A hybrid ML-physical modelling approach for efficient approximation of tsunami waves at the coast for probabilistic tsunami hazard assessment

Naveen Ragu Ramalingam, Kendra Johnson, Marco Pagani, and Mario Martina
PhD Student at IUSS Pavia, Italy

IUSS Scuola Universitaria Superiore Pavia

GEM GLOBAL EARTHQUAKE MODEL working together to assess risk

EGU General Assembly 2022, Austria 23–27 May 2022

This presentation participates in OSPP
EGU
Outstanding Student & PhD candidate Presentation contest

A hybrid ML-physical modelling approach for efficient approximation of tsunami waves at the coast for probabilistic tsunami hazard assessment

Naveen Ragu Ramalingam, Kendra Johnson, Marco Pagani, and Mario Martina
PhD Student at IUSS Pavia, Italy

IUSS Scuola Universitaria Superiore Pavia

GEM GLOBAL EARTHQUAKE MODEL working together to assess risk

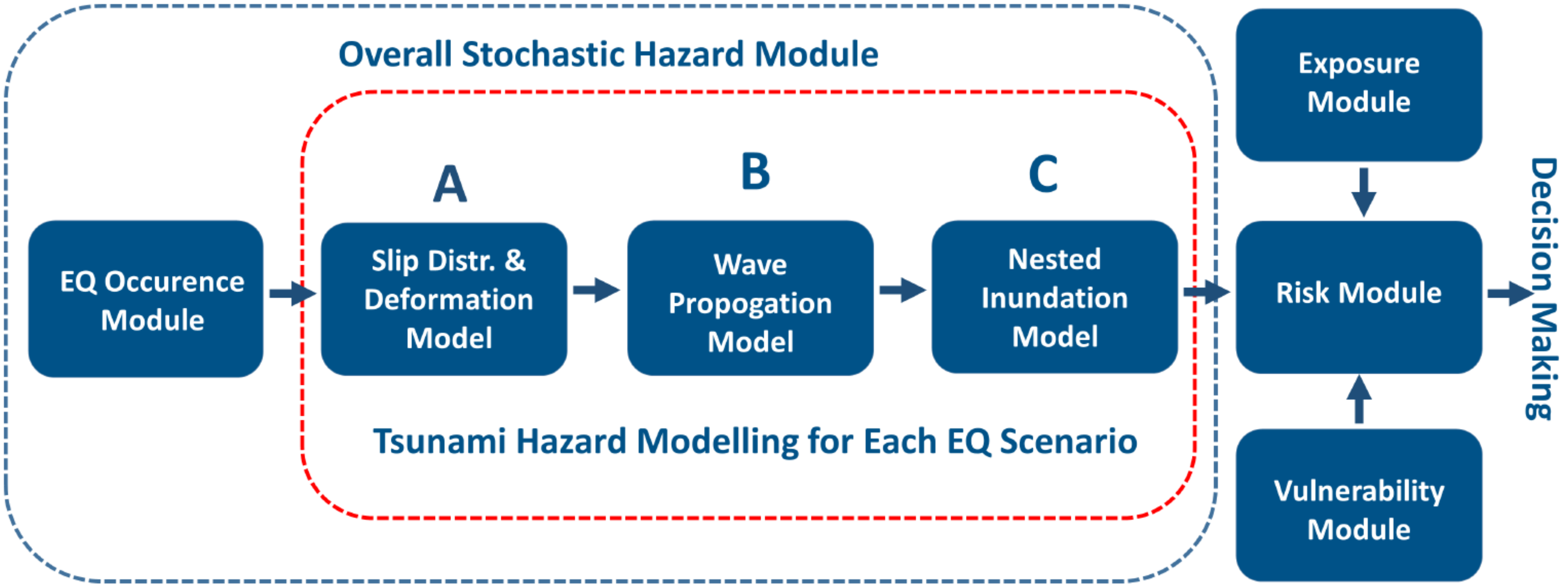
EGU General Assembly 2022, Austria 23–27 May 2022

This presentation participates in OSPP

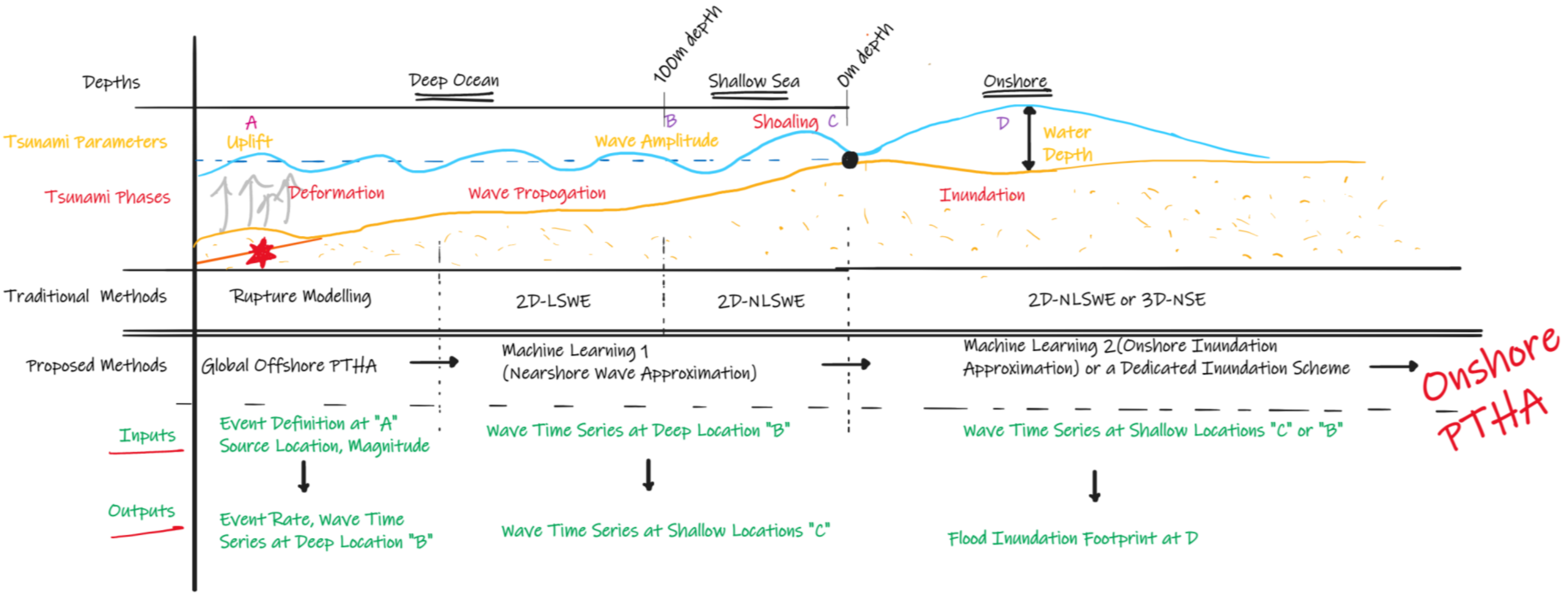
EGU

Outstanding Student & PhD candidate Presentation contest

Workflow of probabilistic tsunami risk assessment(EQ sources)

