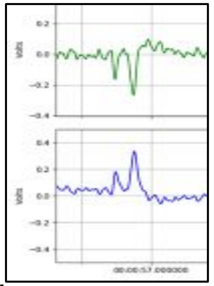
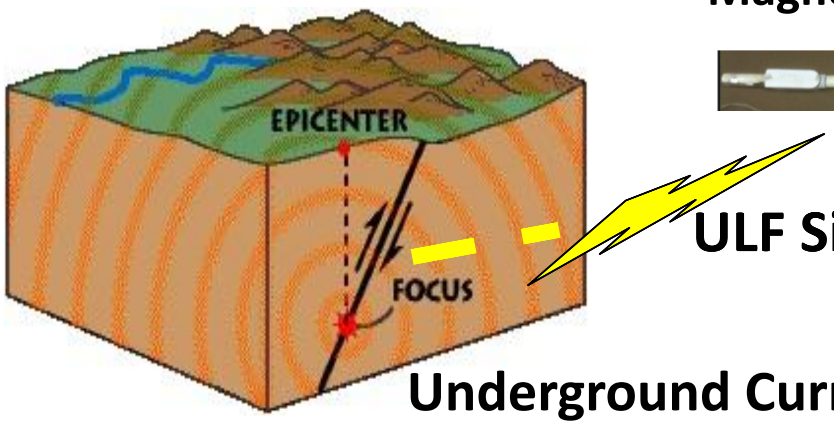
Are there Geomagnetic Precursors to Earthquakes?
— Two Statistical Studies from California

Presented by Karl N. Kappler
Spring 2023



Induction Magnetometers

Pulse Counting



ULF Signals

Underground Currents?

Features

Warning lights

A flickering glow in the sky seems to accompany some earthquakes. Could this point to a way of predicting these disasters? **Nathaniel Scharping** investigates

THE resort of Acapulco in Mexico has long been known for its attractions: gorgeous mountains, upmarket hotels, crystal clear waters. But on 7 September 2021, something happened that was on nobody's wish list – a magnitude-7.0 earthquake rocked the city's sandy beaches and seafront high-rises.

Along with trembling buildings and shaking trees, those caught in the quake also witnessed something substantially more eerie. A barrage of blue lights, like flashes of cerulean lightning, lit up the night sky, apparently right above the fault line. This strange display was an example of what are known as "earthquake lights", a semi-mythical phenomenon that has cropped up in reports of tremors for centuries.

The idea that these blue flashes are caused by an earthquake is often dismissed by scientists. Indeed, after Acapulco, some suggested the flickering lights may have come from damaged power lines. But a small group of researchers now claim to have evidence for an alternative hypothesis. It says that when tectonic faults rupture, electrical currents are created. And whether these currents produce lights or not, there should be telltale electromagnetic signals produced by them that would be detectable in advance.

If they are right, we could potentially use these signals as a warning of disaster. It is a long shot: the search for ways to predict earthquakes has frustrated us for decades. But new evidence linked to these uncanny, dancing lights in the sky is shaking up the field.

Predicting major tremors is currently just about impossible. Scientists, including those at the United States Geological Survey (USGS), a national agency, compile long-term seismological data that can tell us the chance of an earthquake hitting a given area, but only across a window of time that spans years or decades, rather than anything more precise. Then, there are warning systems like ShakeAlert in the US, which uses seismometers to give people alerts of incoming quakes – but only seconds in advance.

Earthquake lights captured over Mount Kimyo, Japan, in 1968

6 May 2023 | New Scientist | 43

06 May,2023; New Scientist features QuakeFinder Array

Contact: QuakeFinder.com
Dan Schneider dschneider@quakefinder.com
Tom Bleier tbleier@quakefinder.com

Outline

- Project Background
- Research objectives
- [QuakeFinder Paper]
- [Google Paper]
- Summary
- State of QuakeFinder ULF research

Computers & Geosciences 133 (2019) 104317

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

ELSEVIER

Check for updates

An algorithmic framework for investigating the temporal relationship of magnetic field pulses and earthquakes applied to California

K.N. Kappler [a,b], D.D. Schneider [a,*], L.S. MacLean [a], T.E. Bleier [a], J.J. Lemon [a]

[a] QuakeFinder, 250 Cambridge Avenue, Suite 204, Palo Alto, CA, 94306, United States
[b] DataCloud International, 100 South King St, Suite 100-710, Seattle, WA, 98104, United States

ABSTRACT

An end-to-end algorithm is described wherein field-collected magnetometer time series data were processed and analyzed for potential statistical correlation with pre-seismic activity. The process included windowing the data, extraction of statistically-determined anomalies via a short term average - long term average (STA-LTA) signal processing technique, collating and ranking the anomalous windows as precursory behavior, and testing the results via a Receiver Operating Characteristic (ROC) formulation. The algorithm was employed on a large dataset of over 100 magnetic observatories in California totaling hundreds of thousands of station-days. Using the ROC curve to evaluate its performance, this implementation of the algorithm obtained a 2.20 z-score. This number improved with the preliminary attempt at removing a severe cultural noise source. This work emphasizes an analytic framework more than parametric exploration or optimization, nevertheless there appears to be some suggestion of predictive power in the magnetic field time series.

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2022JB024109

Case-Control Study on a Decade of Ground-Based Magnetometers in California Reveals Modest Signal 24–72 hr Prior to Earthquakes

William D. Heavlin[1], Karl Kappler[2,3], Lusann Yang[1], Minjie Fan[1], Jason Hickey[1], James Lemon[3], Laura MacLean[3], Thomas Bleier[3], Patrick Riley[1], and Daniel Schneider[3]

[1]Google Research, Applied Science Team, El Granada, CA, USA, [2]Imdex Technology USA LLC, San Luis Obispo, CA, USA, [3]QuakeFinder, Palo Alto, CA, USA

Key Points:

- Frequency domain analysis of ground-based magnetometer data shows a modest change in days leading up to intermediate-large ($M \geq 4.5$) earthquakes
- One novel part of the analysis is the use of cross-power signals, which combines the signals from instruments separated by tens of km
- A supplementary analysis of the data with first order global geomagnetic effects subtracted increased the measured effect size significantly

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Schneider,
dschneider@quakefinder.com

Citation:

Heavlin, W. D., Kappler, K., Yang, L., Fan, M., Hickey, J., Lemon, J., et al. (2022). Case-control study on a decade of ground-based magnetometers in California reveals modest signal 24–72 hr prior to earthquakes. *Journal of Geophysical Research: Solid Earth*, *127*, e2022JB024109. https://doi.org/10.1029/2022JB024109

Received 27 JAN 2022
Accepted 19 AUG 2022

Abstract Magnetic field changes as earthquake precursors have been the subject of numerous studies and some controversy. Infrequent large earthquakes and sparse magnetometer coverage along fault zones complicate statistical analysis. We present an analysis of ground-based magnetic time-series measurements before 19 earthquakes ≥M4.5 in California drawing from over 330,000 site-days of measurement spanning a decade. To perform a fair existential test for electromagnetic antecedents we applied a pre-specified statistical analysis with two key ideas. First, we combine signals from nearby (≤40 km) sites via spectral cross-power, and then look for large spikes in frequency domain (0.016–25 Hz). The former is only possible with a dense set of sites running over a long period of time. In this statistical case-control study we used the machine learning concept of rigorously separated train and test sets of earthquakes which were generated via a rule-based query of the USGS earthquake catalog. Before each declustered earthquake, we constructed one period 24–72 hr before (the "precursor" or "*p*-period") and a series of seven equally-sized preceding periods ("quiescent" or "*q*-periods"). We distilled the data in each period to a frequency-dependent feature—the 98th percentile of spectral cross power. We trained a model based on Linear Discriminant Analysis and applied the discriminator to the test set revealing a modest effect in the days leading up to an earthquake. While the observed effect size is not directly useful for earthquake prediction (long a scientific goal), it suggests a relationship which should be further investigated for a physical link.

Plain Language Summary We identified changes in the magnetic field near intermediate-large earthquakes in California in the days before the earthquakes happened. The significant signal is of modest size, which means that we can not directly provide a prediction that can be used to alert the public. This study provides evidence that there is a physical change that can be observed in the days before an earthquake, but further scientific study is needed to understand this process.

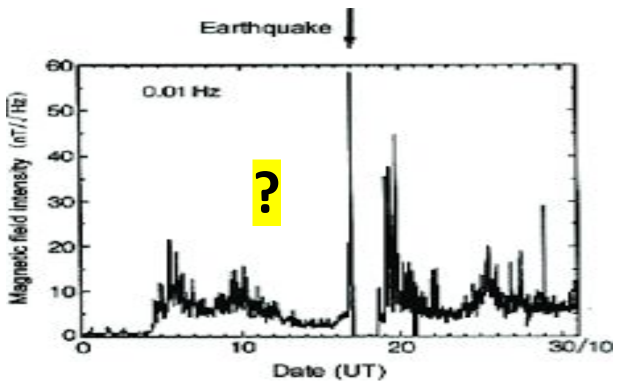
Research Team

- **QuakeFinder:** Project under Stellar Solutions Inc.
  - Private Satellite Systems Engineering company
  - Humanitarian R&D project, using System Eng.
  - 20 year effort, $25M (private+NASA+), 1-15 people (avg. 5)

- **Google Research, Applied Science Team**
  - Provided independent analysis of same QuakeFinder data
  - Use Google's big data tools and vast computing power
  - 2 years, 1B CPU hours, 1-5 people

Objective:

- Try a different approach to earthquake forecasting research
  - **Electromagnetic (EM)** rather than seismic monitoring
  - Motivated by reports of EM anomalies, i.e. Earthquake lights, Loma Prieta, myriad reports in the literature…
- **Are Short-term (days) forecasts even possible?**



Today's Presentation:

- Describe results of two recent publications leveraging the **QuakeFinder EM dataset in California**
- NOT to show that we can operationally forecast earthquakes
- NOT to address possible physical processes to generate ultra-low frequency (ULF) magnetic signals
- NOT to give a review of other international efforts underway

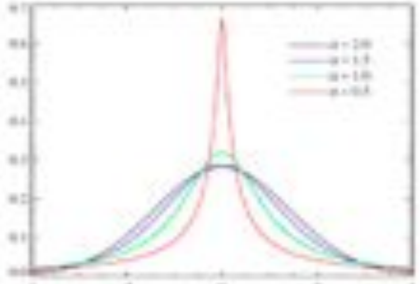
Current methods provide:

Seconds of warning

- Earthquake “Early” Warning (EEW) systems Based on seismic detection of a quake after it has occurred

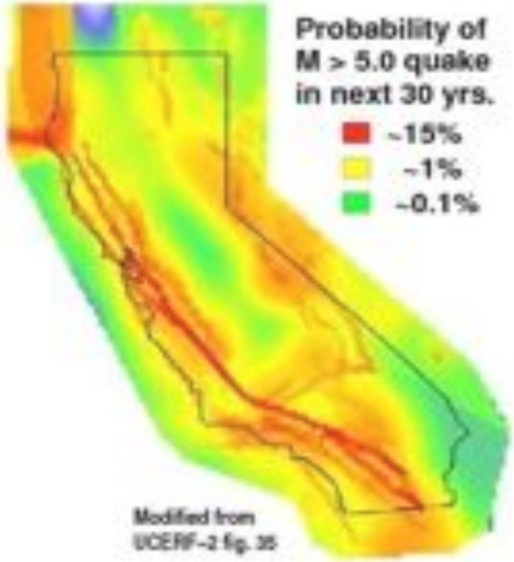
ShakeAlert

Days of warning ?



Decades of warning

- 30-year probabilities
- Based on statistical analysis



USGS

QuakeFinder Instruments and Network

QF Station installation with Solar Panels, and Communications.

- Continuous 50Hz Sampling
- Deployed along fault zones
- Data relayed by cellular network to QF datacenter
- Magnetometers buried 15 cm below surface
- 3 orthogonal coils at each site

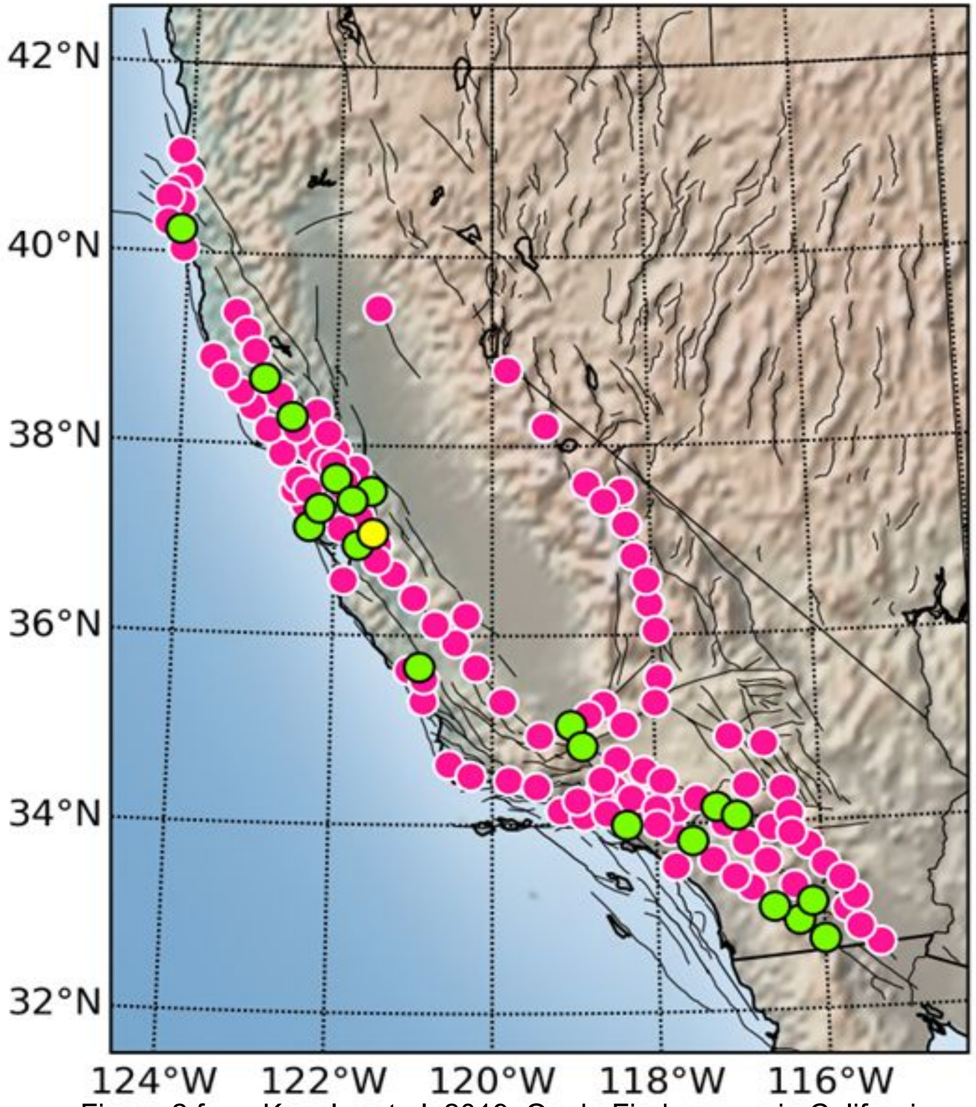
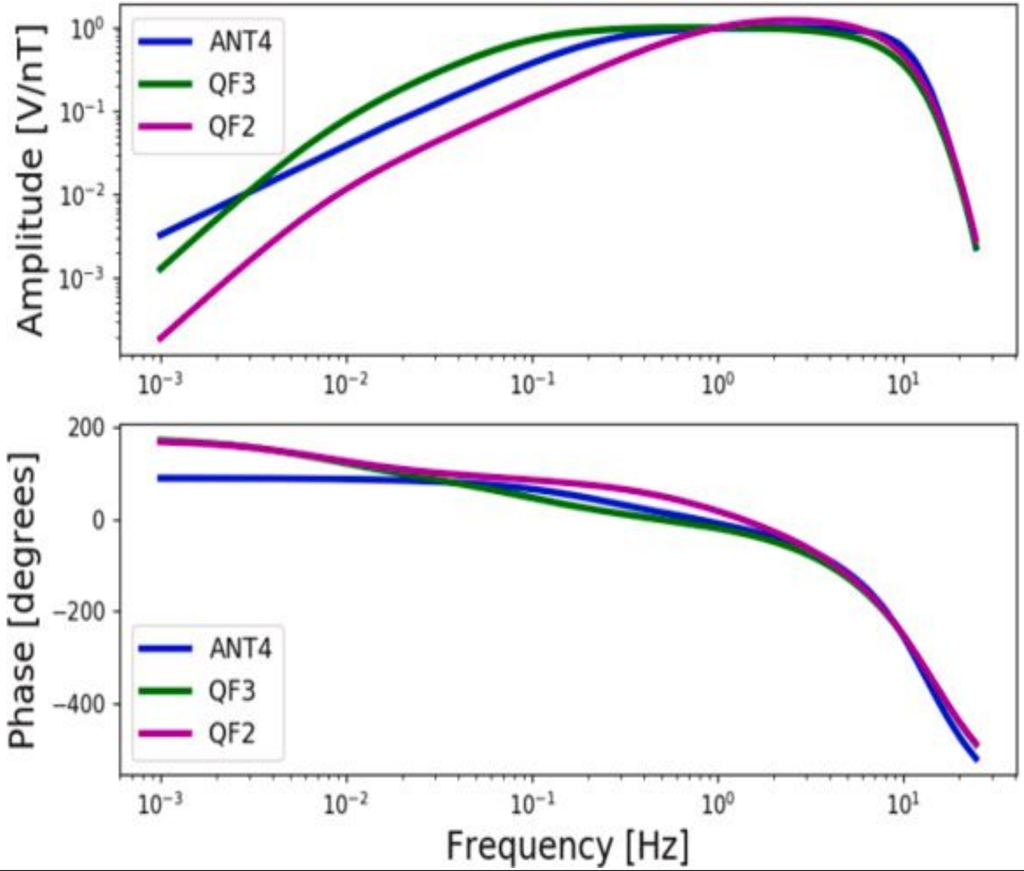
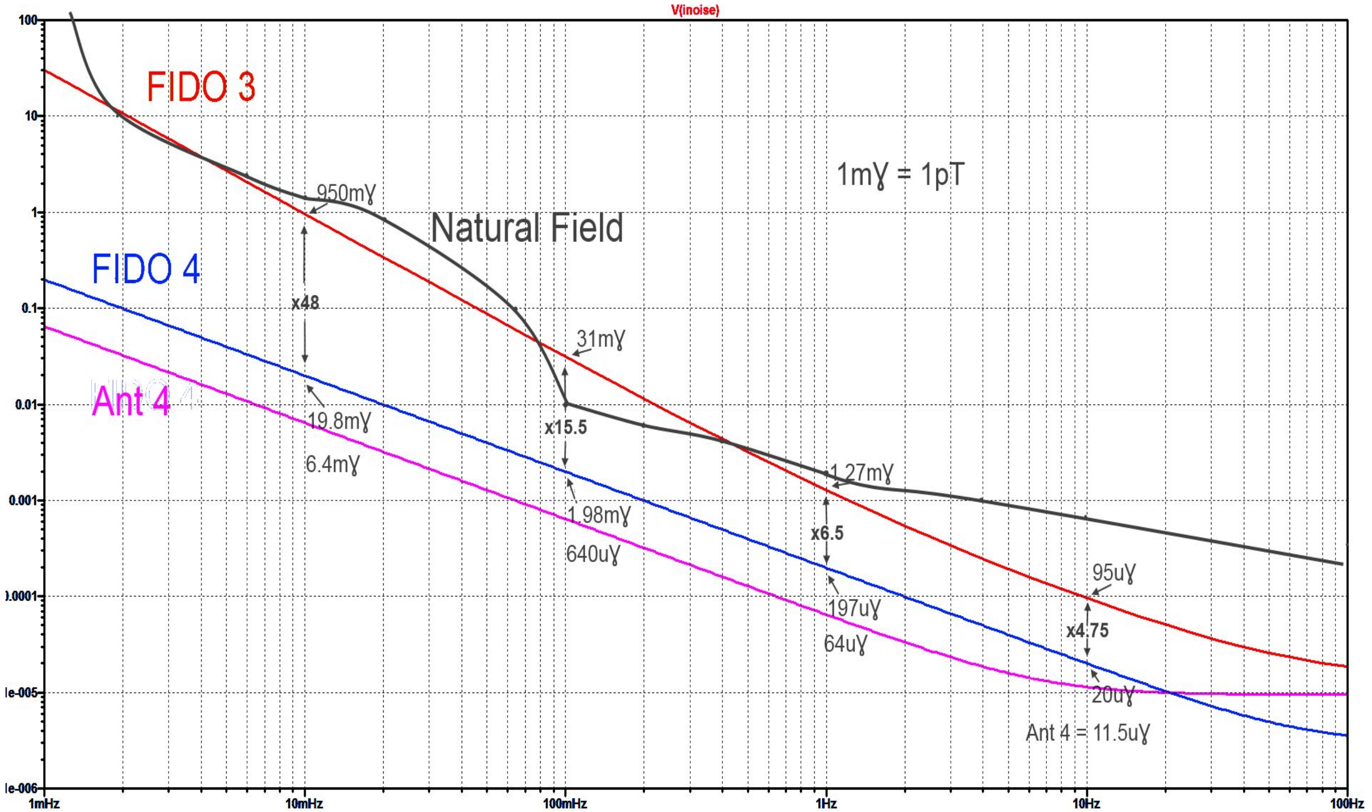


Figure 2 from Kappler et al, 2019: QuakeFinder array in California with 121 stations shown in magenta, 21 historical stations shown in green which were decommissioned, and station 859 shown in yellow which supplied data _in the manuscript_

QF Induction Magnetometer Total Response



Total response functions above including circuit board and digitizer.



