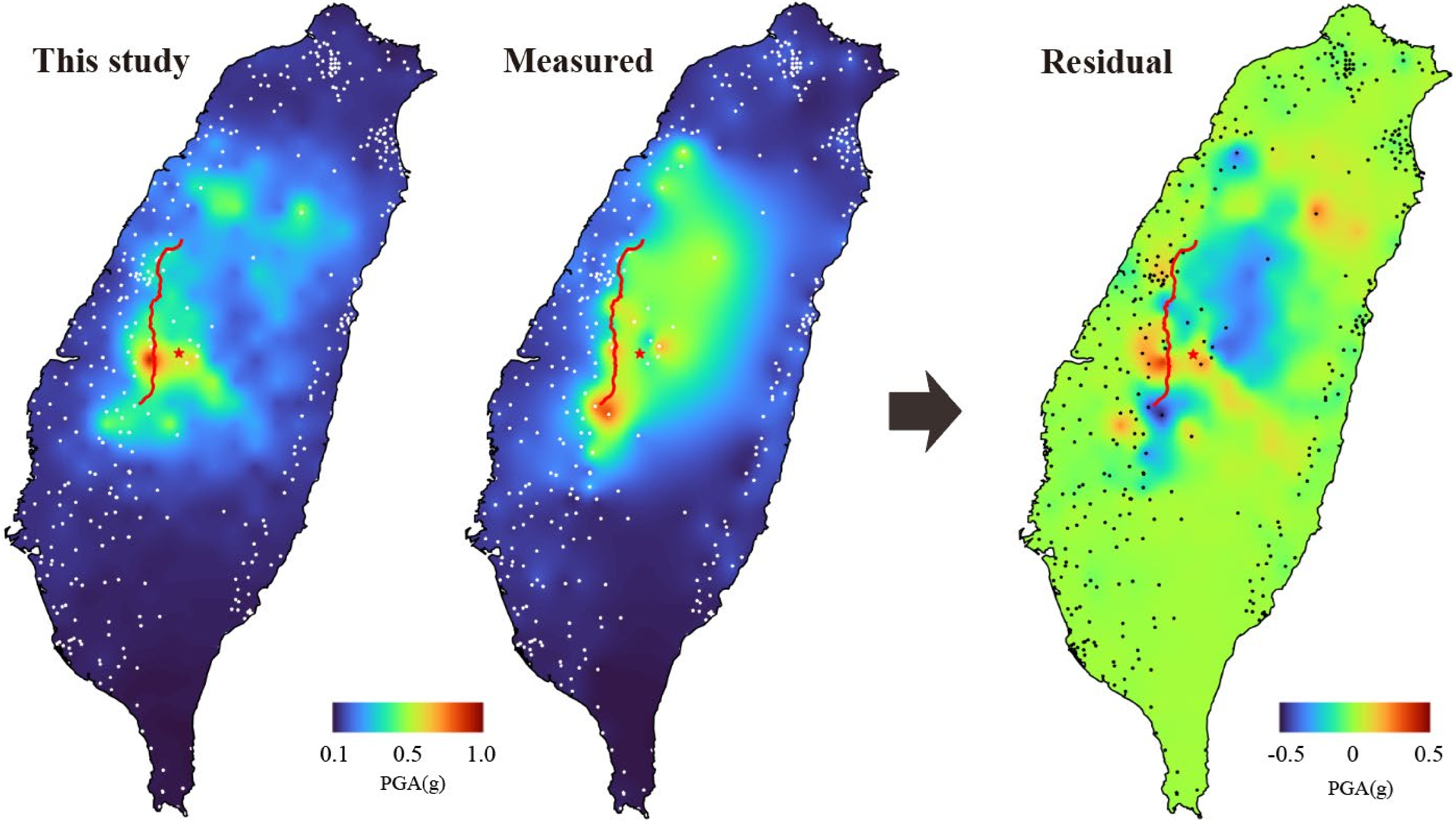
Development of a Ground Motion Model for Taiwan Using XGBoost and Its Implementation in Seismic Hazard Assessment