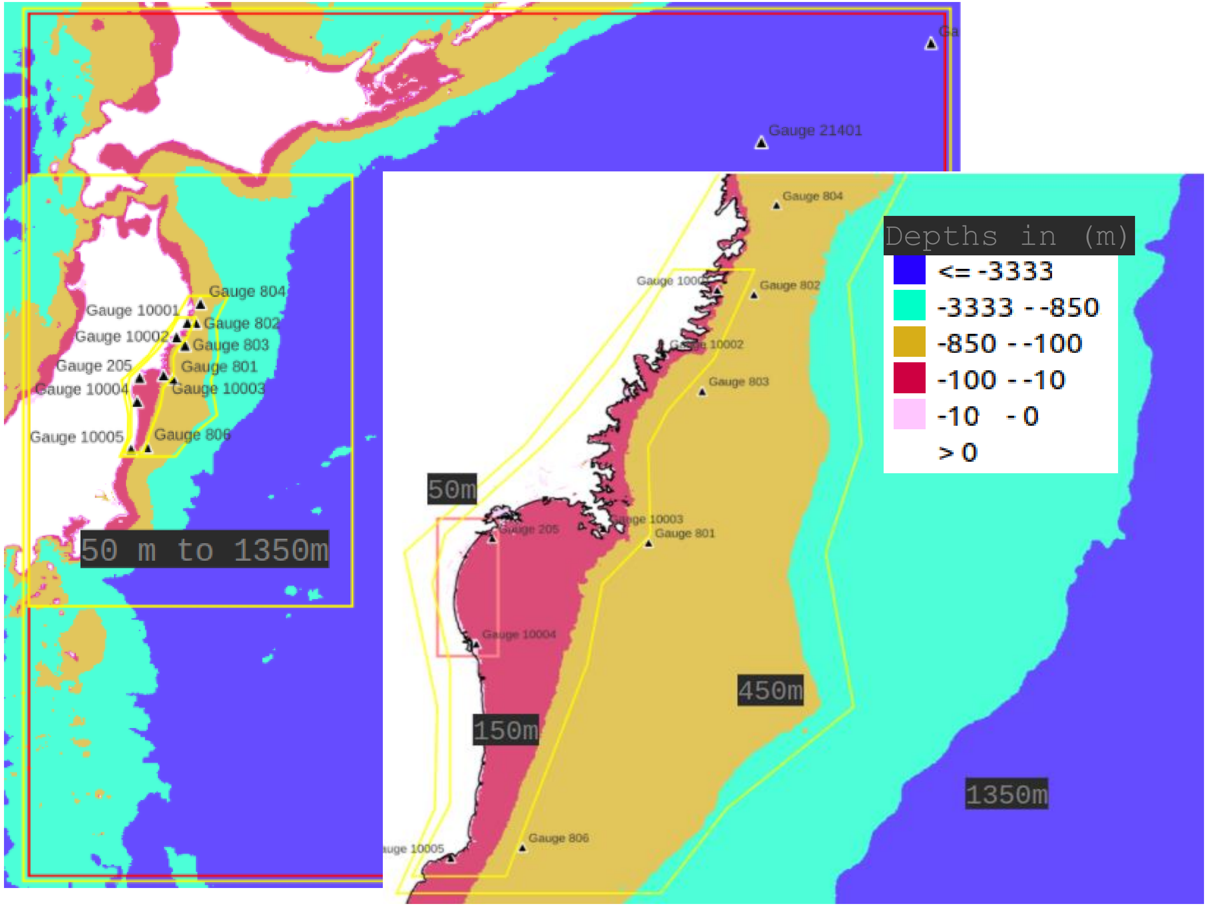


Hazard Module

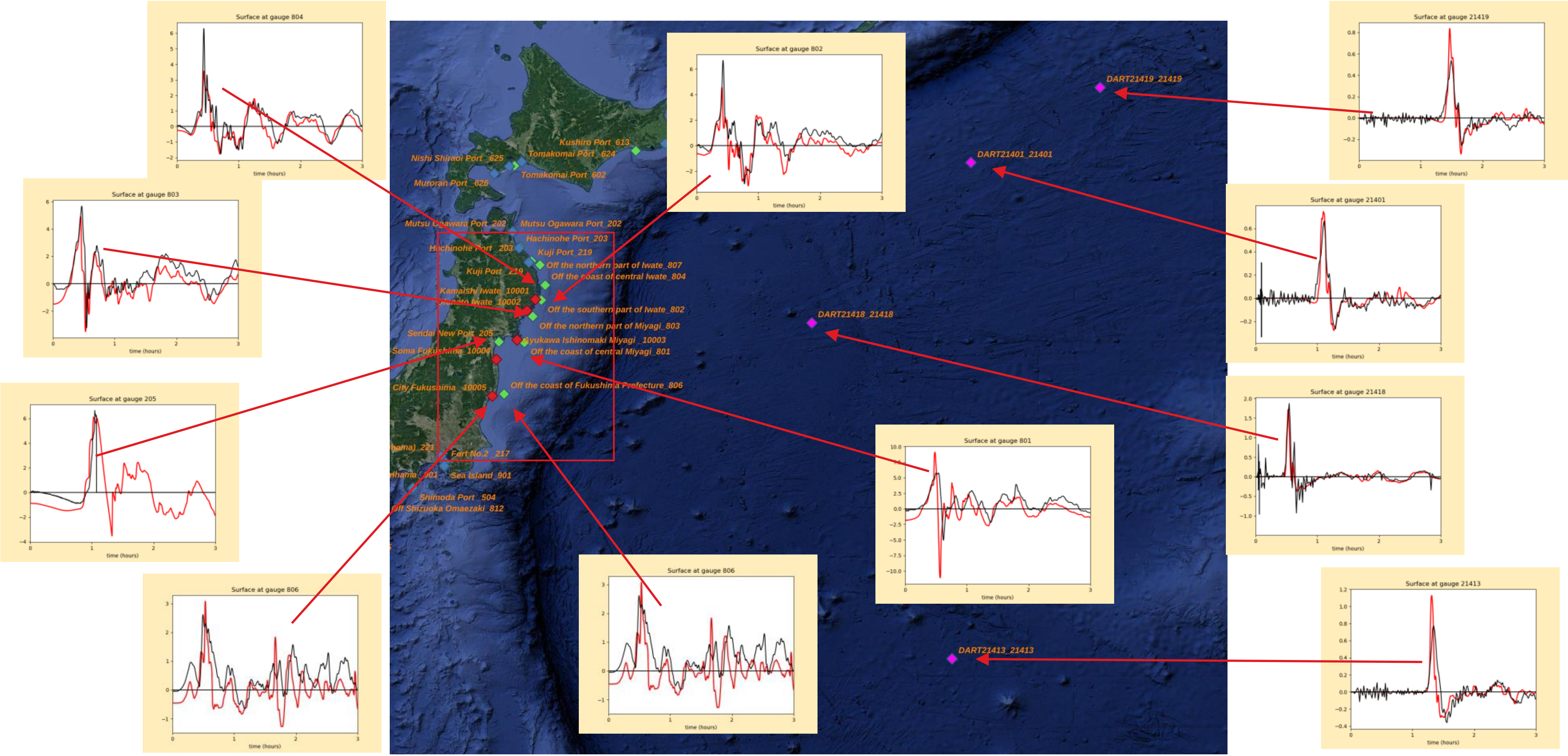


A Hydrodynamic Model (to capture the nearshore and onshore dynamics)

