Effects of inventory bias on landslide susceptibility calculations

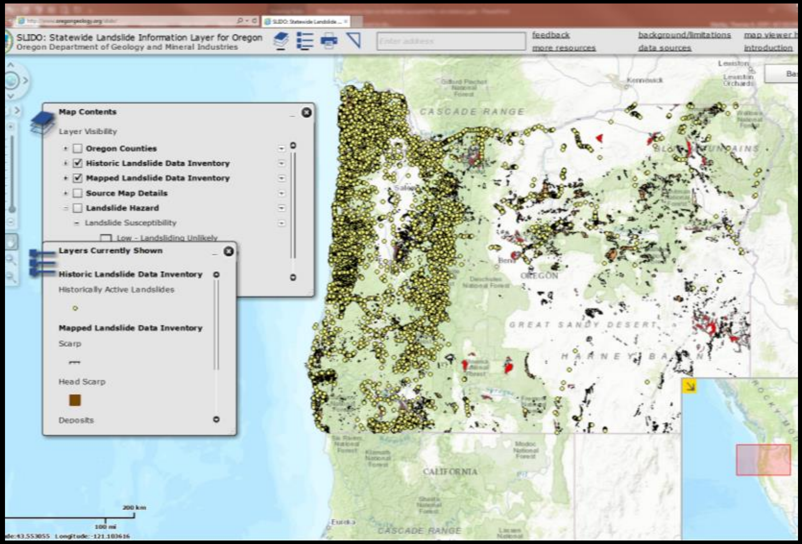
Thomas Stanley, USRA/GESTAR/NASA GSFC
Dr. Dalia B. Kirschbaum, NASA GSFC

"Multi-temporal landslide information is essential to new approaches for the generation of quantitative landslide probability maps."

-Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview
Cees J. van Westen, Enrique Castellanos, Sekhar L. Kuriakose

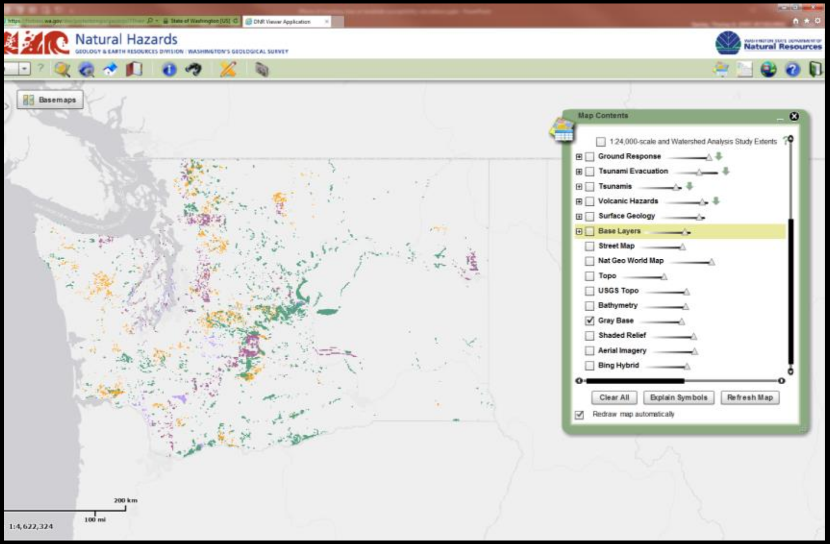
Pacific Northwest Inventories

Oregon



http://www.oregongeology.org/slido

Washington



https://fortress.wa.gov/dnr/protectiongis/geology/?Theme=natural_hazards

Landslide susceptibility workflow

Landslide susceptibility. A quantitative or qualitative assessment of the classification, volume (or area), and spatial distribution of landslides which exist or potentially may occur in an area. Susceptibility may also include a description of the velocity and intensity of the existing or potential landsliding. Although it is expected that landsliding will occur more frequently in the most susceptible areas, in the susceptibility analysis, time frame is explicitly not taken into account.

- Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., and Savage, W.Z., 2008, Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning: Engineering Geology, v. 102, no. 3–4, p. 99–111, doi: 10.1016/j.enggeo.2008.03.014.

