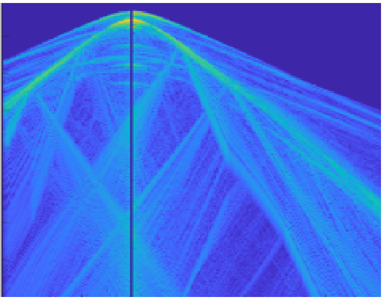
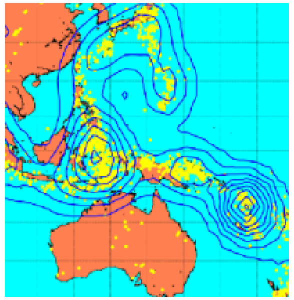
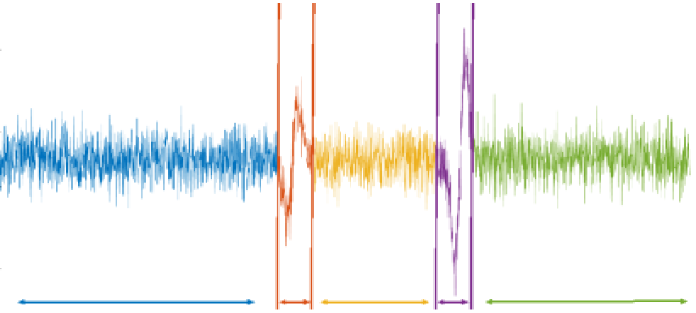
SAND2021-8341PE







Sandia National Laboratories

Exceptional service in the national interest

Feature-Based Bayesian Inference for Seismic Event Monitoring

Thomas A. Catanach

July 20-21, 2021

WAVEFORMS

U.S. DEPARTMENT OF ENERGY

National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

This research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, MSTS, PNNL, and SNL.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Overview: Challenge

- **Challenge:**
  - Detecting and Locating very weak seismic signals requires sensor fusion and utilizing more information for signal waveforms
  - Uncertainty quantification is essential since decisions will have to be made based on limited knowledge about the complexities of the models, sensors, and data
  - Historic data or simulations will need to be used to understand these complexities and synthesize them into simpler/tractable models that we can use for monitoring

- **Potential Impact:**
  - Facilitate high consequence decision making by providing event information with well calibrated confidences for very weak signals in domains with little historic data
  - Provide a framework for data fusion to integrate multi fidelity and phenomenology data
  - Enable experimental design methods to quantify a monitoring network's ability to detect events and test improvements to the network/processing system

Overview: Approach

- **The Bayesian Perspective:**
  - Probability distributions quantify uncertainty due to insufficient information
  - Bayesian methods for identification and estimation are critical to robust decision-making
- **Target Contribution:**
  - Take a Bayesian approach to waveform processing to detect and identify weak seismic events while integrating various sources of uncertainty.
  - Use a unique statistical framework and novel computational methods to make waveform-based Bayesian inference tractable
- **General Approach:**
  - Formulate an inference problem based upon predicting waveform features instead of the waveforms themselves
  - Simulate waveforms to build a statistical model of waveform features with uncertainty
  - Use Sequential Tempered Markov Chain Monte Carlo to efficiently identify events

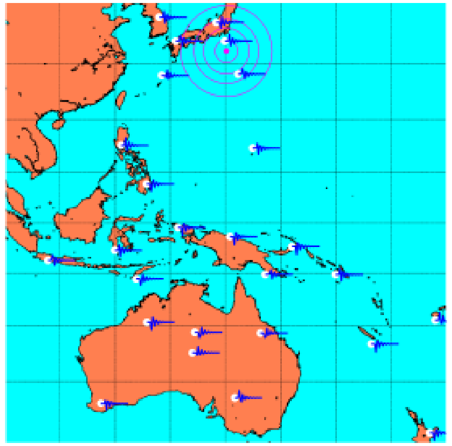
Talk Outline

1. **Formulation of Seismic Monitoring as a Bayesian Inference Problem**
2. **Feature-Based Bayesian Inference**
3. **Building the Feature-Based Inference Workflow**
4. **Example with synthetic data**
5. **Challenge with Real data**
6. **Future work and Conclusion**

BAYESIAN INFERENCE FOR SEISMIC EVENT LOCALIZATION

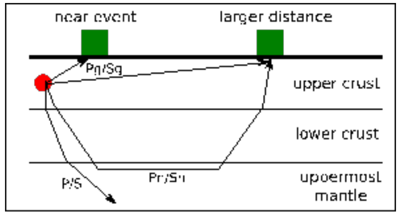
The Bayesian Framework

Data: $\mathcal{D}$

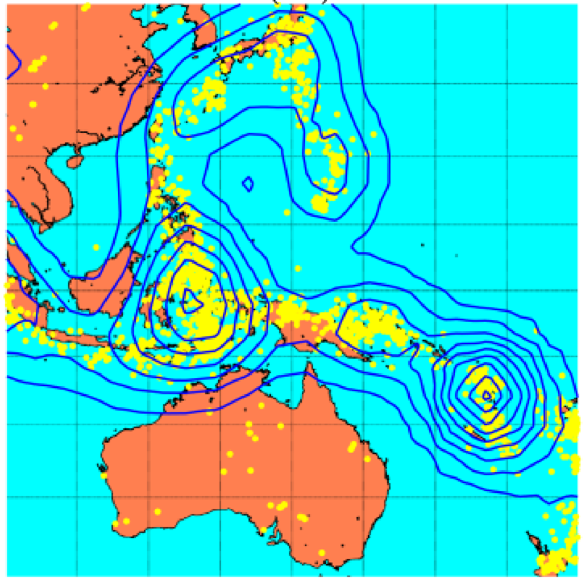


Likelihood: $p\left(\mathcal{D} \mid \theta\right)$

Physics Model
Sensor Model
Uncertainty Model



Prior: $p\left(\theta\right)$

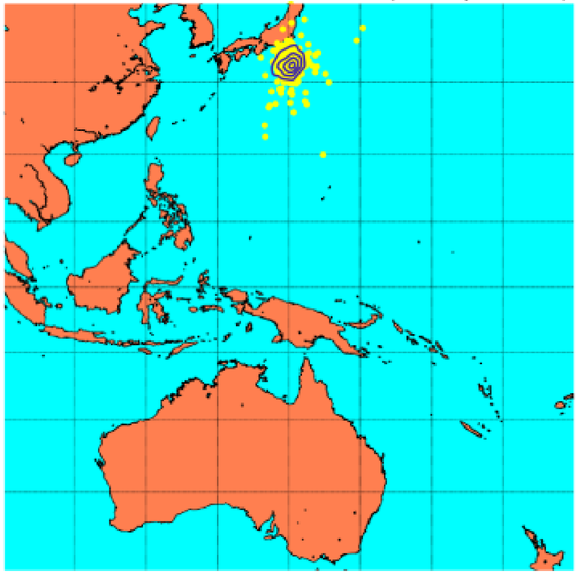


Knowledge about where events are likely to occur

Bayes' Theorem:

$$p\left(\theta \mid \mathcal{D}\right) = \frac{p\left(\mathcal{D} \mid \theta\right) p\left(\theta\right)}{p\left(\mathcal{D}\right)}$$