Noise Floor Ant4: 100 fT per root Hz at 1 Hz

Three models of Induction coil in QF Array
- ~15% ANT4 (exploration grade sensors)
- ~85% QFIDO3: Noise level around the natural field amplitude

What do the data time series look like?

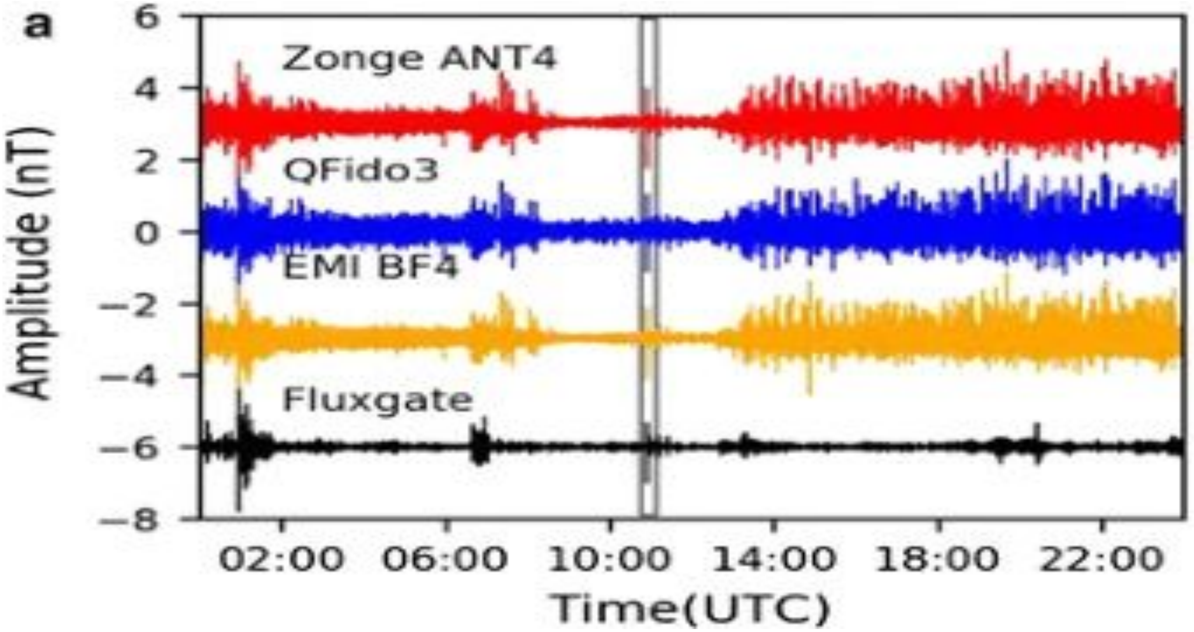


*Fig. **a** 24h time series with 200-s high-pass filter. Induction coils at Jasper Ridge, fluxgate at FRN.*

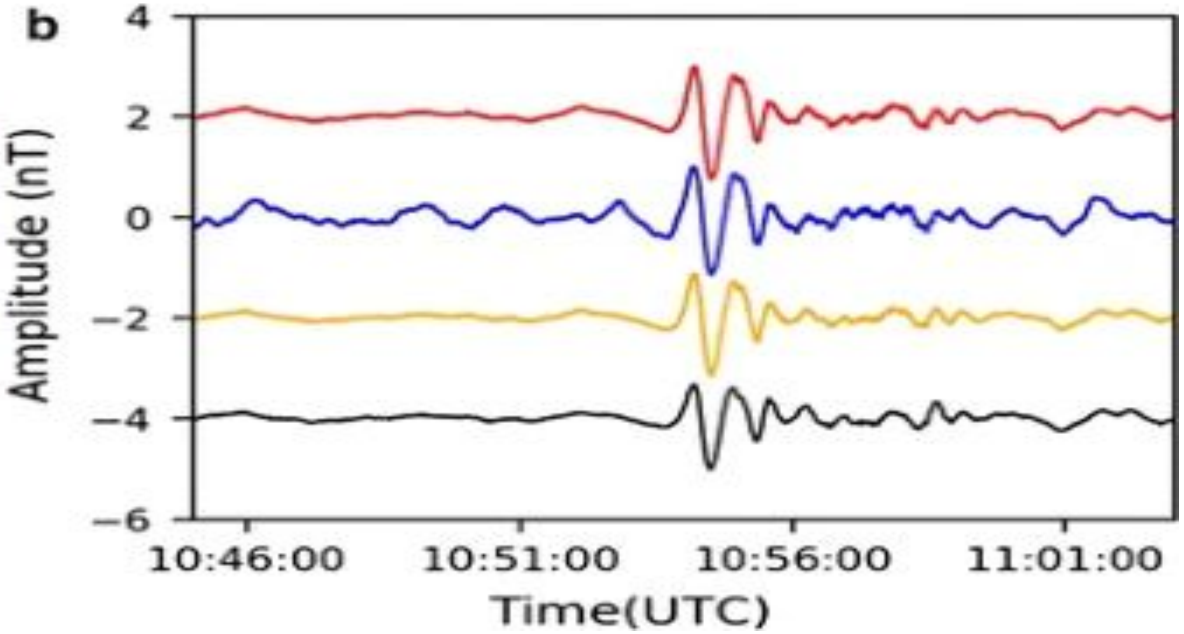
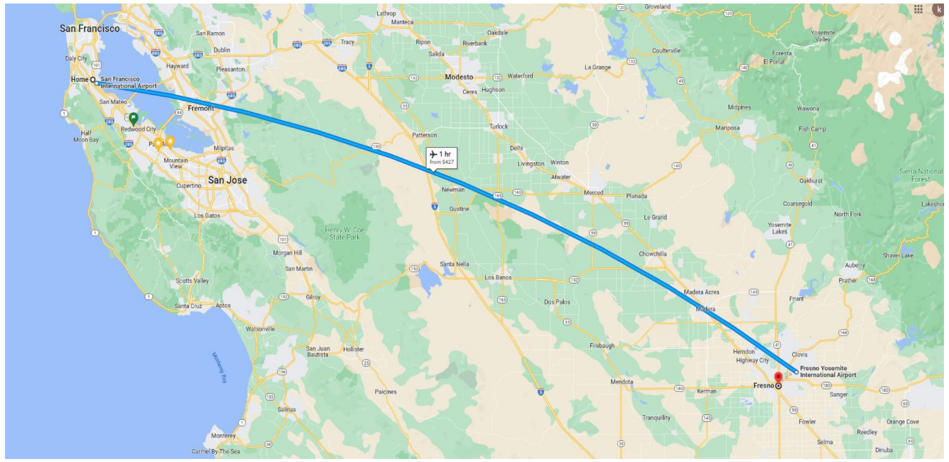


*FIg **b** Expanded view of gray box in **a**, ~ 20 min of data, showing a Pi2 irregular geomagnetic pulsation spanning ~ 3 min*

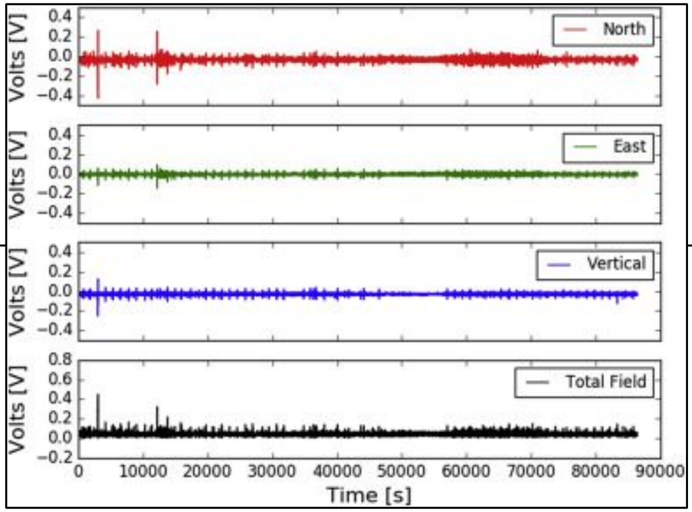
- Recorded fields dominated by:
  - cultural noise in urban areas
  - natural fluctuations in remote areas
    - driven by space-weather
    - coherent over hundreds of kilometers



SF Bay Area and Fresno California, separated by ~250 km

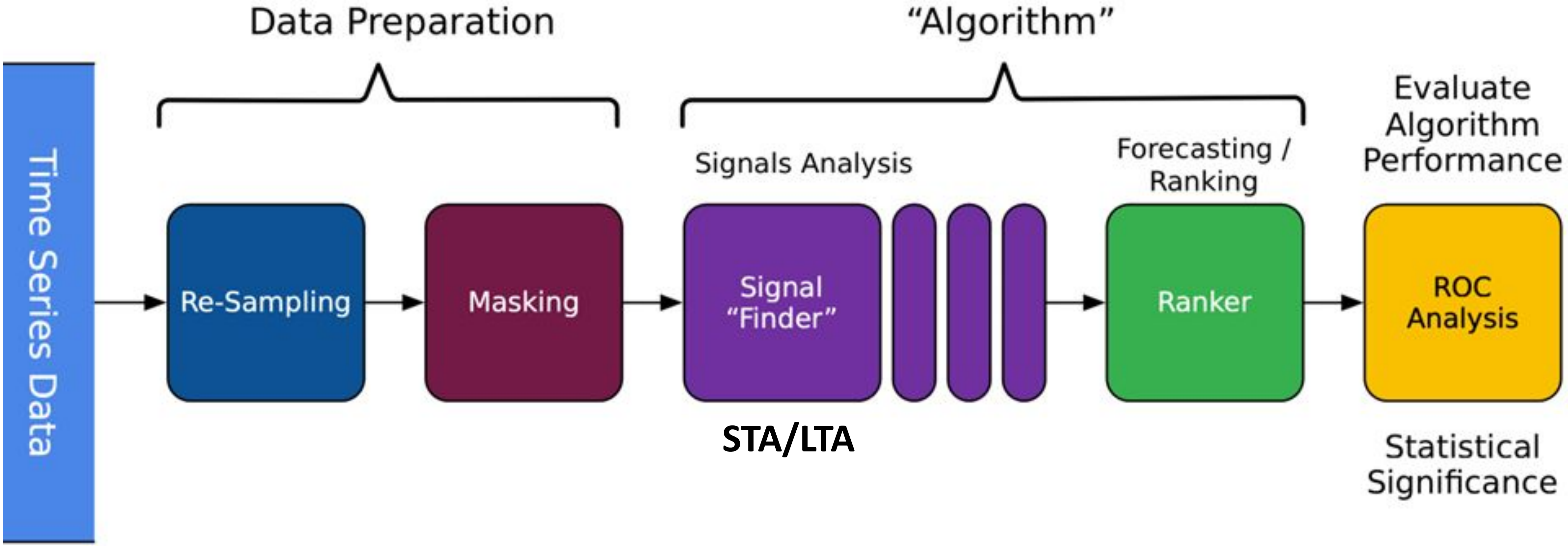
Wang, C., Bin, C., Christman, L. E., Glen, J. M. G., Klemperer, S. L., McPhee, D. K., Kappler, K. N., Bleier, T. E., & Dunson, J. C. (2018). Cross-validation of independent ultra-low-frequency magnetic recording systems for active fault studies. Earth, Planets and Space, 70(1), 57.

Data and Processes



- **Data**:
  - 14 years of 3-axis induction magnetometer data
    - 125 stations in CA.
    - 70 TB+ (32 and 50 sps)
  - Reduced to “total field amplitude”
- **Processes:** Detect Signals from Noise:
  - Short Term Average/Long Term Average (**STA/LTA**) for single station (QF)
  - 2-Station cross-spectral multiplication (Google)
- **Goal**: **<u>Statistically significant ULF magnetic signal prior to quakes?</u>**
- **Constraints**: Within Magnetic Signal Limits:
  - Greater than a quake minimum threshold
  - Within a threshold distance from the instrument sites

“An algorithmic framework for investigating the temporal relationship of magnetic field pulses and earthquakes applied to California”*



STA/LTA

*2019 Computers and Geosciences

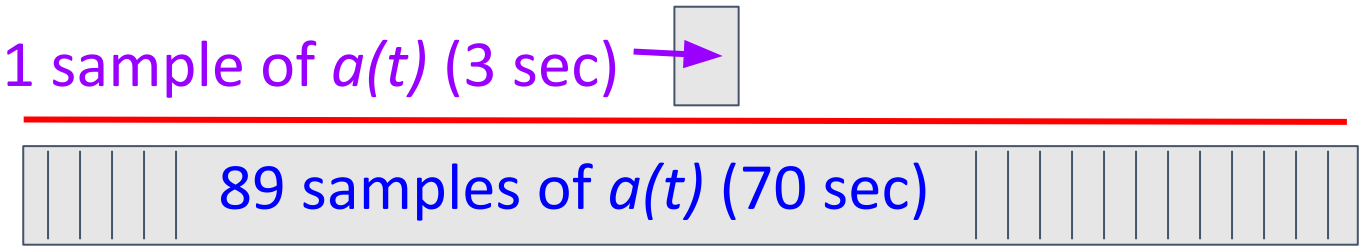
Key Concepts

- **Station-Day**
  - Unit of data reduction
  - After imputation, ~200,000 station days (from California stations)
- **Magnetic field “Pulse”**
  - A spurious transient signal in the magnetic field,
  - Stands out against the background time series
- **Pulse Counting**
  - Used STA/LTA filter for pulse counting, (short term/long-term → ~3s/70s)
  - A time series of daily “pulse counts” was created for each station
- **Normalized Pulse Counts**
  - Pulse Counts per station-day normalized by median of counts over previous 100 station-days
    - (allows inter-station comparison)
- **Ranking**
  - Each station-day was assigned a scalar value, its “Rank” (R)
  - R = Number of normalized magnetic field pulses over prior 4-12 days
  - Allows an ordering over all of the station-days to be applied
- **Hypothesis to Test**:
  - Increased normalized pulse counts for previous 4-12 days is a risk factor for earthquakes “nearby”

Processing Flow – 1 of 2

DATA REDUCTION & STA/LTA

- Each "Station-Days" transformed to "Total Field" (from NED, 3x reduction)
- Sliding Window Variance (3s window, 75% overlap)
  - 40x reduction of data, Results in "Magnetic Activity" *a(t)*