1 Compile inventory

↓

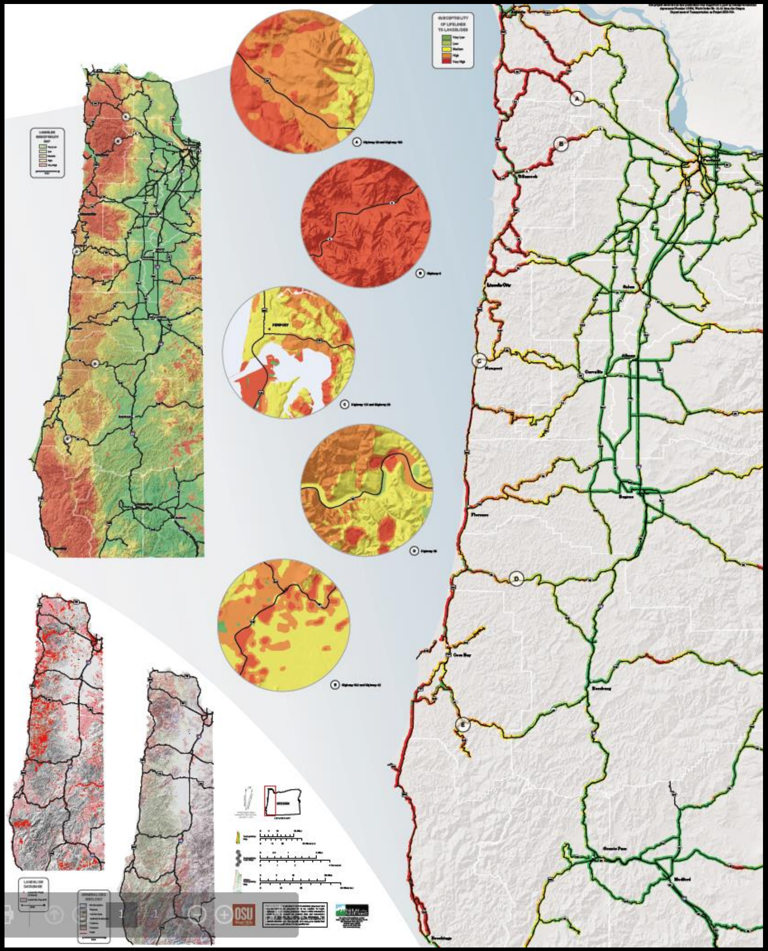
2 Fit empirical model or apply physical model

↓

3 Validate map with unused portion of inventory

Landslide susceptibility analysis of lifeline routes in the Oregon Coast Range

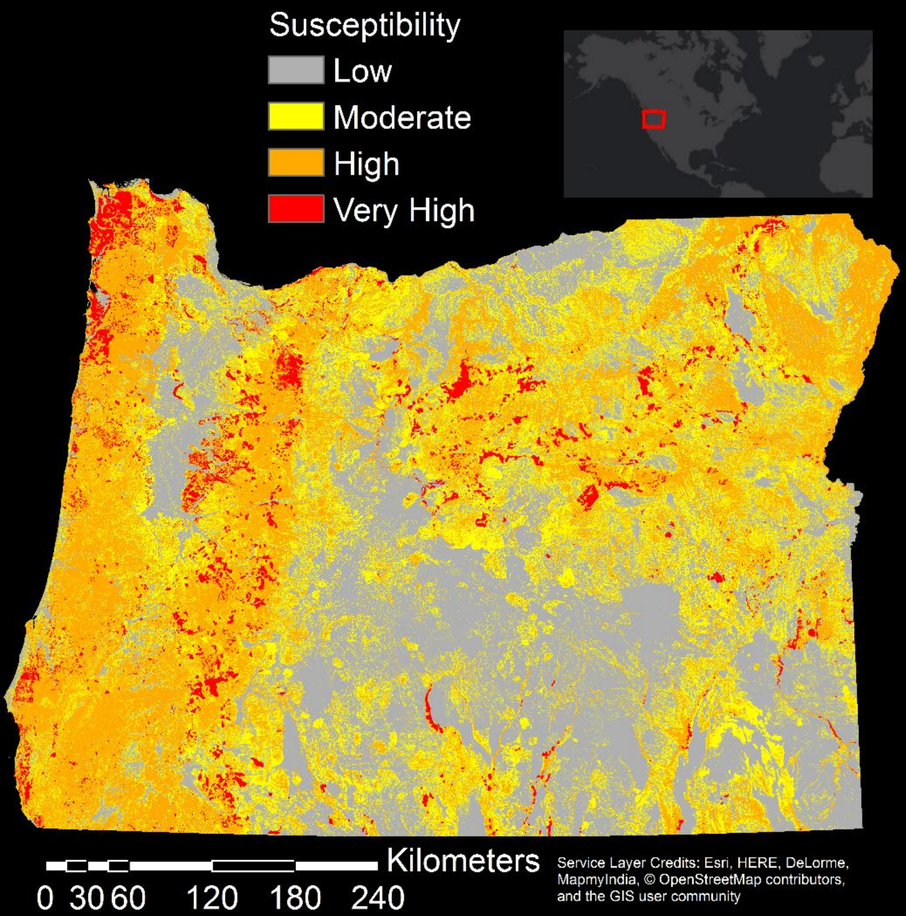
- Predictors: PGA, PGV, Slope, Precipitation (PRISM 30-year MAP)
- Empirical susceptibility model fitted with landslide polygons and points
- Random assignment of non-landslides
- Classification of probability into 5 bins