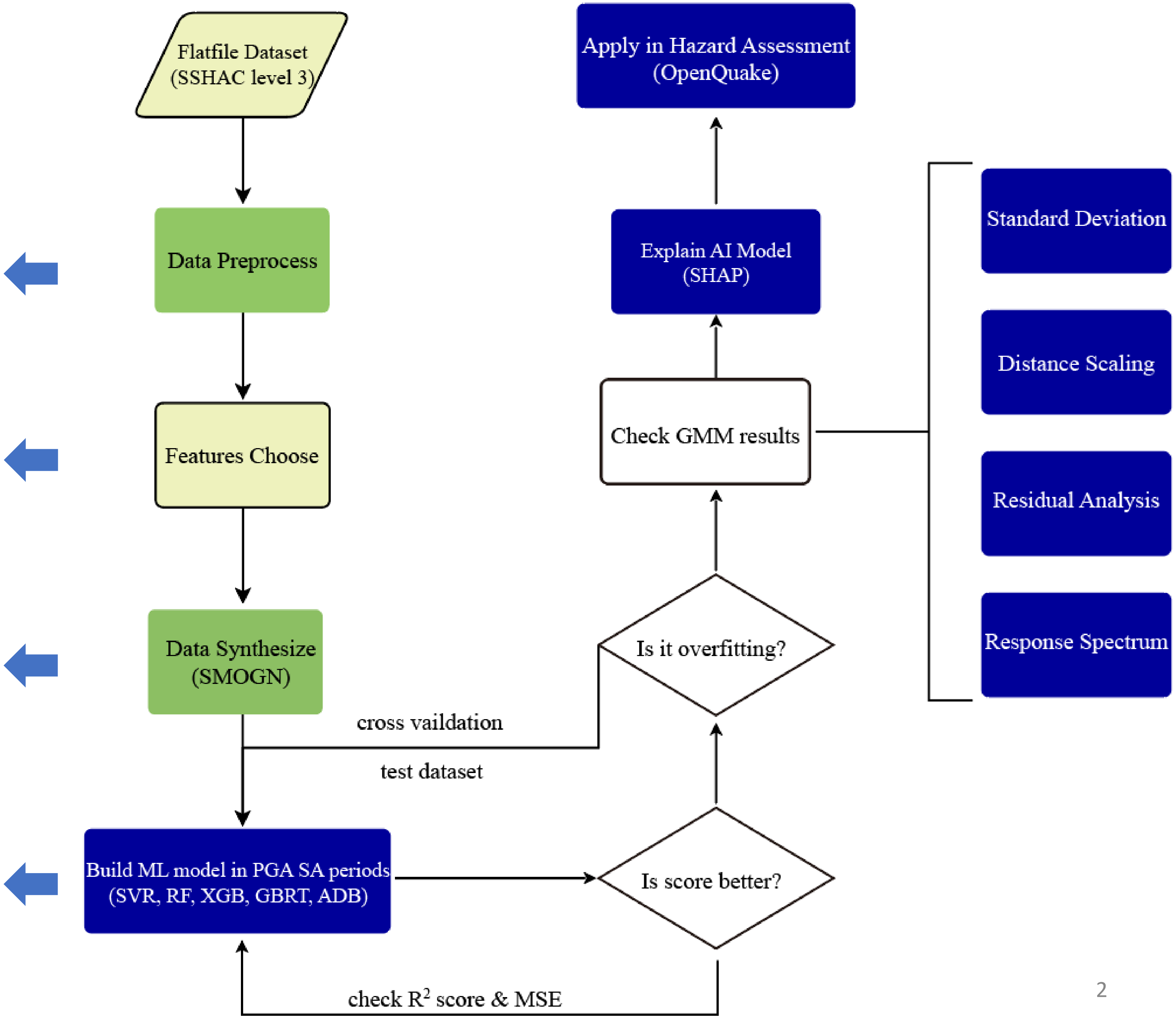
Workflow

- Choose crustal events
- Filter unuseful periods

- Rrup、Mw 、Vs30、Focal type、Station ID

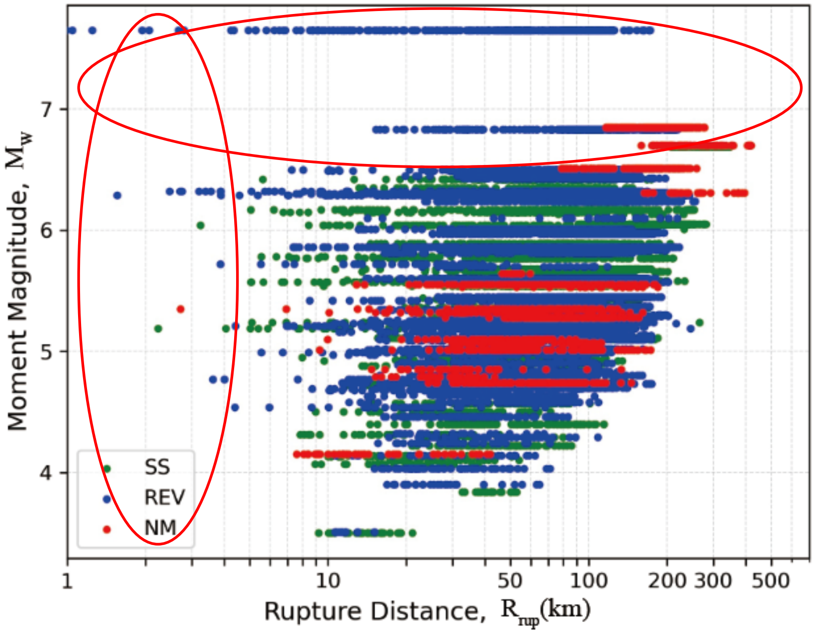
- Oversampling
- Undersampling

- Build model for each SA periods
- Compare their performance

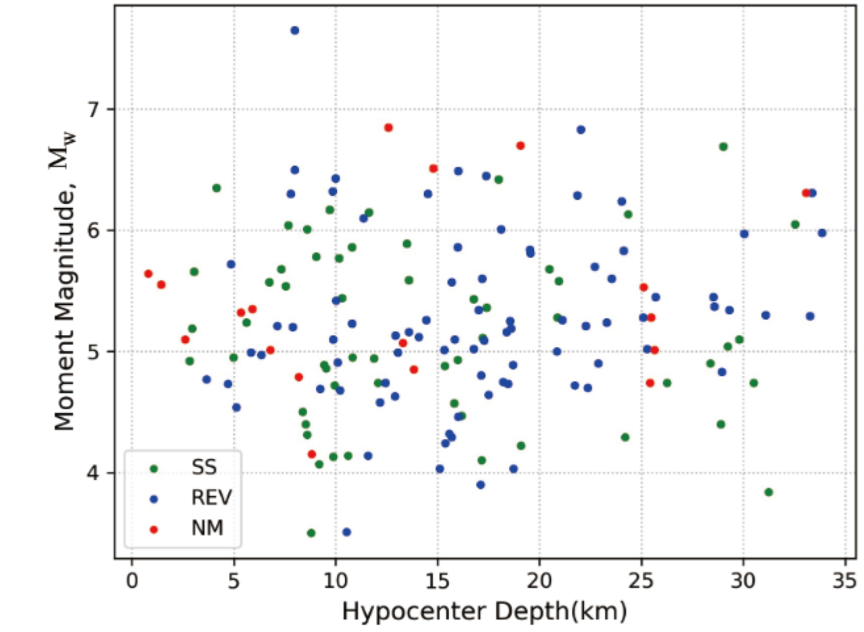


Data Distribution

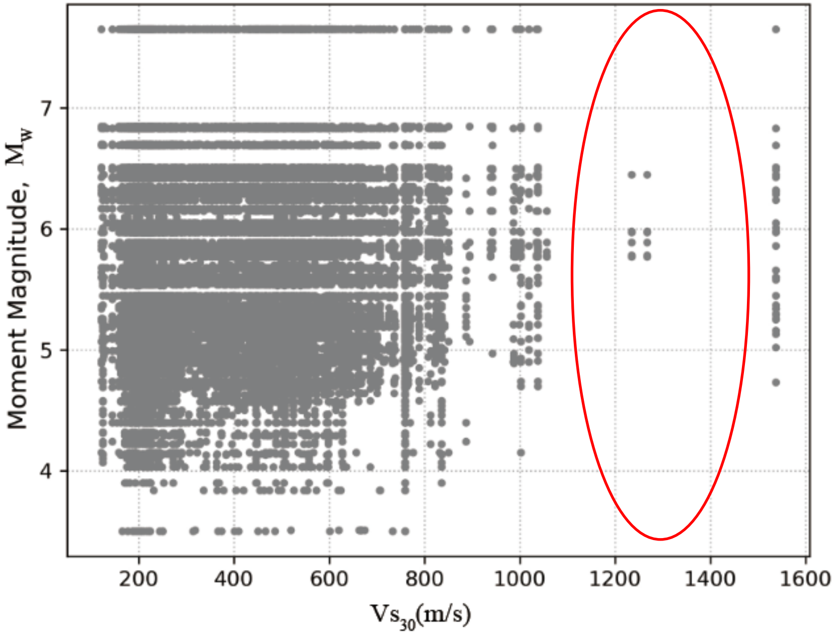
- TWSSHAC dataset
- TSMIP network records
- Crustal earthquake event



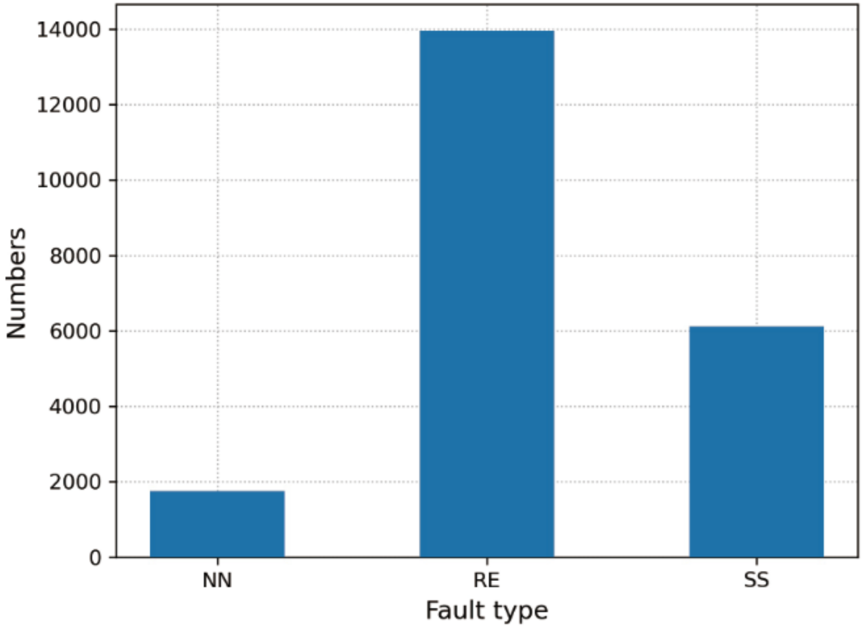
(a)



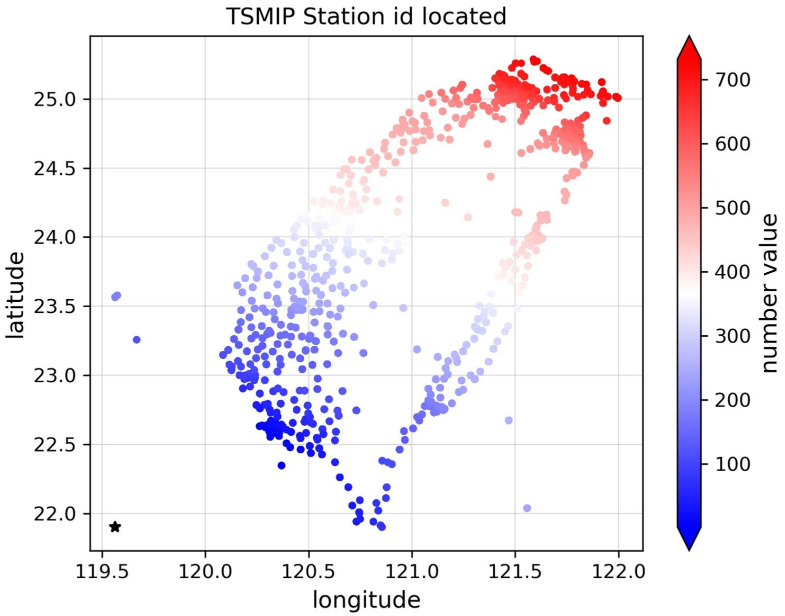
(b)



