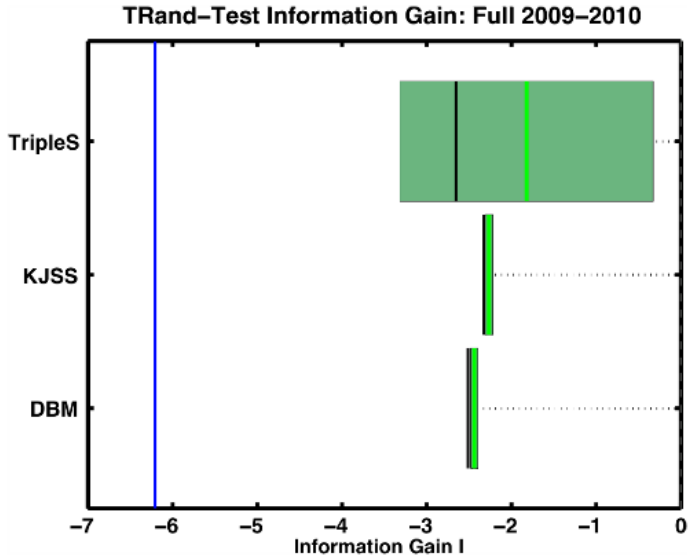
Consistency Tests

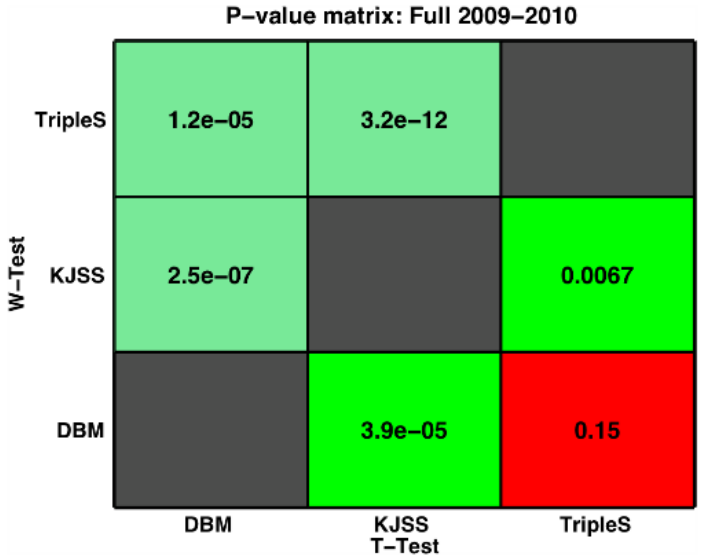






Comparison Tests





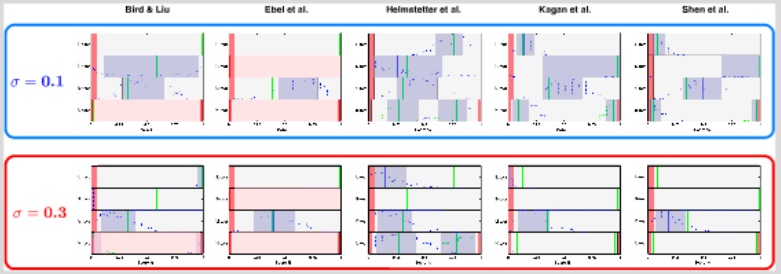
Uncertainties in RELM

Setup

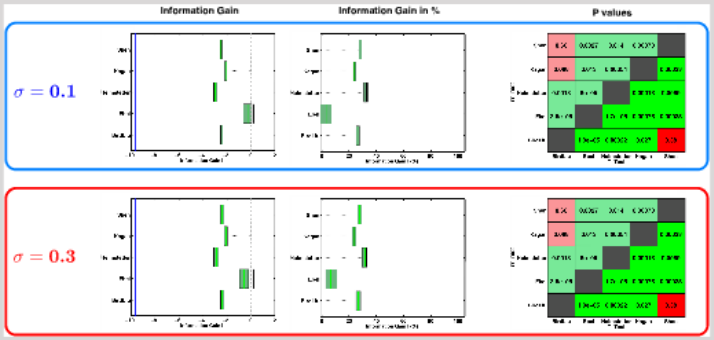
- Rework of the RELM experiment with uncertainty estimates
- 5 forecast models
  - Bird & Liu
  - Ebel et al.
  - Helmstetter et al.
  - Kagan et al.
  - Shen et al.
- Used catalog uncertainties
  - Location uncertainties: 5 km
  - Magnitude uncertainties: Laplace distributed with sigma = 0.1 and 0.3
- The ANSS Catalog did change between the first evaluation and this rework