Mahalingam, R., Olsen, M.J., Sharifi-Mood, M., and Gillins, D.T., 2015,
Landslide susceptibility analysis of lifeline routes in the Oregon Coast Range

Landslide susceptibility overview map of Oregon

- Predictors: Slope, Geologic Unit
- Empirical susceptibility model fitted with landslide polygons
- Validated with historic landslide points
- Classification into 4 bins
- Very high susceptibility category consists solely of SLIDO polygons



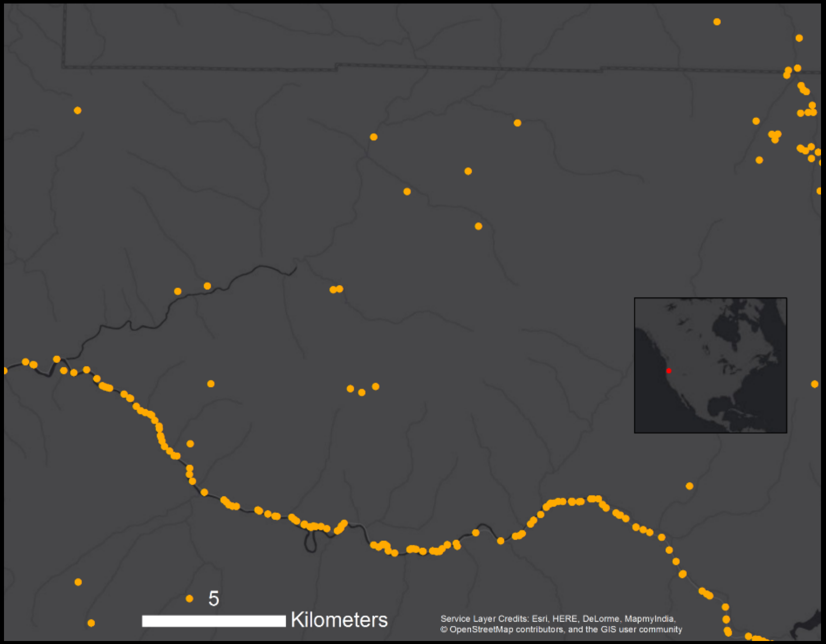
Burns, W.J., Mickelson, K.A., and Madin, I.P., 2016, Landslide susceptibility overview map of Oregon.

Inventory bias

Not only can inventory bias alter the results of landslide susceptibility calculations, it may even ***improve*** the validation statistics, giving false confidence in the map.

-Steger, S., Brenning, A., Bell, R., and Glade, T., 2016, The impact of systematically incomplete and positionally inaccurate landslide inventories on statistical landslide susceptibility models, *in* EGU General Assembly Conference Abstracts.

