A Non-Ergodic Ground Motion Model for the Groningen, Netherlands: Merging Empirical Observations and Physics-Based Simulations

CRESCENT: GMM virtual session #2

Inputs and Methodologies"

Grigorios Lavrentiadis[1 2]

Elif Oral[2]

Jinyan Zhao[2]

Domniki Asimaki[2]

Houman Owhadi[2]

Jouni Susiluoto[3]

[1]State University of New York at Buffalo

[2]California Institute of Technology

[3]Jet Propulsion Laboratory

January 29, 2026

Caltech mce

University at Buffalo The State University of New York

Jet Propulsion Laboratory California Institute of Technology

Financial support provided by:

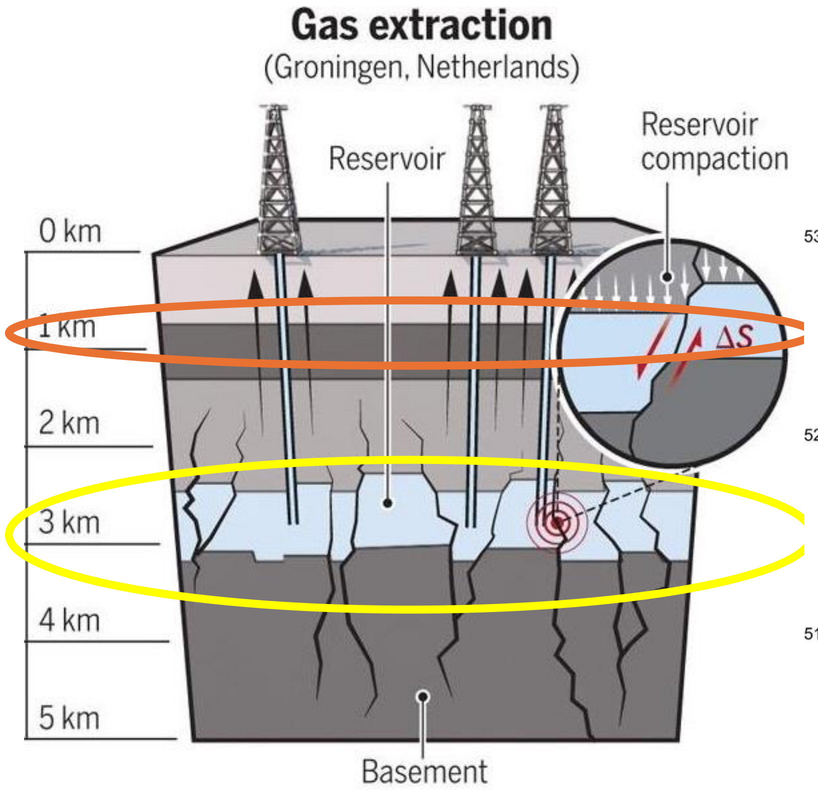
Geomechanics and Mitigation of Geohazards

NSF

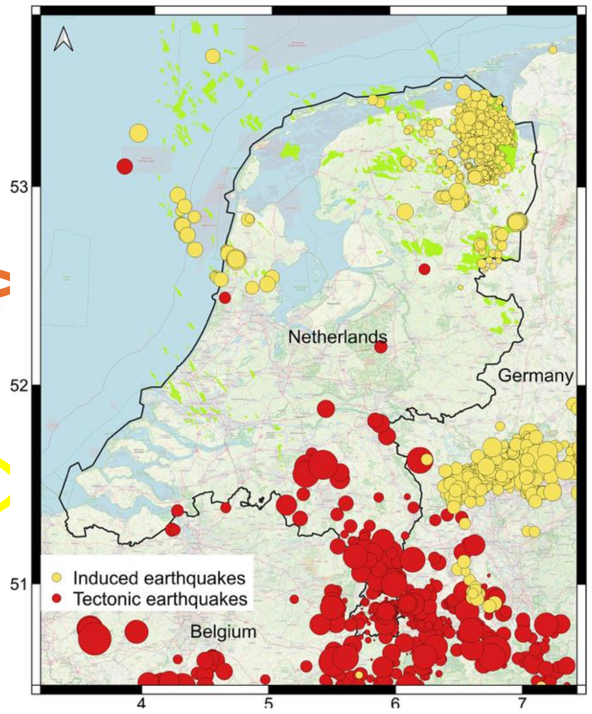
Present Day Challenges for Methodologies and Putting It All Together

- GMM Expressiveness of Complex Physical Phenomena
  - Capture intricate physical effects (e.g., wave propagation in complex media) while preserving physical scaling.
- Scalable Performance on Large Ground Motion Datasets
  - Methodologies need to scale to $> 1e6$ recordings produced by simulations
- Learning from Inhomogeneous Data Sources
  - Recognize differences between empirical and simulated datasets
  - Mitigate simulation bias due to larger dataset size

Study Area: The Groningen Gas Field



https://www.science.org/doi/full/10.1126/science.aat2776



From Bommer et al (2022)

