Exploring Improvements on Magnitude Estimates for Paleo-Earthquakes Based on Data Type and Volume Availability

Yajaira De Haro[1]; Allison Schiffmaier[2]; Alba Rodriguez Padilla[3]
[1]San Jose State University , [2]University of California at Davis, [3]Utah State University

Abstract

Estimating paleoearthquake magnitudes often uses simple regressions based on displacement or rupture extent. As more data becomes available in the form of additional displacement measurements or rupture extent constraints from new trenches, can paleo-magnitudes be improved by leveraging these additional datasets? We tested whether a neural network could improve accuracy using modern data from 75 surface-rupturing earthquakes. Comparing the neural net to three standard methods, we down-sampled data to simulate sparse conditions like those from trenches and geomorphic offsets. Our preliminary analysis shows that the neural net, using displacement and rupture length, outperforms other methods until data is downsampled 60%, at which point rupture is the best predictor. Last, we explore application of the methods to geomorphic offset data from the Coyote Creek fault.

Introduction

Modern surface-rupturing earthquakes yield dense displacement profiles and accurate rupture-length measurements, but paleoearthquake magnitudes rely on sparse trench/offset observations. Current paleoseismic magnitude estimates use simple regression on mean slip or rupture length, without leveraging large modern datasets. So, how does systematic within-event down-sampling of modern data impact four magnitude estimation methods? We train and test a neural-net based model (**poster #111**) on data from 75 modern events, and compare this model to the performance of typical regression estimates using in paleoseismic magnitude studies.