Historic points from SLIDO

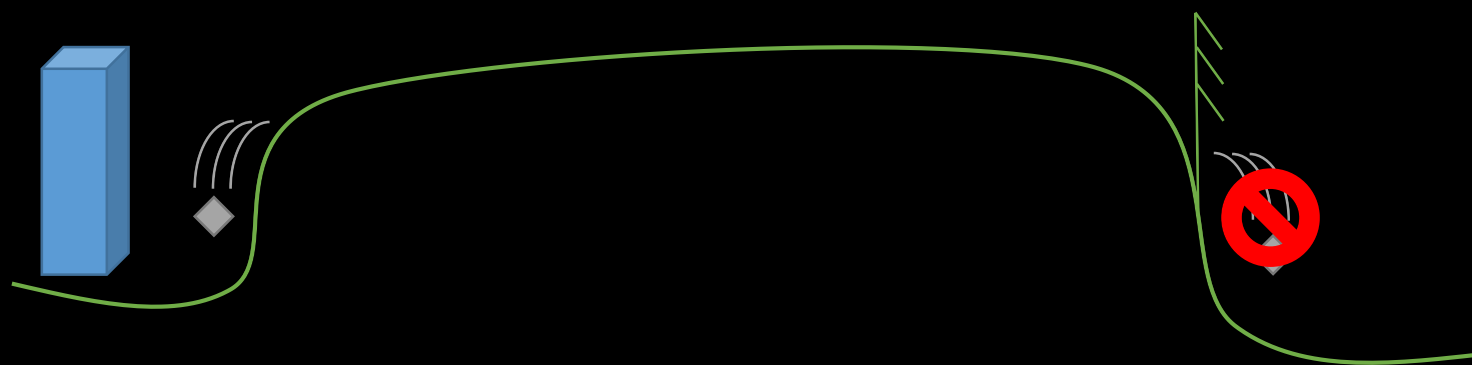


How it works

Cathy’s Condo
Mean slope = 30
Forest cover = 0%

Millie’s Meadow
Mean slope = 0
Forest cover = 0%

Bob’s Bluff
Mean slope = 30
Forest cover = 50%



	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-23.57	79460.0	0	1
slope	1.57	3746.0	0	1
forest	0.00	224800.0	0	1

Revised susceptibility workflow

In order to train an accurate landslide susceptibility model, we must correct for the numerous false negatives implied by our biased inventory.