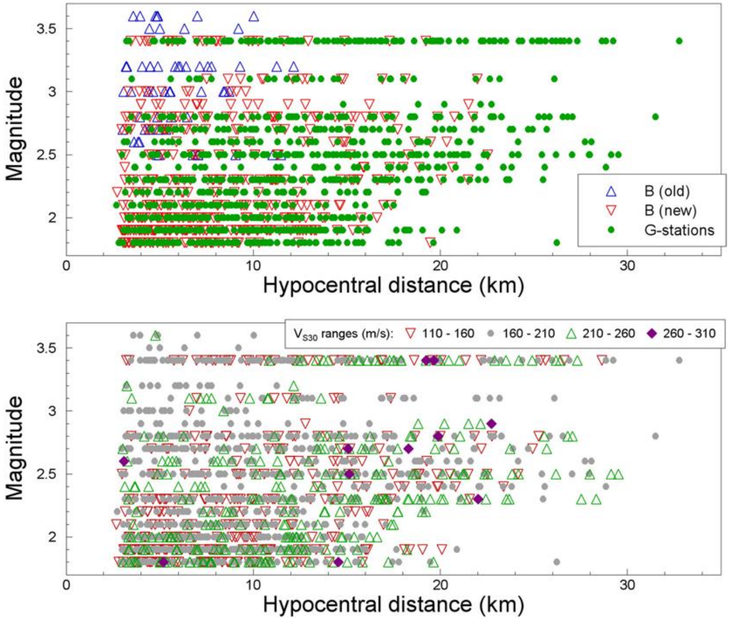
Why Groningen:

- Detailed 3D velocity model
  - -> wave-propagation simulations
- Building and occupancy data
  - -> risk analysis (next phase)

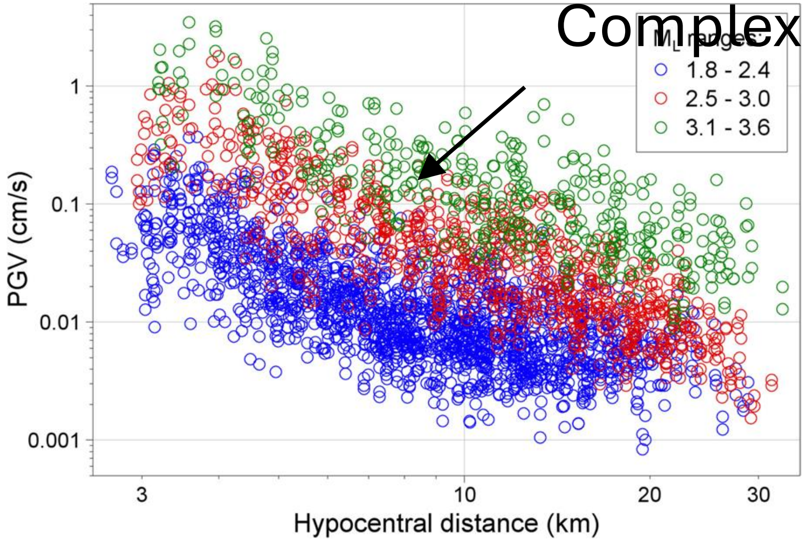
Framework is transferable to other regions with varying data types

Ergodic Backbone Model

Followed ergodic backbone approach, adopting Bommer et al (2022) to capture median source, path, and site scaling