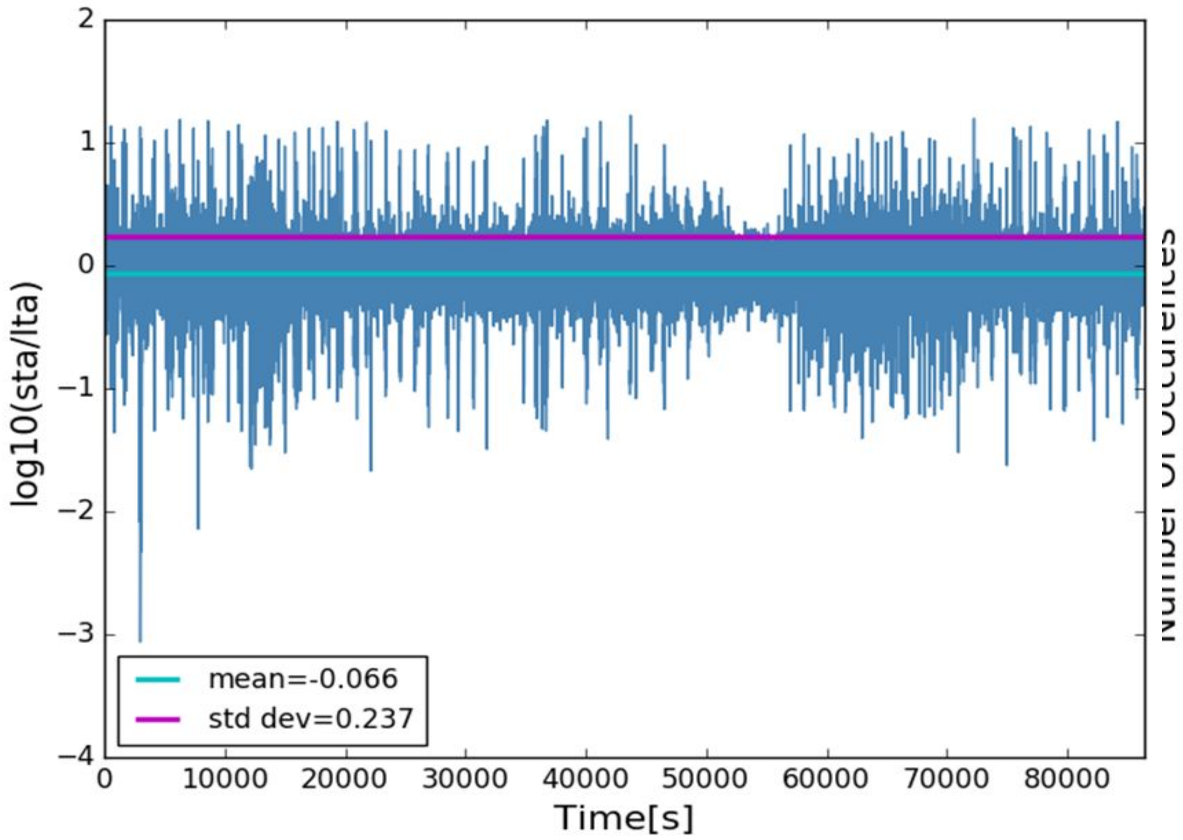


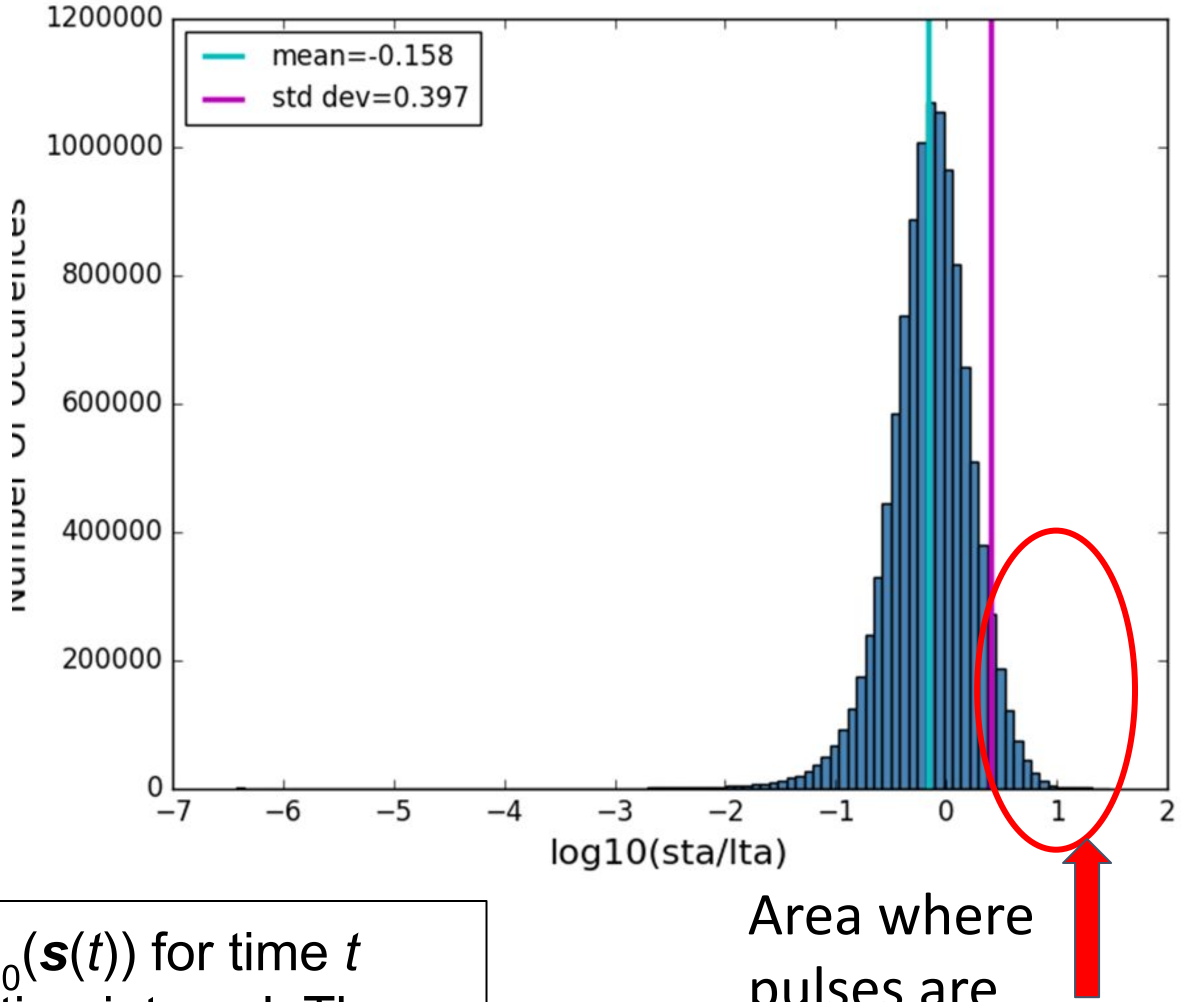
STA/LTA
Activity that stands out above background ("pulses")

FEATURE EXTRACTION

- Apply STA-LTA to ***a(t)***, obtaining ***s(t)***
- worked with distributions of $\log_{10}$(***s(t)***) (has a bell-shaped distribution)
- Apply threshold to the daily $\log_{10}$(***s(t)***) histograms to "count pulses"

Setting the pulse counting criteria





Area where pulses are counted

The threshold was based on the distribution of $\log_{10}(\boldsymbol{s}(t))$ for time t over the first 100 days at the beginning of each station-interval. The threshold computation required at least 98 days of valid data

Processing Flow – 2 of 2

PULSE COUNTS

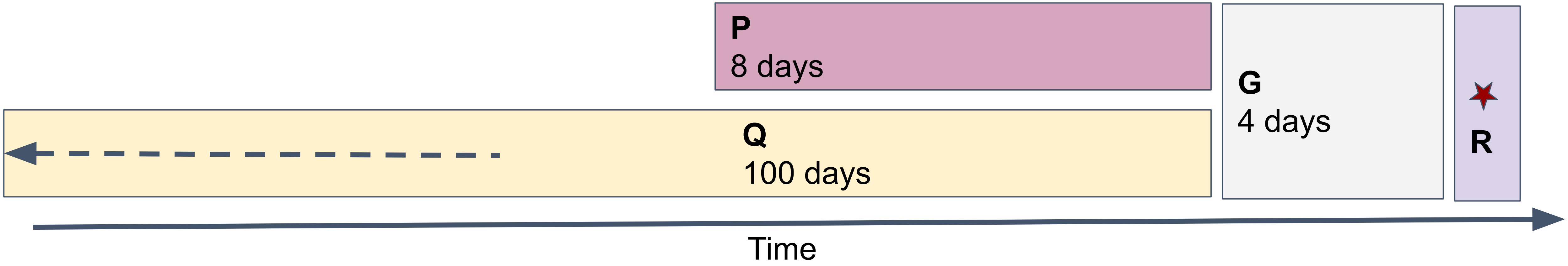
- Different background values for different stations
  - site-specific noise environments, urban vs. rural, etc.

- **NORMALIZED COUNTS PER STATION-DAY:**
  - normalize by the 100-day moving average
    - *allows inter-station comparison of counts*
  - Fundamental input feature to the algorithm

RANKING

- Associate with each Station-Day, a number (R) that represents the number of normalized pulse counts from the previous 4-12 Days

Comparison of Precursory (P) vs Quiescent (Q) Periods and Ranking



P: Mean “normalized pulse counts” per day in 8-day window shown above

Q: Median “normalized pulse counts” per day in the 100 days leading up to P

“Ranking” is a ratio: R = P/Q

R generated for *every* station-day, regardless of earthquakes

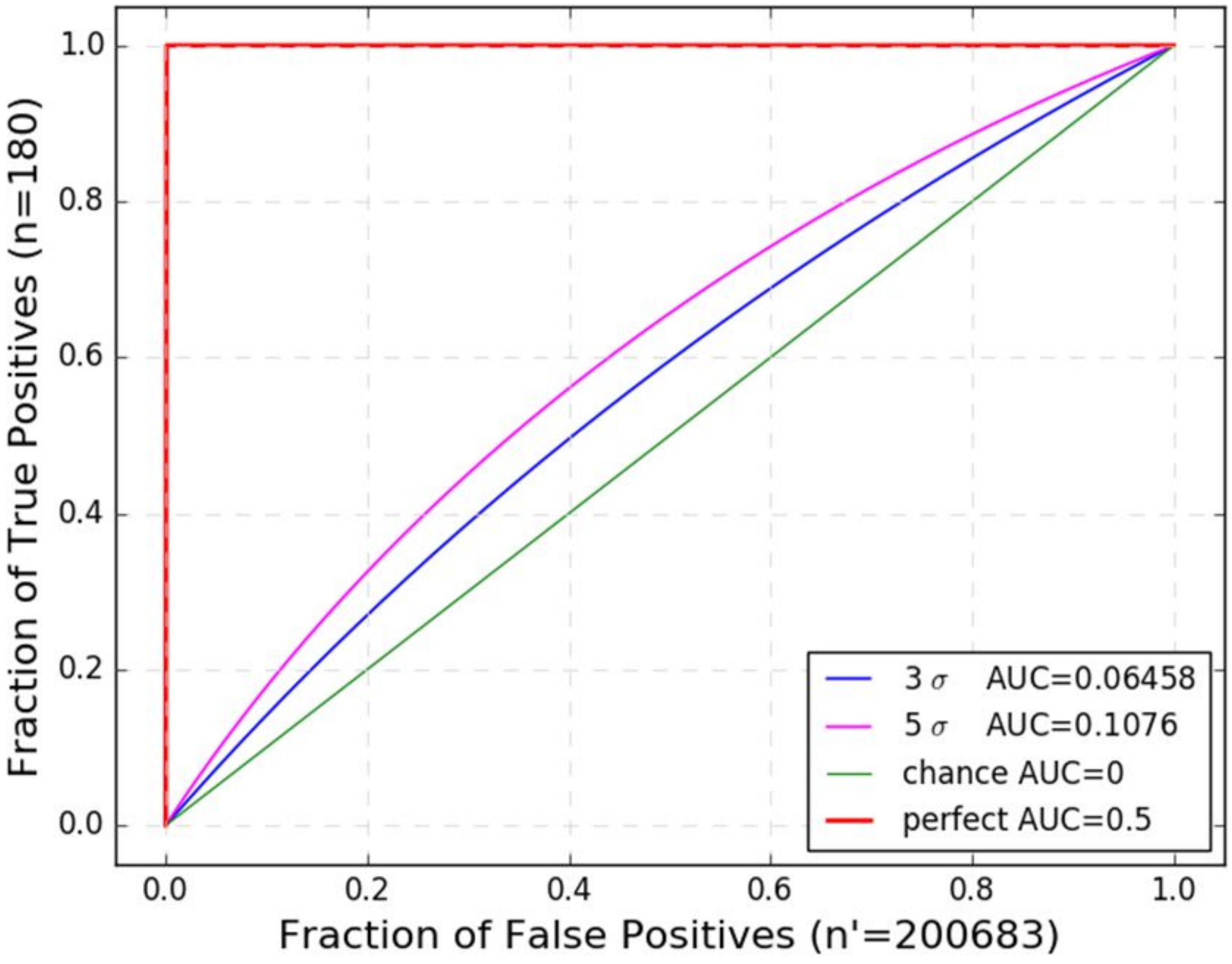
Are high Ranked station days more likely to have earthquakes?

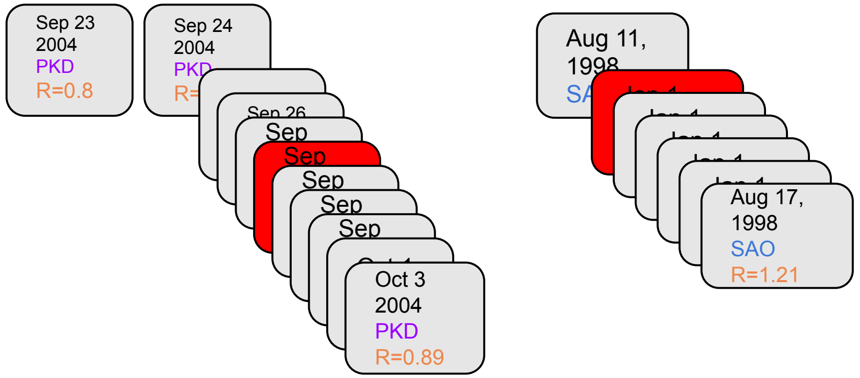
Resultant Data Packets

Oct 3, 2014,
Station 858
R=0.897

Mar 2, 2018,
Station 914
R=1.08

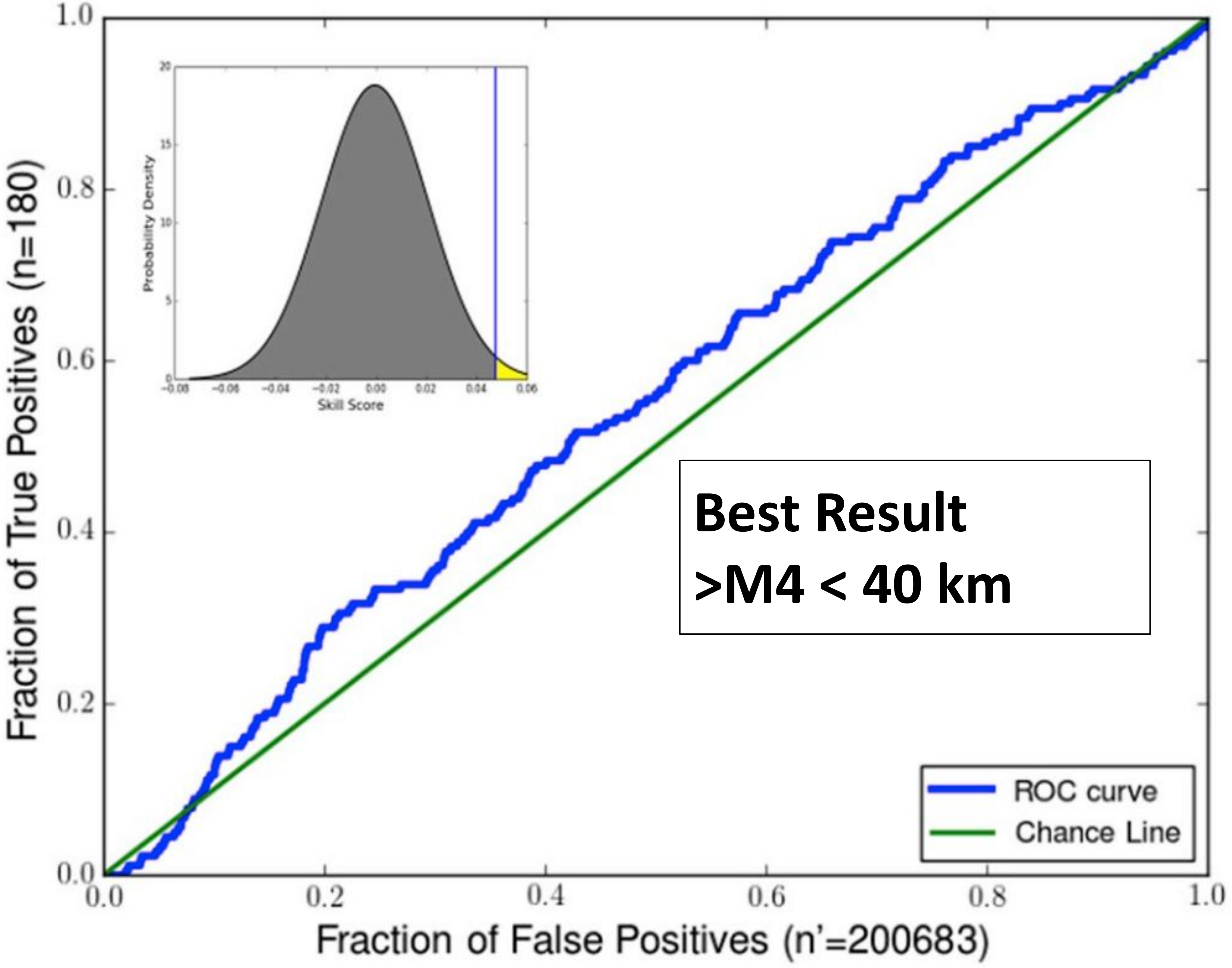
Hypothesis Testing:
Receiver Operating Characteristic - ROC Diagram





- One card per station-day
- "Mark" (red) cards if M>4.0, within 40 km
- Merge into a single "deck",
- Sort by ranking (R)

- **Null Hypothesis**: Station-day Rankings carry no information about future earthquakes nearby
- Thresholds on Magnitude and hypocentral distance used to associate earthquakes with stations
- **If Rankings tend to be high on days with earthquakes near station, ROC area under curve (AUC) increases**
- AUC directly maps to a "Z-score" (σ).
- AUC → Z depends on number of earthquake and non-earthquake samples (station-days) in dataset



Initial z-score: 2.20σ

Noisy stations removed: 2.86 σ

<u>Improvements: 3.06</u>σ

(after publication)