1 Compile inventory

↓

2 Mitigate inventory bias

↓

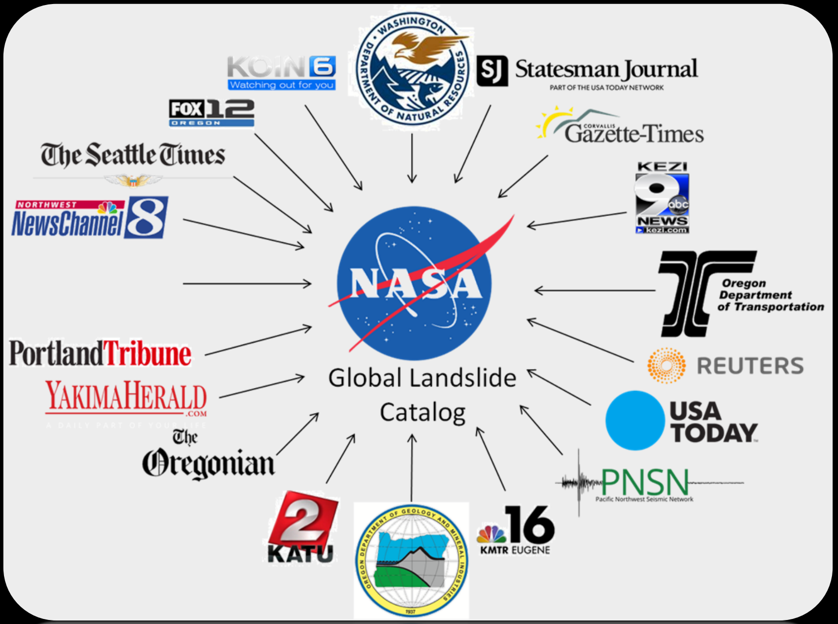
3 Fit empirical model

↓

4 Validate map with unused portion of inventory

1 Compile Pacific Northwest Landslide Inventory (PNLI)

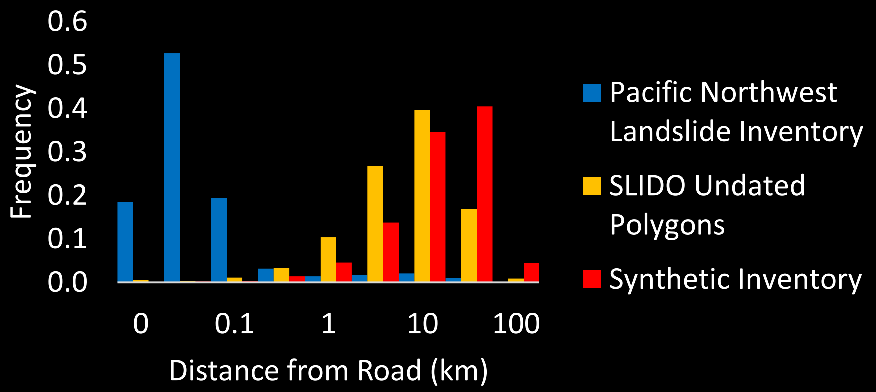
- 7,454 landslides with a known date, with 3,373 in Oregon and 4,081 in Washington.
- Year of occurrence was known for an additional 7,967 landslides.
- Year of occurrence was not known for 58,780 landslides. These were not used to generate the initial landslide susceptibility map.

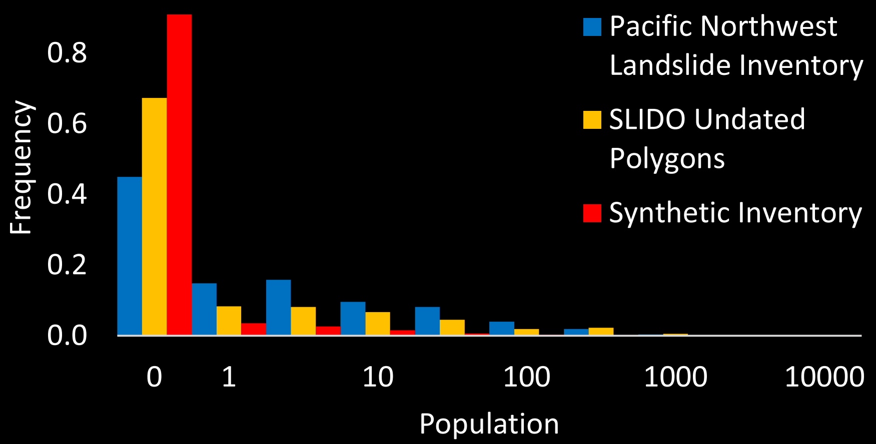


Kirschbaum, D., Psaltakis, J., and Stanley, T., 2016, Spatiotemporal properties of landslides in the Pacific Northwest, in Abstracts with Programs, Geological Society of America, Denver, Colorado, USA.

PNLI reporting biases

- A synthetic inventory was generated for this analysis from the Landslide susceptibility overview map of Oregon by randomly creating landslide initiation points with the same frequency of occurrence documented for each category (low-very high).





The synthetic inventory was generated for this analysis from the Landslide susceptibility overview map of Oregon.