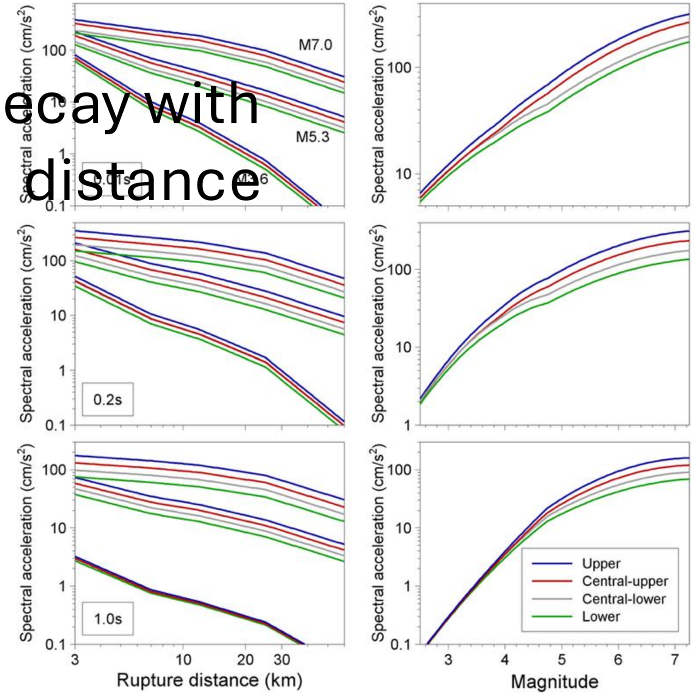


Empirical ground motion catalog



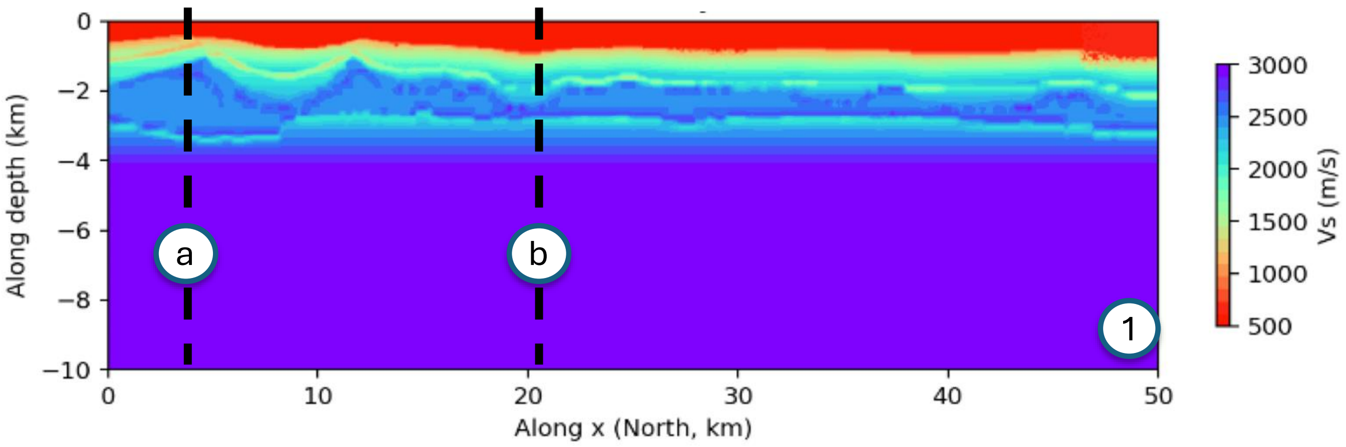
Complex decay with distance

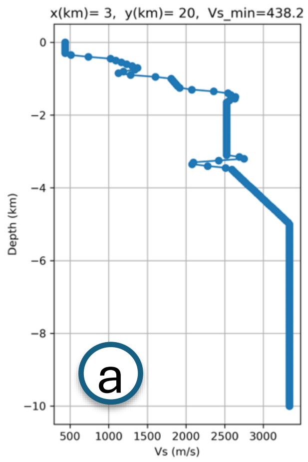
Ground motion attenuation with distance

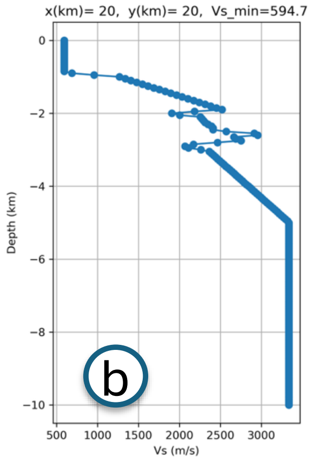


GMM Distance and Magnitude scaling

3D Velocity Model and Simulations: Illuminating Path Effects





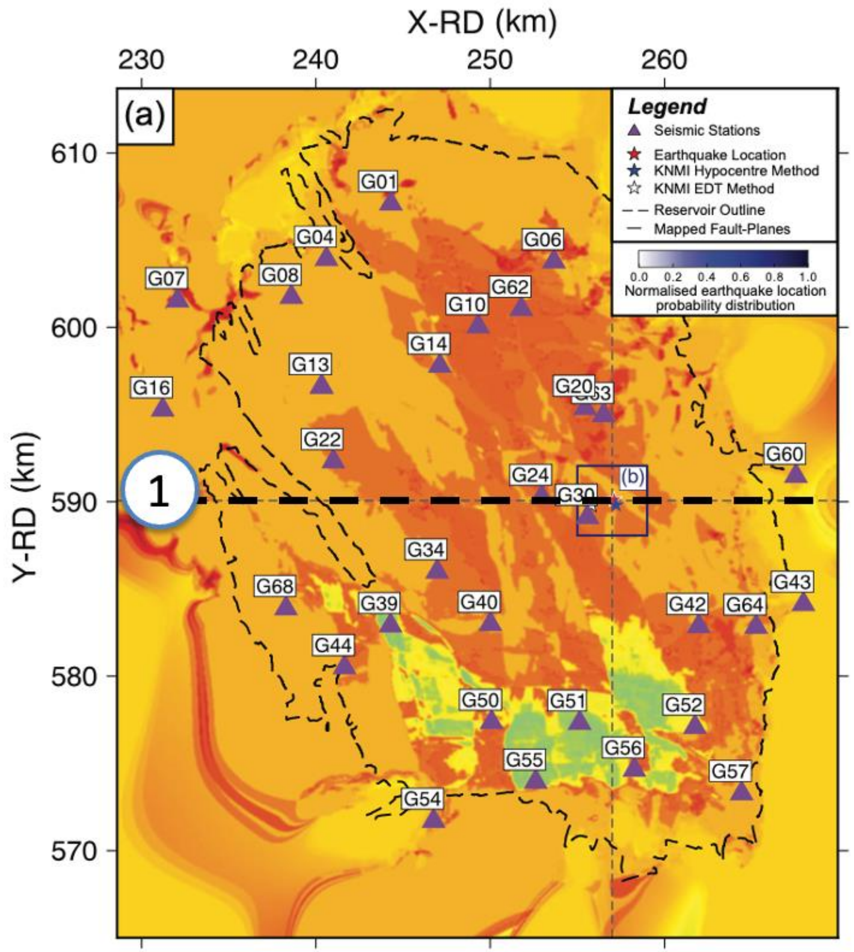


Sources in reservoir at 3km

Strong impedance at ~800m

→ Mode conversion between body and surface waves

Shear-wave velocity profiles (Kruiver et al., 2017)