Posterior: $p\left(\theta \mid \mathcal{D}\right)$



Updated knowledge about where a specific event occurred

Background on existing Bayesian Seismic Monitoring Methods

Detection-Based

- **Description**
  - Each station pre-processes their observed waveforms to extract arrival “picks”
  - The likelihood of an event (or events) is based upon how well the observed arrival times correspond to arrivals from seismic waves generated by the event hypothesis
  - Arrival time and detection uncertainty can be integrated into the model
- **Examples**: BayesLoc[1], NET-VISA[2]
- **Advantages**
  - Requires only a model of travel times and not the waveform
- **Disadvantages**
  - Events that produce weak signals below the pick threshold cannot be detected, even when many sensors are combined

[1]Myers, S. C., Gardar Johannesson, and Robert J. Mellors. “BayesLoc: A robust location program for multiple seismic events given an imperfect earth model and error-corrupted seismic data” (2011)

[2]Arora, Nimar S., Stuart Russell, and Erik Sudderth. "NET-VISA: Network processing vertically integrated seismic analysis" (2013)

Background on existing Bayesian Seismic Monitoring Methods

Signal-Based

- **Description**
  - The likelihood of a candidate event (or events) is based upon comparing predicted waveforms given the event hypothesis, noise process, and other modeled uncertainty to the observed waveforms
- **Example**: SIG-VISA[3]
- **Advantages**
  - Can integrate many sensors to detect low magnitude signals
  - Waveform characteristics can contain useful information for event identification
- **Disadvantages**
  - Requires learning and evaluating a generative model of the full waveform to compute the likelihood of the observed signal

[3]Moore, David A., and Stuart J. Russell. "Signal-based Bayesian seismic monitoring" (2017)

FEATURE-BASED INFERENCE

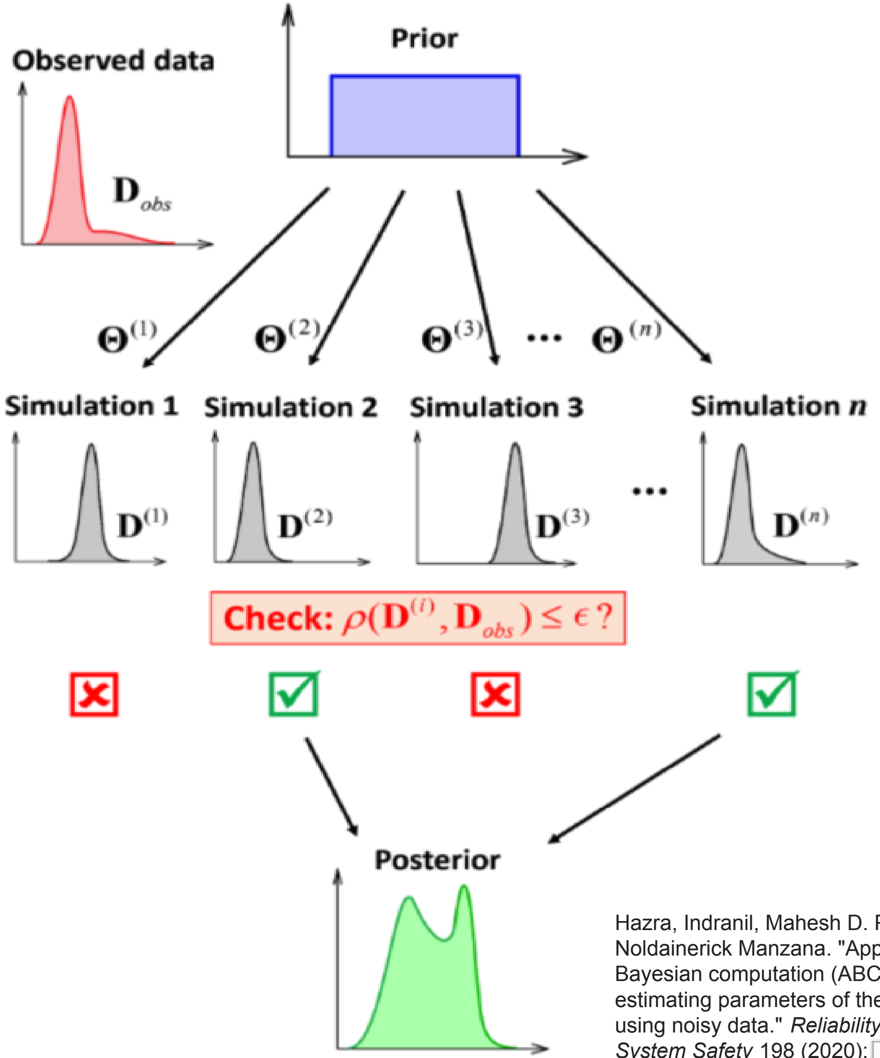
Sandia National Laboratories

Traditional feature-based inference (Approximate Bayesian Computation)

Approximate Bayesian Computation:

- Likelihood free inference (i.e. no full generative model) can be done approximately if you have a model for a data statistic (D_{obs}).
- Requires that the method for computing D_{obs} is not Θ dependent e.g. mean of the data.

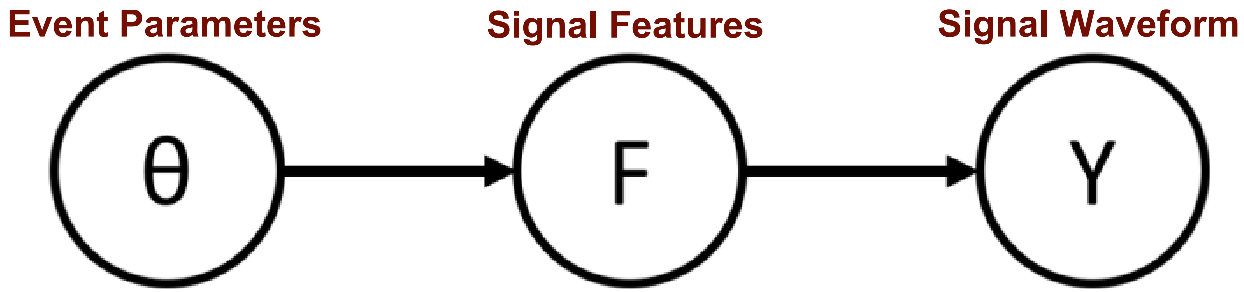
$$\underbrace{p\left(\theta \mid Y\right)}_{\textbf{Posterior}} = \frac{\overbrace{p\left(Y \mid \theta\right)}^{\textbf{Likelihood}}\overbrace{p\left(\theta\right)}^{\textbf{Prior}}}{\underbrace{p\left(Y\right)}_{\textbf{Evidence}}}$$



Hazra, Indranil, Mahesh D. Pandey, and Noldainerick Manzana. "Approximate Bayesian computation (ABC) method for estimating parameters of the gamma process using noisy data." *Reliability Engineering & System Safety* 198 (2020): 0.