Test Results

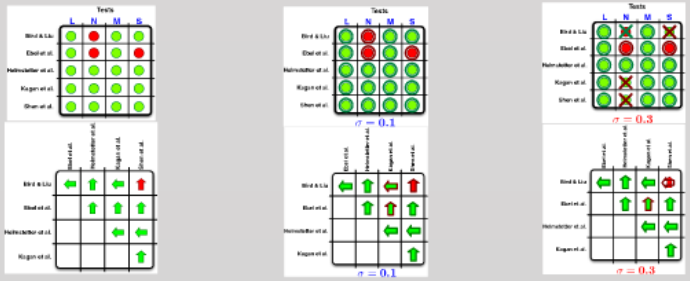
Consistency



Comparison



"Original" vs Mean



Setup

- Rework of the RELM experiment with uncertainty estimates

- 5 forecast models
  - Bird & Liu
  - Ebel et al.
  - Helmstetter et al.
  - Kagan et al.
  - Shen et al.

- Used catalog uncertainties
  - Location uncertainties: 5 km
  - Magnitude uncertainties: Laplace distributed with sigma = 0.1 and 0.3

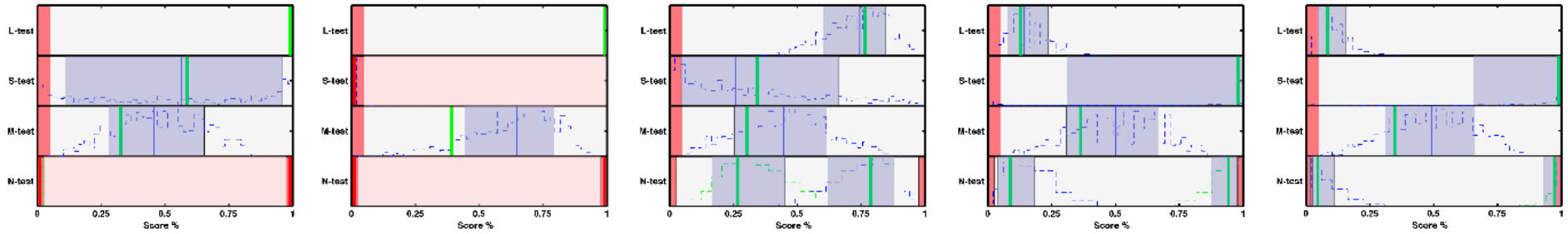
- The ANSS Catalog did change between the first evaluation and this rework

Test Results

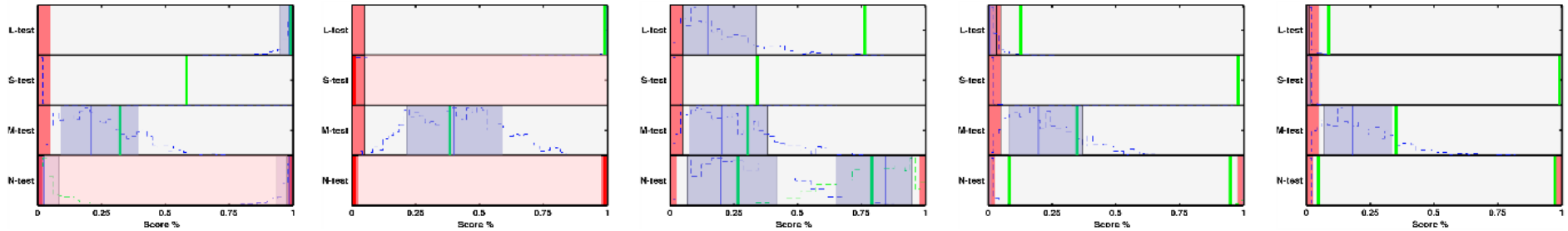
Consistency

Bird & Liu | Ebel et al. | Helmstetter et al. | Kagan et al. | Shen et al.