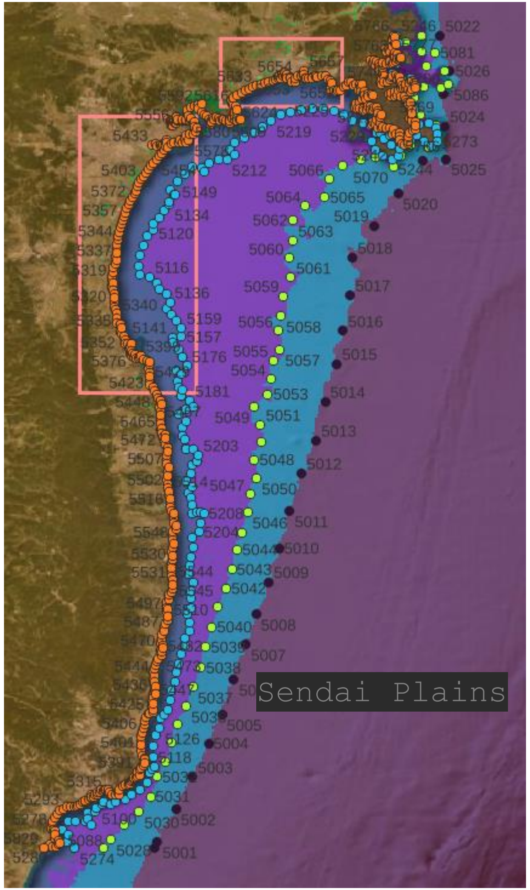
- Tsunami model for Tohoku region of Japan with GeoClaw
- Uses topo-bathy-defense data from global and local sources - (JP Cabinet project data + GEBCO 2021 + Copernicus DEM)
- Calibrated with offshore and onshore observation data for Tohoku 2011 event and different source models
- For validation of wave approximation ML model simulated other historical events

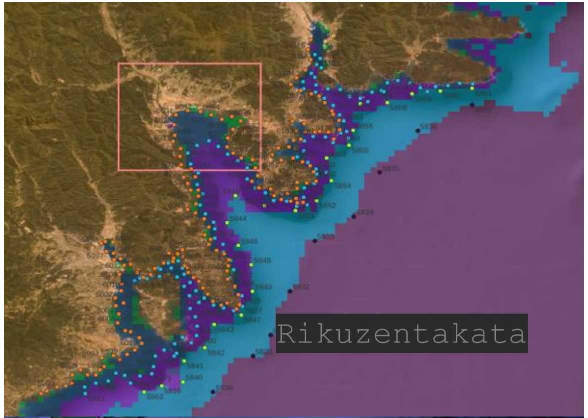


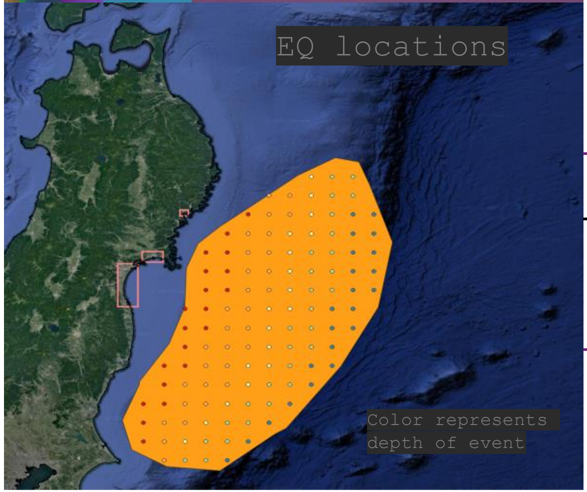
Typical validation with Tohoku 2011 offshore(Fuji 2011)



Propagation/Inundation database for ML







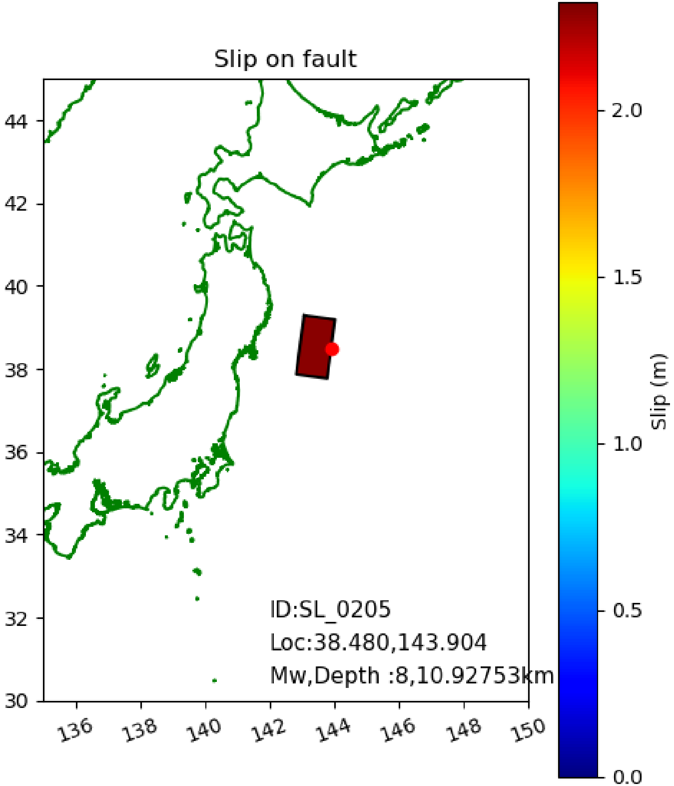
- 1133 locations for observing waveform at different depths (5,25,50 and 100m)
- 3 selected AOI selected to record the max flood depths highlighted by pink rectangles.
- 594 events of varying location and magnitude are simulated for 6 hours of duration
- Homogenous slip events for rectangular fault whose dimensions are scaled based on Mw(7.5, 8, 8.5, 9,9.5)
- Fault parameters(angles) defined using SLAB.2 data and deformation modelled using Okada solution