Training, Validation, and Testing Data

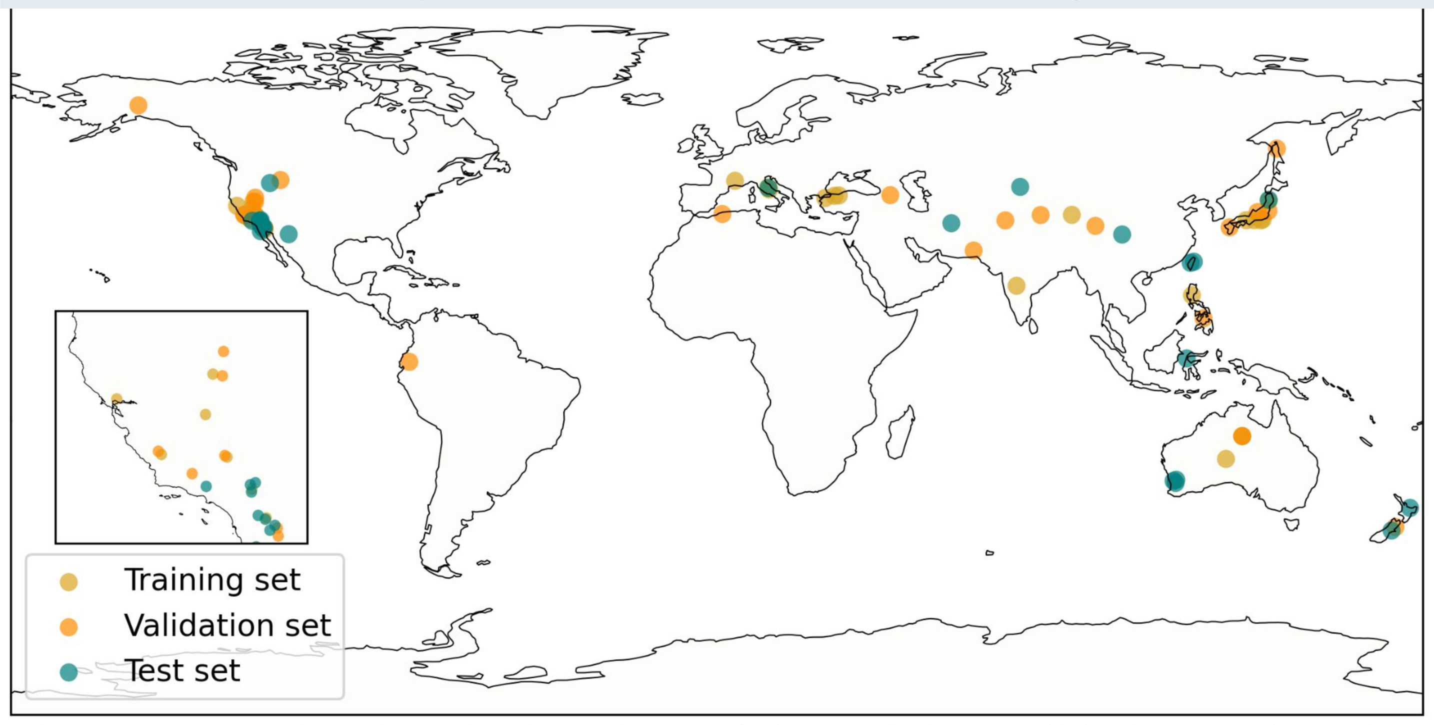

Figure 1: Map showing the locations of the training, validation, and testing datasets. Event data from Sarmiento et al. (2024)

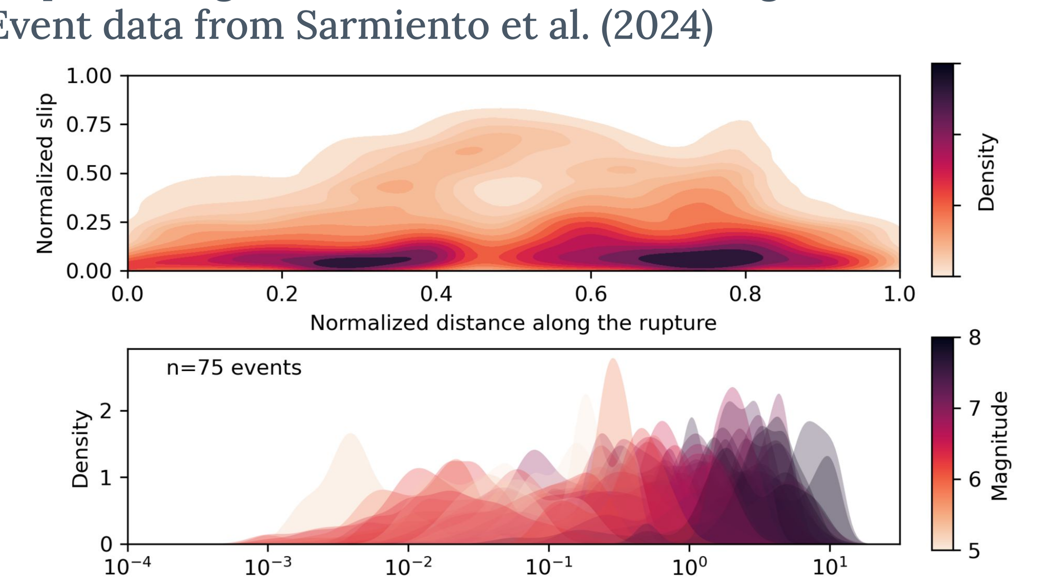

Figure 2: **Top**: Normalized slip vs. distance along the rupture. The colors indicate data density, with darker colors indicating larger number of points. **Bottom**: Kernel density estimates of the distribution of slip per event, color-coded by magnitude. Event data from Sarmiento et al. (2024).