Updated quake catalog

Removed Noisy Stations near Pacific DC Intertie 1M volt power line nearby

As Published

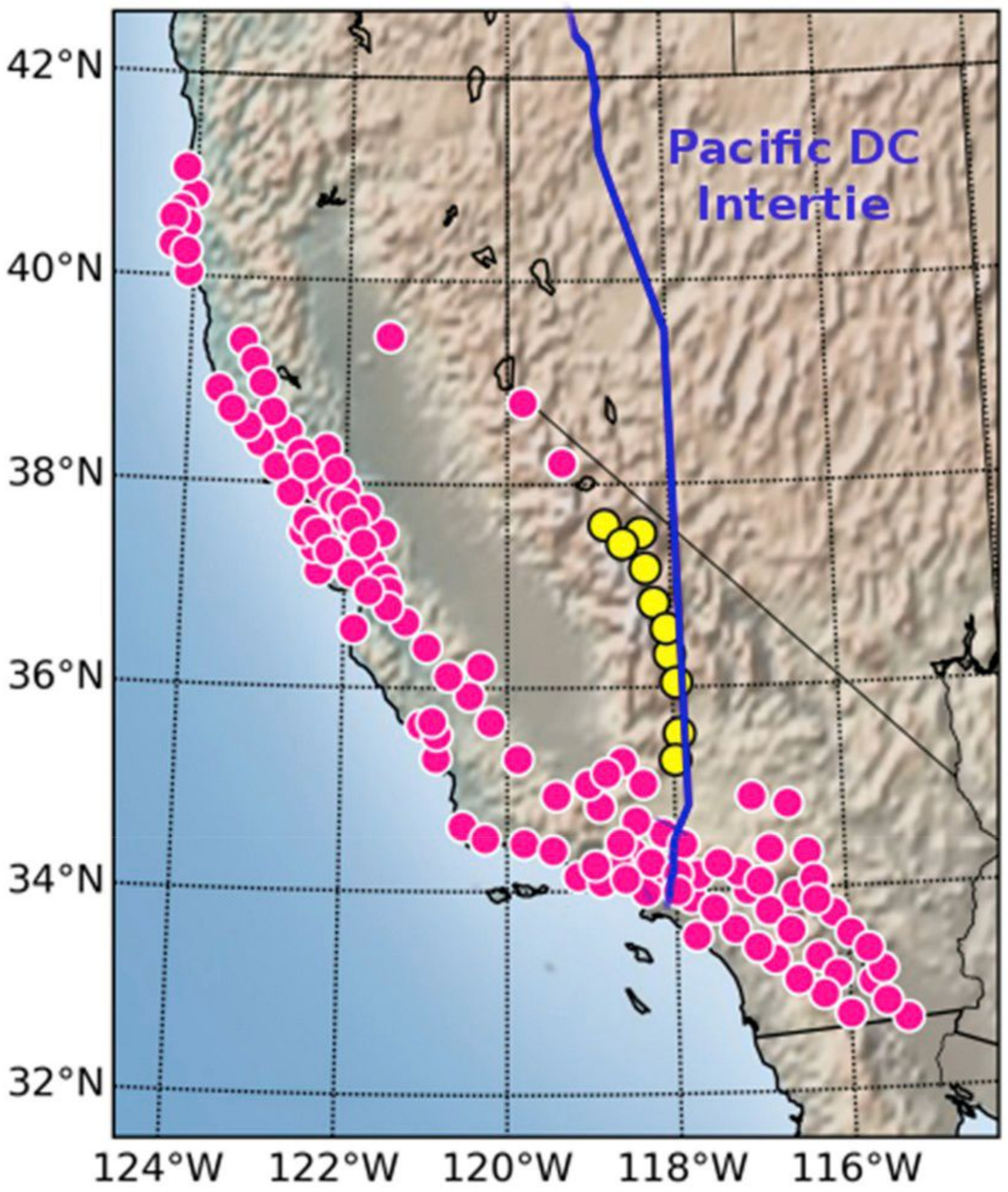
2.2σ / 2.86σ

with /without PDCI stations

Updated USGS Catalog:

2.4σ / 3.06σ

with /without PDCI stations

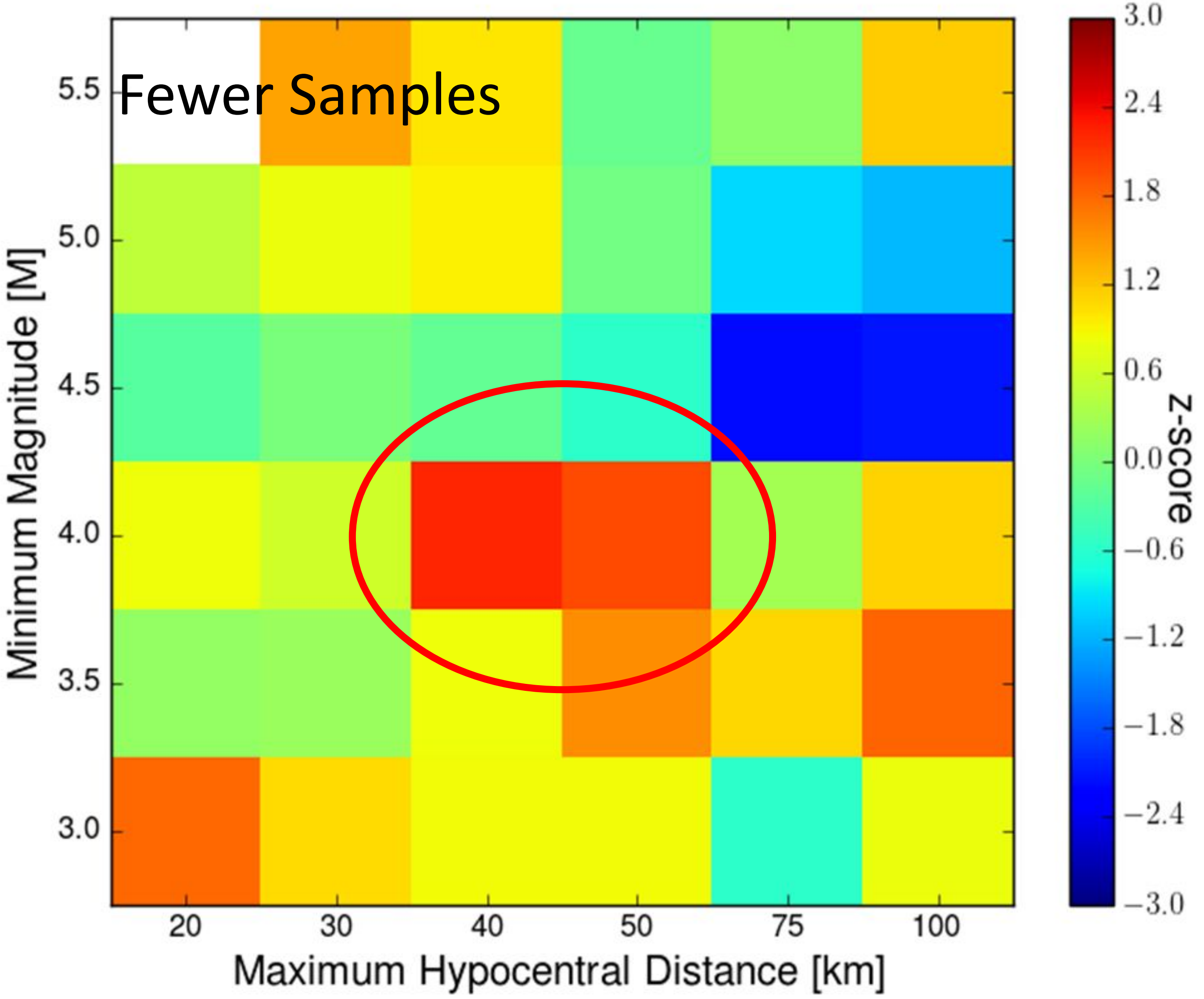


Distance & Magnitude Sensitivity

- PDCI not removed
- Dark blue attenuates significantly when excluding PDCI stations
- Best Score >M4, <40km

Table 1
Number of positive, 'earthquake-station-days', for each colored cell in Fig. 10.

Minimum Magnitude (M)							
	5.5		3	4	10	17	29
	5	3	11	14	28	59	96
	4.5	6	20	30	58	124	205
	4	47	98	180	295	587	899
	3.5	197	383	691	1051	2031	3061
	3	620	1273	2142	3179	6150	9287
		20	**30**	**40**	**50**	**75**	**100**
		Maximum Hypocentral Distance (km)					



Summary: 1st paper "Framework…"

- Total field, 1 station at a time, STA/LTA Pulse finder, ROC Analysis
- Station rankings provides 4 day lookahead
- Results: 2.2**σ**, updated to 3.06**σ** with noisy sites removed and updated quake catalog
  - Suggesting that a **4-day forecast** based on this algorithm does have predictive power, i.e. it is a valid risk identifier
- Need for independent validation

JGR Solid Earth

RESEARCH ARTICLE
10.1029/2022JB024109

Case-Control Study on a Decade of Ground-Based Magnetometers in California Reveals Modest Signal 24–72 hr Prior to Earthquakes

William D. Heavlin[1], Karl Kappler[2,3], Lusann Yang[1], Minjie Fan[1], Jason Hickey[1], James Lemon[3], Laura MacLean[3], Thomas Bleier[3], Patrick Riley[1], and Daniel Schneider[3]

[1]Google Research, Applied Science Team, El Granada, CA, USA, [2]Imdex Technology USA LLC, San Luis Obispo, CA, USA, [3]QuakeFinder, Palo Alto, CA, USA

Key Points:

- Frequency domain analysis of ground-based magnetometer data shows a modest change in days leading up to intermediate-large ($M \geq 4.5$) earthquakes
- One novel part of the analysis is the use of cross-power signals, which combines the signals from instruments separated by tens of km
- A supplementary analysis of the data with first order global geomagnetic effects subtracted increased the measured effect size significantly

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Schneider,
dschneider@quakefinder.com

Citation:

Heavlin, W. D., Kappler, K., Yang, L., Fan, M., Hickey, J., Lemon, J., et al. (2022). Case-control study on a decade of ground-based magnetometers in California reveals modest signal 24–72 hr prior to earthquakes. *Journal of Geophysical Research: Solid Earth*, *127*, e2022JB024109. https://doi.org/10.1029/2022JB024109

Received 27 JAN 2022
Accepted 19 AUG 2022

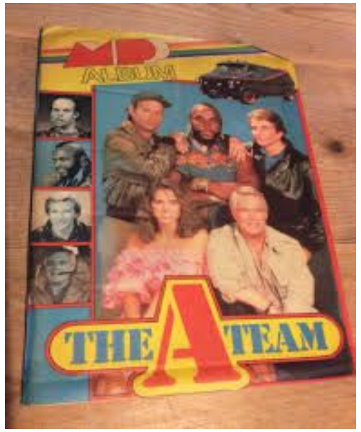
Abstract Magnetic field changes as earthquake precursors have been the subject of numerous studies and some controversy. Infrequent large earthquakes and sparse magnetometer coverage along fault zones complicate statistical analysis. We present an analysis of ground-based magnetic time-series measurements before 19 earthquakes ≥M4.5 in California drawing from over 330,000 site-days of measurement spanning a decade. To perform a fair existential test for electromagnetic antecedents we applied a pre-specified statistical analysis with two key ideas. First, we combine signals from nearby (≤40 km) sites via spectral cross-power, and then look for large spikes in frequency domain (0.016–25 Hz). The former is only possible with a dense set of sites running over a long period of time. In this statistical case-control study we used the machine learning concept of rigorously separated train and test sets of earthquakes which were generated via a rule-based query of the USGS earthquake catalog. Before each declustered earthquake, we constructed one period 24–72 hr before (the "precursor" or "*p*-period") and a series of seven equally-sized preceding periods ("quiescent" or "*q*-periods"). We distilled the data in each period to a frequency-dependent feature—the 98th percentile of spectral cross power. We trained a model based on Linear Discriminant Analysis and applied the discriminator to the test set revealing a modest effect in the days leading up to an earthquake. While the observed effect size is not directly useful for earthquake prediction (long a scientific goal), it suggests a relationship which should be further investigated for a physical link.

Plain Language Summary We identified changes in the magnetic field near intermediate-large earthquakes in California in the days before the earthquakes happened. The statistical signal is of modest size, which means that we can not directly provide a prediction that can be used to alert the public. This study provides evidence that there is a physical change that can be observed in the days before an earthquake, but further scientific study is needed to understand this process.

Second Manuscript

Timeline:

-2017 QF meets with GAS & presents research



- 2017-2019 Analysis
- 2020 Writing
- 2022 Publication

Key Concepts

- Same data, but work in **Frequency domain**
- Analyse only earthquakes that were close to **2 stations**
  - Spectral **Cross-power** amplifies common signals
- Employ **Train/Test Split** to avoid overfitting
- **Case-Control Framework**
  - Each earthquake defines 1 **P**recursor & 7 **Q**uiescent periods
- Feature Extraction Defines **P-features**, **Q-features**
- **Linear Discriminant Analysis (LDA)** separates **P** from **Q** on Training Set
- Directly apply same classifier on the Test Set
- **Results for Test Set**: Initial z-score 2.1**σ** (Modest)
- Recognize **natural fields could influence results**
  - detrend with respect to global geomagnetic activity index (Ap)
  - Results: **3.7, 4.4, 4.9 σ**

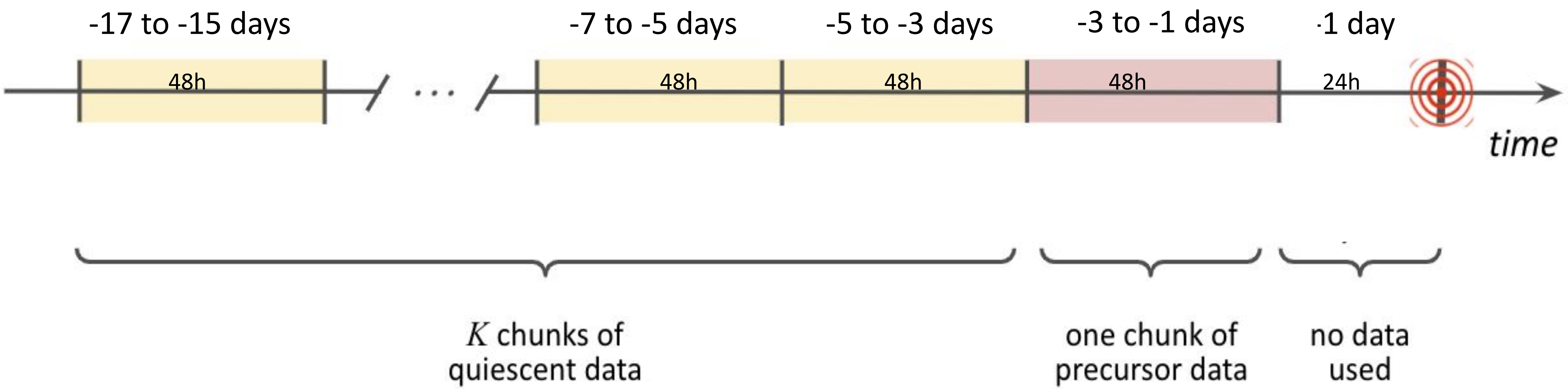
The Case-Control Framework

How The Problem was Approached

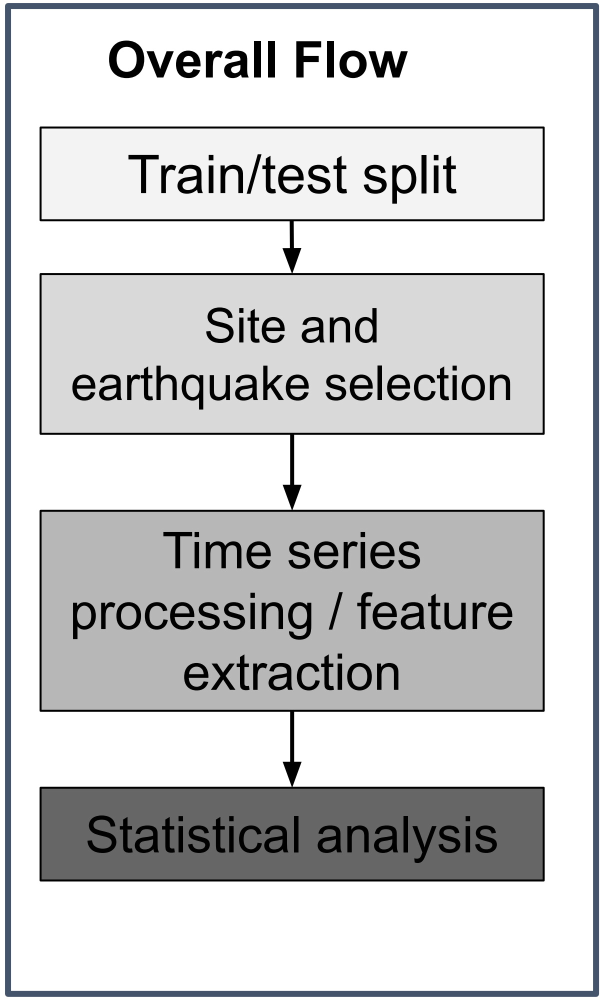
- Given an earthquake, was there a change in the magnetic field “just before” (24-72h) it occurred?
- We hypothesize a “Precursor” period before each earthquake in the study
  - For each **P**recursor (case), we hypothesize 7 “**Q**uiescent” (control) Periods
- Statistically measure the difference in the data between the **P** and **Q** periods

Advantages of the Case-Control Approach

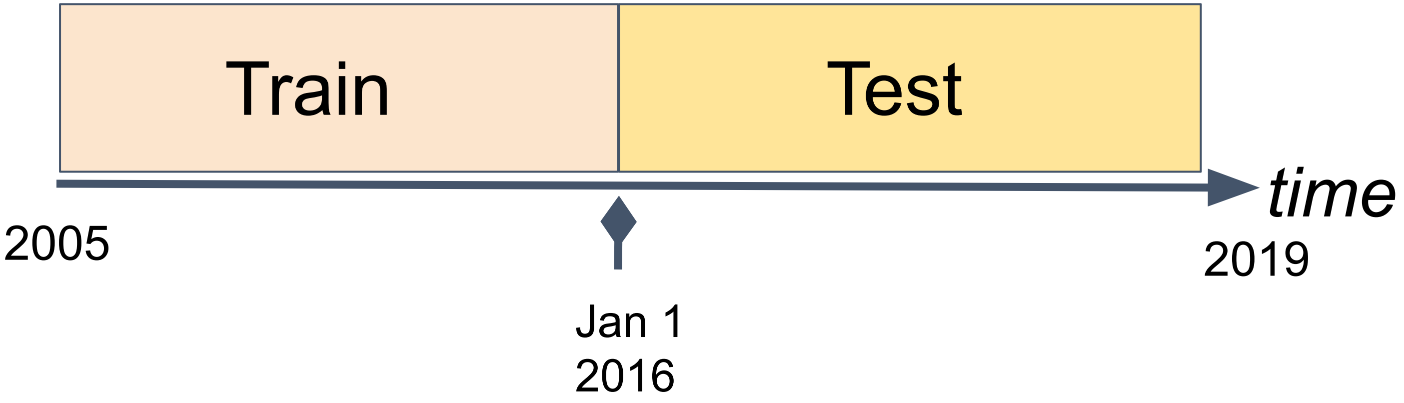
- **Data Reduction**: Orders of magnitude less data to process/analyse
- **Controllability**: Focus on relatively short term before earthquake, control for geography and long-term effects
- **Rare Events**: desensitizes the analysis to the relative rarity with which earthquakes occur
- **Visualizable**: Resulting data structure permits visual analysis of data