(c)



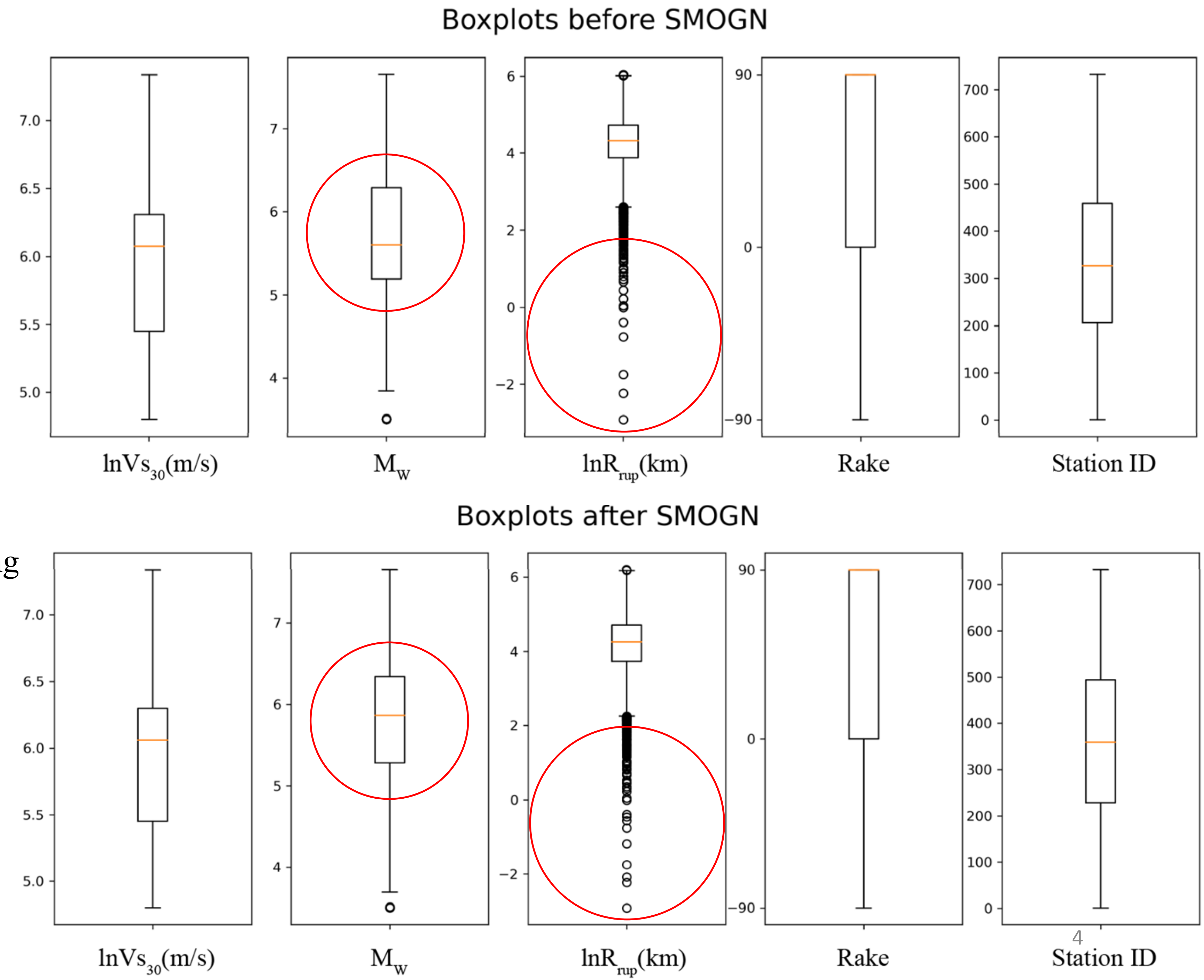
(d)



Reference: https://sshac.ncree.org.tw/

Data Preprocess

- Include over / under sampling
- Make dataset more balance
- Take synthesized dataset as training