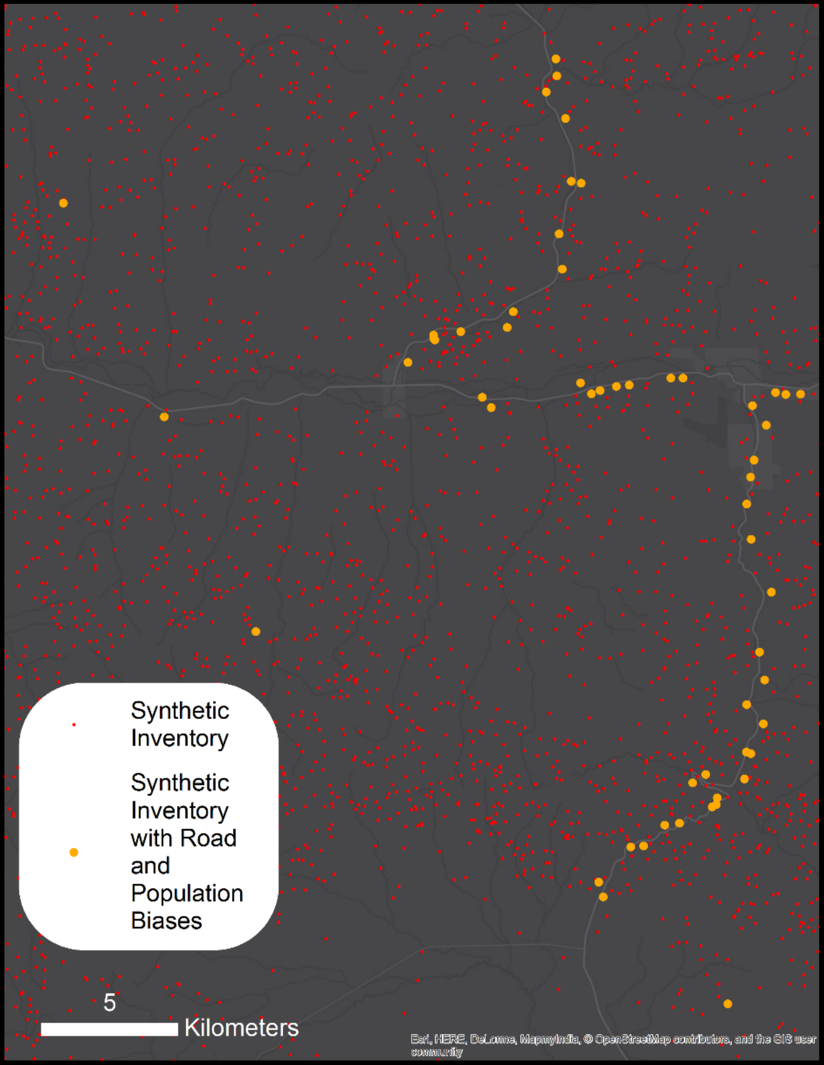
Simulate PNLI reporting biases

a) Start with two relatively unbiased inventories for Oregon

b) Truncate each inventory to match biases (roads, population, roads + population)

c) Fit logistic regression models

d) Compare biased susceptibility maps to results for initial inventories

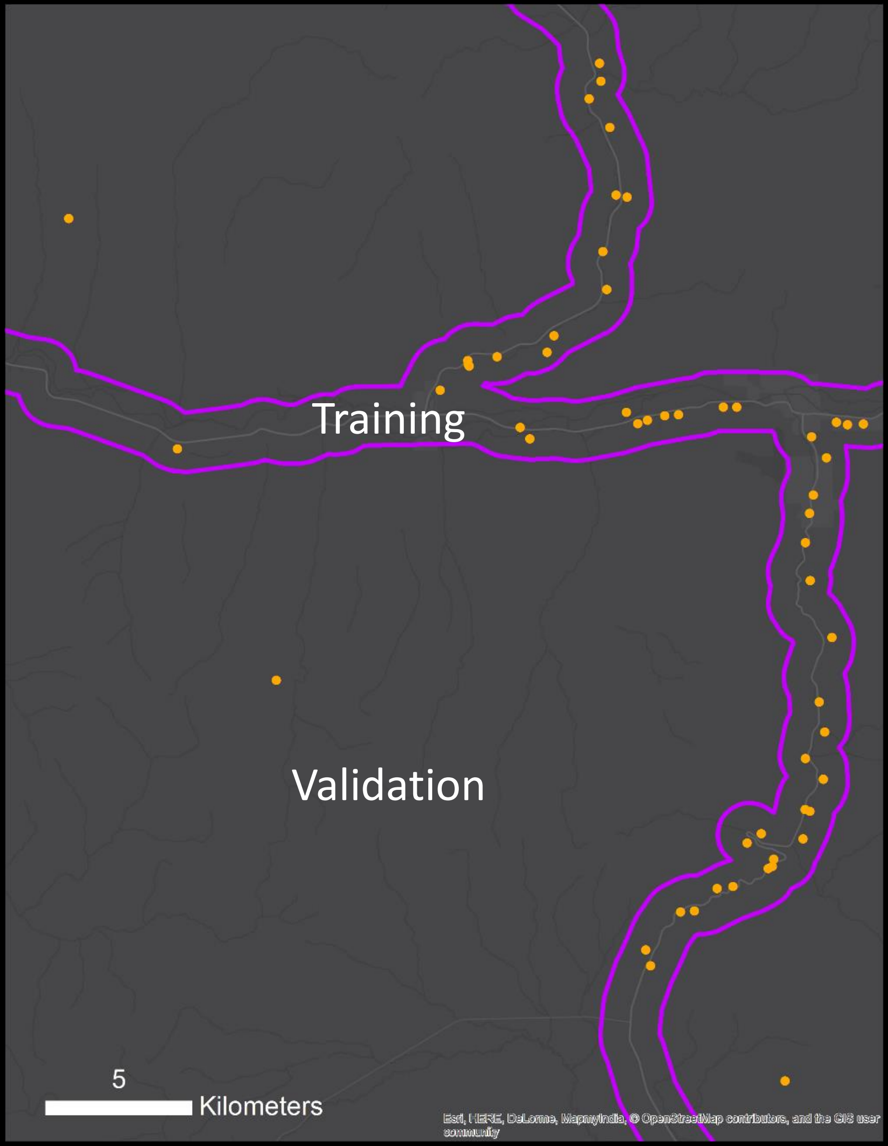


Synthetic inventory derived from:

Burns, W.J., Mickelson, K.A., and Madin, I.P., 2016, Landslide susceptibility overview map of Oregon.