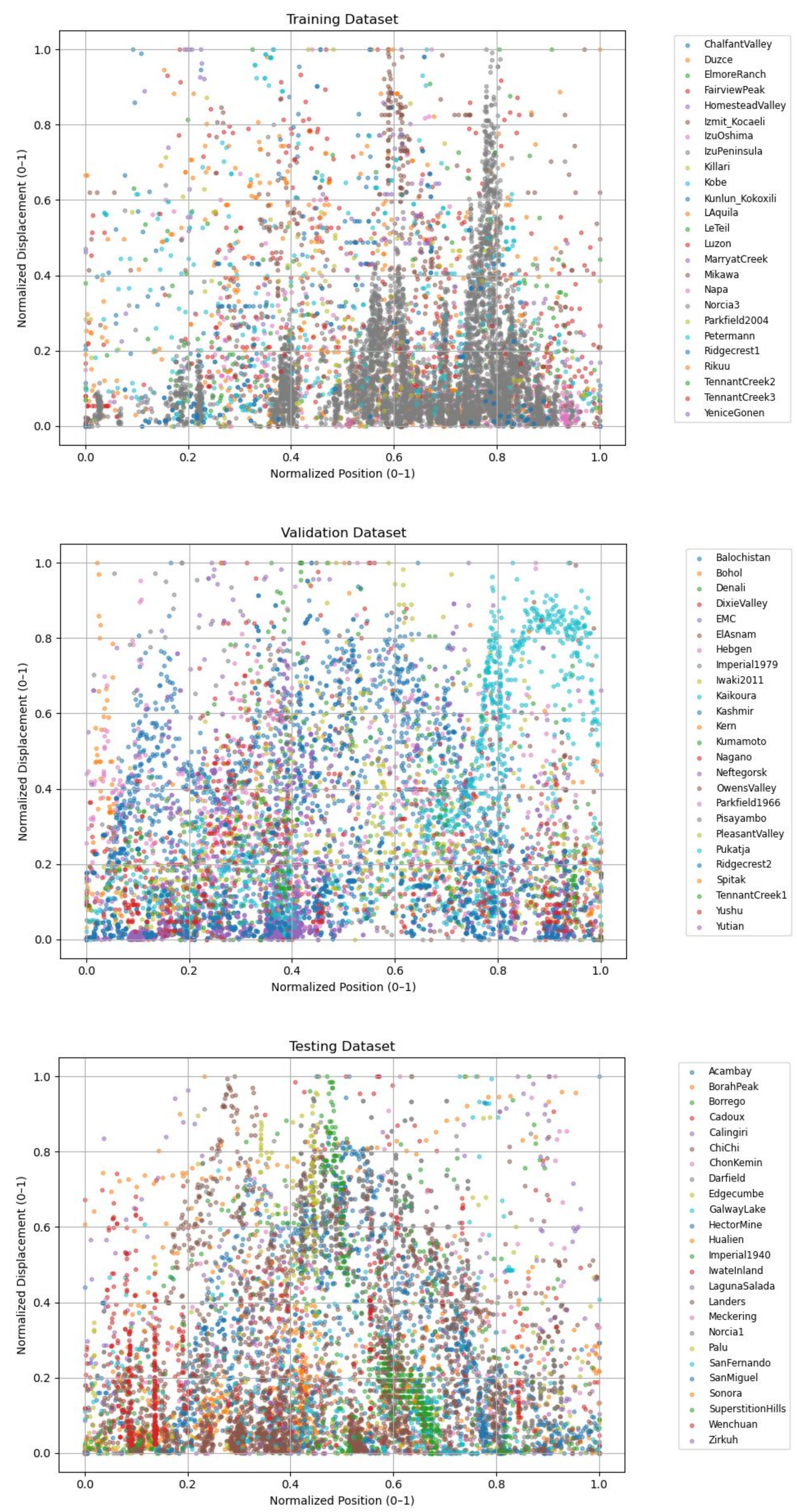

Figure 3: Graphs showing normalized displacement vs. position for events within training, validation, and testing dataframes, after distributing data points more evenly within each dataframe. Event data from Sarmiento et al. (2024).

Methodology

Dataset: Cleaned 75 modern surface-rupture events

Down-sampling: Randomly removed 0-90% of rows per event in 10% steps, 20 replicated per level.

Algorithms implemented in Python:

- **Neural network** (feed-forward on displacement profiles and rupture lengths)
- **Moment-magnitude from seismic moment:** $M_w = ⅔(\log_{10}M_0 - 9.05)$
- **Wells & Coppersmith (1994), rupture-length regression:** $M_w = 5.08 + 1.16\log_{10}(L_{SRL})$
- **Wells & Coppersmith (1994), mean-slip regression:** $M_w = 6.93 + 0.82\log_{10}(D_m)$