$\sigma = 0.1$

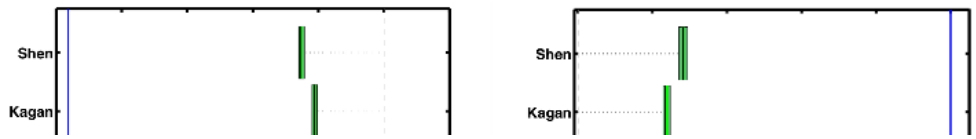


$\sigma = 0.3$

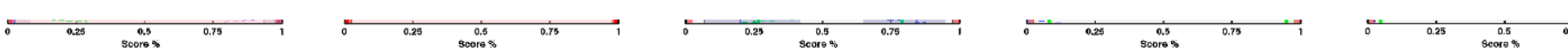


Comparison

Information Gain | Information Gain in % | P values



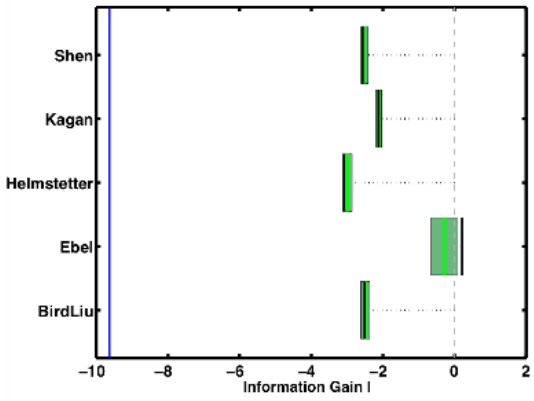
Shen	0.66	0.0027	0.014	0.00078	
Kagan	0.088	0.013	0.00054		0.00033