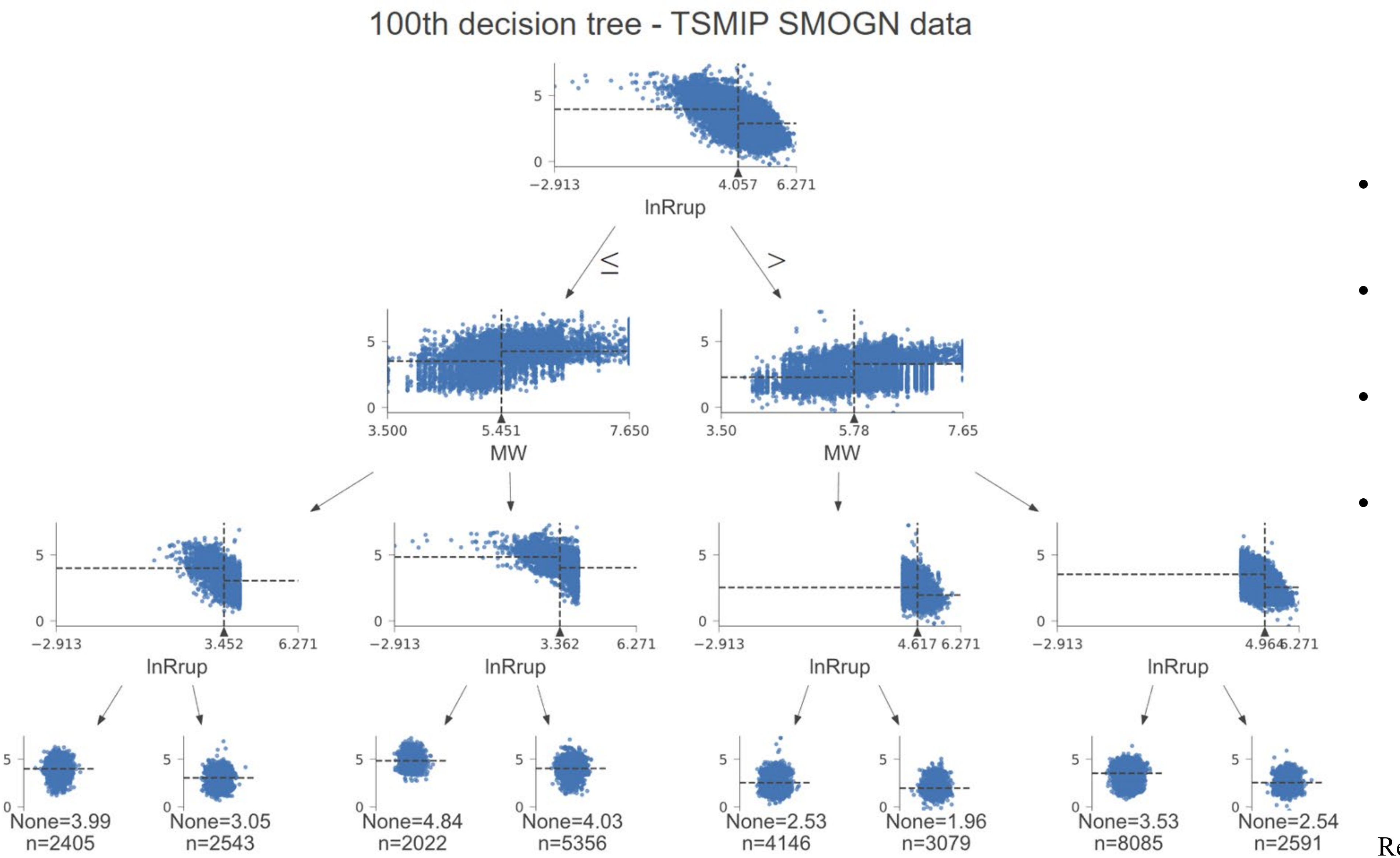


Take PGA dataset as example ▶

Reference: Branco et al., 2017

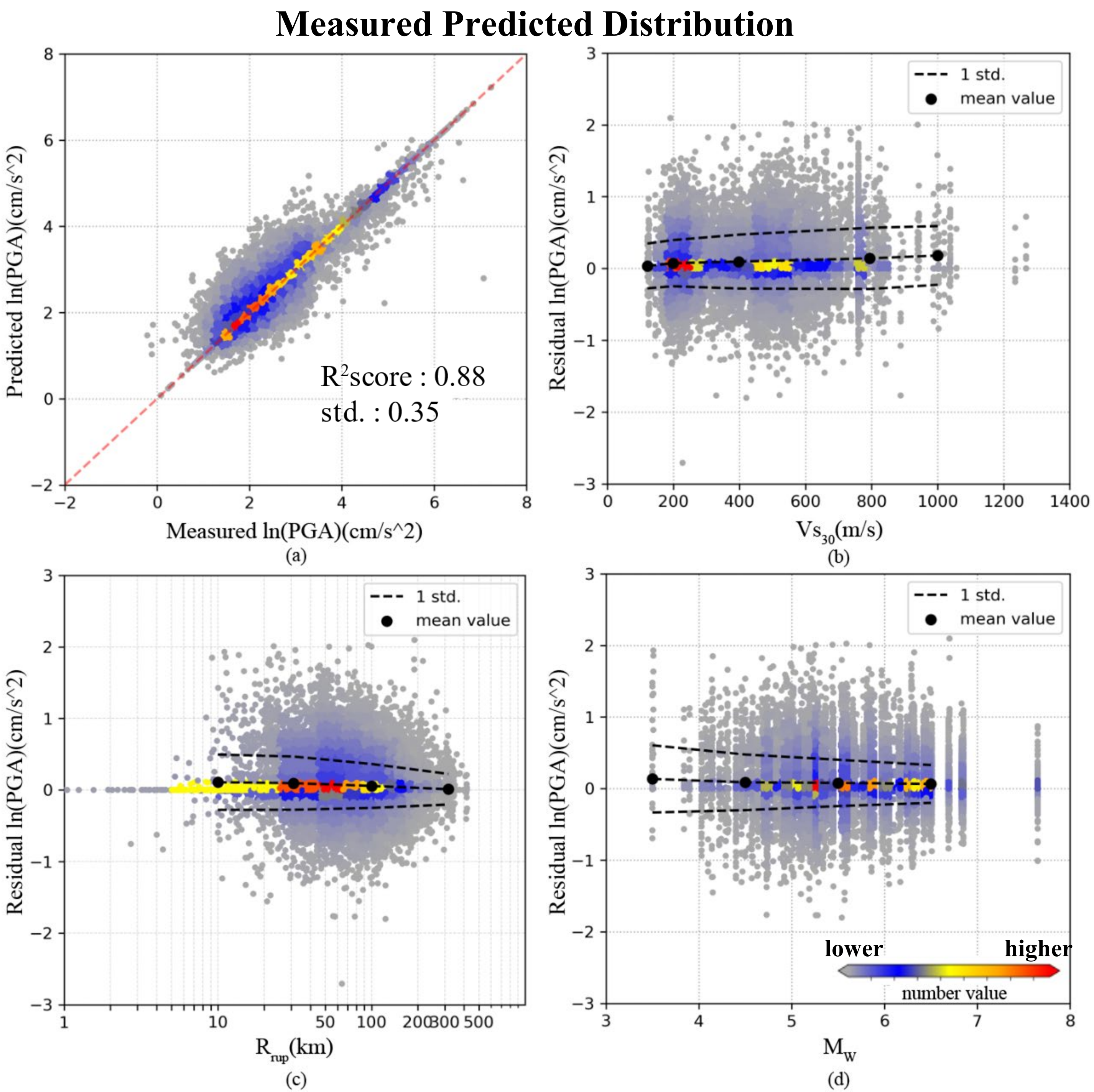
Ground Motion Model

Input parameters : Rrup、Mw、Vs30、 Focal type、Station ID



- XGBoost model
- Decision tree based algorithm
- Boosting based algorithm
- Ensemble method

Reference: Chen and Carlos, 2016



Model Validation - Residual

- Take PGA model as example
- Check if model have some bias
- Compare result in different features

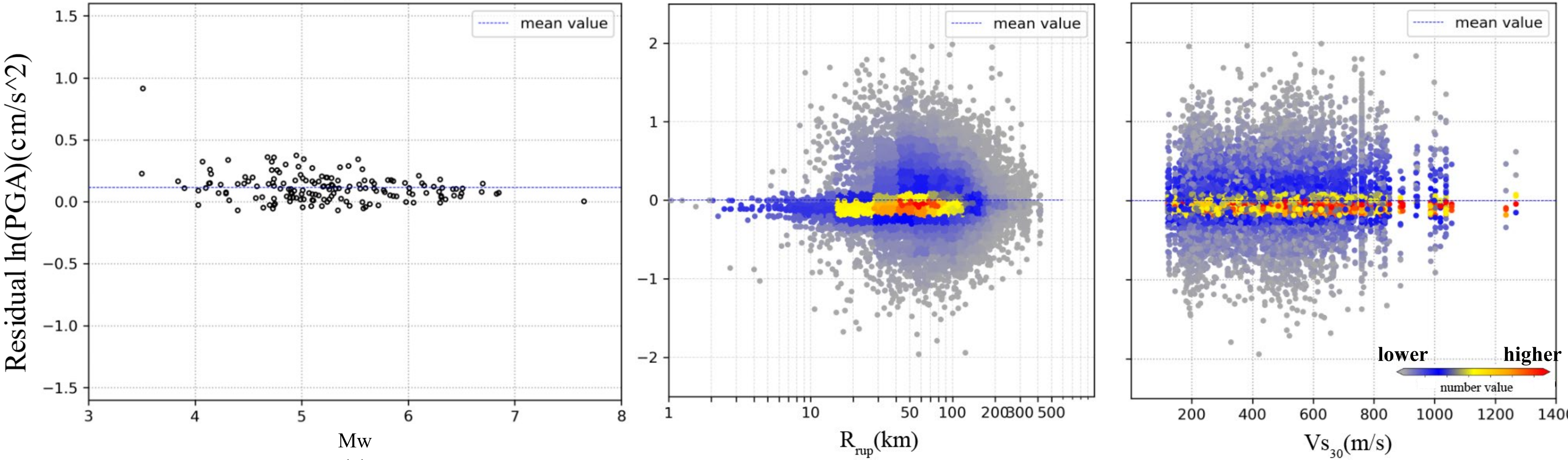
Model Validation – Inter-event, Intra-event

- Check the inter-event residual
- Check the intra-event residual
- Make sure don’t have some bias