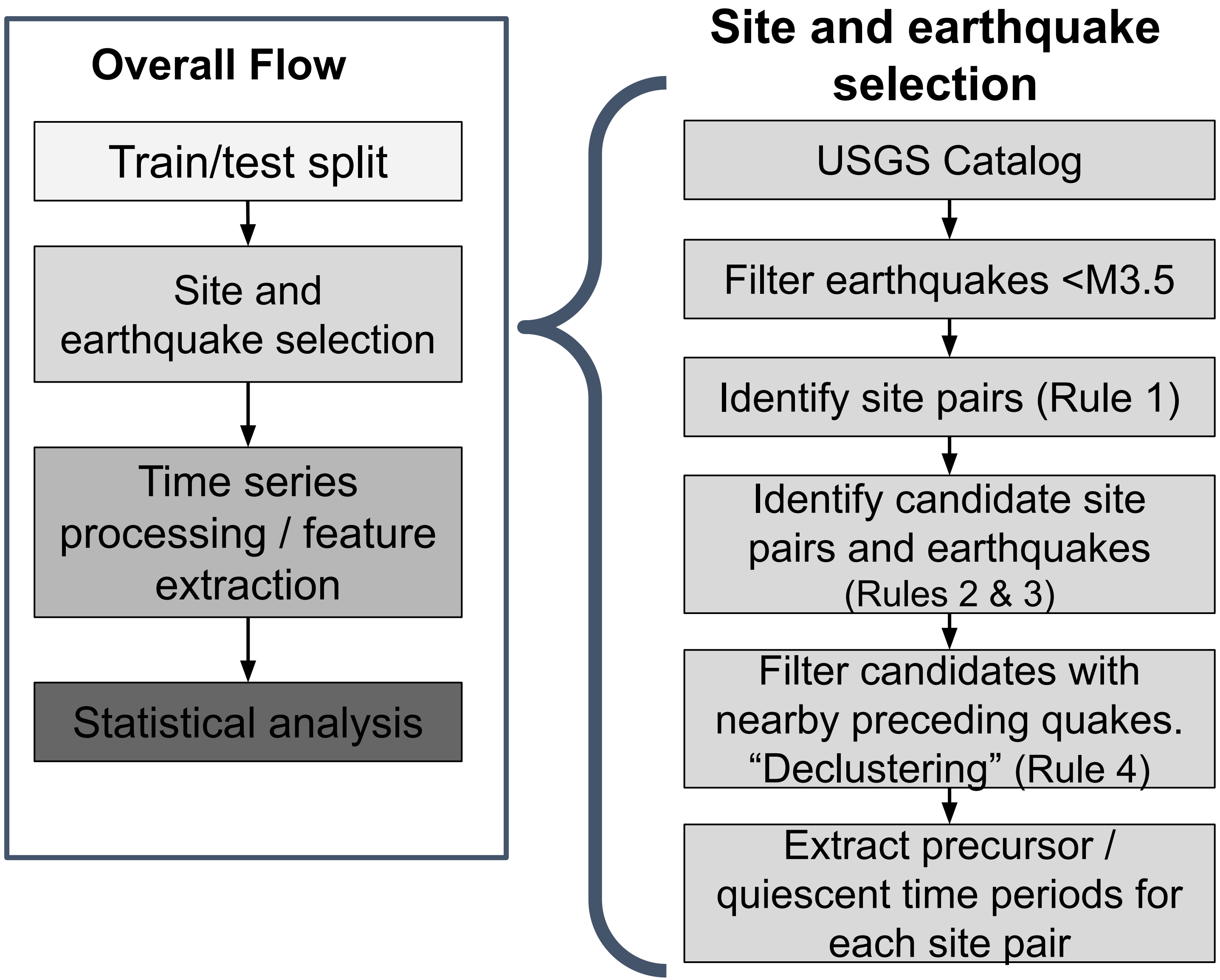
Data Processing Flow



- Approximately split the data in half (volumetrically)
- Train on data *before*, **January 1 2016**, Test on data *after*
- “Natural” split
    - previous data used to train a model, applied to future observations
- Used USGS Earthquake Catalog & QF station locations (but no QF data)
- Reporting results for *only the test set* helps to prevent model overfitting



Site & Earthquake Selection Workflow



Tuning parameters

θ, M_0, Δ_M

M_0	**The minimum magnitude threshold (tuning parameter)**
Δ_M	**Threshold applied to limit the magnitude of the maximum earthquake allowed in a quiescent period**
θ	**Characteristic qualifying distance between earthquake and stations, or between earthquake and station-pair midpoint**

Rules 1-4 are defined set theoretically in Supporting Table S10

“Tuning Parameters” for each of the three “tuning conditions” considered in the study

			Selected Tuning Conditions		
Parameter	**Symbol**	**Values considered**	**Blue**	**Channel**	**Flathead**
period length	λ	1-3 days	48h	48h	48h
buffer period	β	1h, 24h	24h	24h	24h
number of quiescent chunks	K	7	7	7	7
maximum characteristic distance	θ	20 km, 30km, 40km	30km	40km	40km
minimum magnitude	M_0	M3.5, M4.0, M4.5, M5.0	M4.5	M5.0	M4.5
magnitude threshold for interference	Δ_M	M0.0, M0.5	M0.0	M0.0	M0.0
percentile threshold	q	98%, 99%	98	98	98
Amount of data used		Total unique			
#SSE (training)		55	18	23	54
#earthquakes (training)		10	6	3	9
#SSE(test)		60	20	22	59
#earthquakes (test)		9	7	4	9

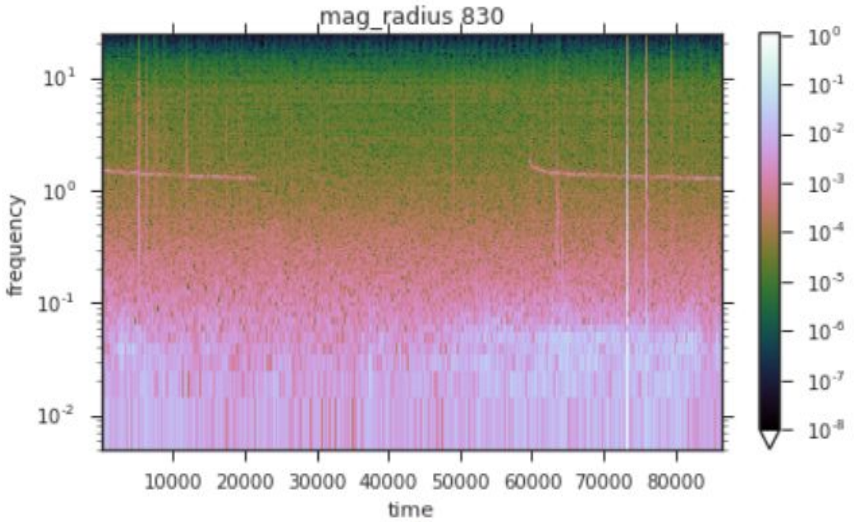
Overall Flow

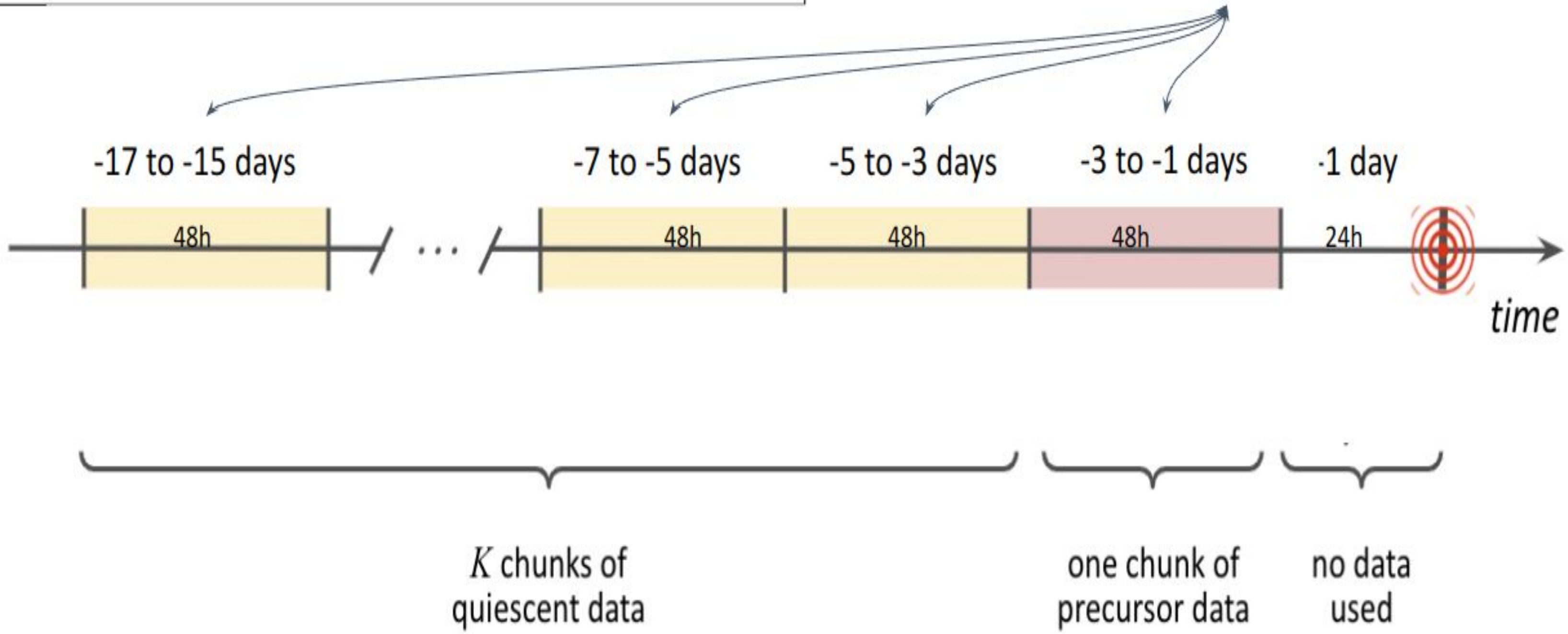
1. Train/test split
2. Site and earthquake selection
3. Time series processing / feature extraction
4. Statistical analysis

Feature Extraction

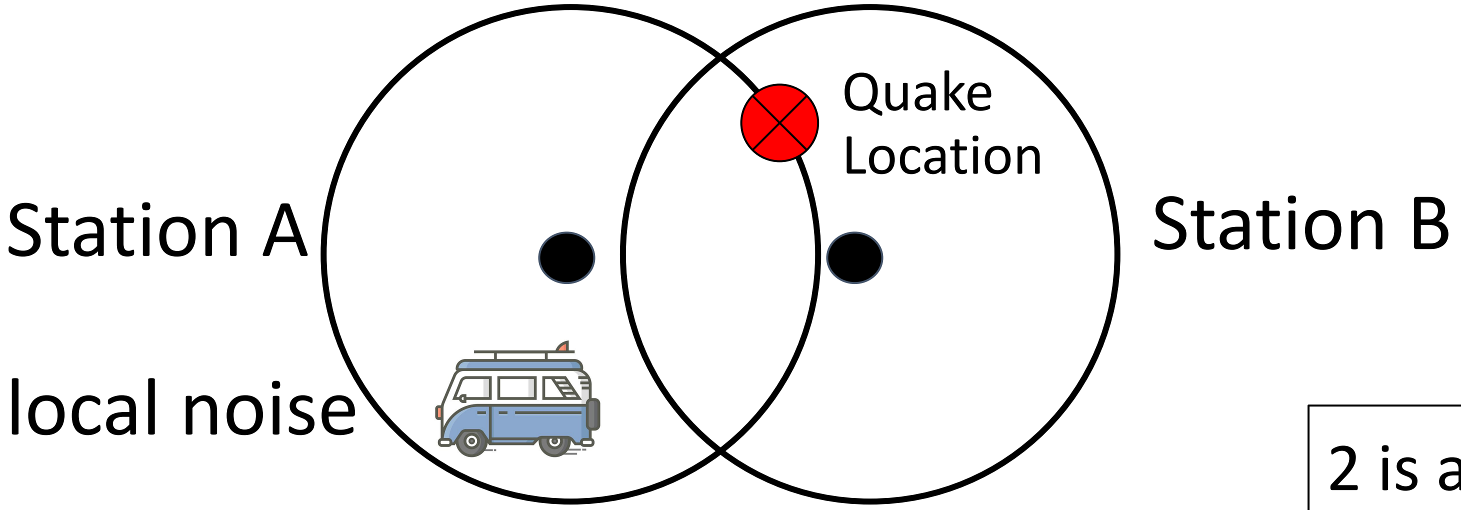
For each period ("p-period", and "q-period"):

- Compute the **spectrogram** (85 frequency bins)
  - within each frequency bin:
    - extract the 98th percentile of the log of the spectral cross-power amplitude.





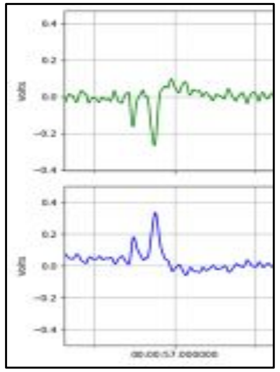
Example of Spectral Cross Power (Multiplication)
(Amplify Simultaneous Signals)



Spectral Amplitude (@ Hz)

Normal noise	2	x	2	4
One pulse	7	x	2	14
Simultaneous pulses	7	x	7	49

2 is avg background
7 is sample pulse level



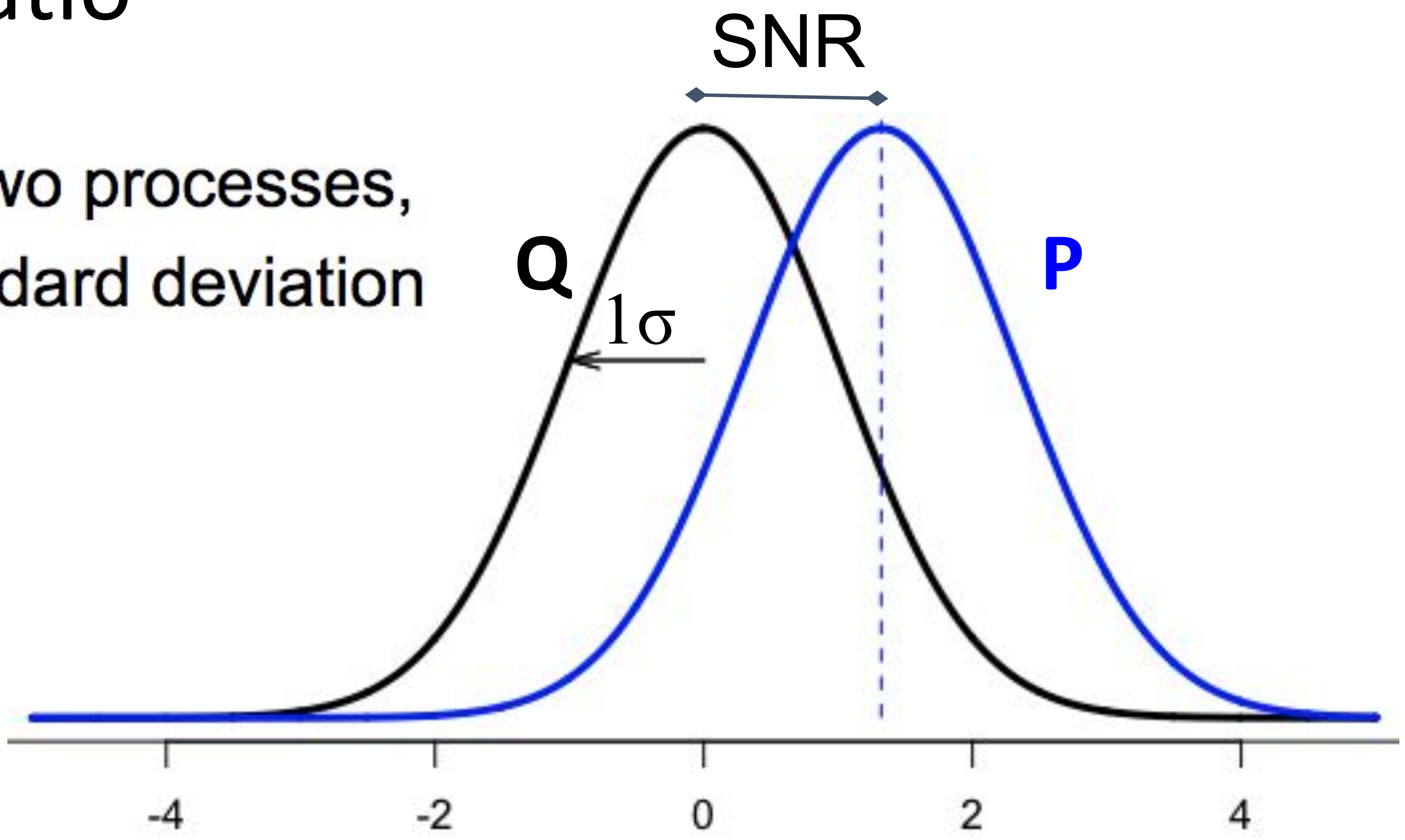
Slide courtesy of tbleier@quakefinder.com

signal-to-noise ratio

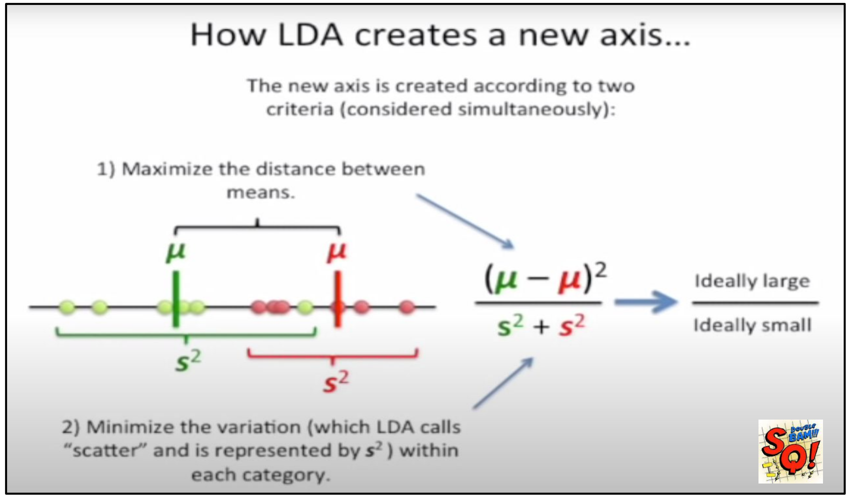
μ_1, μ_2, σ : means of two processes, common standard deviation

$$\text{SNR} = \frac{\mu_2 - \mu_1}{\sigma}$$

in standard deviation units, the change in means



Linear Discriminant Analysis (LDA)



Above Image from StatQuest

LDA seeks to optimize the **SNR**

$$\mathrm{SNR} = \boldsymbol{b}^{\top}\boldsymbol{d} / [\boldsymbol{b}^{\top}\mathbf{S}\,\boldsymbol{b}]^{1/2}$$

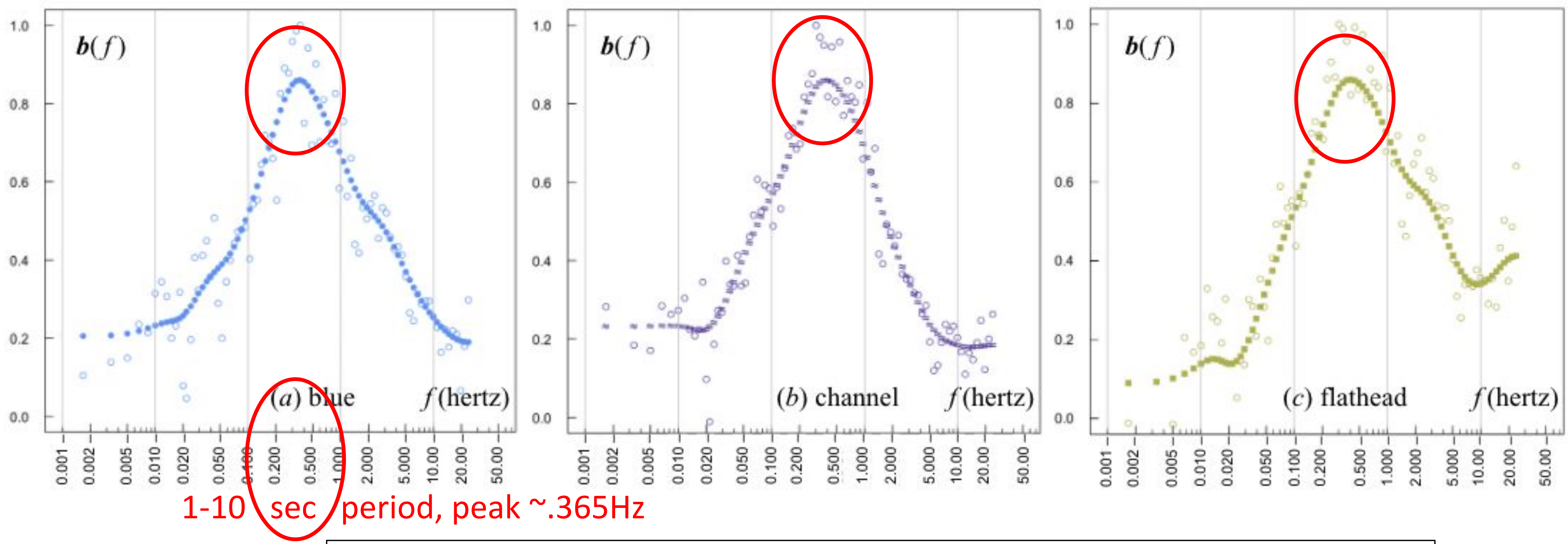
- ***d*** denotes an 85-vector of average differences between p- and q- periods
- ***S*** is the 85x85 pooled within-SSE covariance matrix

- Classical LDA coefficients are estimated as proportional to $\mathbf{S}^{-1}\boldsymbol{d}$.
- Requires regularization, e.g. $\mathbf{S} \leftarrow \mathbf{S} + \mu\mathbf{D}$, for some positive semidefinite matrix **D**, to successfully invert the matrix **S**.