General feature-based inference

Graphical Model:



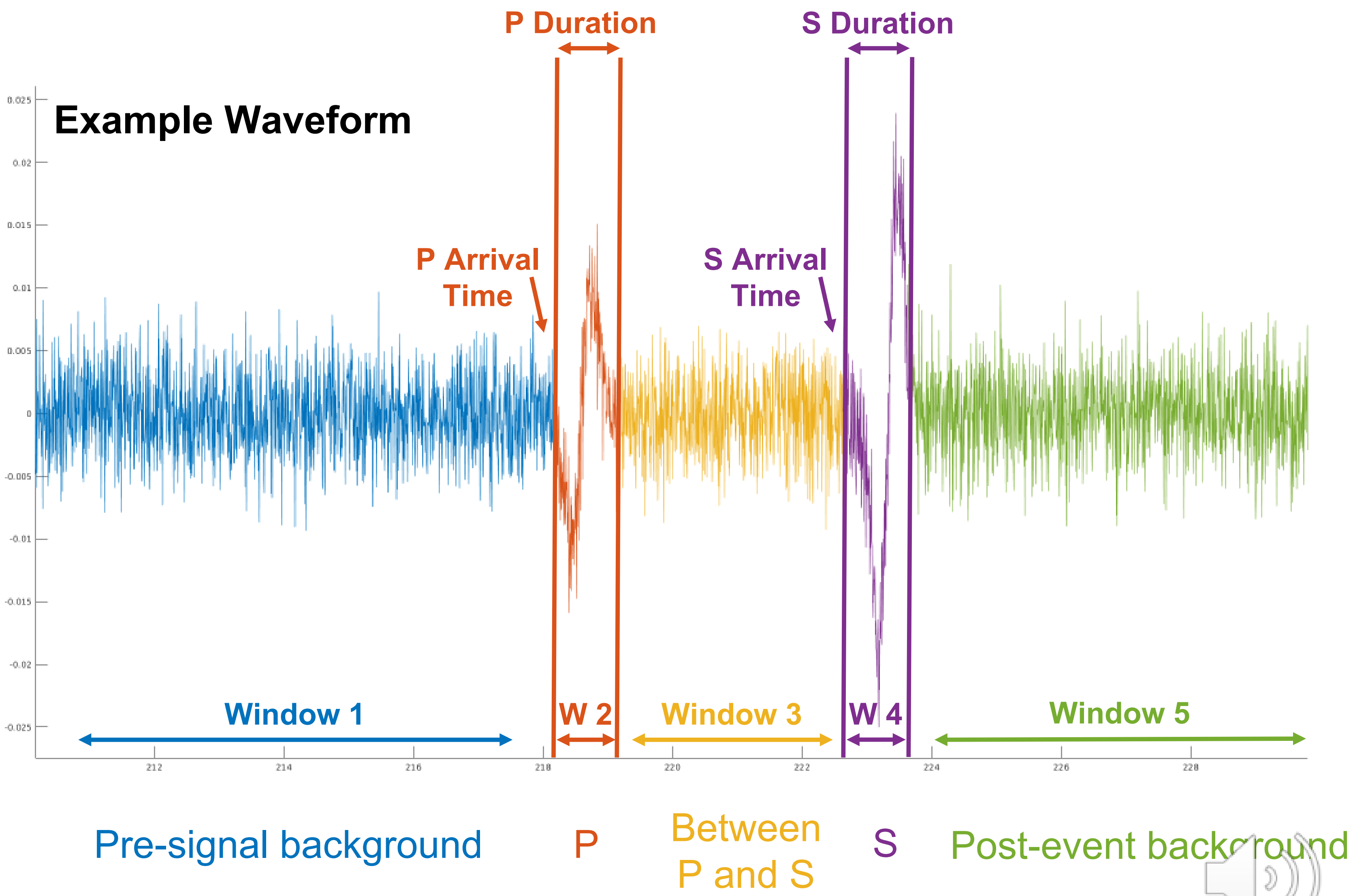
Bayesian Inference:

- Feature-based inference requires building statistical models for the likelihood of a signal given certain features and the likelihood of those features given an hypothesized event parameterization.
- Features can be θ dependent.

$$\underbrace{p\left(\theta \mid Y\right)}_{\text{Posterior}} = \frac{\overbrace{p\left(Y \mid \theta\right)}^{\text{Likelihood}}\overbrace{p\left(\theta\right)}^{\text{Prior}}}{\underbrace{p\left(Y\right)}_{\text{Evidence}}} = \underbrace{\left(\int \overbrace{p\left(Y \mid F\right)}^{\text{Signal Likelihood}}\overbrace{p\left(F \mid \theta\right)}^{\text{Feature Likelihood}} dF\right)}_{\text{Marginalize over features}} \frac{p\left(\theta\right)}{p\left(Y\right)}$$

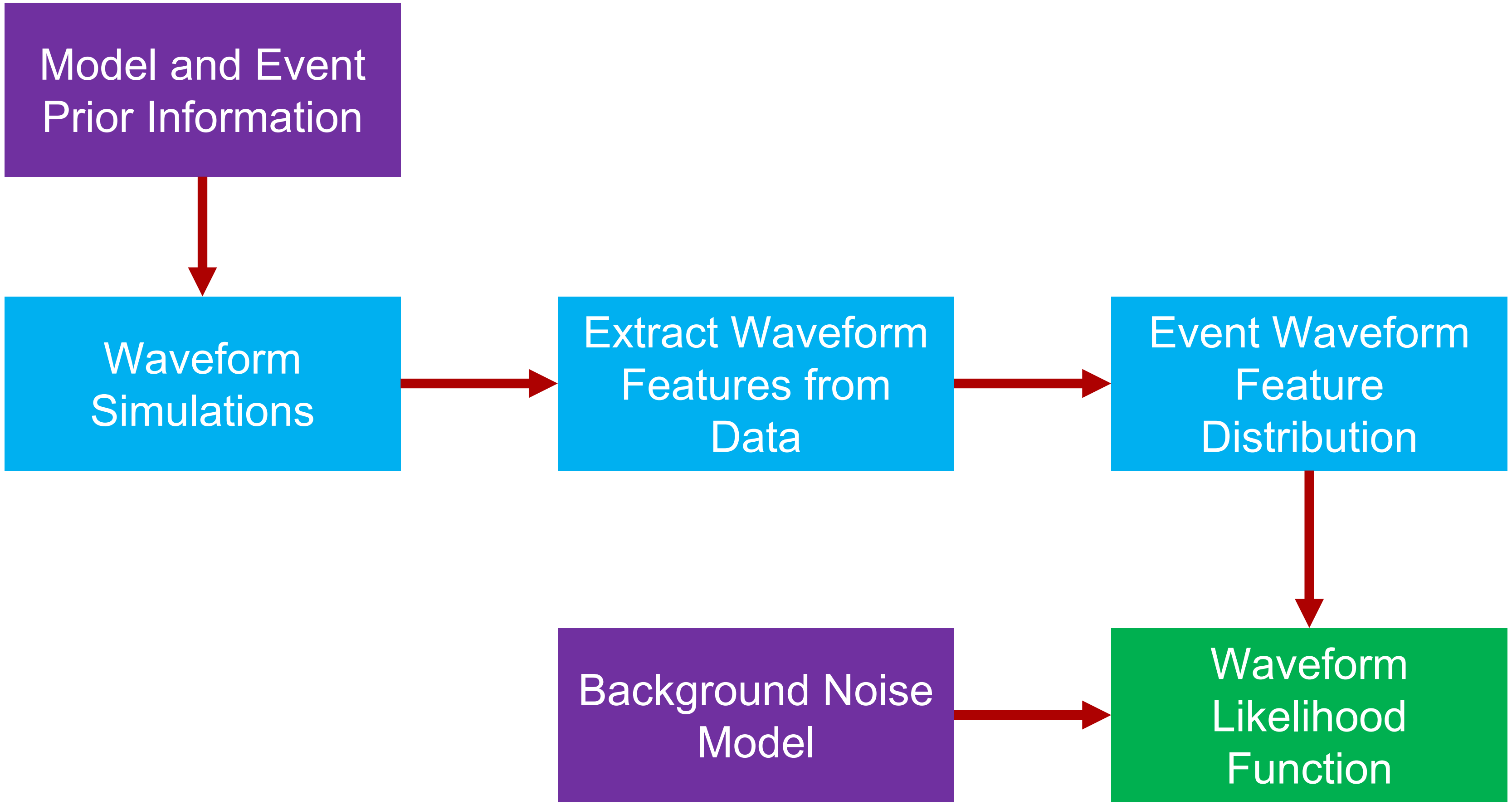
Feature Based Inference for Seismic Monitoring