Std. ➡ Event effect + Station effect

Inter-event Residual Mean :0.11 Std :0.12

Intra-event Residual Mean :0.0 Std :0.34

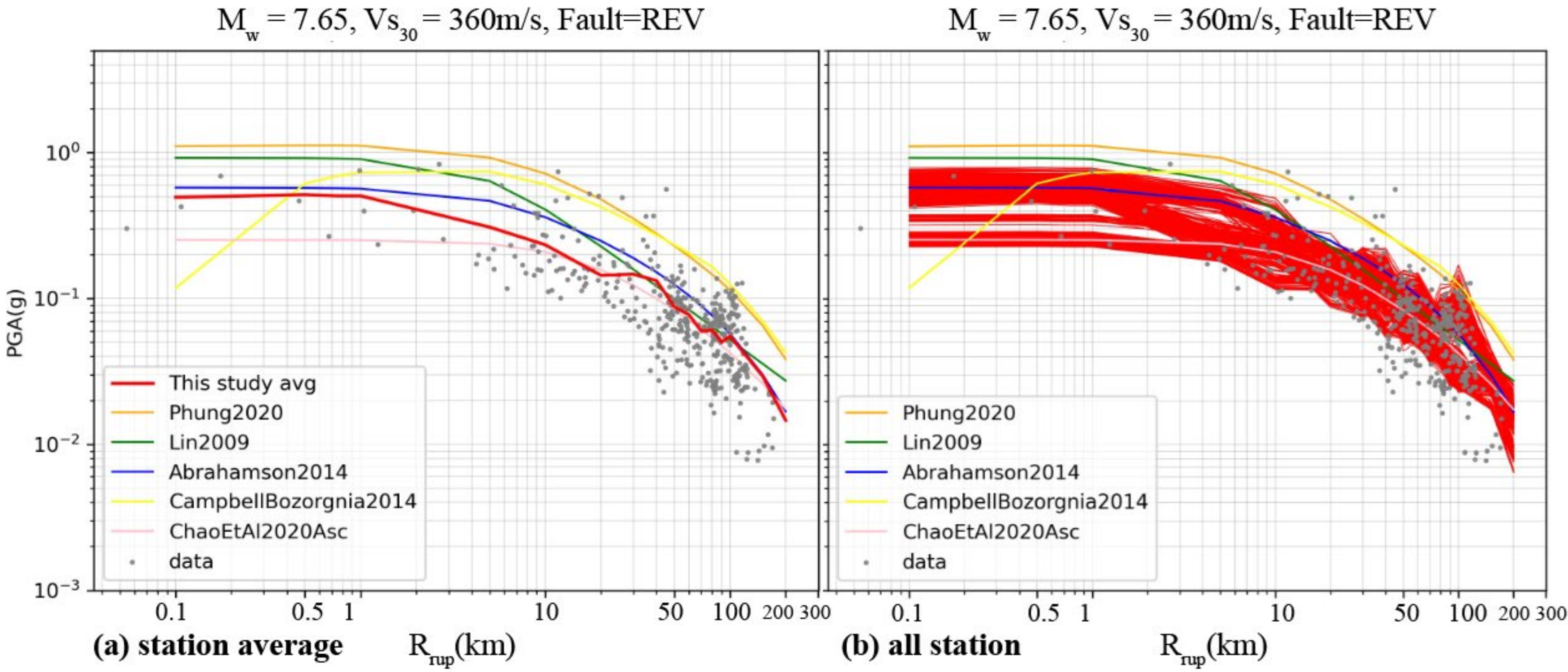


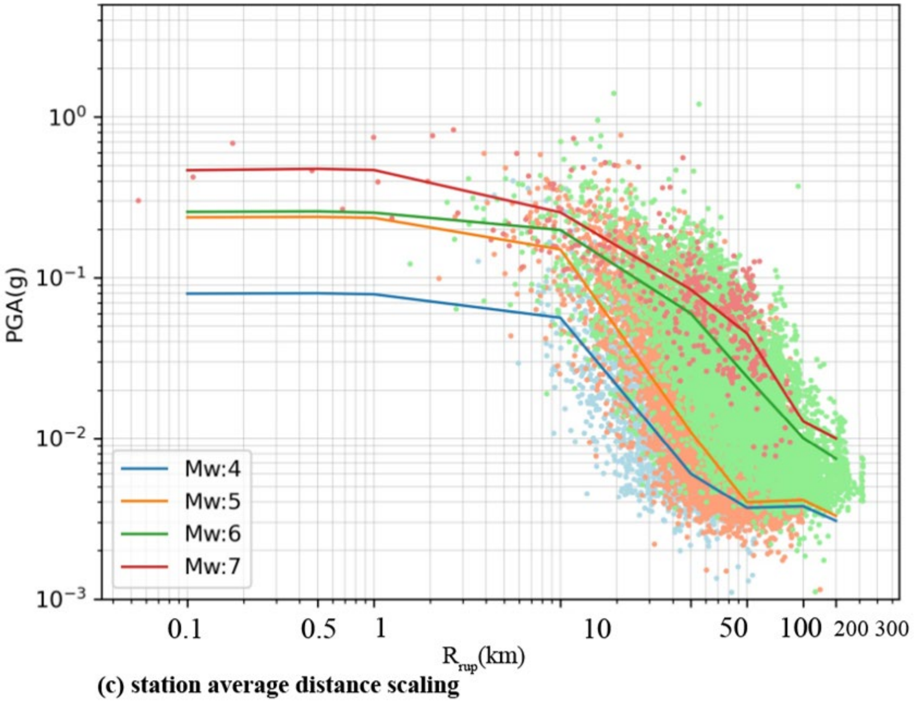
(a) (b) (c)