Instead, we choose this simplification: take as the coefficients ***b*** as the difference vector ***d*** itself.

- This choice makes our approach more easily replicable
  - less dependent on the details of an algorithm constructing a matrix **D** and choice of constant μ.
- The curves implied by the two vectors ***d*** and $(\mathbf{S}+\mu\mathbf{D})^{-1}\boldsymbol{d}$ look quite similar (see Figure S9);
- This choice certainly reduces the signal-to-noise ratio on the training set, so in this sense, it can be considered both scientifically conservative and statistically suboptimal.

Training Set defines LDA coefficients



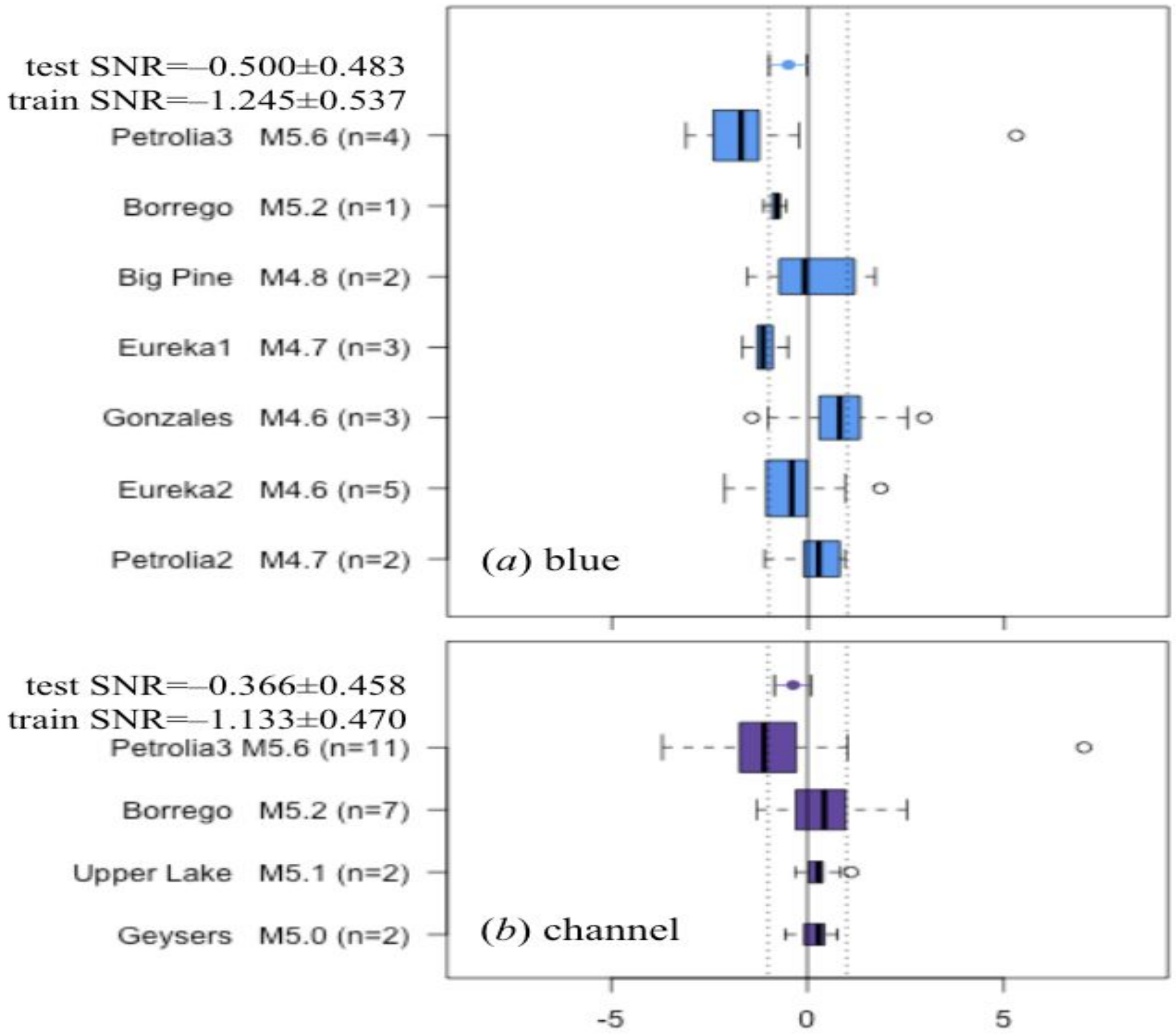
These coefficients (weights) are the pooled average Q-P differences over the training dataset

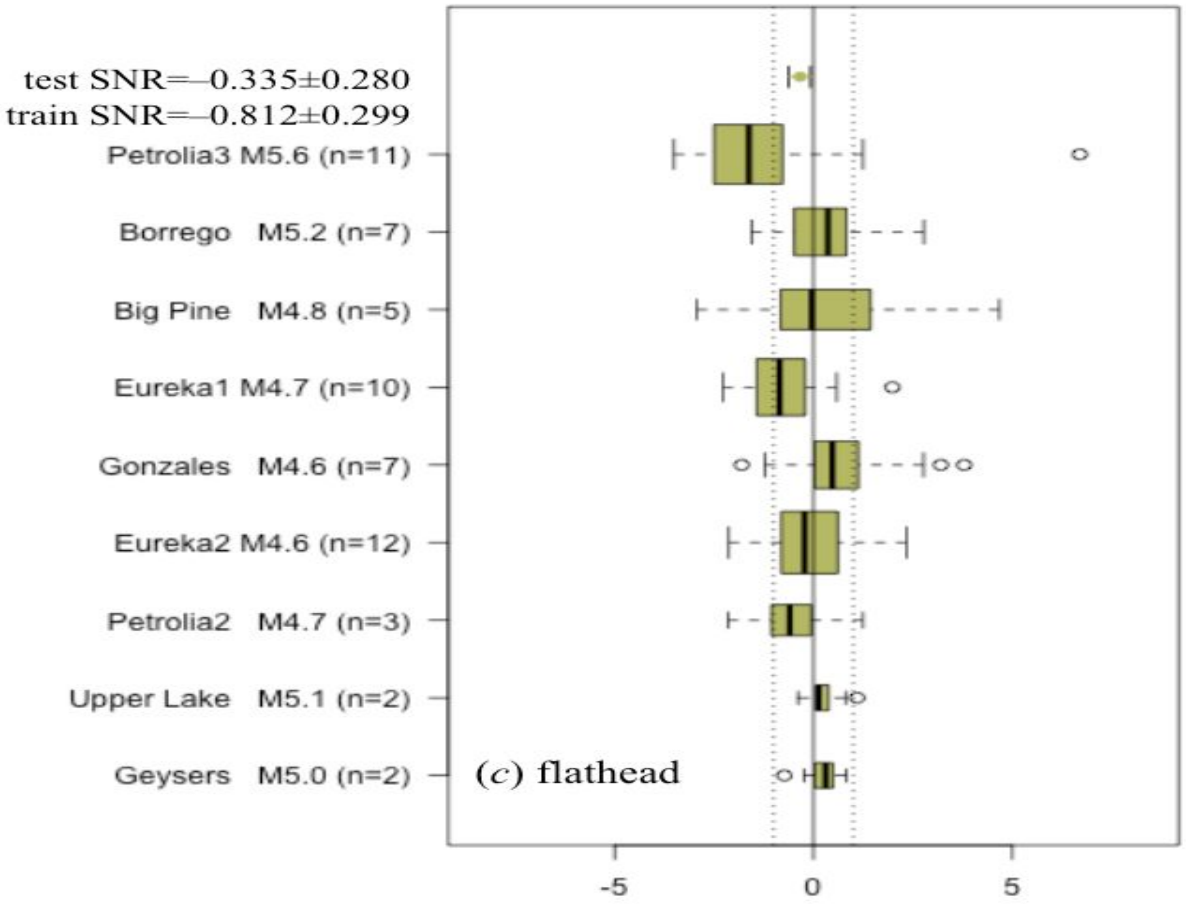
- Larger weights correspond to frequencies where separation of P and Q was better

Sidenote: Makes physical sense

- Maximum weight corresponds to the "MT Dead Band" - where natural fields tend to be smallest amplitude

Test Set Results





Sidenotes:

- Topmost interval (Tie-Fighter) is test SNR ± 2 standard errors
- x-axis units are SNR
- Height of box is proportional to square root of number of observing station pairs
- Box Left of center: P > Q
- Box Right of center: Q < P

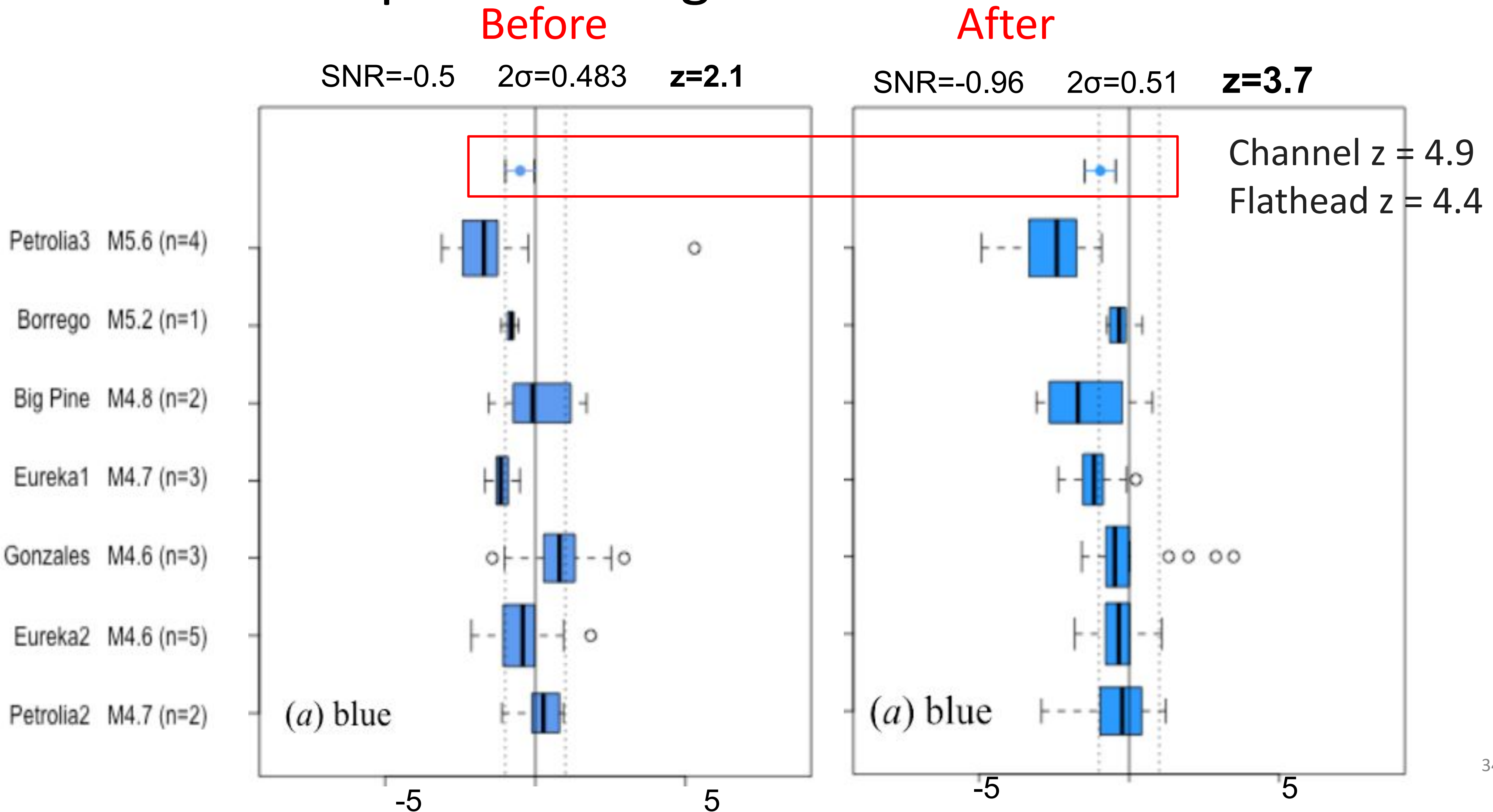
Interpretation

- SNR of 0.5 is not *very* significant, but *does* support rejection of the null hypothesis at a **2.1σ** level
- Consistent with the history of anecdotal observations

Are there first order effects that we may be able to correct for?

- **Global geomagnetic variations are common to all stations in the array**
- Cross-spectra tends to amplify these effects during times of high Ap
- **Ap detrending** can partially compensate for this effect
  - Suggests regression of features against average Ap-value during the period

The effect of Ap detrending on the ***test*** set



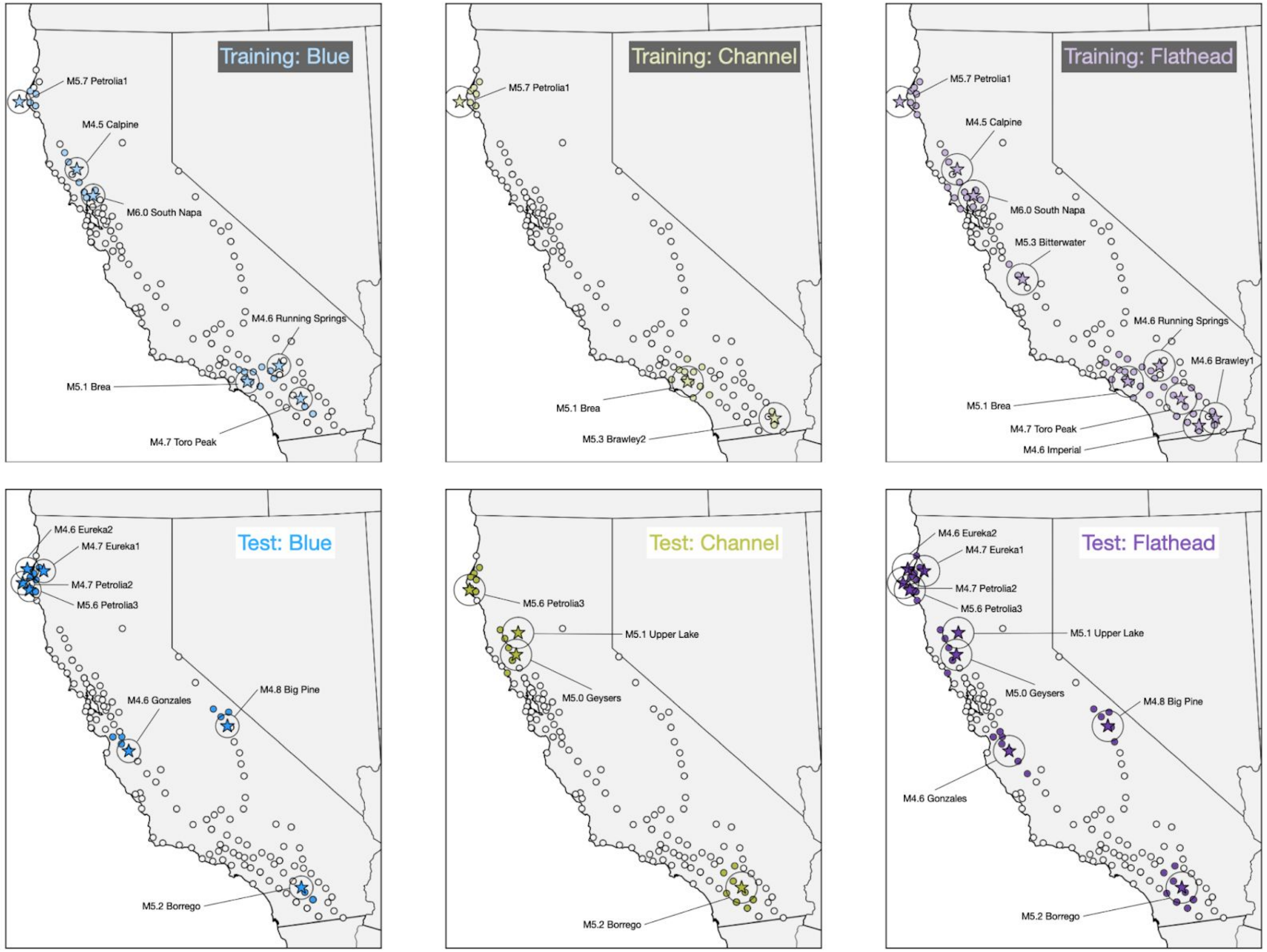
Caveat from last slide:

- There are some more details required to fully describe the difference between the two plots

  - We discuss these in detail in the manuscript
    - And describe all variations in Table S8

  - They involve the use of “k-detrending” which is a linear time trend
    - we applied this to ensure against a “pocketwatch effect”
    - desensitize the analysis to the fact that the Precursor was always latest in time
      - i.e. guard against drift in the measurements incorrectly being interpreted as an effect