- Waveform Features
  - P and S arrival time
  - Waveform feature within window e.g. total signal power or power within a band.
- P and S arrival times and uncertainty can be found using models like AK135
- We can build a statistical model for the signal power using simulations and background models.



BUILDING FEATURE-BASED WORKFLOW

Data driven workflow



Bayesian Inference Problem

- **Parameters**
  - Event Parameters: Latitude, Longitude, Depth, Magnitude, Origin Time
  - Uncertainty Parameters: Travel time uncertainty

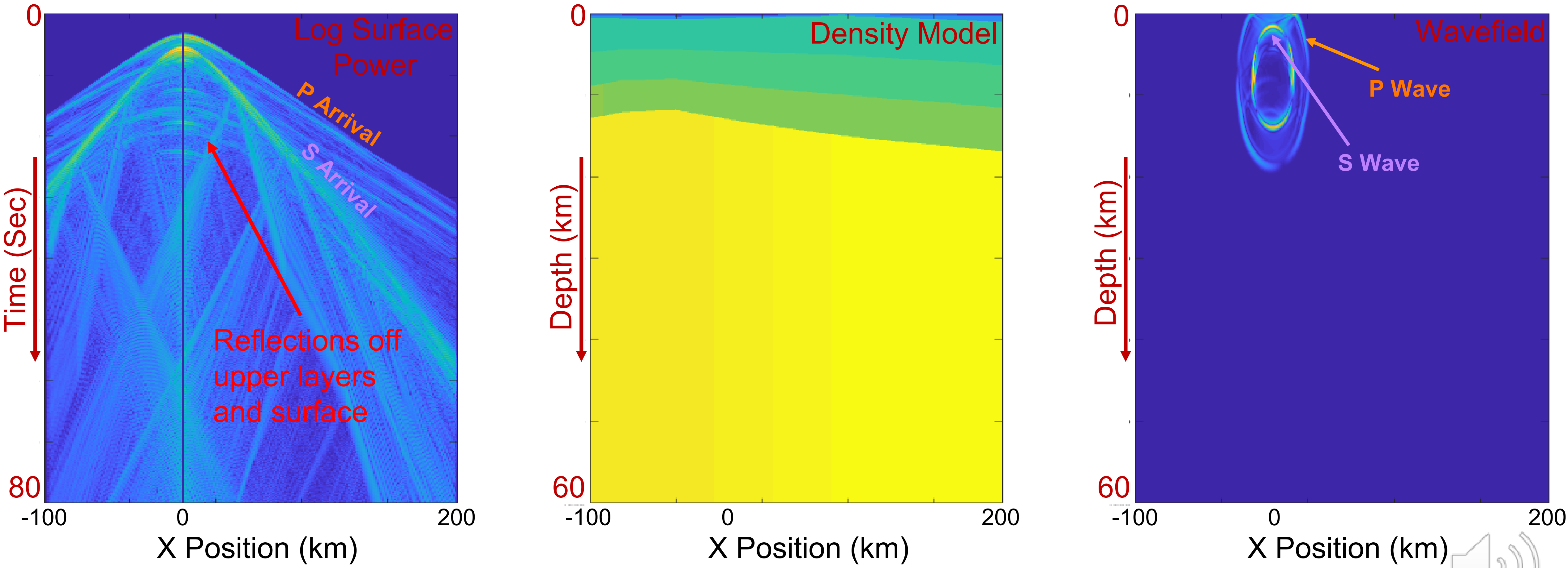
- **Feature Model**
  - AK135 for mean travel time and approximate travel time uncertainty
  - Waveform Simulations build signal power distribution as a function of distance from the source and marginalize over sources of uncertainty like focal mechanism, stochastic earth model.

- **Background Noise Process**
  - Assume a process modeled as a stationary Gaussian process within each window with known covariance
  - Independent of the event signal

Building Feature Model

Simulation Environment

- 2D waveform simulations[4] on 300 km x 60 km domain from Crust 1.0 cross-sections of Utah
- Simulated 1k events at 10 sensors with uniformly distributed event and focal mechanism parameters



[4]Li, Dunzhu, et al. "Global synthetic seismograms using a 2-D finite-difference method." (2014)

Building Feature Model

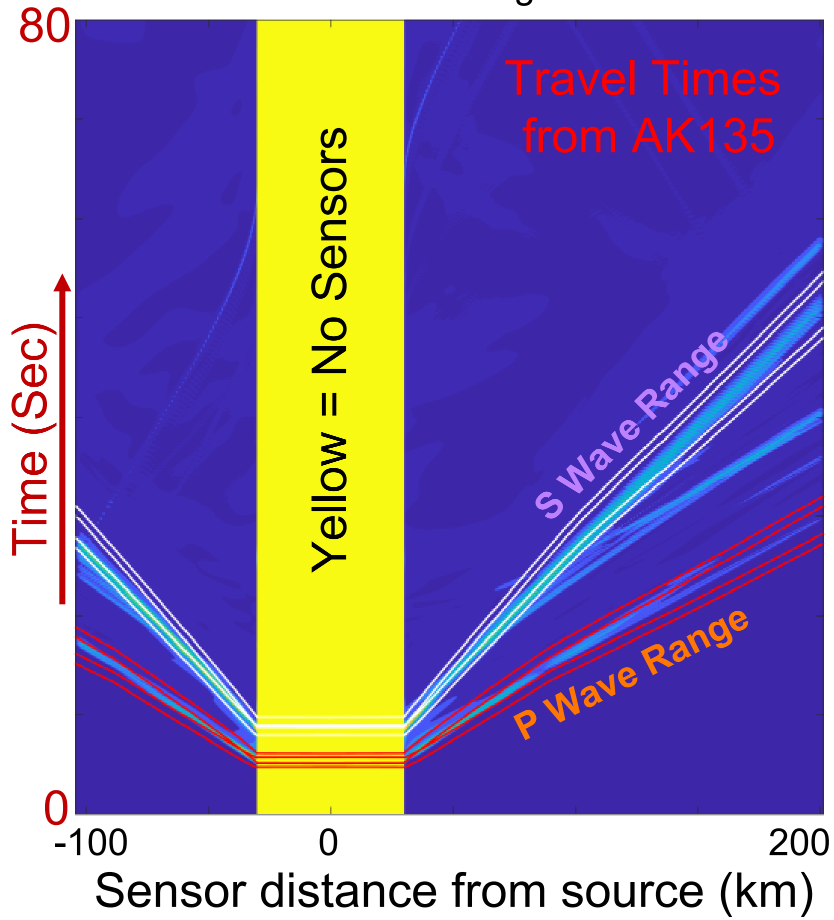
Extracting Waveform Features

- P and S travel times, uncertainty, and assumed duration define possible window arrangements