Method Performance on Modern Data

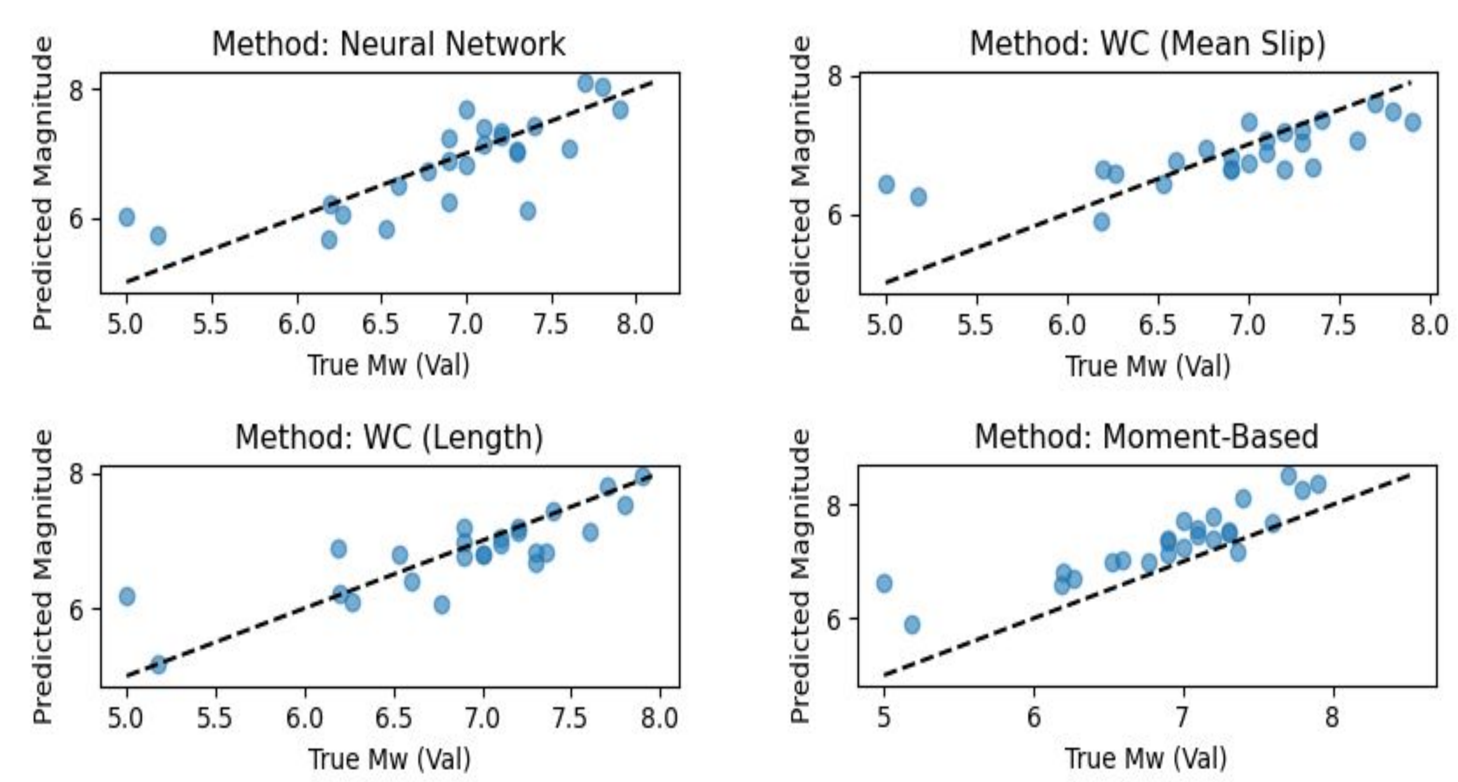

Figure 4: Predicted magnitude vs. true magnitude for each method within the validation dataframe. Event data from Sarmiento et al. (2024).