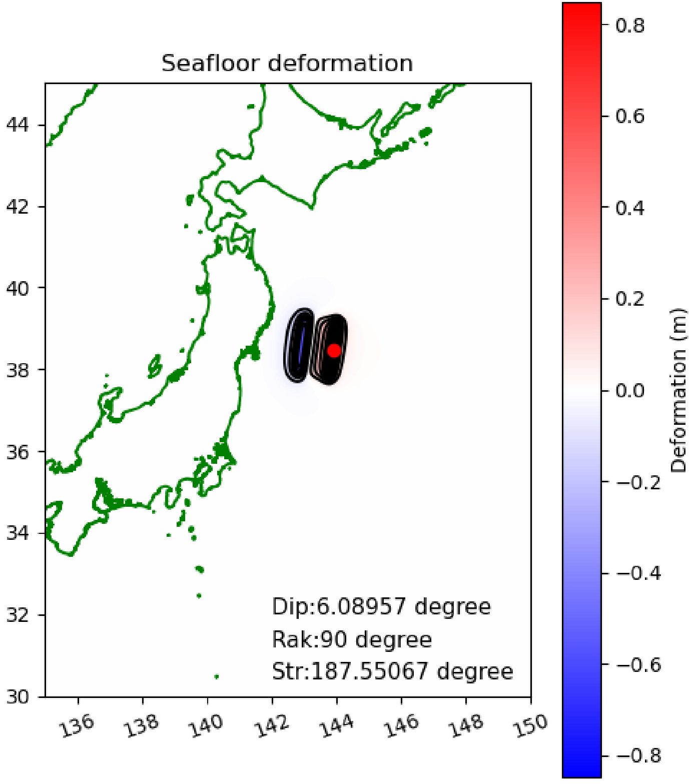
EQ Source Parameters

Range	Mw	Lat	Lon	Dep	Rak	Str	Dip
min	7.5	35.73	141.15	10.2	90	187.20	5.54
max	9.5	39.48	143.90	45.7	90	225.78	17.0

Preprocessing Events – Feature design for ML

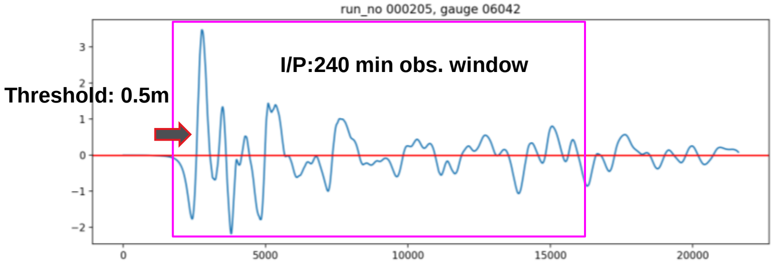
Event 205



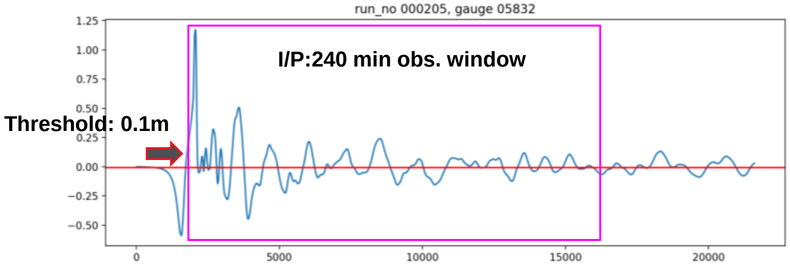


Preprocessing Events – Feature design for ML

Event 205

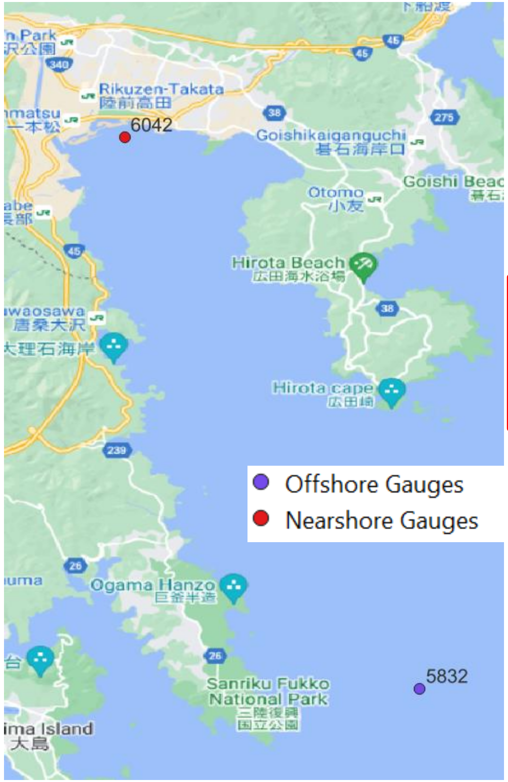


Nearshore Gauge



Offshore Gauge

Possible ML and training configuration



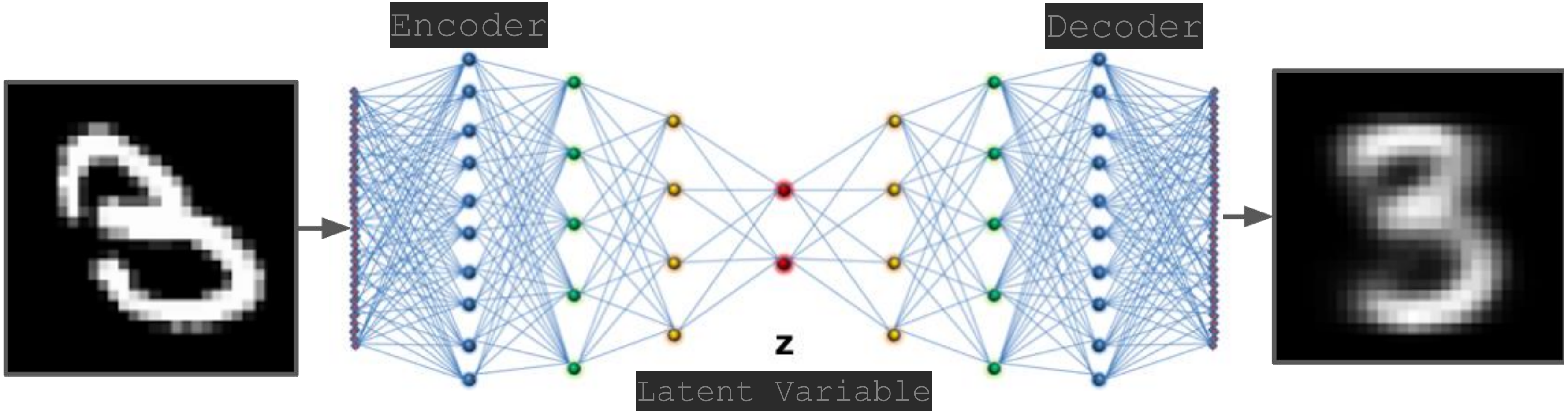
Offshore Parameter → ML Methods → Nearshore Parameter

Option 1:
Tsunami wave time series → Support Vector Machines → Max tsunami amplitude

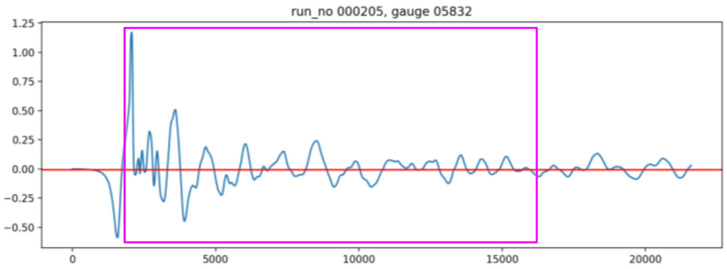
Option 2:
Tsunami wave time series → Variational Autoencoder → Tsunami wave time series