- The window arrangement which contains the maximum event power is extracted

Extracting Waveform Power

Simulated Signal

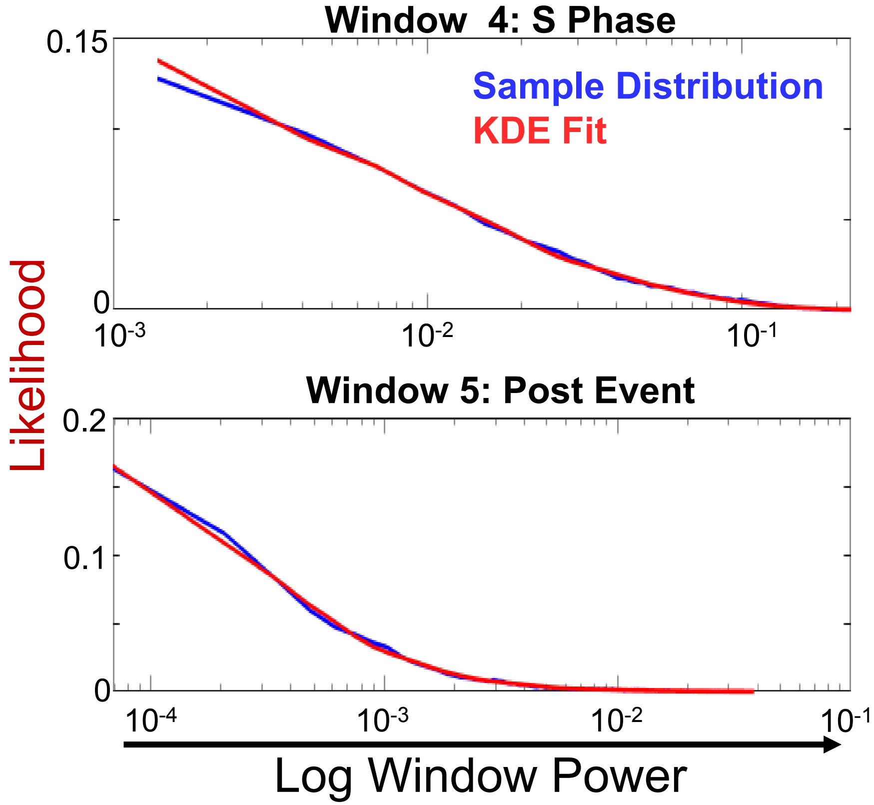


Building Feature Model

KDE Model for p(F | θ)

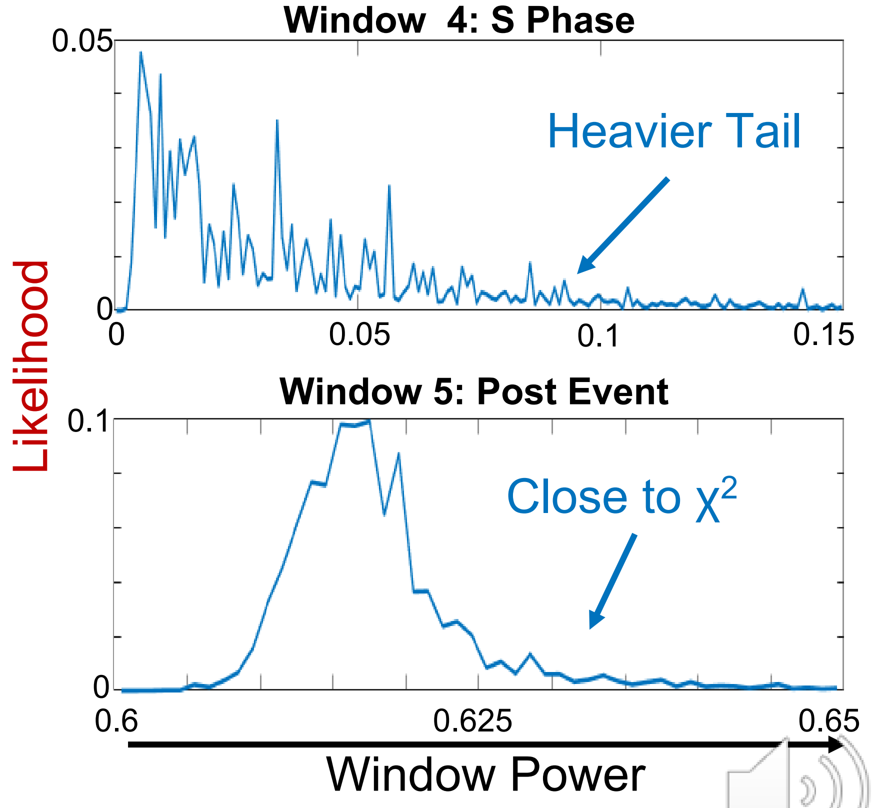
- Use the maximum event power to fit a Kernel Density (KDE) model for events at similar distances