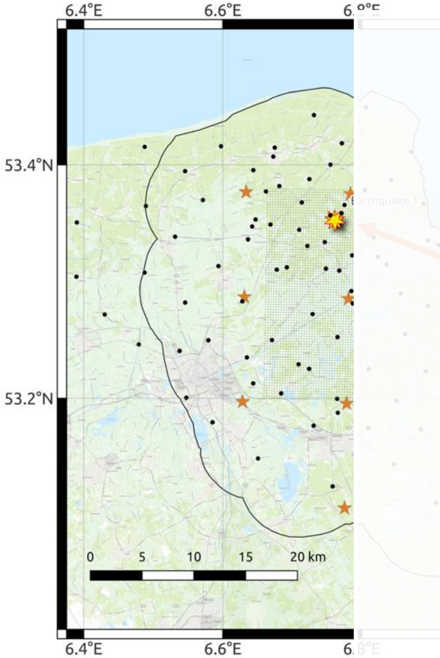


Shear-wave velocity plan view

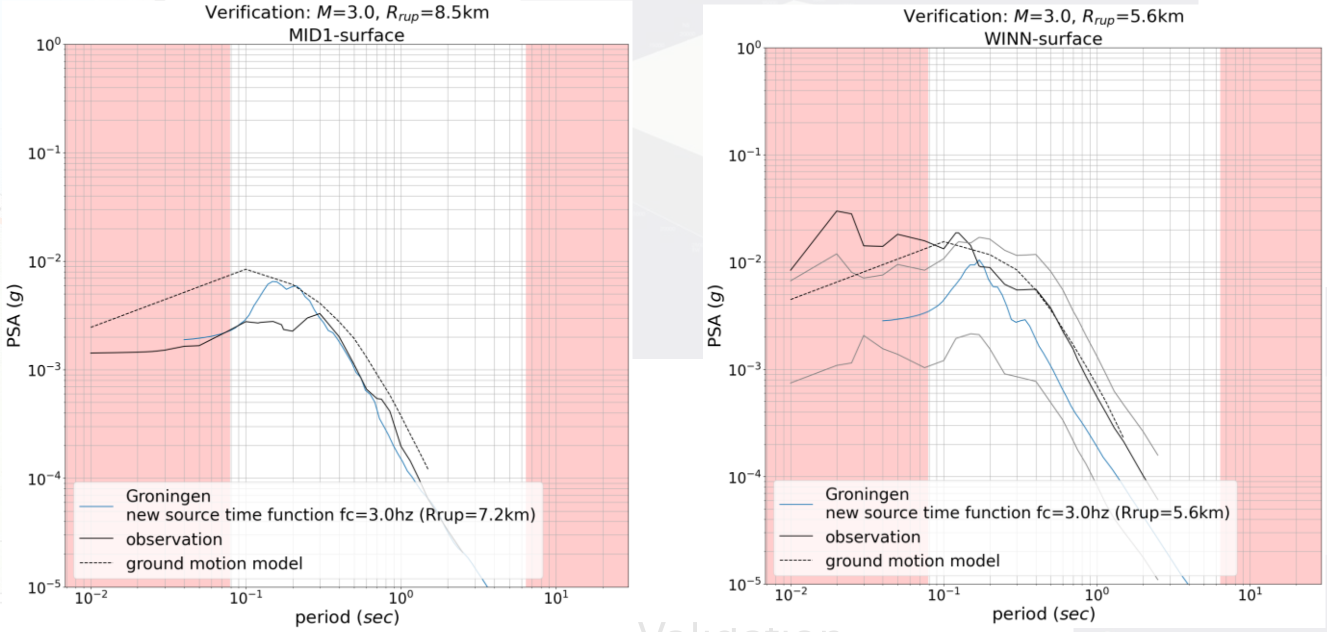
Simulation Validation

Simulations performed in SPECFEM3D with point sources

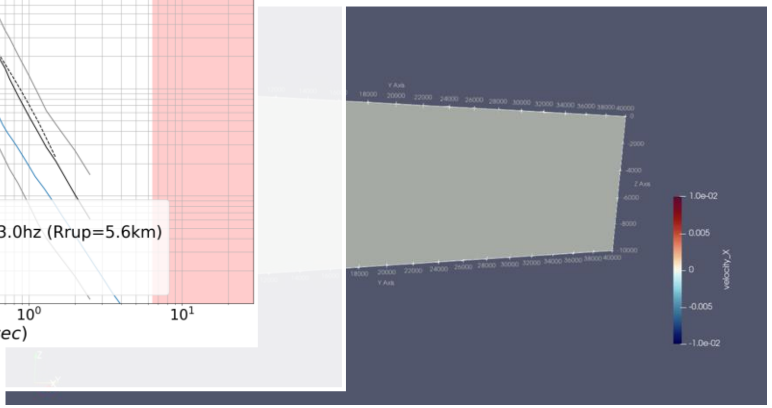
- Validate input parameters with well-recorded real events
  - Adjustment of corner-frequency, source time function, and model