Recap



Step 1: Split data into train and test

Step 2: Find pairs of stations near earthquakes (each “case”)

Step 3: Compute spectral cross power

Step 4: Define quiescent and precursor periods

Step 5: Use **training data to learn** to distinguish between periods

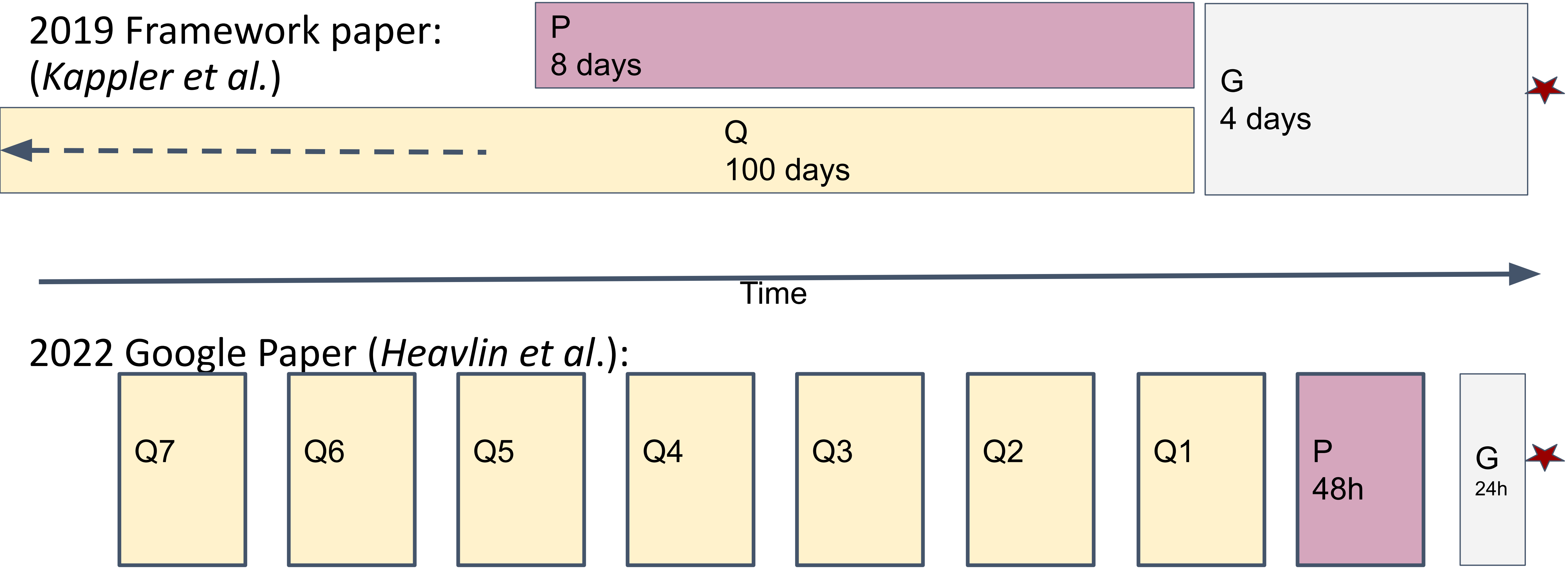
Step 6: Test if the differences **hold true for the test** dataset?

Summary:

- Two independent statistical analyses of QuakeFinder magnetic field data
- Time Domain Study suggests a shift in the STA/LTA energy histogram 4-12 days before
- Frequency Domain Study suggests increased 98th percentile of cross-spectral power 1-3 days before
- Both studies support null hypothesis rejection > 2σ
- Both improve > 3σ with simple, reasonable algorithm enhancements

- Not sufficient for practical forecasting, but does point at the *existence* of an effect that should be studied further

Comparison of Hypothetical Precursory vs Quiescent Periods from Both Studies



State of QuakeFinder (QF) ULF research

- Despite these intriguing results, there is no plan for follow up research
- Going forward, maintenance of the array, and analysis of the data is a larger task than Stellar Solutions can handle*
- QF Network to be decommissioned in the next two years
- This fate could perhaps be changed if there were alternative financial resources available for the project
- Perhaps there are other applications for the array data – space weather, magnetotelluric monitoring, … your suggestions?

Contact: QuakeFinder.com
Dan Schneider dschneider@quakefinder.com
Tom Bleier tbleier@quakefinder.com

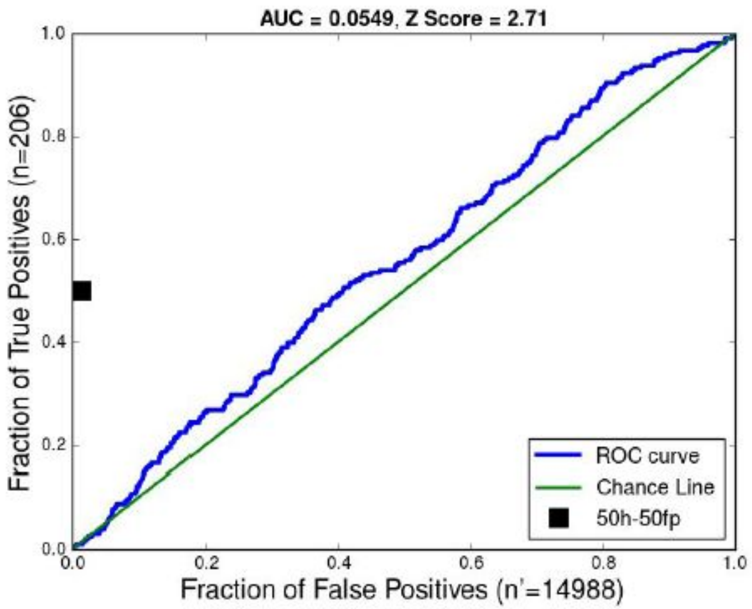
*****_despite the Herculean efforts of Jon Riley to keep it running_

Thank you for your attention

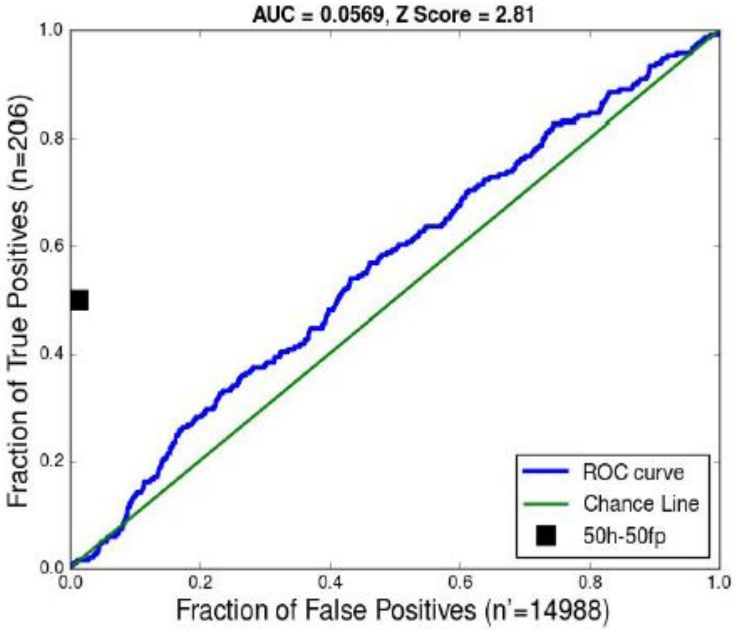
Both papers can be found on www.quakefinder.com

Backup Slides

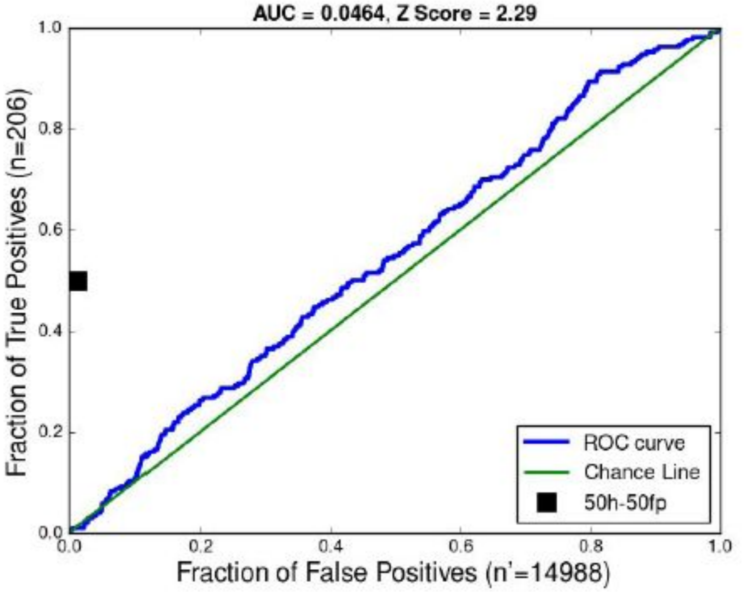
Ranker Optimization (new baseline)



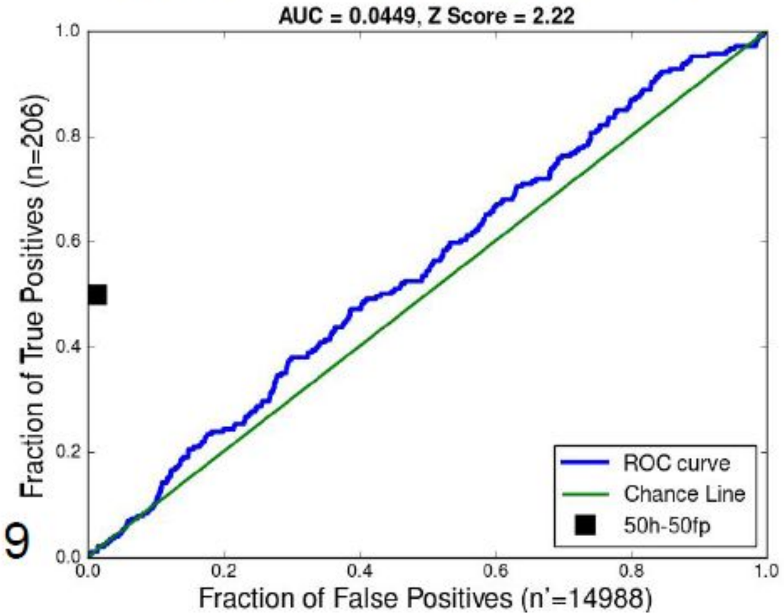
High Frequency Band



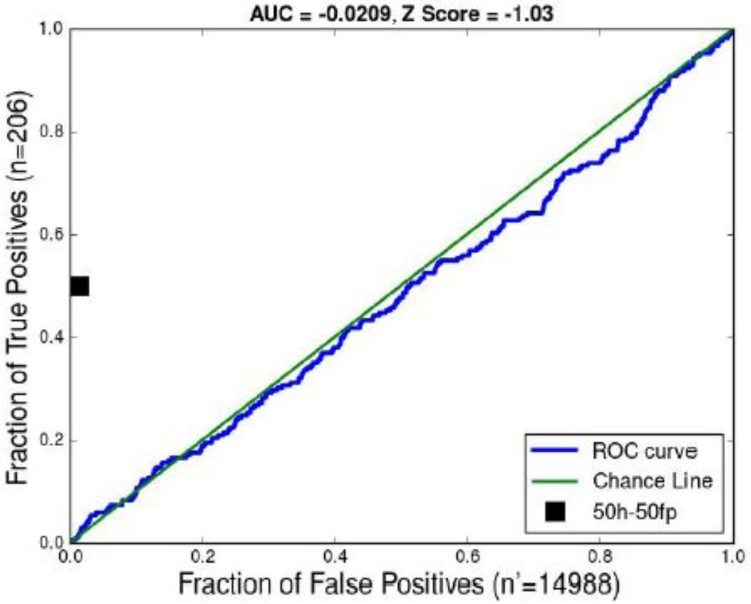
Z-cutoff (1.0-2.6)



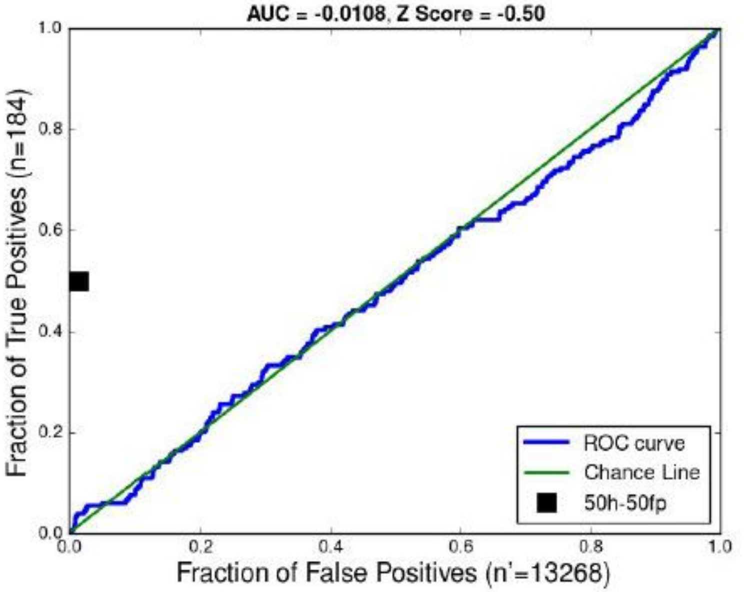
Vehicle and lightning filtered



BART Frequency Band



Z-cutoff (2.6-inf)



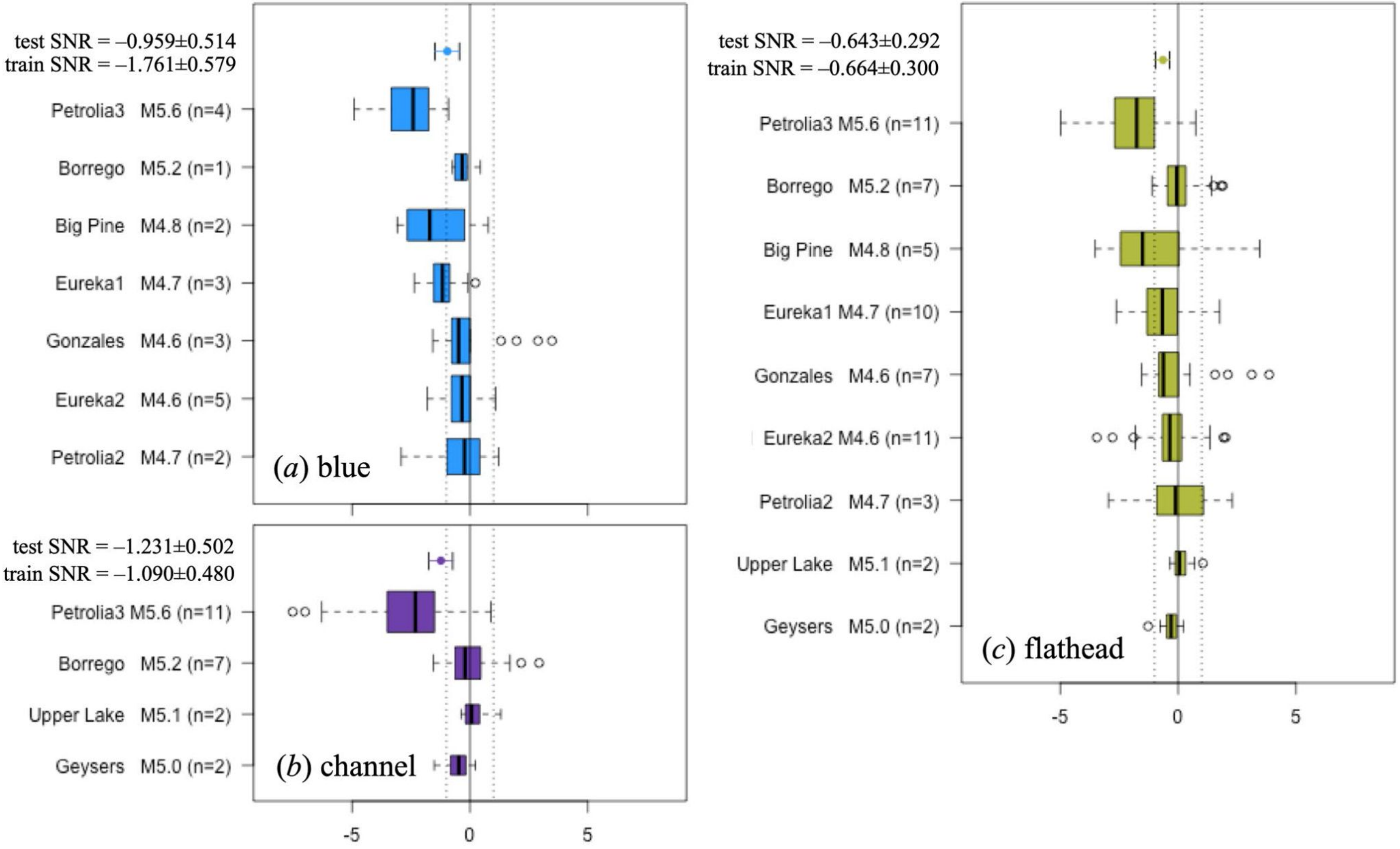


Figure 6. Boxplots of the test set for (a) blue, (b) channel, and (c) flathead tunings when the q-periods are only included if they are in the 14 days before the p-period. Linear detrending uses Āp only. This figure uses the same conventions as Figure 5.