Option 3:
Tsunami wave time series → Variational Autoencoder → Max tsunami Inundation

Time to feed the ML model - VAE(variational autoencoder)

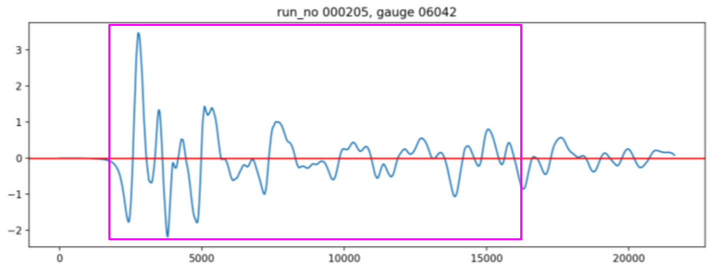


Offshore Gauge



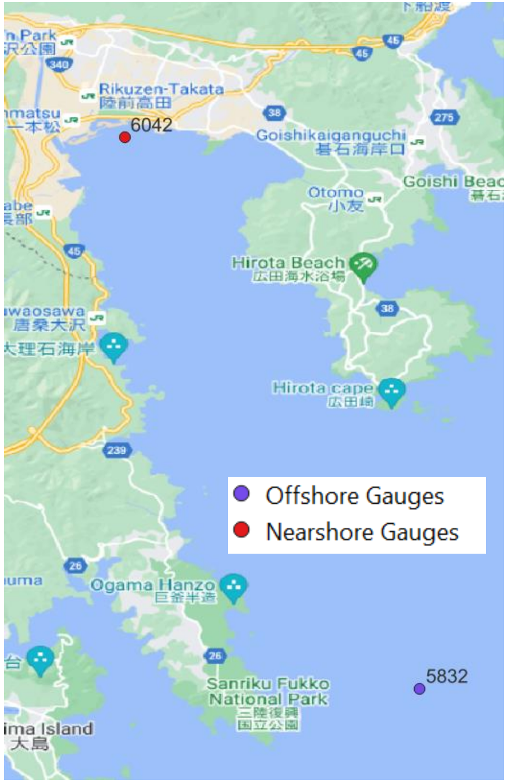
VAE Transformer

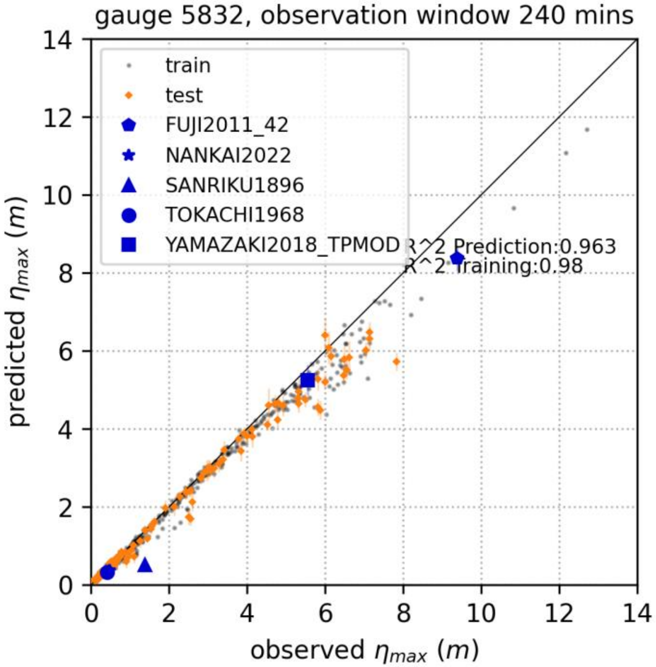
Nearshore Gauge



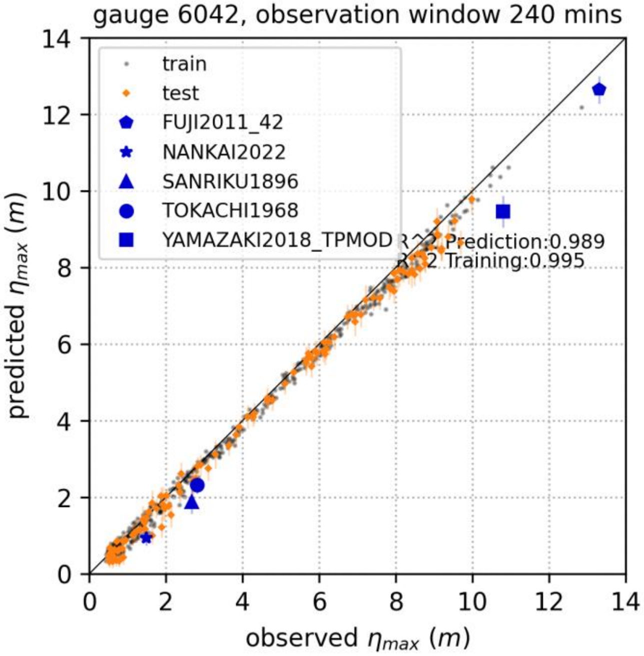
Liu et al 2021

But does it work? Testing at Rikuzentakata





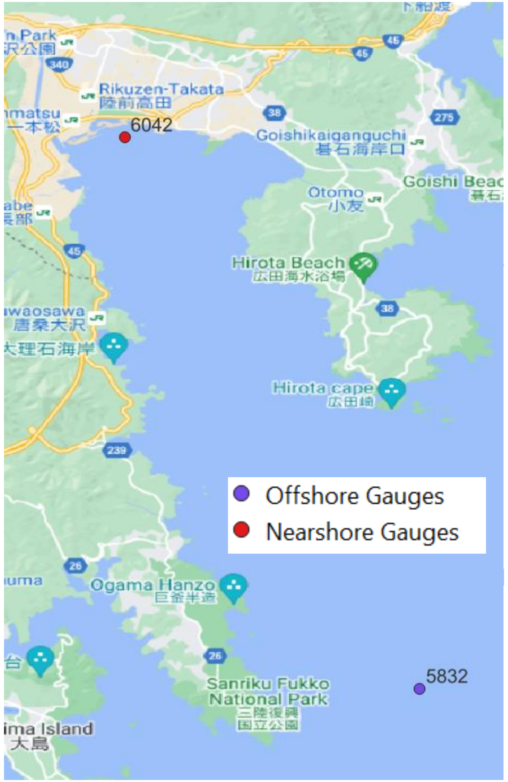
Offshore Gauge