Distribution of Window Power
Simulated Signal



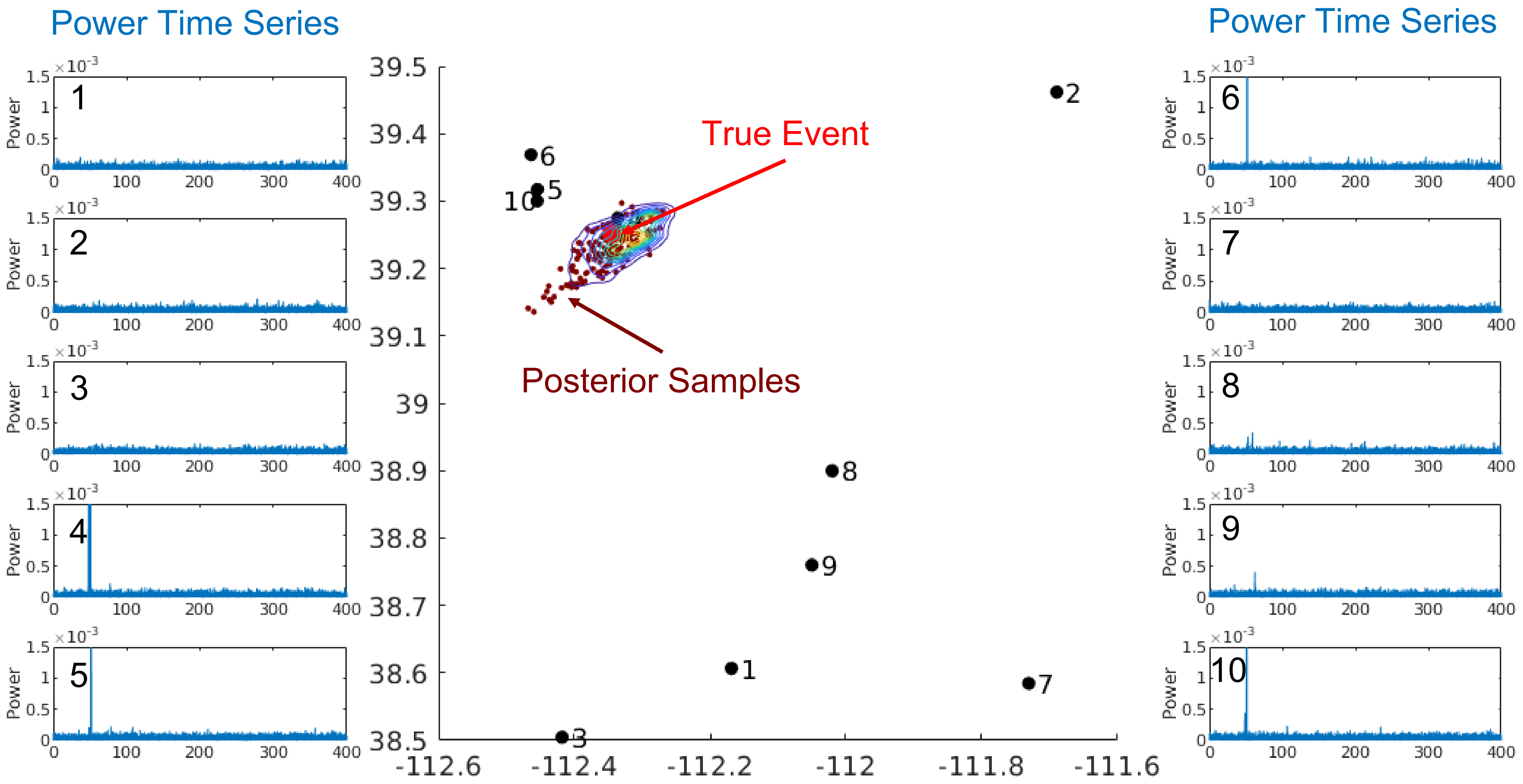
Factor in event magnitude and background process