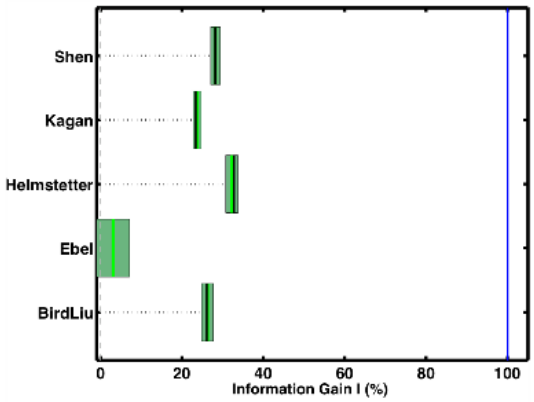


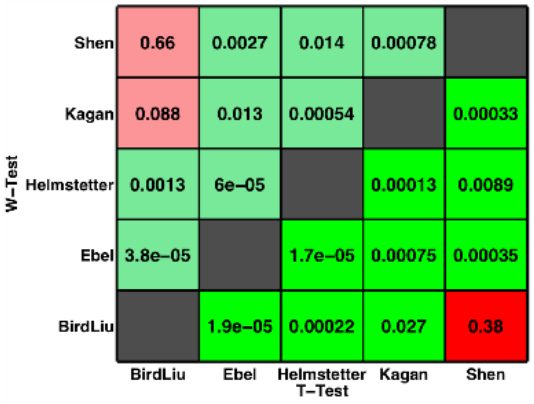
Comparison

Information Gain | **Information Gain in %** | **P values**

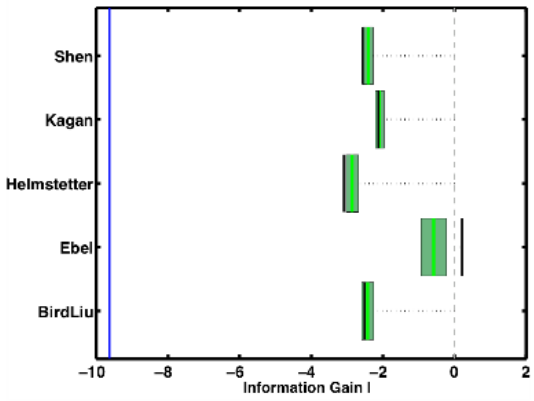
$\sigma = 0.1$

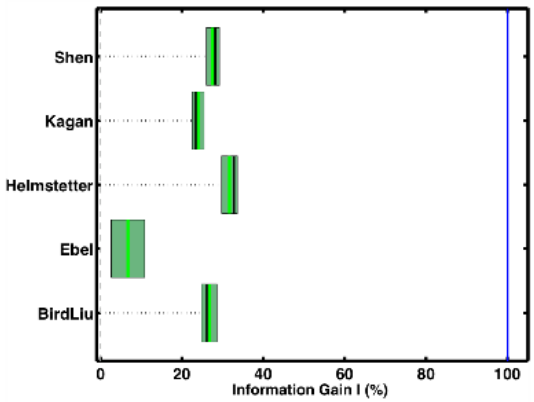


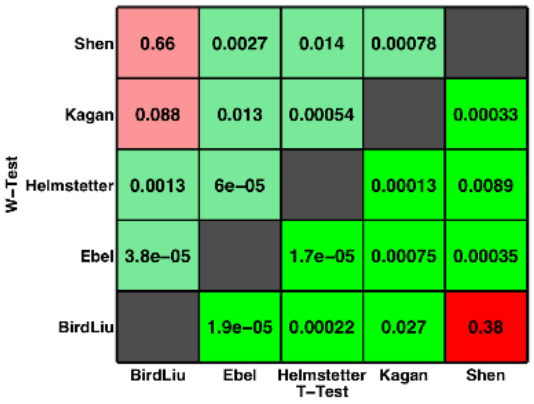




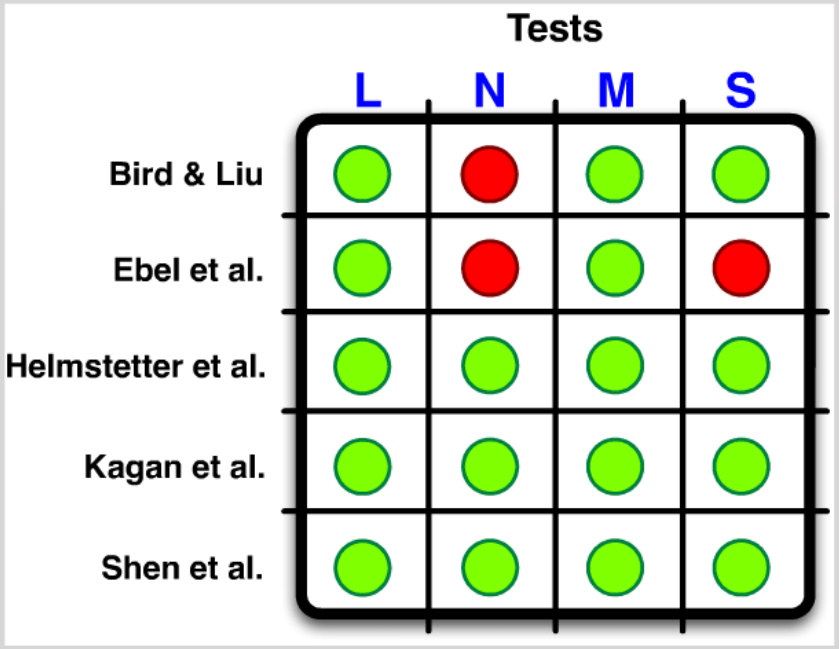
$\sigma = 0.3$

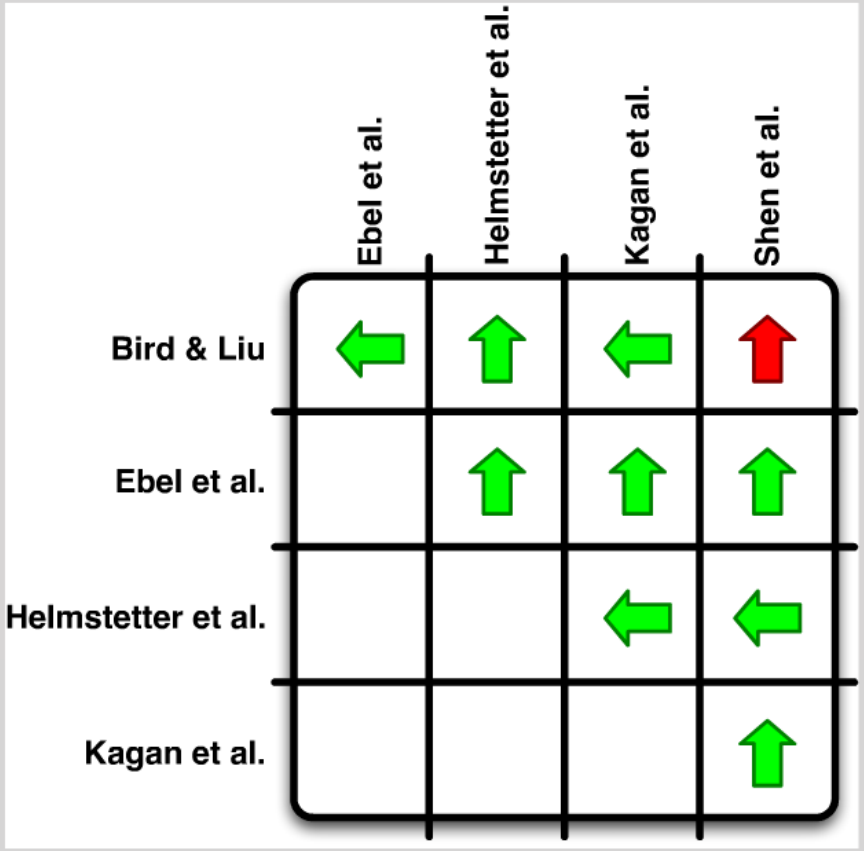


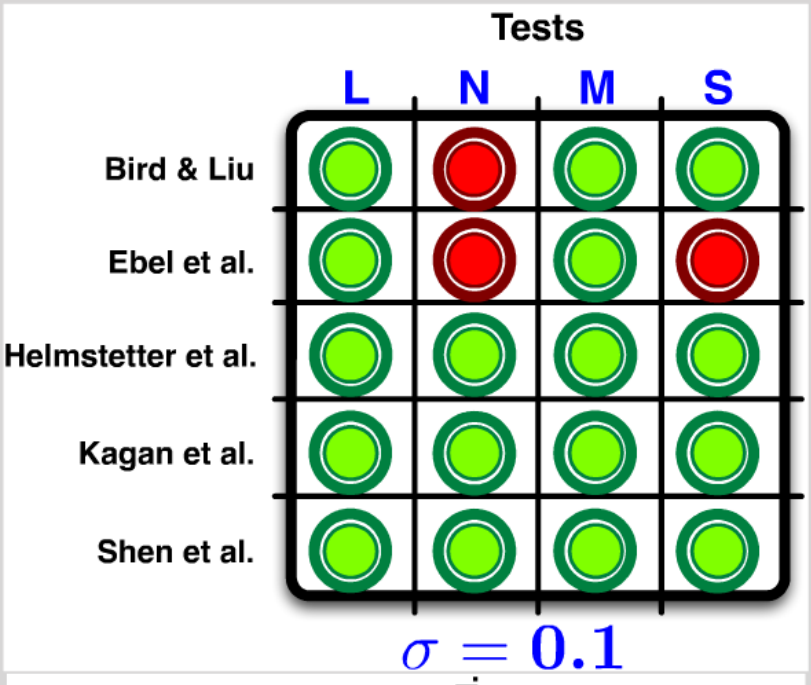


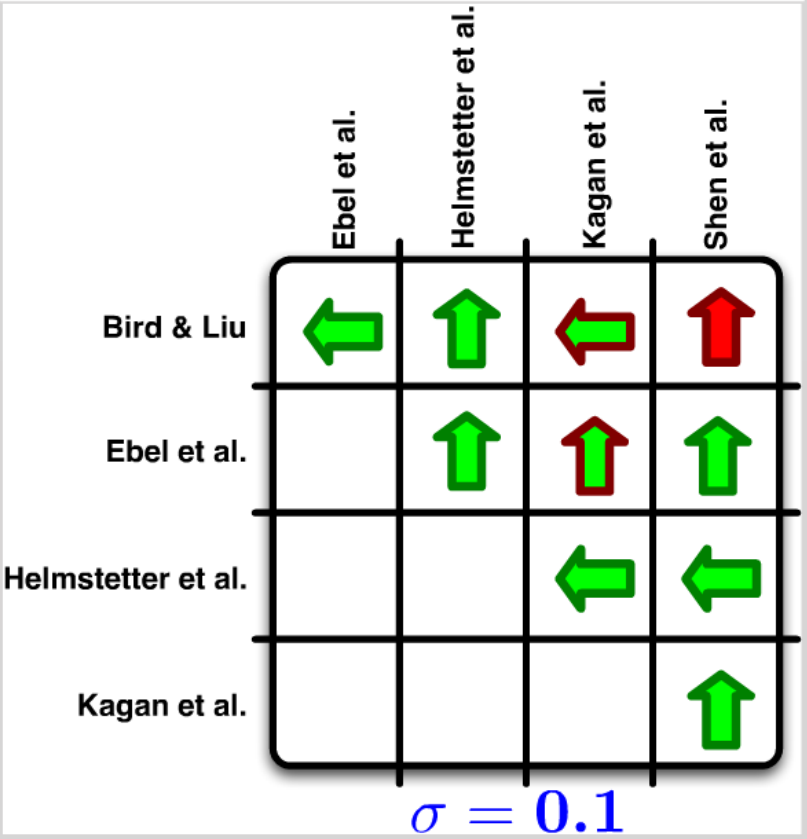


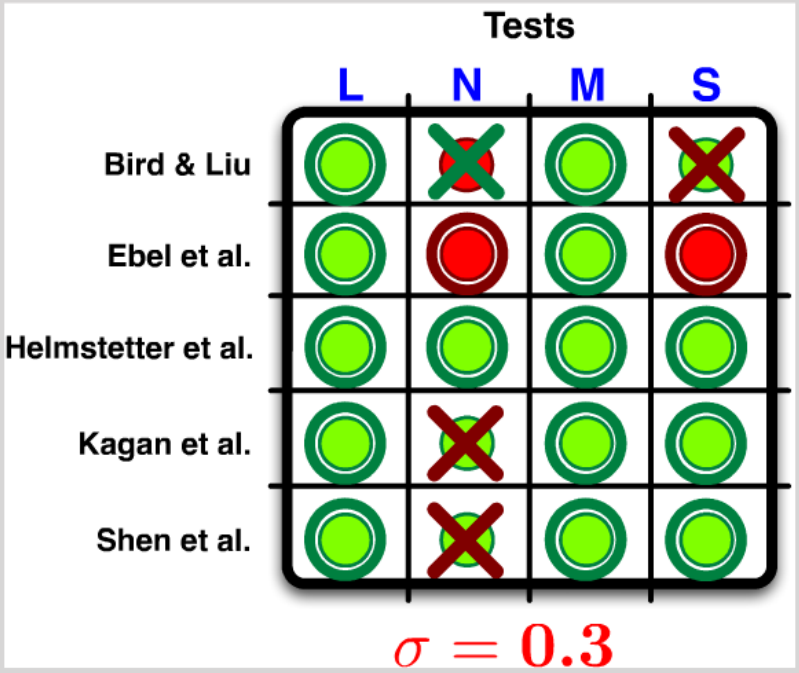
"Original" vs Mean