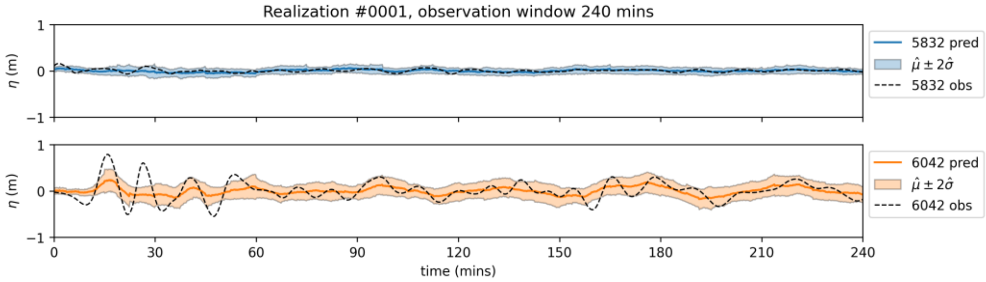


Nearshore Gauge

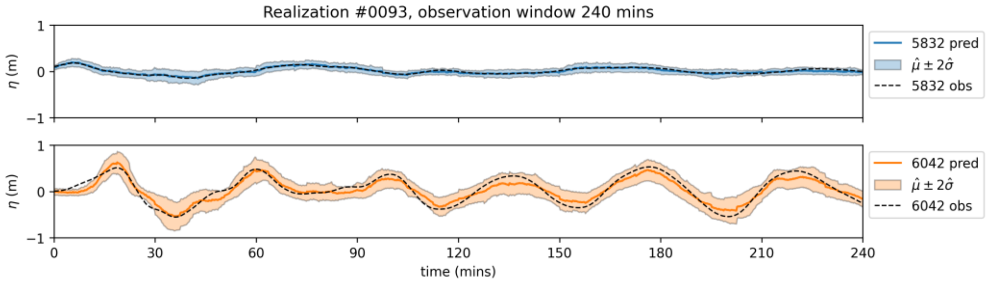
Sample size: Events passing threshold – 523, Training set – 418, Test set – 105, Historical Set - 5

Predicting the test events – are similar to training events but unseen

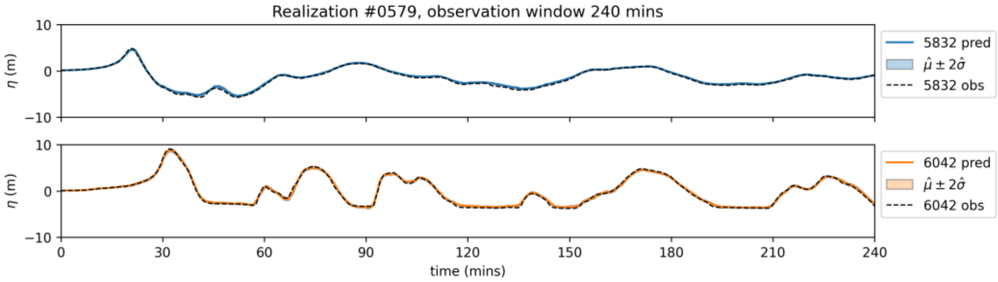




Small mag events, large spread ML doesn't work very well(5%)

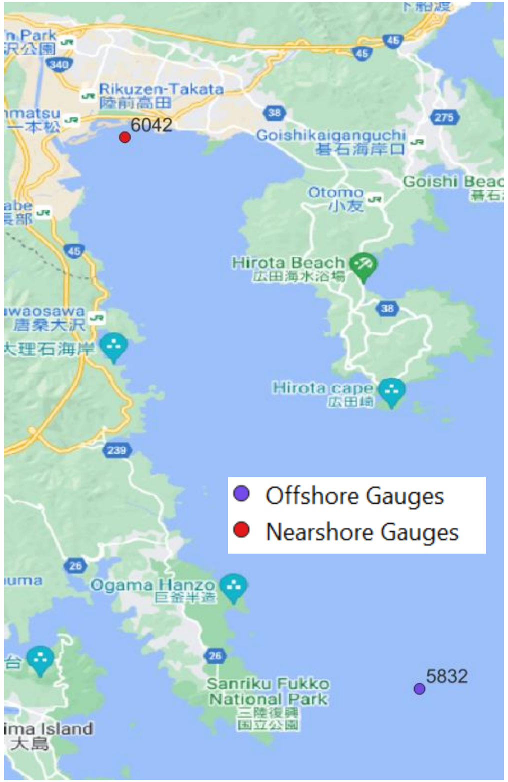


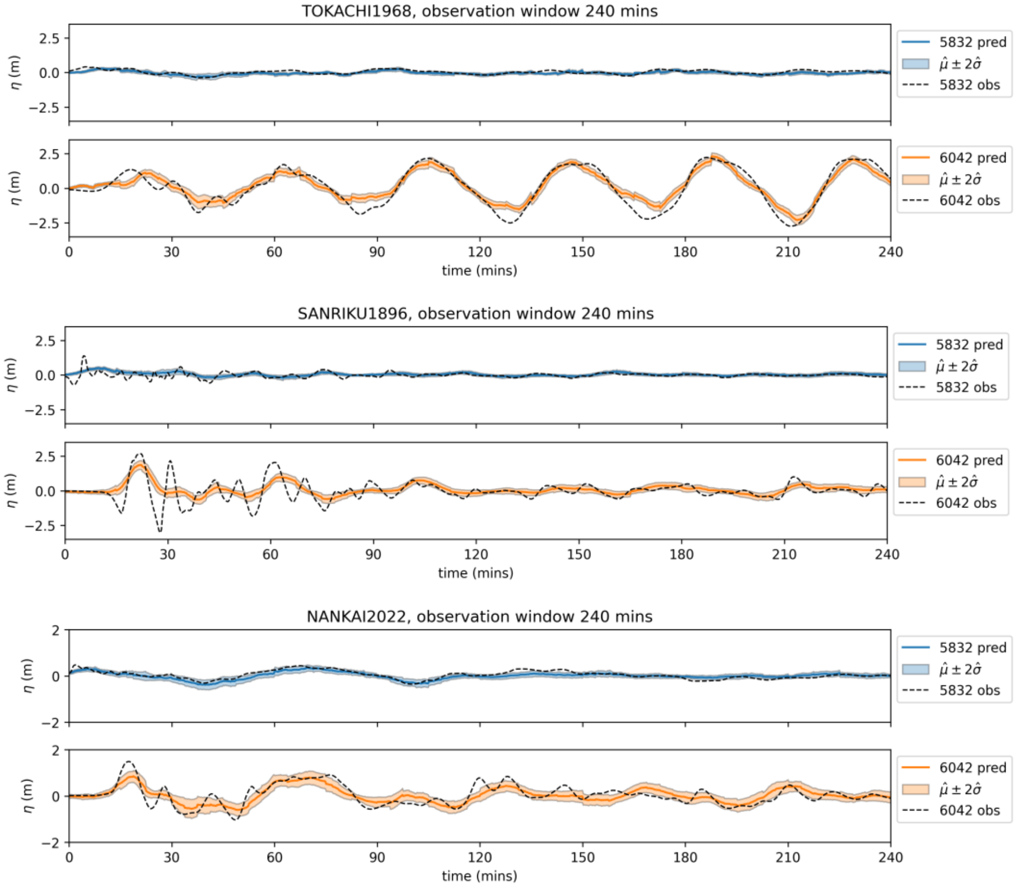
Small/mid mag events, ML works and has some uncertainty(10%)