Validation event



Comparison with recorded ground motions

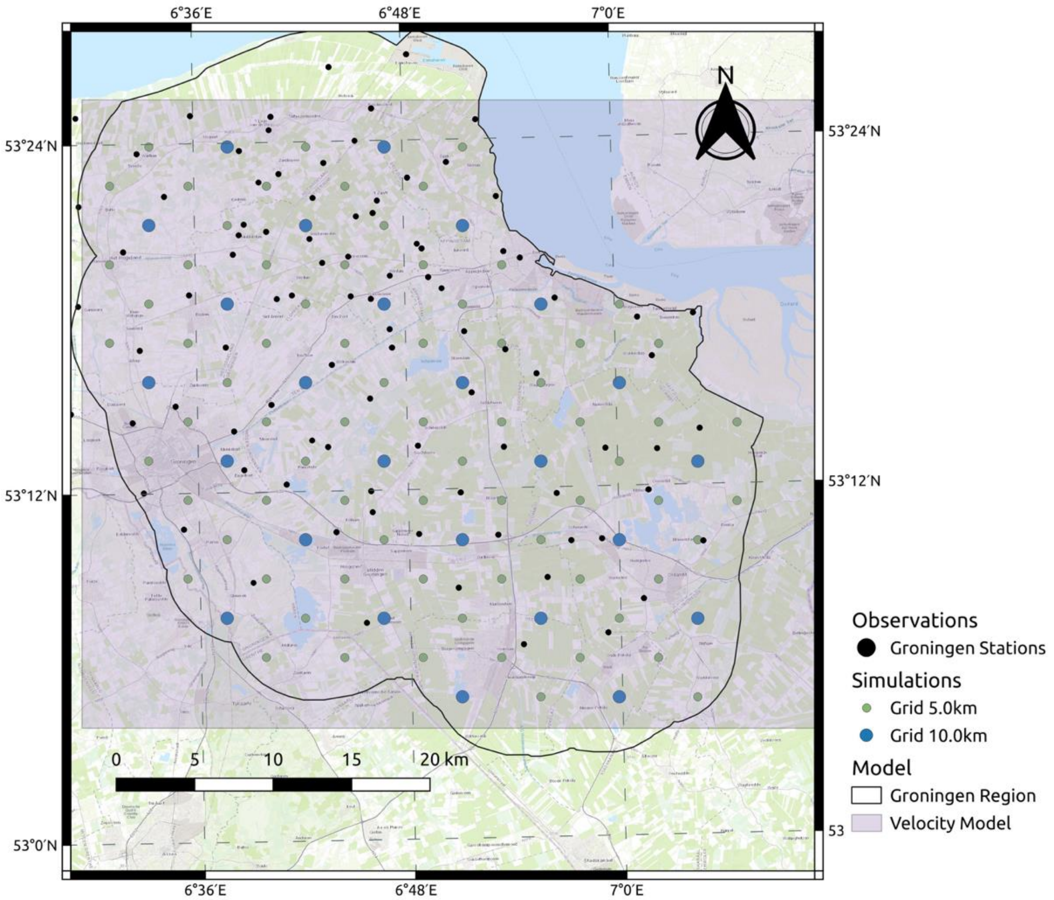


Validation simulations

Production Simulations

- 98 3D point source simulations
- Spaced on 5km grid
- Hypocenter depth: 3km
- Max frequency: 7hz
- 800,000 synthetic ground motions

  (Dense covariance 4.7 TB)

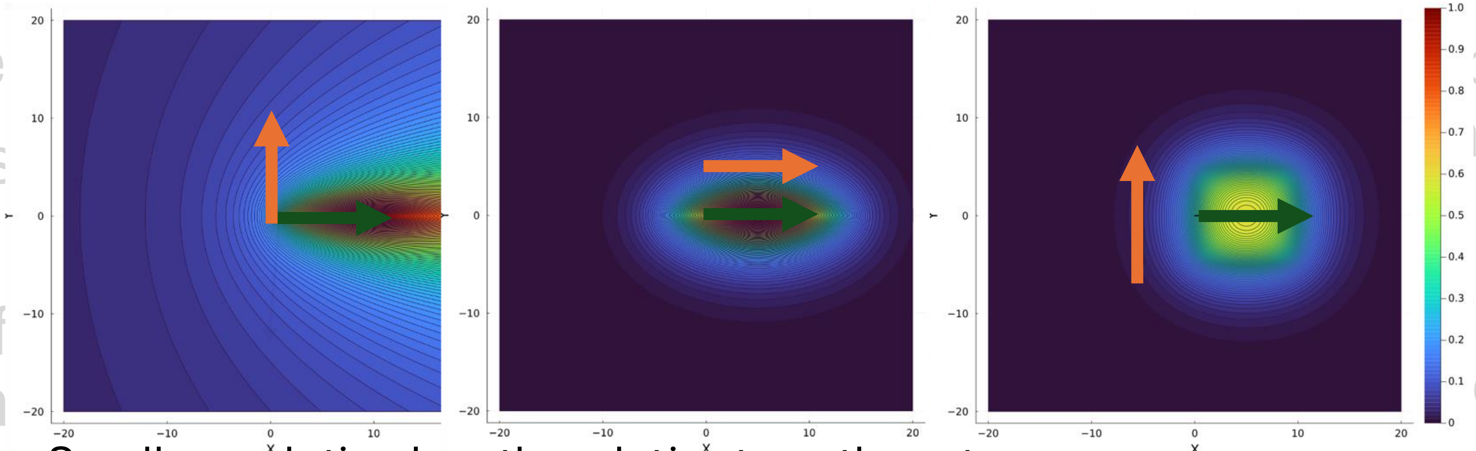


Development of Path Kernel

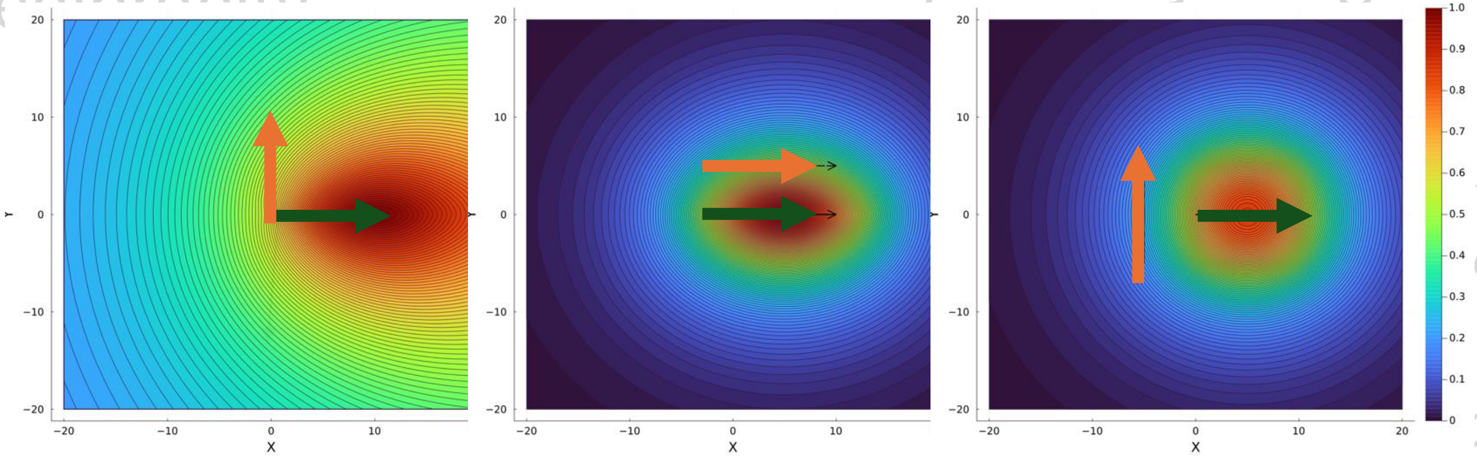
Requireme ... path field ... riability:

- Capture s ... locations
- Account f ... (direction
- Enforce re...