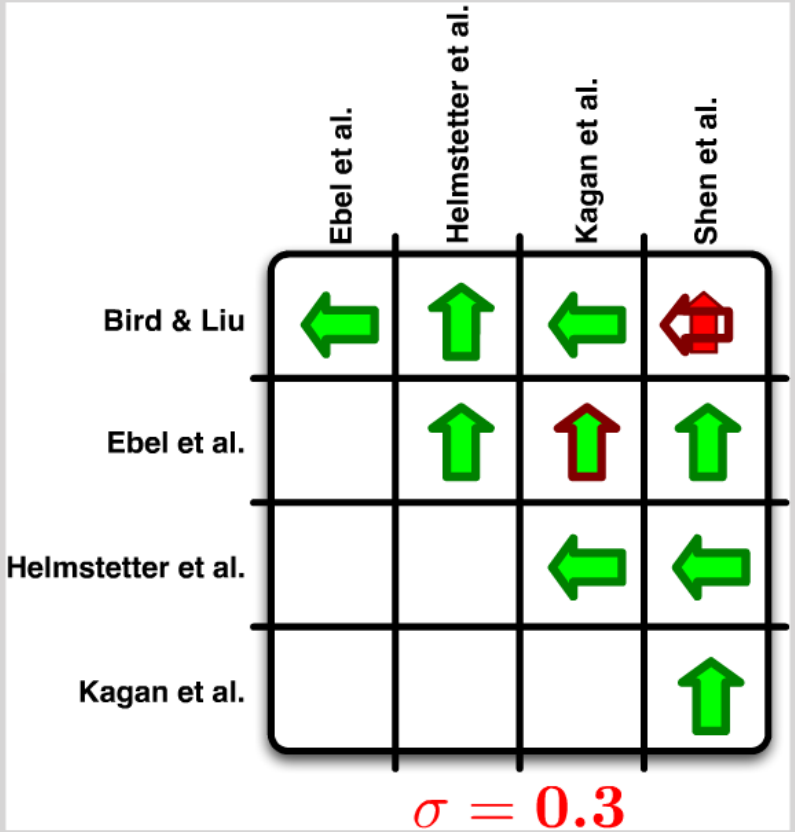

Tests
L
N
M
S
Bird & Liu
Ebel et al.
Helmstetter et al.
Kagan et al.
Shen et al.


Ebel et al.
Helmstetter et al.
Kagan et al.
Shen et al.
Bird & Liu
Ebel et al.
Helmstetter et al.
Kagan et al.


Tests
L
N
M
S
Bird & Liu
Ebel et al.
Helmstetter et al.
Kagan et al.
Shen et al.
$\sigma = 0.1$


Ebel et al.
Helmstetter et al.
Kagan et al.
Shen et al.
Bird & Liu
Ebel et al.
Helmstetter et al.
Kagan et al.
$\sigma = 0.1$


Tests
L
N
M
S
Bird & Liu
Ebel et al.
Helmstetter et al.
Kagan et al.
Shen et al.
$\sigma = 0.3$


Ebel et al.
Helmstetter et al.
Kagan et al.
Shen et al.
Bird & Liu
Ebel et al.
Helmstetter et al.
Kagan et al.
$\sigma = 0.3$

Unfinished Business

CSEP West Pacific

- 2 years of data are not enough, addition 1 year periods will be added
- 3 models are a start but not enough
- Magnitudes bins are missing and successor test regions like the global test region should consider them.

Accounting for Uncertainties

- The effect on the test outcome is dependent on the forecast and the location of the earthquakes
- Can help to evaluate the robustness of forecast for a certain test region and time period
- Uncertainty estimation is a valuable addition to the test procedure

- 2 years of data are not enough, addition 1 year periods will be added
- 3 models are a start but not enough
- Magnitudes bins are missing and successor test regions like the global test region should consider them.

- The effect on the test outcome is dependent on the forecast and the location of the earthquakes
- Can help to evaluate the robustness of forecast for a certain test region and time period
- Uncertainty estimation is a valuable addition to the test procedure