Model Validation –

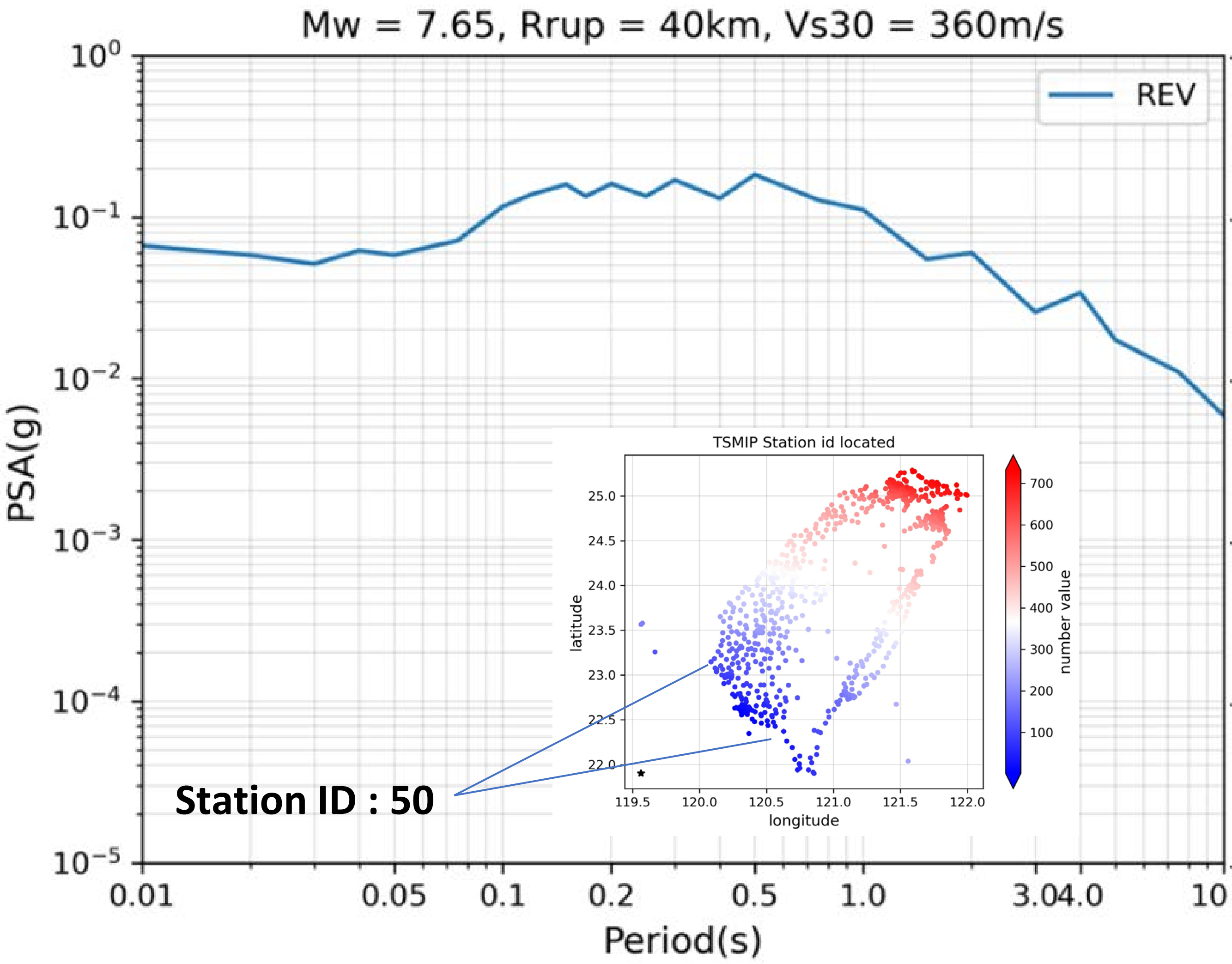
Distance Scaling

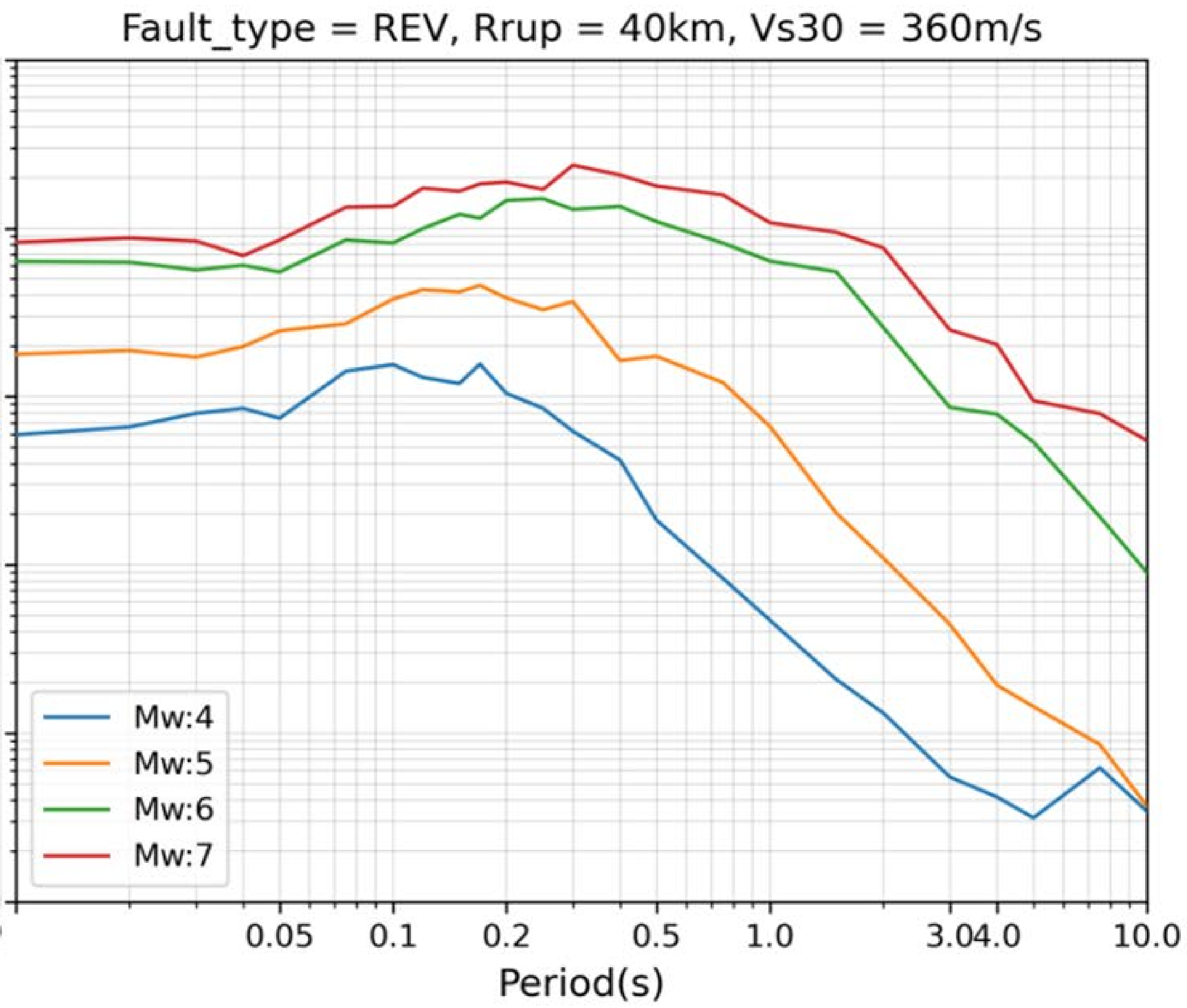
- Compare with other GMMs
- Average all station id in same condition
- Compare with different magnitude





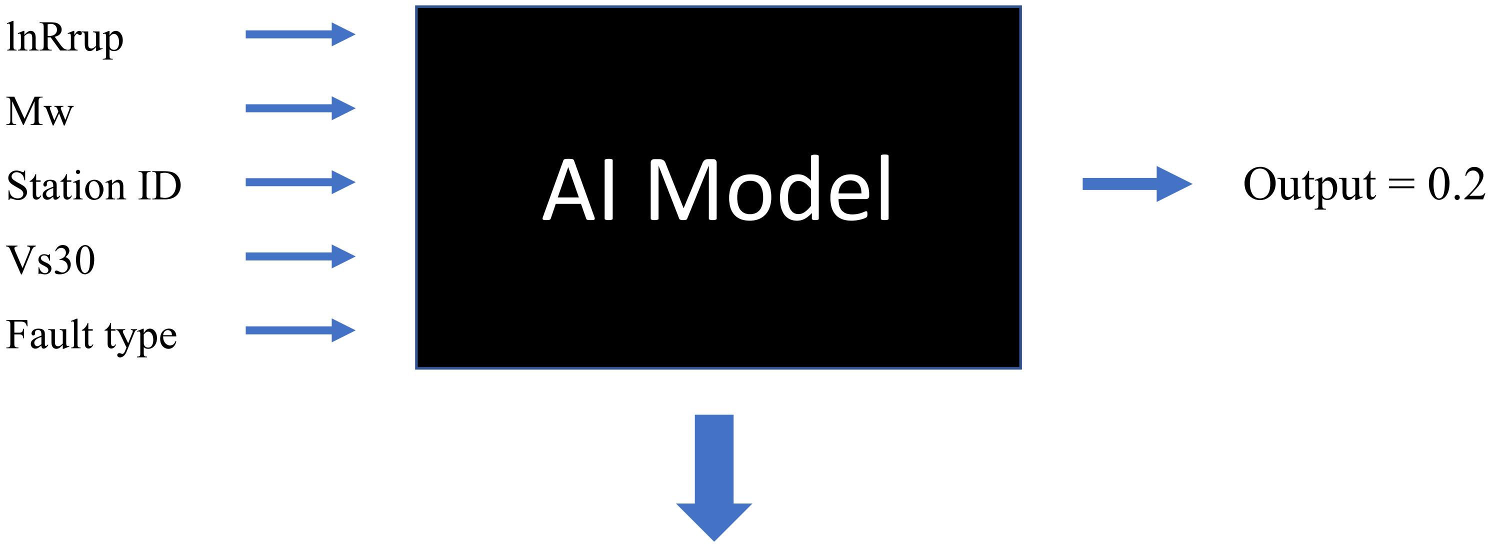
Model Validation – Respond Spectrum for Engineering