... en path is computed by integrating along



Small correlation lengths relative to path vectors



Large correlation lengths relative to path vectors

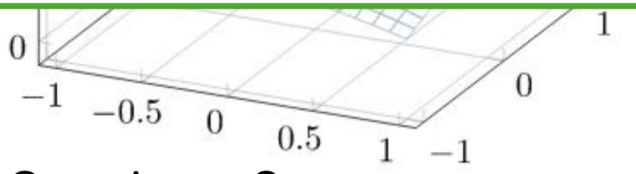
Scalability to Large Datasets

In Gaussian Processes, ground motion predictions are given by:

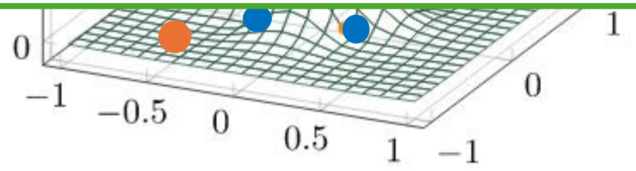
$$\vec{y}_p = \vec{a}_i \cdot \vec{y}_t$$

Key Message:

Far away ground motions do not mater for non-ergodic effects at site of interest (i.e., negligible contribution -> sparse approximation)



Covariance Structure



$\vec{a}_i$ structure

Approximation Scheme

Step 1 **Screening** Identify n_c recordings with highest correlation

- May contain redundant information

Step 2: **Conditional Selection:** Identify n_i mutually most informative

recordings using KL divergence

Complexity:

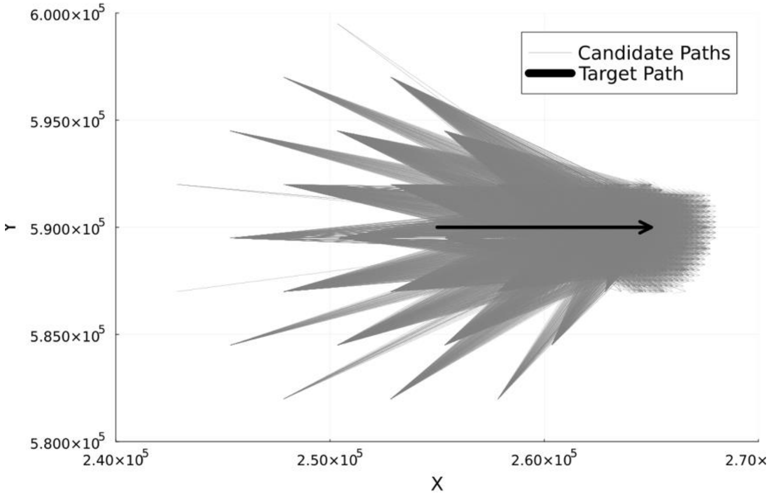
- Memory: $O(N) + O(n_c \log(n_c)) + O(n_i^2)$
- Computation: $O(N) + O(n_c^2 \log(n_c)) + O\left(n_i^3\right)$