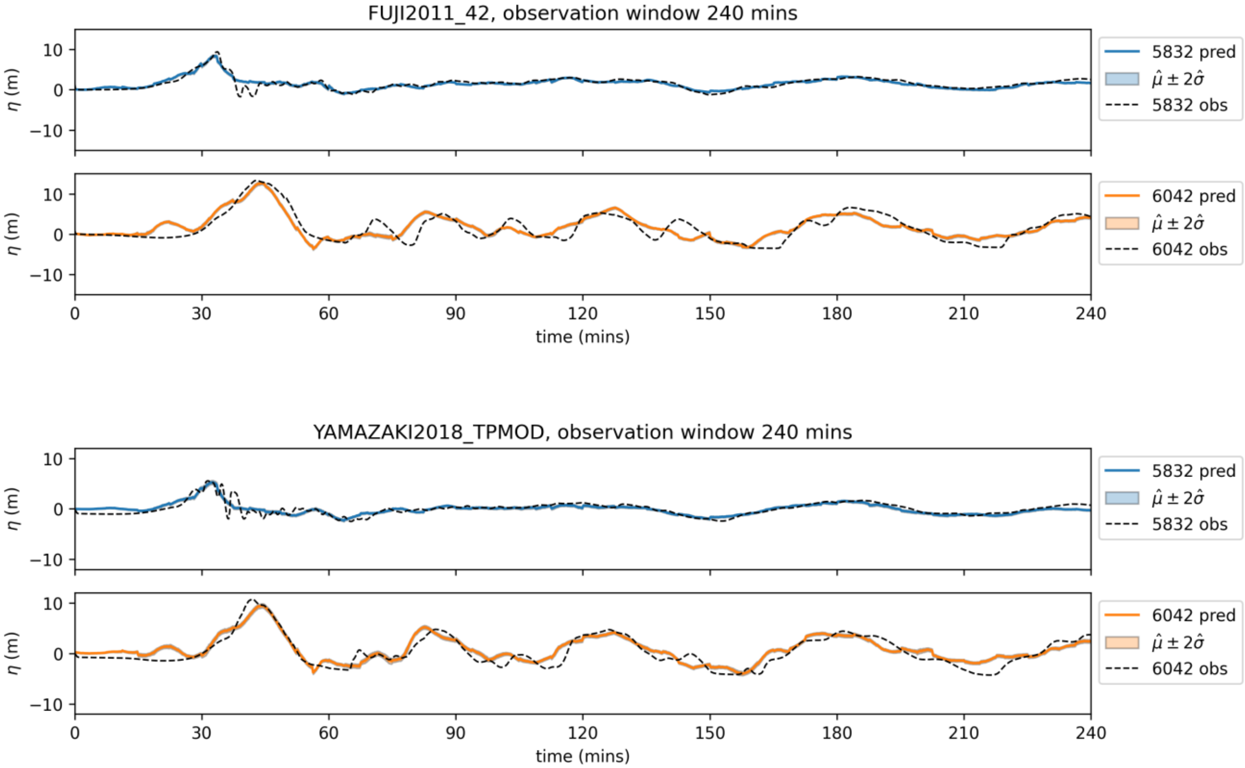
Large mag events, ML works very well(85%)