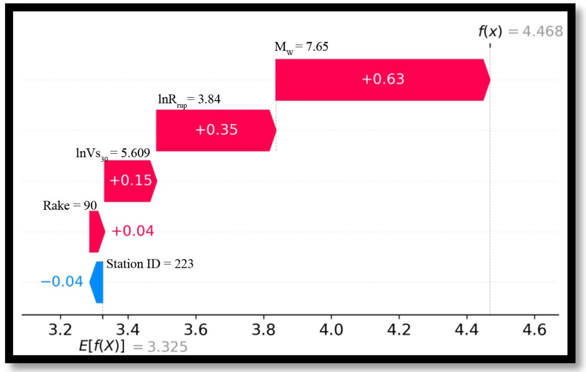
Explainable AI –

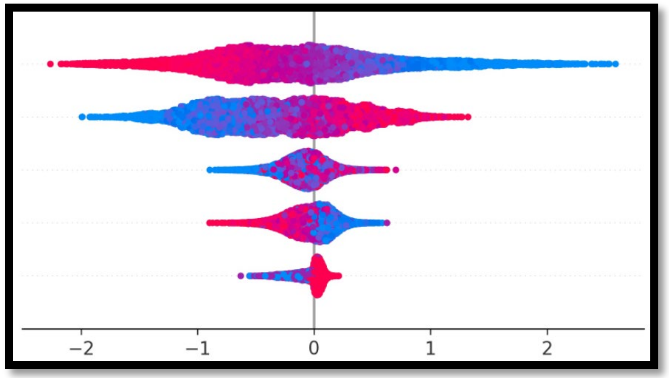
SHAP tool



Local Explanation

Global Explanation

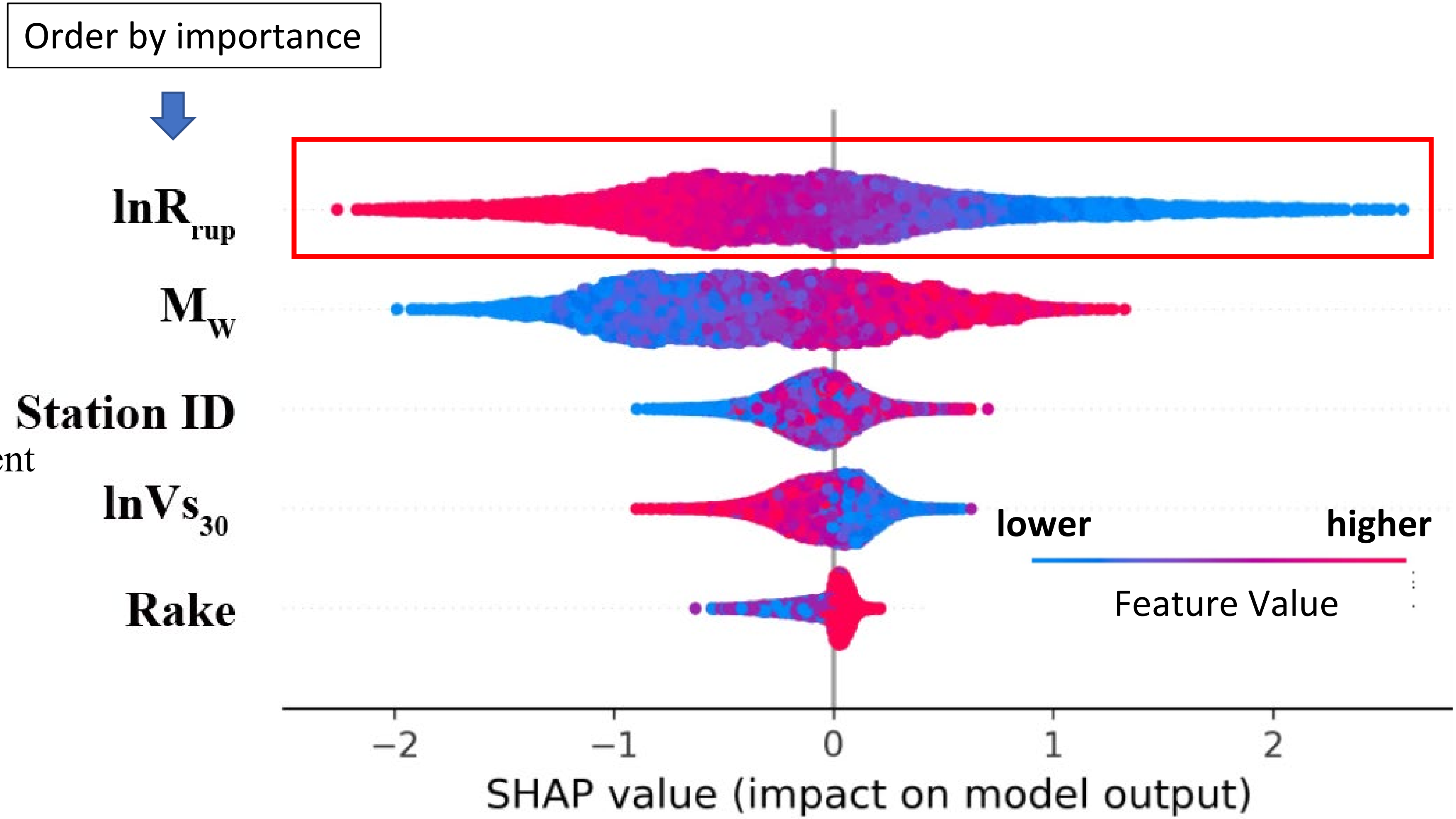




Ref : Lundberg and Lee, 2017

SHAP Framework – Global Explanation

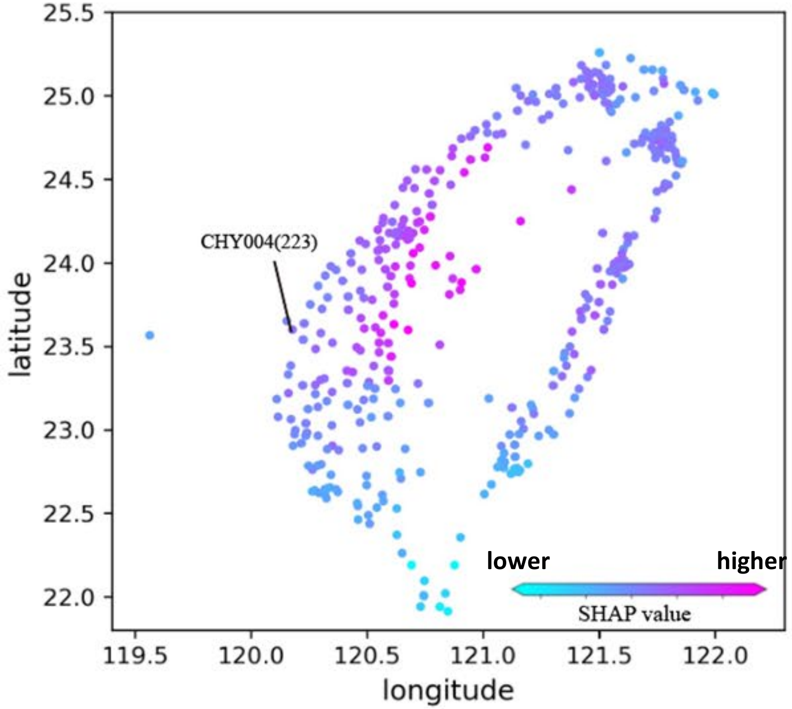
- Check predictions by ML model
- Check the feature importance
- Explain model in global component



(a) global explanation

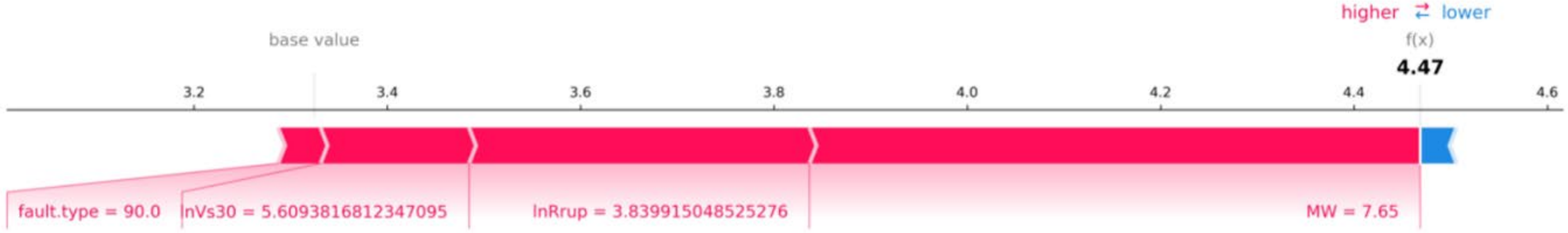
SHAP Framework – Local Explanation

Take Chi-Chi earthquake and CHY004 station as example



(a)