Distribution of Window Power
Background Distribution Added and Scaled

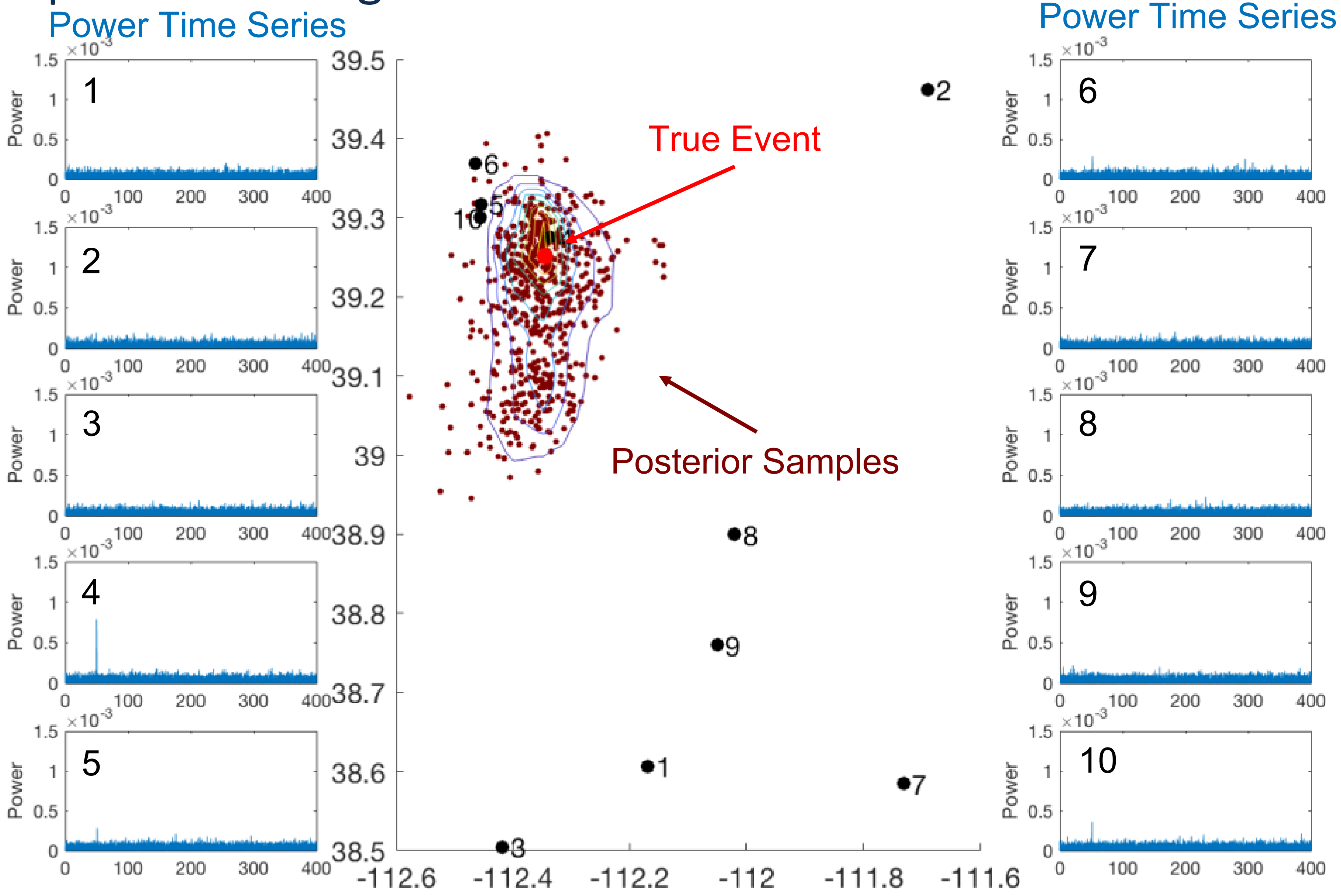


SYNTHETIC EXAMPLE

Example 1: Well identified strong signal



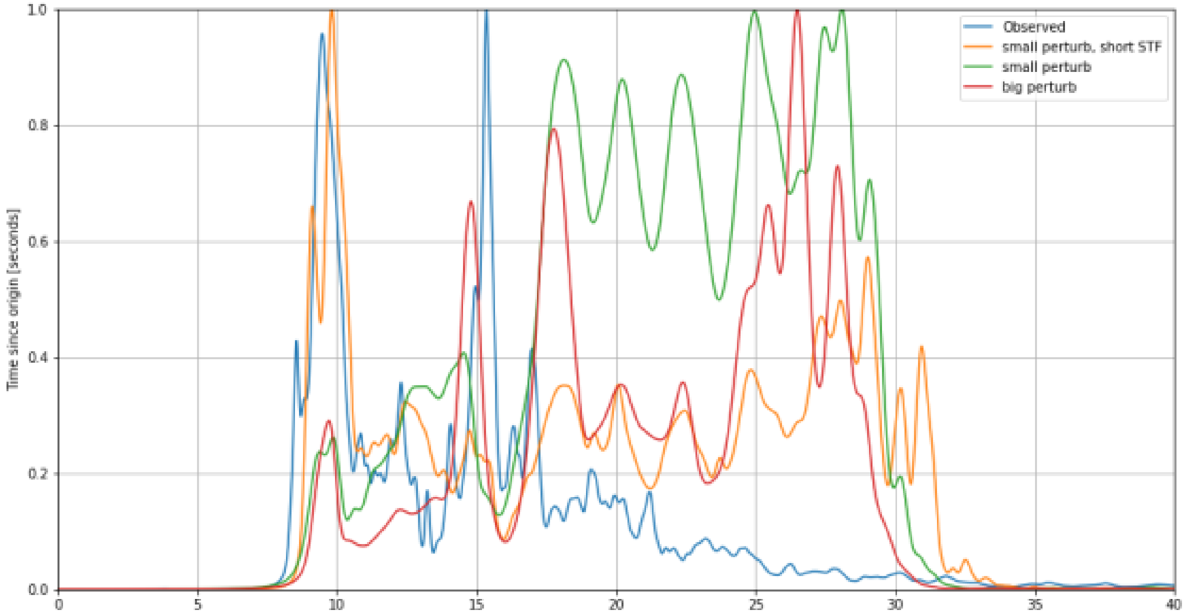
Example 2: Weak signal with more variance