Some historical events – unseen events

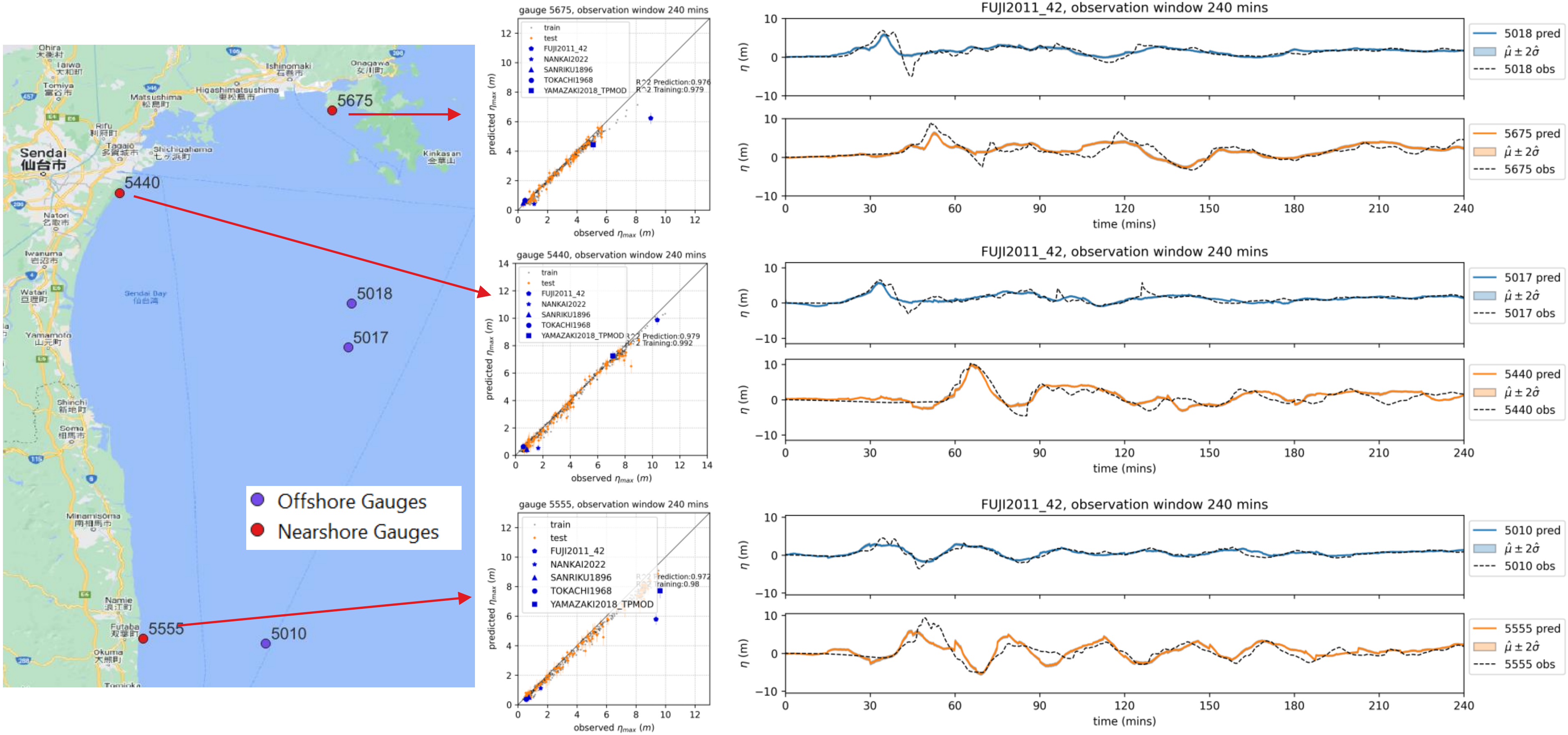




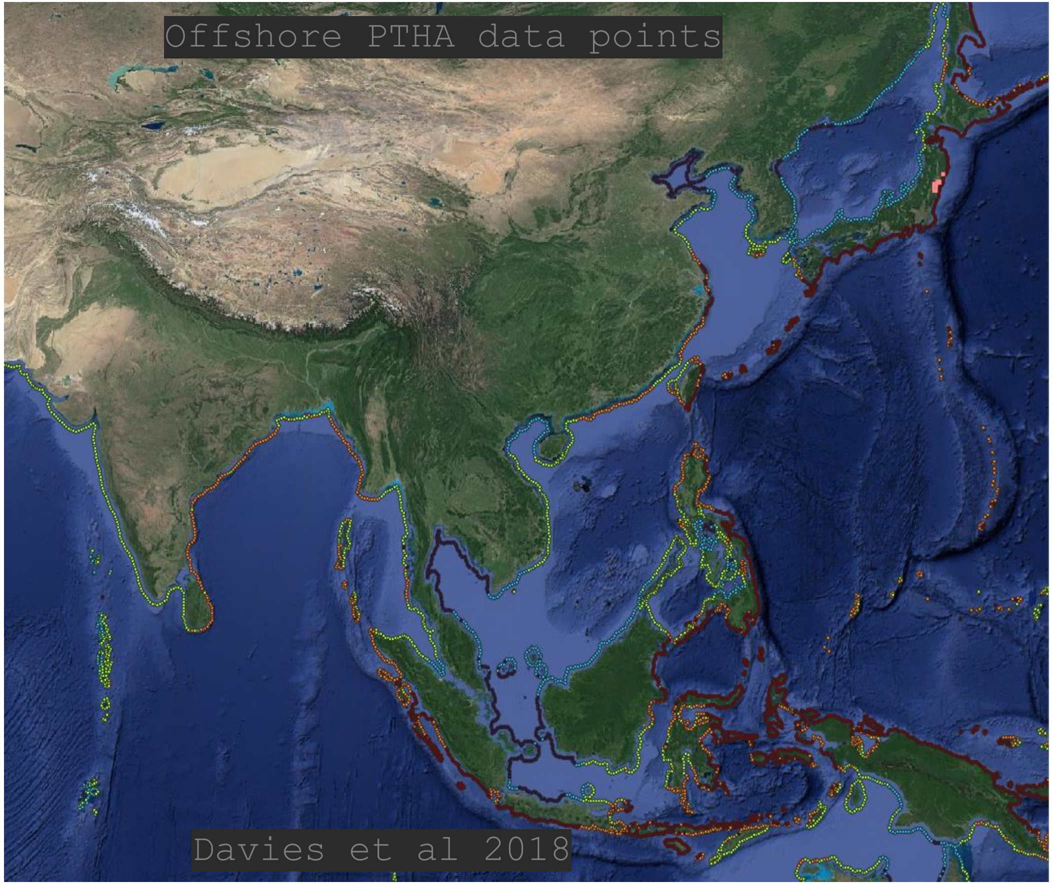
Tohoku 2011 event



But does it work?

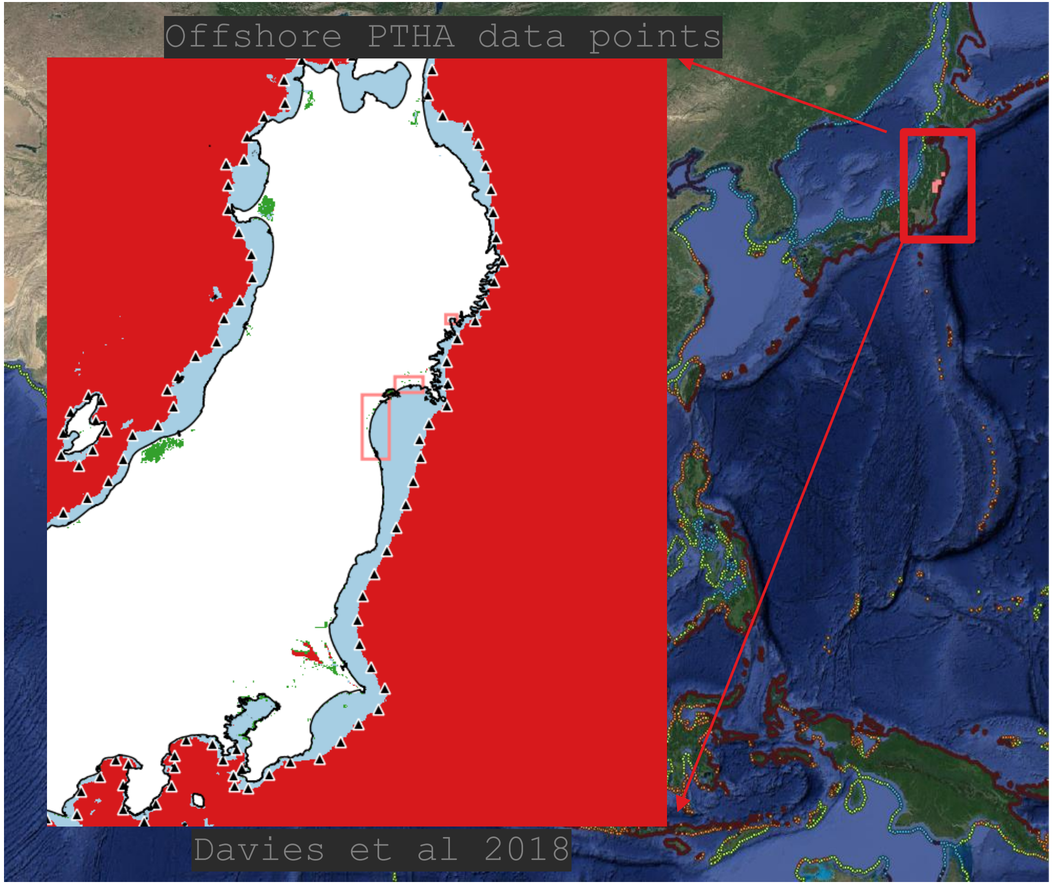


Challenges and Opportunities



- Sparse dataset(balance number of simulation vs accuracy of ML model)
- Fine tuning of ML and transferability – lots of hyperparameters, training configuration, model architecture
- Expand work to multi-input architecture, model inundation footprint directly
- Implement smart feature design and training(clustering, batch etc)
- Probabilistic wave or inundation database can be used as BC
- Link with available PTHA model which provides hazard offshore and convert them to hazard or risk onshore

Challenges and Opportunities



- Sparse dataset(balance number of simulation vs accuracy of ML model)
- Fine tuning of ML and transferability – lots of hyperparameters, training configuration, model architecture
- Expand work to multi-input architecture, model inundation footprint directly
- Implement smart feature design and training(clustering, batch etc)
- Probabilistic wave or inundation database can be used as BC
- Link with available PTHA model which provides hazard offshore and convert them to hazard or risk onshore

Thank you for your attention!

naveenraguramalingam@iusspavia.it

naveenragur.github.io

This presentation participates in OSPP

EGU

Outstanding Student & PhD candidate Presentation contest