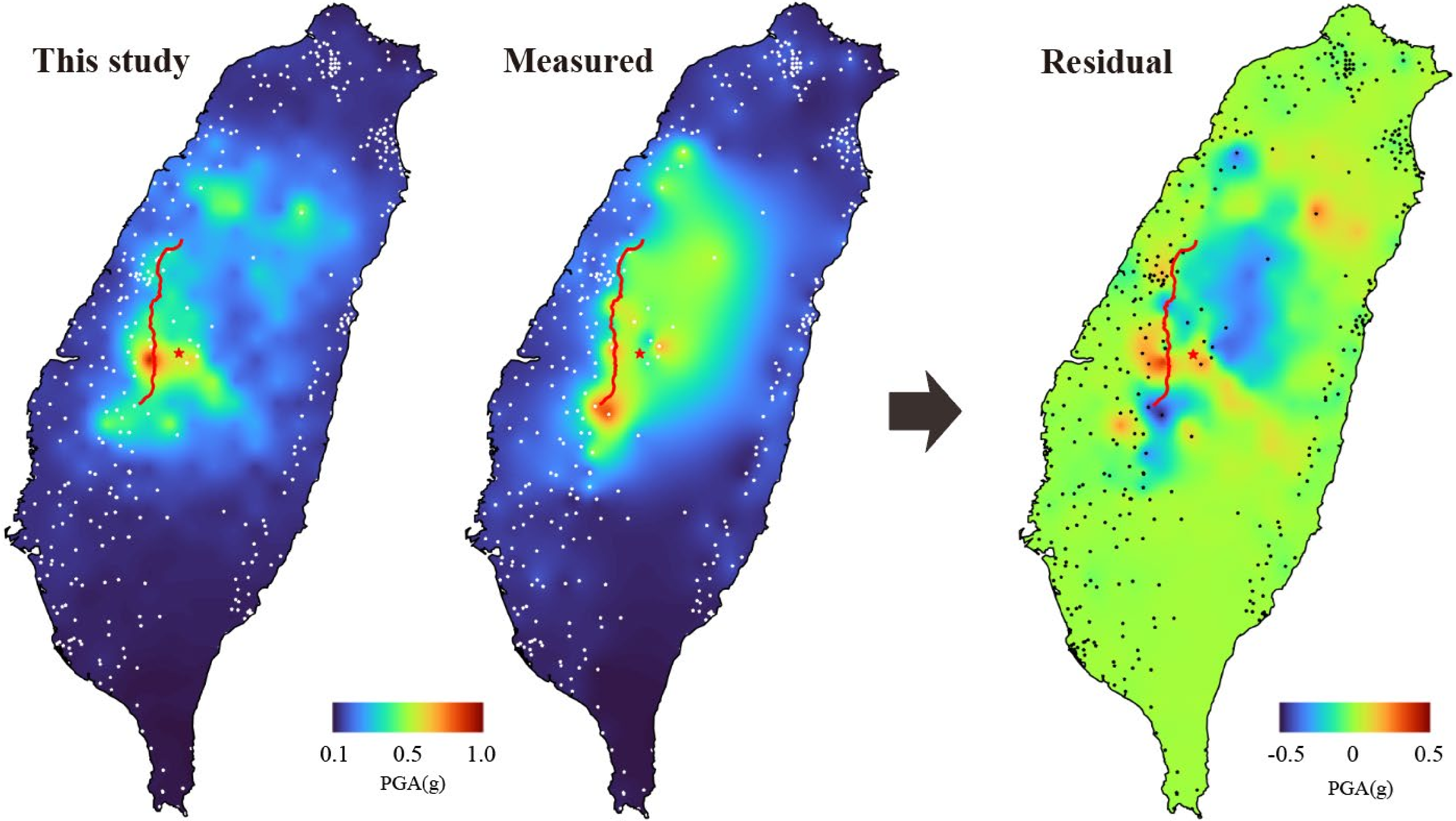
(b)



(c)

Application to Hazard Assessment – OpenQuake

(fork from oq-engine) ▶ https://github.com/badguy361/oq-engine



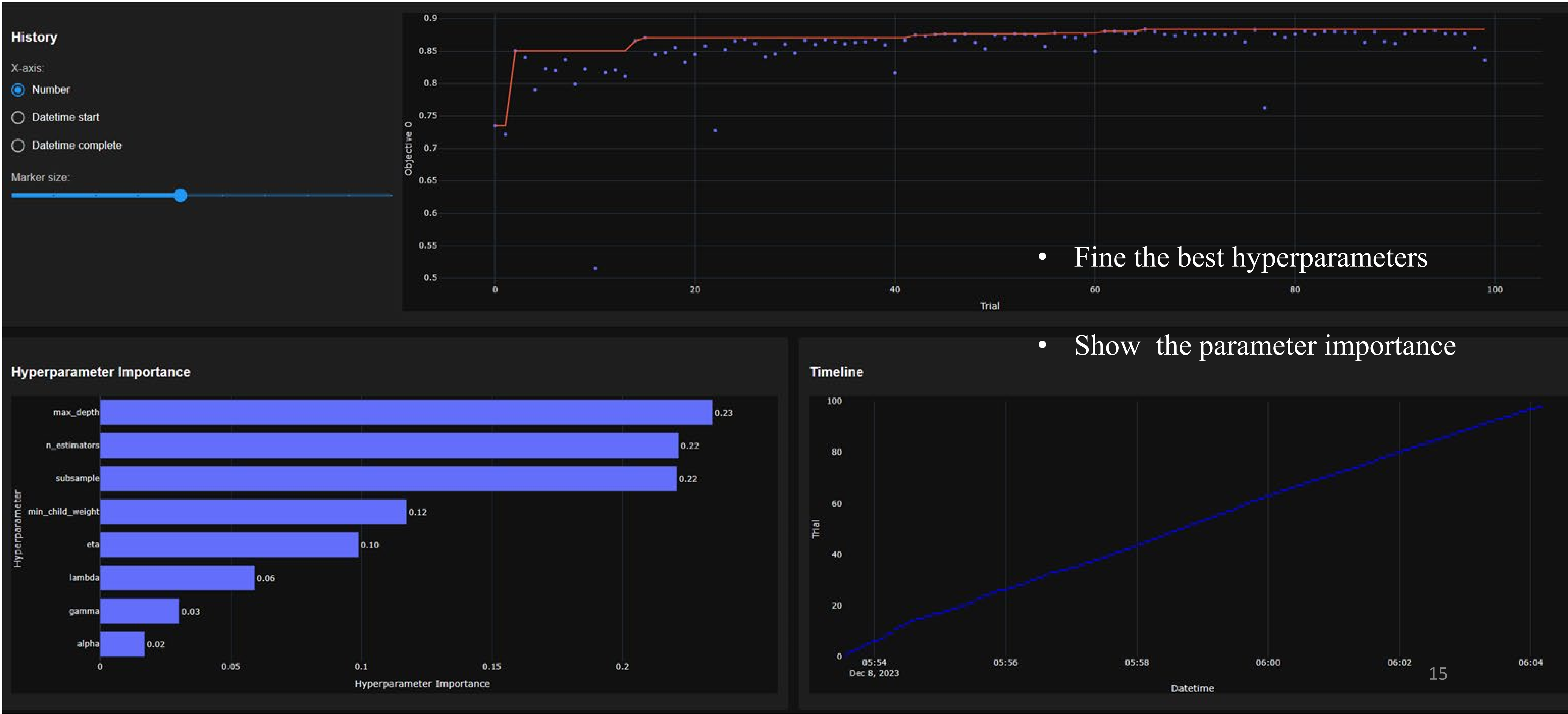
Conclusions

- The XGBoost model can perform well in the ground motion model question.
- Take advantage of residual or another statistical figure to make sure it would not have bias.
- Explain ML base mode by SHAP value to understand model details.
- Apply in the open source platform – OpenQuake to ensure our model is usable.

Connect with me to discuss more study details~ Email : t1616joy@yahoo.com.tw ; Name : Chih-Yu Chang

Optimize Hyperparameters

Reference: Takuya Akiba et al,. 2019 ▶ OPTUNA



- Fine the best hyperparameters
- Show the parameter importance

Vs30 - Station ID = SHAP value residual