For $N \gg n_c$: Memory & Computation → $O(N)$

Memory requirement for simulations dataset 10MB

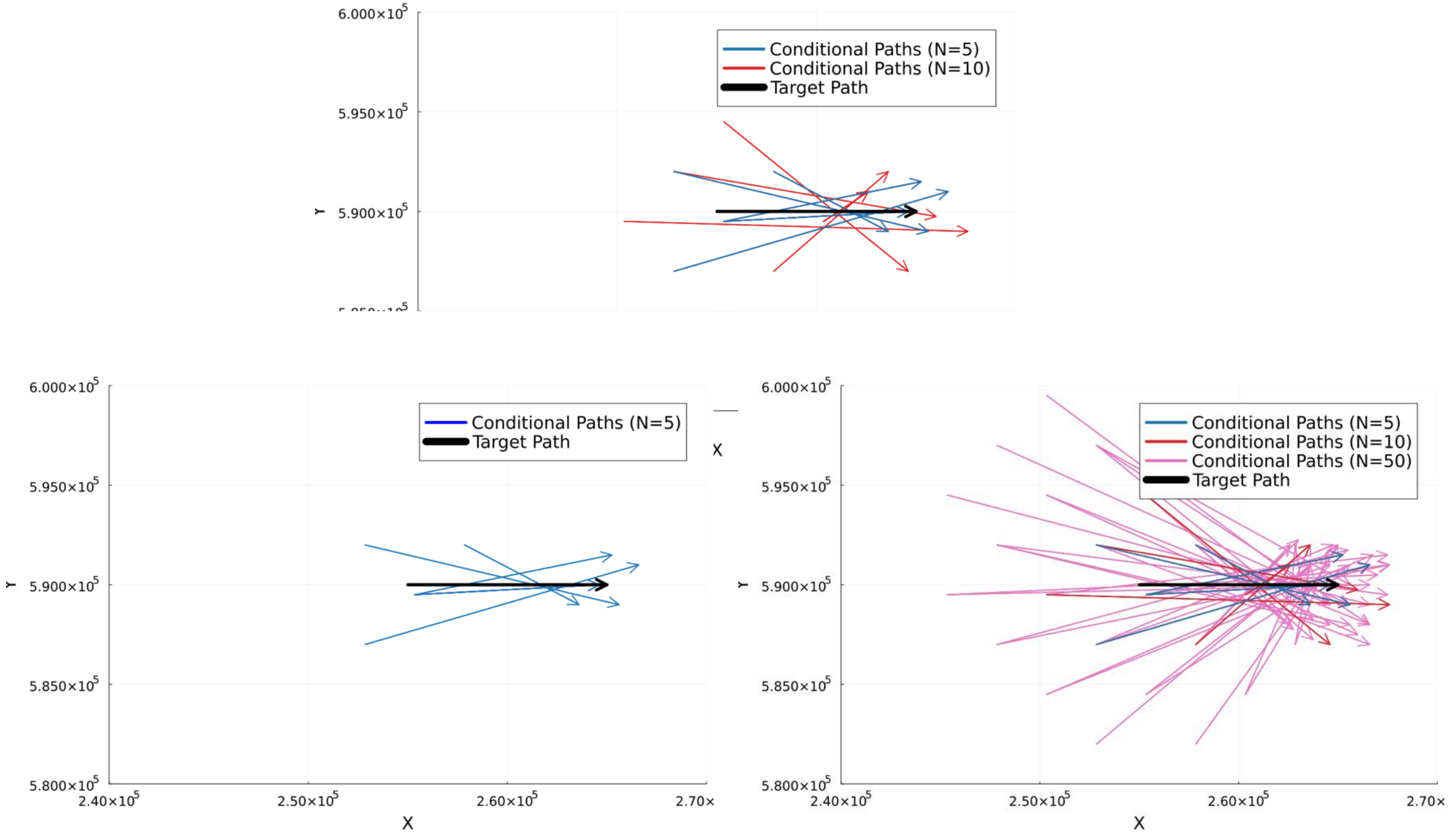
Example of Ground Motion Selection

Step 1 **Screening**



Screened ground motions
(5000 selected ground motions)

Step 2: **Conditional Selection:**



Conditional ground motions

Hybrid Dataset Regression Formulation

Non-ergodic effect decomposition (example for path effects)

- Empirical: $\delta_{P2P} = \hat{\delta}_{P2P} + \dot{\delta}_{P2P}$
- Simulation: $\tilde{\delta}_{P2P} = \hat{\delta}_{P2P} + \ddot{\delta}_{P2P}$

Additional real effects not included in simulations

Simulations artifacts (don't want to propagate to predictions)

Real effects simulations can capture

Hybrid Dataset Covariance

Implied Non-ergodic Variance/Covariance structure:

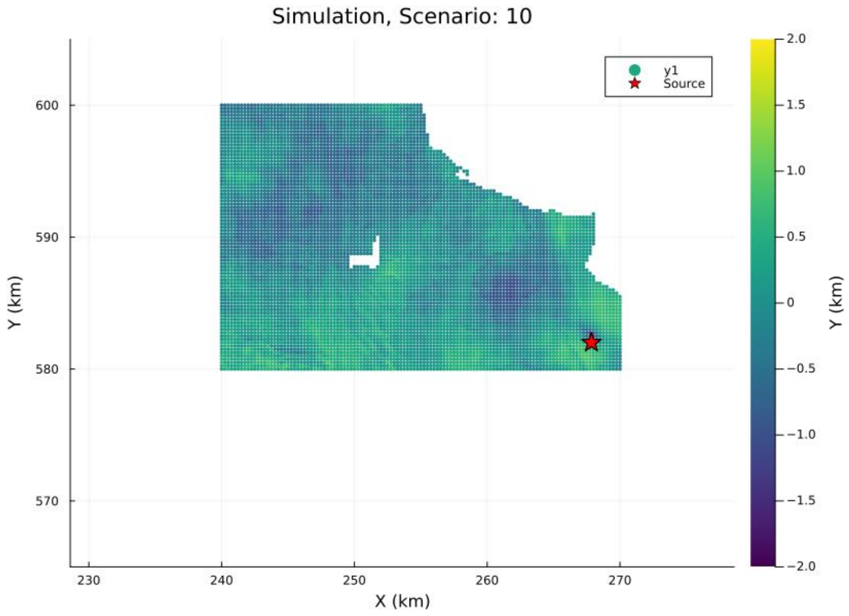
- $\mathrm{Var}\left(\delta_{P2P}(\vec{x})\right) = \omega_{P2P}$
- $\mathrm{Var}\left(\tilde{\delta}_{P2P}(\vec{x})\right) = \widetilde{\omega}_{P2P}$
- $\mathrm{Cov}\left(\delta_{P2P}(\vec{x}), \tilde{\delta}_{P2P}(\vec{x})\right) = \dot{\omega}_{P2P}$

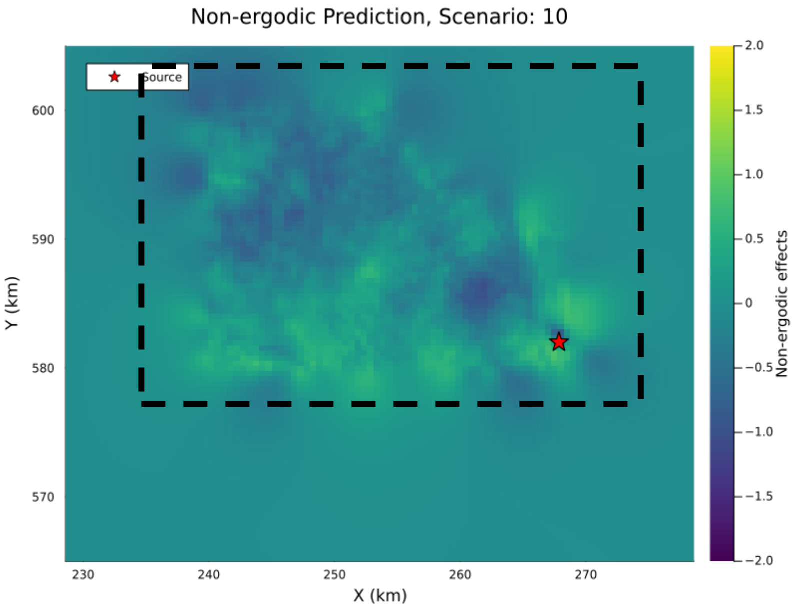
We need to determine three scales instead the traditional one