Method Performance on Downsampled (Paleoseismic Proxy) Data

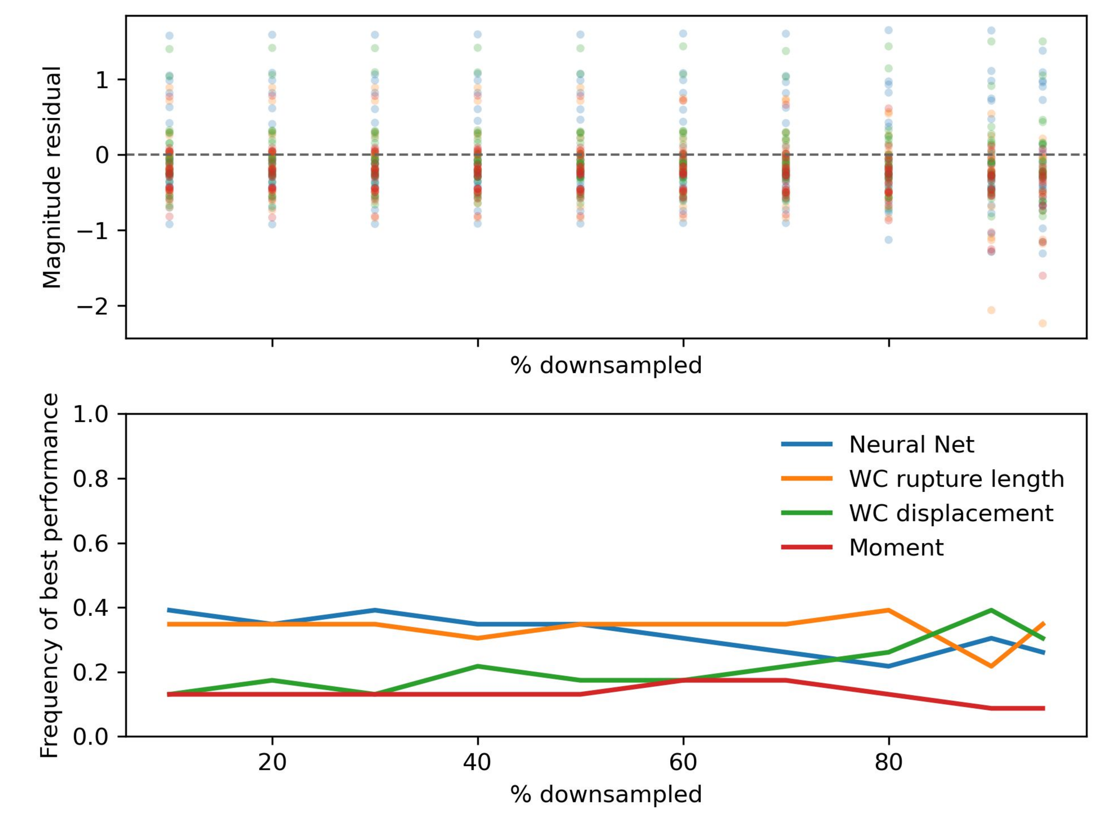

Figure 5: Graphs show that more data favors Neural Net, while less data favors WC rupture length. Event data from Sarmiento et al. (2024).

Conclusion/Discussion

Our analysis shows that the Neural Net performs best when data are abundant, making it well suited for modern datasets. Under sparse conditions that mimic paleoseismic trench data, the Wells & Coppersmith rupture length regression is the most reliable method. Moment-based and displacement-based regressions are less consistent overall. These results suggest that Neural Nets are most effective for data-rich cases, such as geomorphic offset distributions, while rupture length remains the most robust predictor when observations are limited. Future work can focus on hybrid approaches that integrate both data-driven models and robust regression to improve paleoearthquake magnitude estimates across varying data conditions.

References

Sarmiento, A., Madugo, D., Shen, A., Dawson, T., Madugo, C., Thompson, S., ... & Visini, F. (2024). Database for the fault displacement hazard initiative project. *Earthquake Spectra*, 87552930241262766.

Rockwell, T. K., Masana, E., Sharp, W. D., Štěpančíková, P., Ferrater, M., & Mertz-Kraus, R. (2019). Late Quaternary slip rates for the southern Elsinore fault in the Coyote Mountains, southern California from analysis of alluvial fan landforms and clast provenance, soils, and U-series ages of pedogenic carbonate. Geomorphology, 326, 68-89.

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bulletin of the seismological Society of America, 84(4), 974-1002.

Acknowledgements

Big thanks to SCEC for the support and funding that made this internship possible. SCEC is funded by NSF Cooperative Agreement EAR-2225216 and USGS Cooperative Agreement G24AC00072. A big thank you to my parents and siblings for supporting me and encouraging me during my internship. Finally, I'd like to thank my mentor, Alba, for her mentorship and guidance over the summer.

Visit www.scec.org/about/acknowledgements for more guidance

SCEC · SAN JOSE STATE UNIVERSITY · UC DAVIS UNIVERSITY OF CALIFORNIA · UTAH STATE UNIVERSITY 1888 · NSF U.S. National Science Foundation