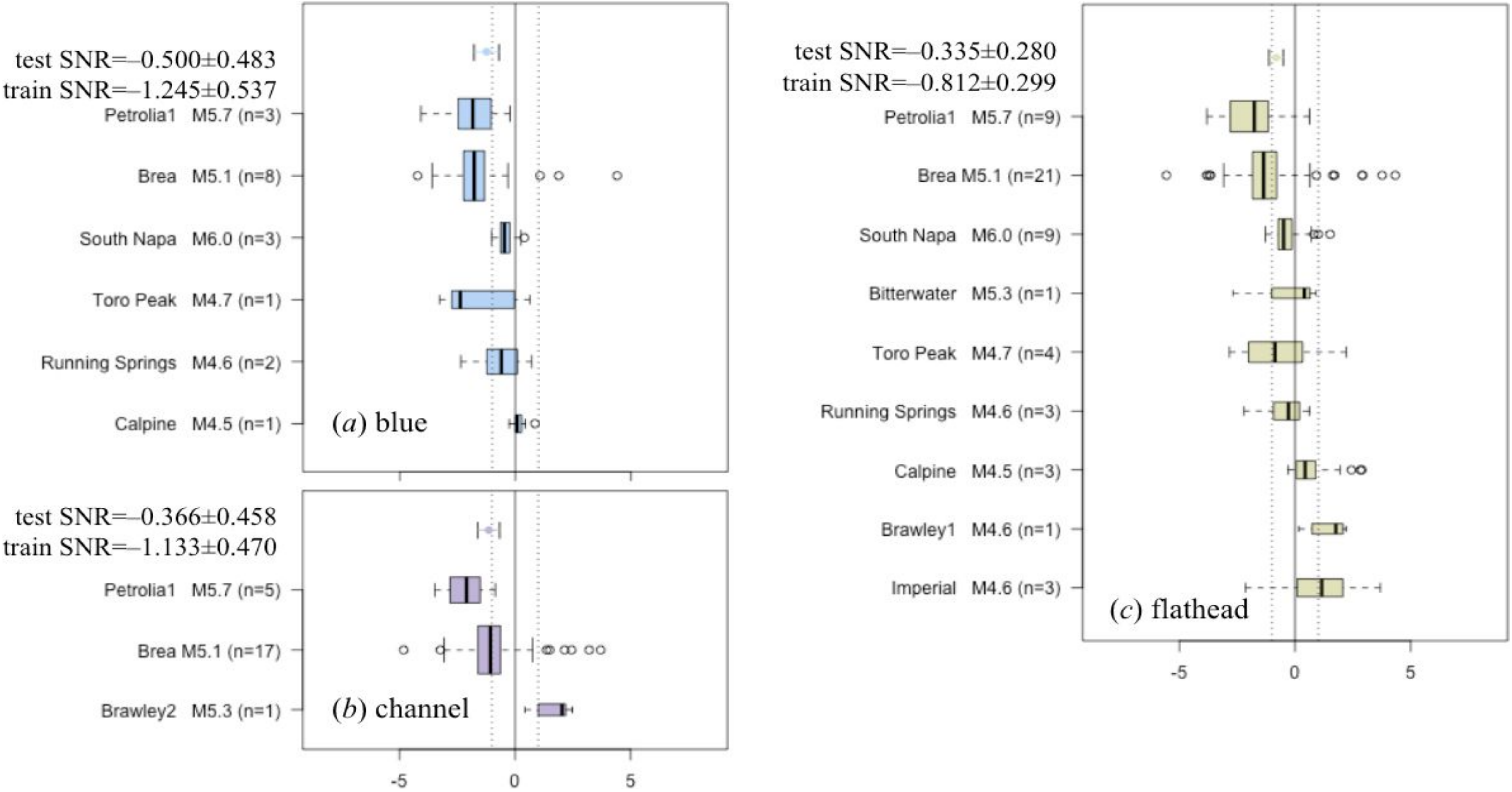


Figure S3. Boxplots of the training set for (a) blue, (b) channel, and (c) flathead tunings. This figure uses the same conventions as Figure 5.

split	tuning	q.cnt	Filter 1	SNR	2 stderr	z		split	tuning	q.cnt	Filter 1	SNR	2 stderr	z
test	blue	$K=7$	k & Ap	-0.755	0.491	3.1		train	blue	$K=7$	k & Ap	-1.62	0.572	5.7
test	blue	$K=7$	Ap	-0.685	0.487	2.8		train	blue	$K=7$	Ap	-1.203	0.535	4.5
test	blue	$K=7$	k	-0.5	0.483	2.1		train	blue	$K=7$	k	-1.245	0.537	4.6
test	blue	$K=7$		-0.656	0.486	2.7		train	blue	$K=7$		-0.815	0.516	3.2
test	**blue**	$K\leq7$	k & Ap	**-0.893**	**0.516**	**3.5**		**train**	**blue**	$K\leq7$	k & Ap	**-2.377**	**0.66**	**7.2**
test	**blue**	$K\leq7$	Ap	**-0.959**	**0.514**	**3.7**		**train**	**blue**	$K\leq7$	Ap	**-1.761**	**0.579**	**6.1**
test	**blue**	$K\leq7$	k	**-0.51**	**0.496**	**2.1**		**train**	**blue**	$K\leq7$	k	**-1.899**	**0.59**	**6.4**
test	**blue**	$K\leq7$		**-0.817**	**0.505**	**3.2**		**train**	**blue**	$K\leq7$		**-1.395**	**0.547**	**5.1**
test	chan	$K=7$	k & Ap	-0.65	0.466	2.8		train	chan	$K=7$	k & Ap	-1.145	0.477	4.8
test	chan	$K=7$	Ap	-0.68	0.465	2.9		train	chan	$K=7$	Ap	-0.422	0.449	1.9
test	chan	$K=7$	k	-0.366	0.458	1.6		train	chan	$K=7$	k	-1.133	0.47	4.8
test	chan	$K=7$		-0.675	0.463	2.9		train	chan	$K=7$		-0.363	0.448	1.6
test	**chan**	$K\leq7$	k & Ap	**-0.981**	**0.496**	**4.0**		**train**	**chan**	$K\leq7$	k & Ap	**-2.882**	**0.645**	**8.9**
test	**chan**	$K\leq7$	Ap	**-1.231**	**0.502**	**4.9**		**train**	**chan**	$K\leq7$	Ap	**-1.09**	**0.48**	**4.5**
test	**chan**	$K\leq7$	k	**-0.489**	**0.471**	**2.1**		**train**	**chan**	$K\leq7$	k	**-2.508**	**0.577**	**8.7**
test	**chan**	$K\leq7$		**-0.922**	**0.482**	**3.8**		**train**	**chan**	$K\leq7$		**-1.359**	**0.487**	**5.6**
test	flat	$K=7$	k & Ap	-0.608	0.283	4.3		train	flat	$K=7$	k & Ap	-1.01	0.307	6.6
test	flat	$K=7$	Ap	-0.519	0.281	3.7		train	flat	$K=7$	Ap	-0.412	0.293	2.8
test	flat	$K=7$	k	-0.335	0.28	2.4		train	flat	$K=7$	k	-0.812	0.299	5.4
test	flat	$K=7$		-0.487	0.281	3.5		train	flat	$K=7$		-0.239	0.292	1.6
test	**flat**	$K\leq7$	k & Ap	**-0.629**	**0.292**	**4.3**		**train**	**flat**	$K\leq7$	k & Ap	**-1.519**	**0.333**	**9.1**
test	**flat**	$K\leq7$	Ap	**-0.643**	**0.292**	**4.4**		**train**	**flat**	$K\leq7$	Ap	**-0.664**	**0.3**	**4.4**
test	**flat**	$K\leq7$	k	**-0.298**	**0.287**	**2.1**		**train**	**flat**	$K\leq7$	k	**-1.258**	**0.316**	**8.0**
test	**flat**	$K\leq7$		**-0.56**	**0.29**	**3.9**		**train**	**flat**	$K\leq7$		**-0.539**	**0.298**	**3.6**

Table S8*

*from Supporting Information

Modified with color and added column for z-score (SNR/1stderr)

Consider:
K<=7 vs. K=7

Significances increases on 23/24 train and test conditions

Test set: Ap vs k,
or
Train set: {K & Ap} vs k
* N.B. Training was done with k-detrending thus train set must c.f. {K & Ap} vs k

12/12 SNRs increase

Takeaways

- ***ALL*** SNR were negative, Precursor (P) is larger than Quiescent (Q)
  - regardless of the permutation of detrending
- No individual treatment of noise / corrupt data was applied
- ***ALL* tuning conditions scored better with Ap correction**
- No retraining with Ap-detrended data has been done
- Significance mostly observed in the "*MT dead band*"
- **The QF instrument mag array data seems to have scientific value**

Simplifying the coefficients:

The classical LDA coefficients are estimated as proportional to $\mathbf{S}^{-1}\,\boldsymbol{d}$.

In practice, this requires regularization, e.g. $\mathbf{S} \leftarrow \mathbf{S} + \mu\mathbf{D}$, for some positive semidefinite matrix **D**, to successfully invert the matrix **S**.

Instead, we choose this simplification: take as the coefficients ***b*** as the difference vector ***d*** itself.

This choice makes our approach more easily replicable, i.e., less dependent on the details of an algorithm constructing a matrix **D** and choice of constant μ. In fact, the curves implied by the two vectors ***d*** and $(\mathbf{S}+\mu\mathbf{D})^{-1}\boldsymbol{d}$ look quite similar (see Figure S9);

This choice certainly reduces the signal-to-noise ratio on the training set, so in this sense, it can be considered both scientifically conservative and statistically suboptimal.

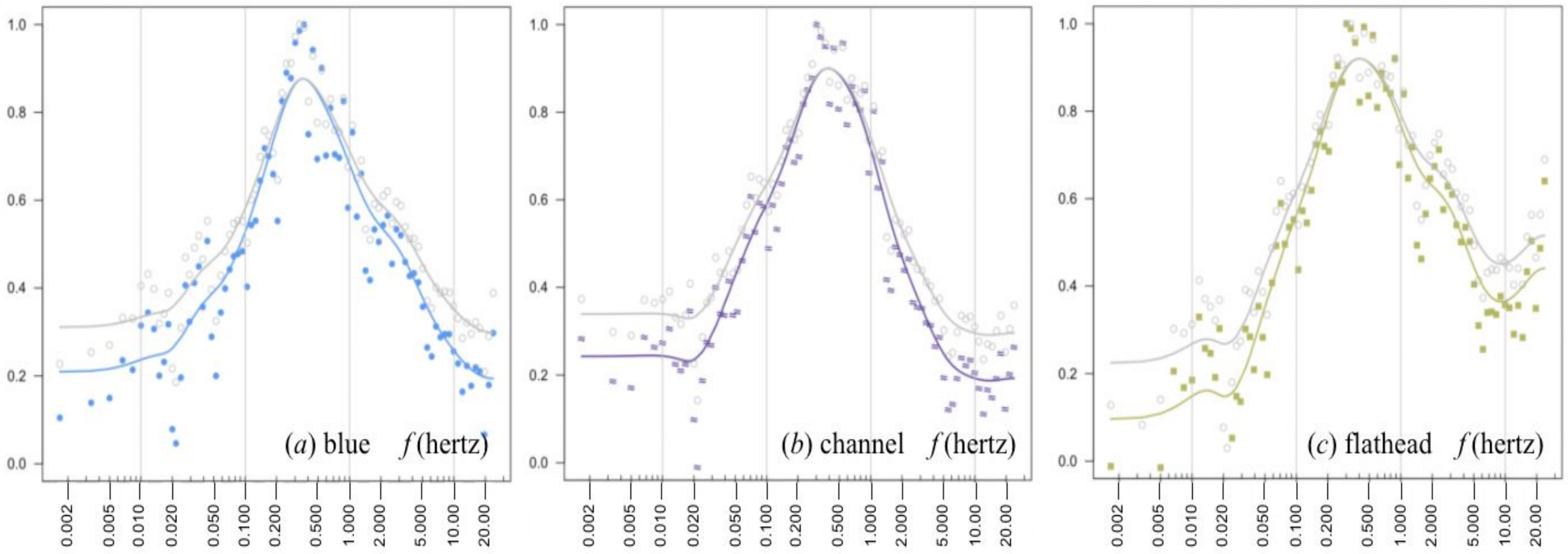


Figure S9. Plots of the ***d***-vectors (more intensely colored symbols and lines) that were actually used compared to L2-regularized LDA coefficients ***b*** (lighter shades); each point is plotted against its corresponding frequency *f* in the final models. X- and y-axes are the same as in Figure 4. The curves result from applying Filter 2 to the points of the same color.

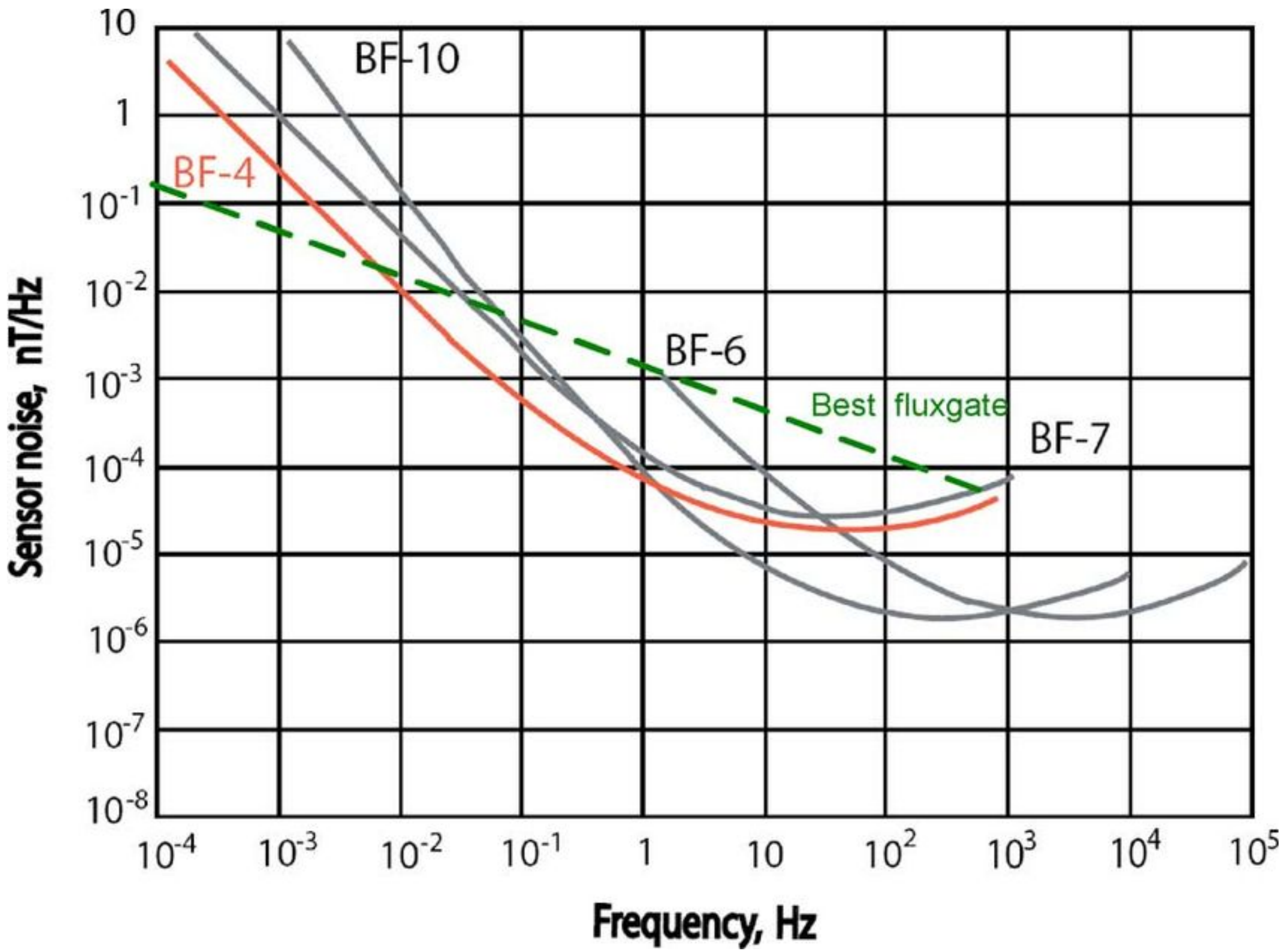
Sites & Earthquake Selection Rules

Table S10:

(1) site-pair inclusion	When $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \le \theta$, then $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ form a site-pair.
(2) earthquake location inclusion	For site-pair $(\boldsymbol{x}_\mathrm{i}, \boldsymbol{x}_\mathrm{j})$, define nearby earthquakes (e) by $\mathcal{E}_{ij}(\theta) \equiv \{e: \|\boldsymbol{x}_i - \boldsymbol{x}_e\| \le \theta$ or $\|\boldsymbol{x}_j - \boldsymbol{x}_e\| \le \theta$ or $\|\boldsymbol{x}_{ij} - \boldsymbol{x}_e\| \le \theta\}$.
(3) earthquake magnitude inclusion	$\mathcal{E}_{\mathrm{M}ij}(\theta, M_0) = \{E: E \in \mathcal{E}_{ij}(\theta)$ and $M_E \ge M_0\}$.
(4) earthquake magnitude exclusion (declustering)	$E \in \mathcal{E}_{\mathrm{M}ij}(\theta, M_0)$ such that the following set is empty: $\{e: e \in \mathcal{E}_{ij}(\theta) \backslash E$ & $t_E - \beta - (K+1)\lambda < t_e \le t_E$ & $M_e \ge M_0 - \Delta_\mathrm{M}\}$.

Table S10: Summary of the four rules described in Section 3.3 which define the SSE test cases. $\boldsymbol{x}_i$, $\boldsymbol{x}_j$, $\boldsymbol{x}_E$ denote longitude-latitude coordinates of sites *i, j*, and epicenter of the earthquake E respectively. The distance between sites $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is denoted by $|\boldsymbol{x}_i - \boldsymbol{x}_j|$. The midpoint between sites $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denoted by $\boldsymbol{x}_{ij}$. Let t_E denote the time of earthquake E and M_E its magnitude.

FluxGate Magnetometer Noise Comparison



Scoring

4.5 Scoring and SNR calculations

Given coefficients $\boldsymbol{b}_{32}$ (from the training set), we can calculate scores on any dataset, training or test. For a given SSE n=(i,j,e), denote these scores by $y_{n0}, y_{n1}, \ldots , y_{nK}$, where as before y_{n0} is the precursor score and $y_{n1}, \ldots , y_{nK}$ are the scores for the corresponding K quiescent periods prior to y_{n0}.

Following the pocketwatch principle, we apply Filter 1 as before:

(a) The slope g_{ij} is calculated using the quiescent periods associated with site-pair (i,j),

(b) where each associated earthquake e receives its own intercept term.

(c) The implemented correction is $\dot{y}_{nk}=y_{nk} - g_{ij}(k–0)$,

the 0 in the latter expression corresponding to the value of the pocketwatch covariate k for the precursor period. Thus, the corrected value $\dot{y}_{nk}$ is modified as if the pocketwatch covariate k for each quiescent were really that of the precursor (k=0).

(d) For the nth SSE, the average quiescent score therefore averages these values $\dot{y}_{nk}$:

$\bar{y}_n = \boldsymbol{\Sigma}_{\mathrm{k=1..K}}\ \dot{y}_{nk}/K.$ Equation 4.5.1 (c.f. Equation 4.3.3)

(e) The associated quiescent-minus-precursor difference is $\bar{y}_n – y_{n0}$, with variance $\sigma^2(1/K+1)$.

(f) The pooled within-SSE variance, a scalar, is

$S^2 = \boldsymbol{\Sigma}_{\mathrm{ij}} \boldsymbol{\Sigma}_{\mathrm{n\in ij}} \boldsymbol{\Sigma}_{\mathrm{k\in n}} (\dot{y}_{nk} – \bar{y}_n)^2/\mathrm{ij}\ DF_{ij}$, Equation 4.5.2 (c.f. Equation 4.3.4),

where the index set $n\in ij$ denotes that set of SSEs with associated sitepair (i,j), and the index set $k\in n$ denotes that set of quiescents associated with SSE n. DF_{ij} are the residual degrees of freedom that remain after applying Filter 1 within sitepair ij. DF_{ij} equals $\Sigma_{n\in ij}$ (K–1)–A, where A is the number of Filter 1 covariates. As with equation 4.3.4, in essence equation 4.5.2 measures the background variation among the quiescents within each SSE. In the same vein, S^2 is similarly invariant to changes in the values of any of the within-SSE precursor scores.