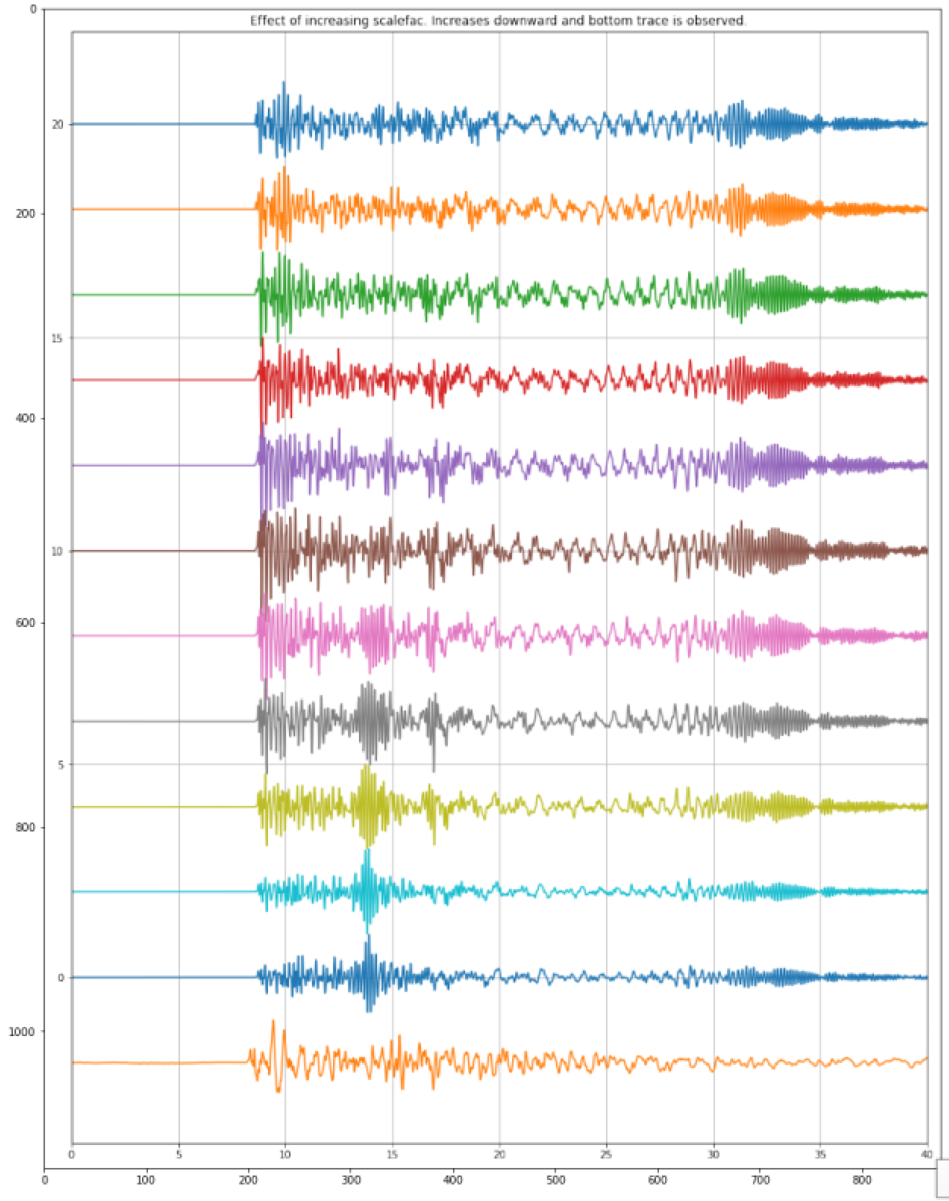


CHALLENGE MOVING TO REAL DATA

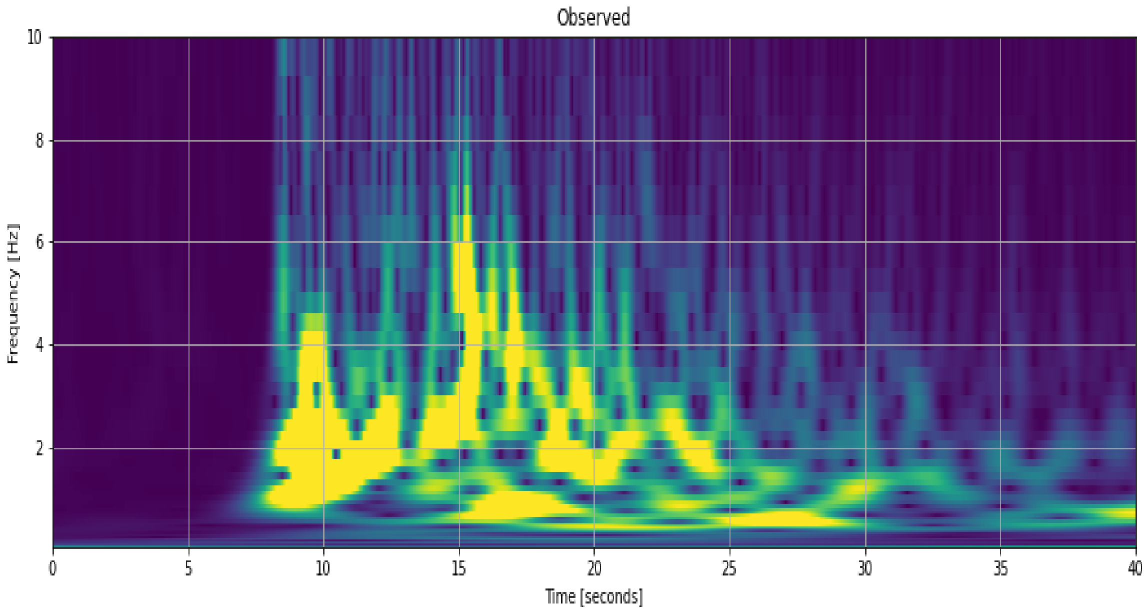
Synthetics vs Reality

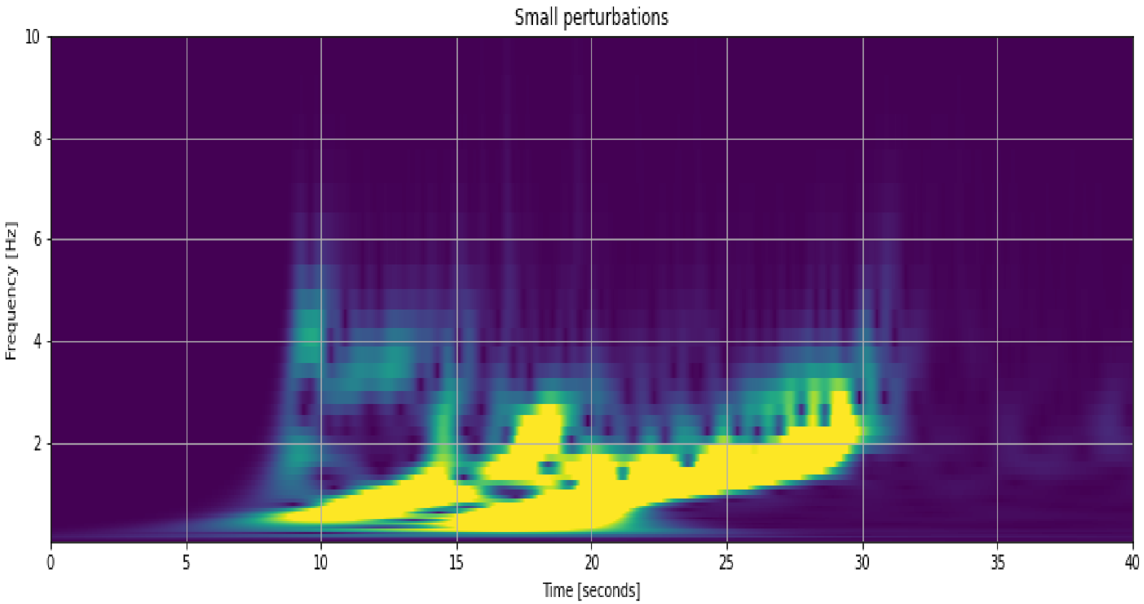
- Identifying features that are robust to modeling uncertainty is very difficult.
- Changing just the scale of stochastic earth model perturbations can significantly change the signal power distribution.

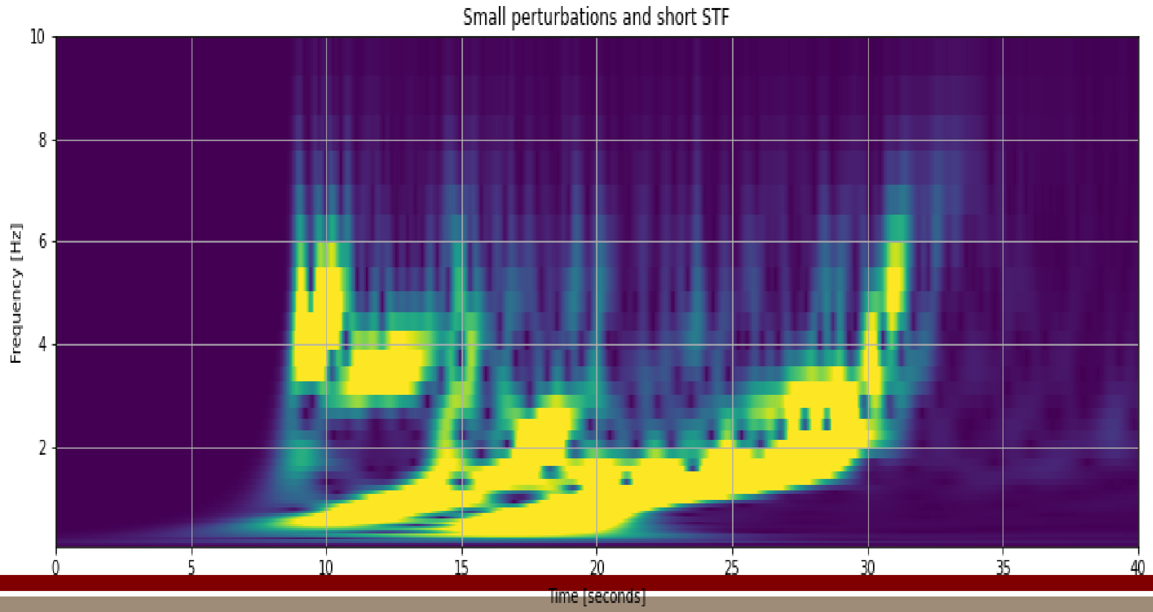


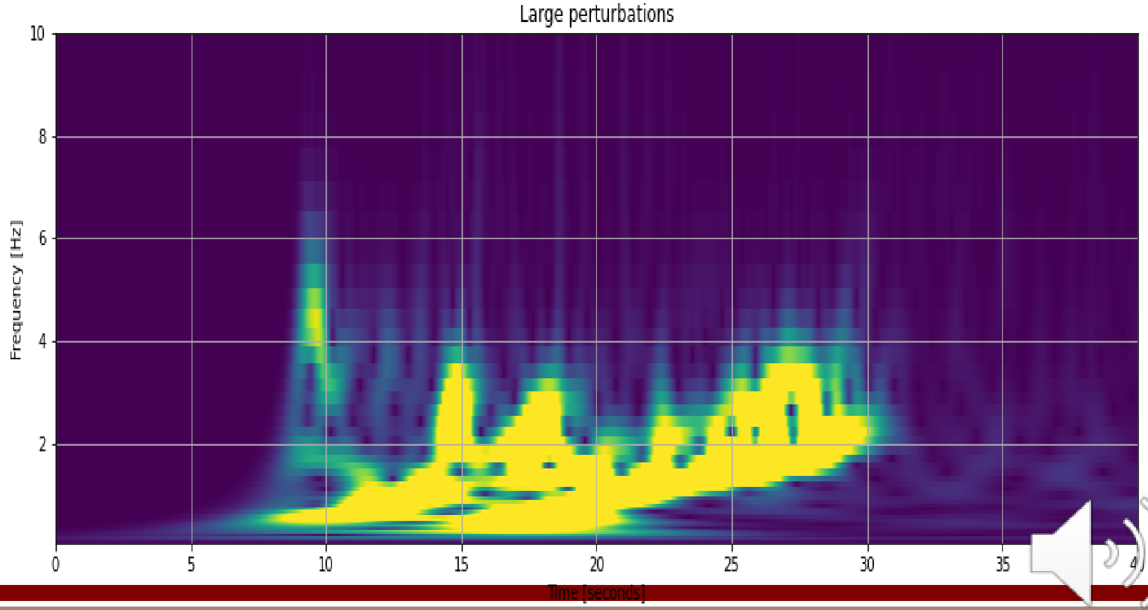


Synthetics vs Reality









FUTURE WORK AND CONCLUSION

Conclusion

- Bayesian inference provides a natural way to express and propagate uncertainty for seismic monitoring and decision-making
- Feature-based inference provides a rigorous way to infer event characteristics from limited data, if we can identify robust features.
- We have not yet identified features that are robust to source and earth model uncertainty.