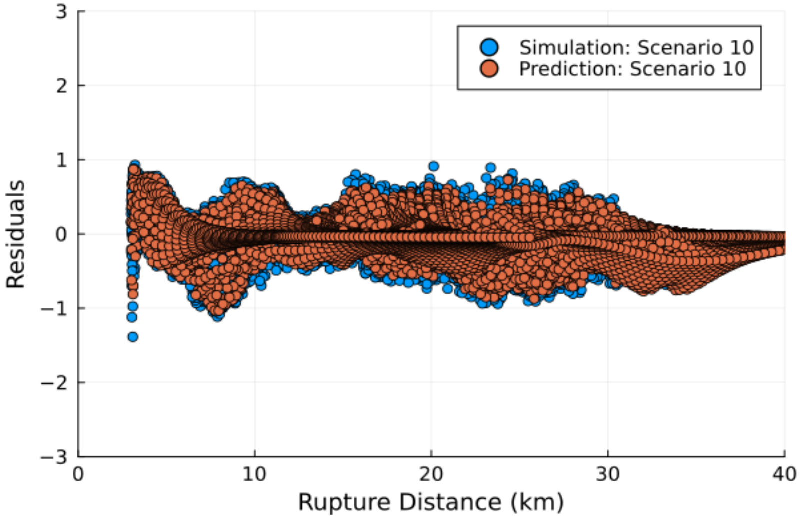
Assumption: stationary simulations' predictive performance

Ground Motion Scenario Prediction

- Comparison of simulations and NGMM prediction for the same scenario

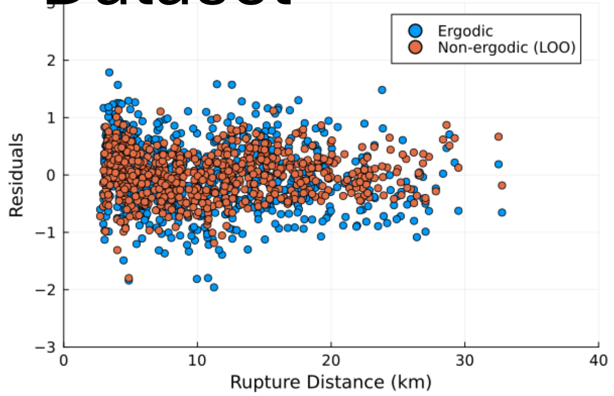


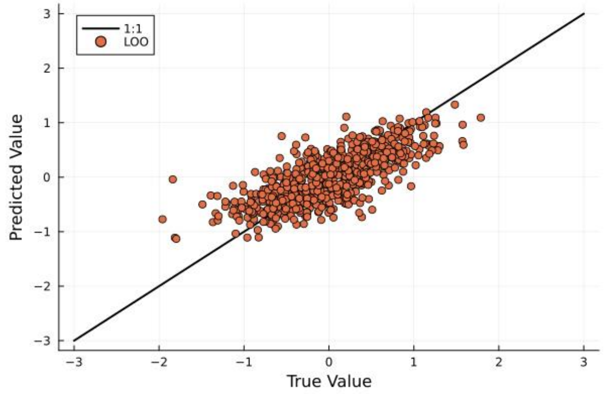




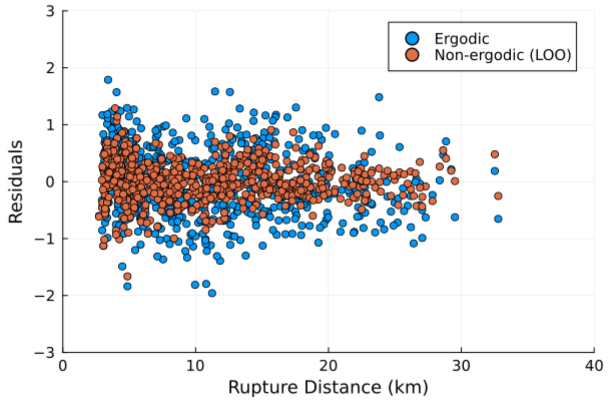
Non-ergodic Model

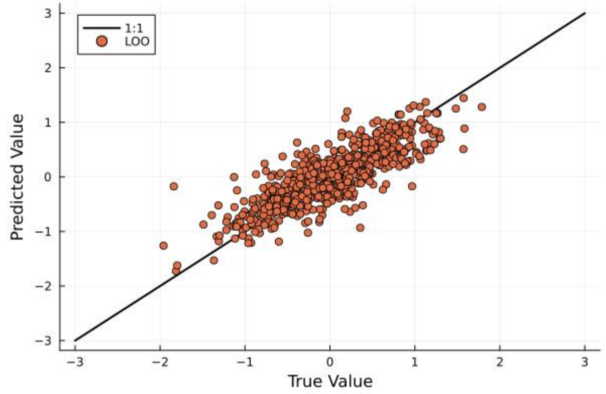
Empirical Dataset





Hybrid Dataset





Approach	RMSE (LOO) $< \sqrt{\phi_0^2 + \tau_0^2}$
Ergodic	0.62
Empirical	0.38
Hybrid	0.34 (45% reduction)

Non-ergodic Effect	Cross-Correlation (Empirical vs Simulations)
Source	0.0
Path	0.64
Site	0.2

Conclusions

- 3D Numerical simulations were able to capture complex wave propagation in the Groningen region
- Proposed kernel function was able to learn systematic path effects from empirical and simulated records
- Proposed approximation scheme significantly improved computational efficiency
- Hybrid regression approach leads to further reduction in aleatory variability