2 Mitigate bias

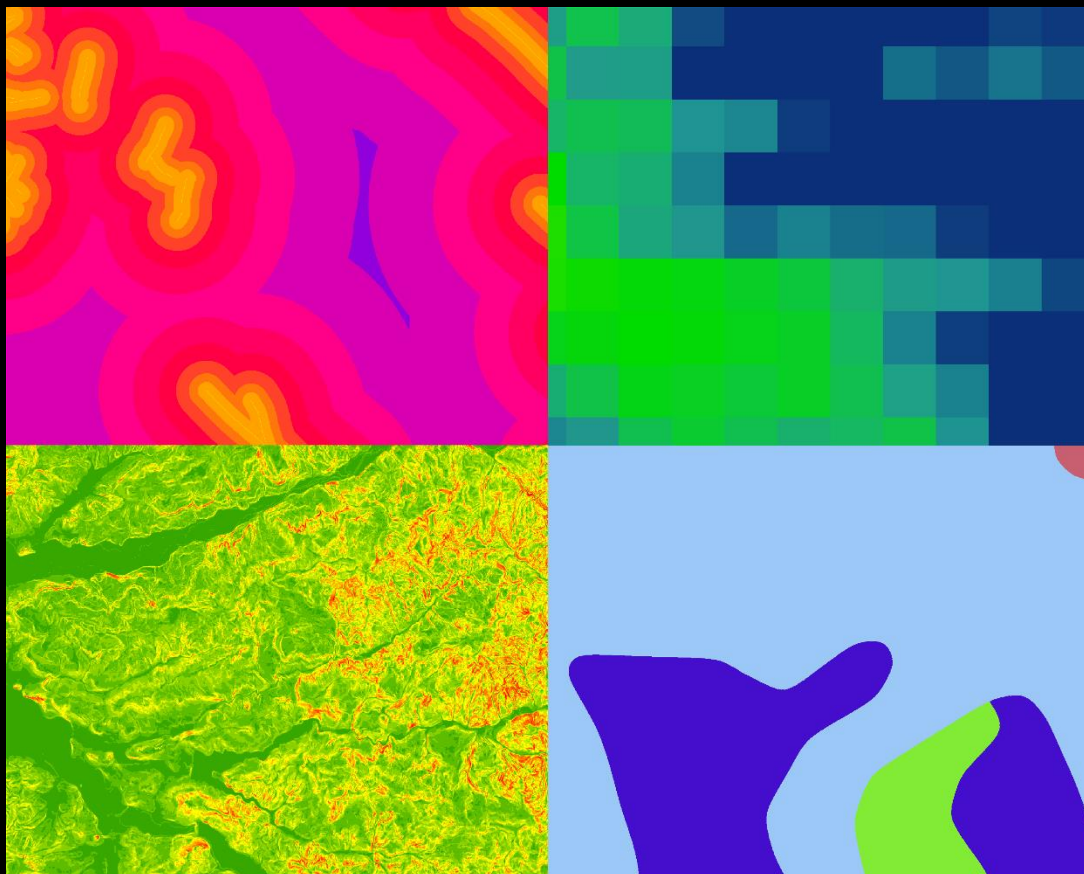
- While the problem could be solved through the use of a predefined method, the large inventory available for the Pacific Northwest would not inform such a model.
- In order to reduce the influence of false negatives, only areas located within 1 kilometer of a major highway were used to fit the model. Landslides in more remote areas were used as a validation dataset.



Center for International Earth Science Information Network, and Information Technology Outreach Services, 2013, Global Roads Open Access Data Set, Version 1.

3 Fit logistic regression model

Distance to fault | Mean Annual Precipitation



Slope | Geologic Unit

Oregon State University, 2016, 1971-2000 Annual Average Precipitation by State.

Smith, R.L., and Roe, W.P., 2015, Oregon Geologic Data Compilation, release 6.

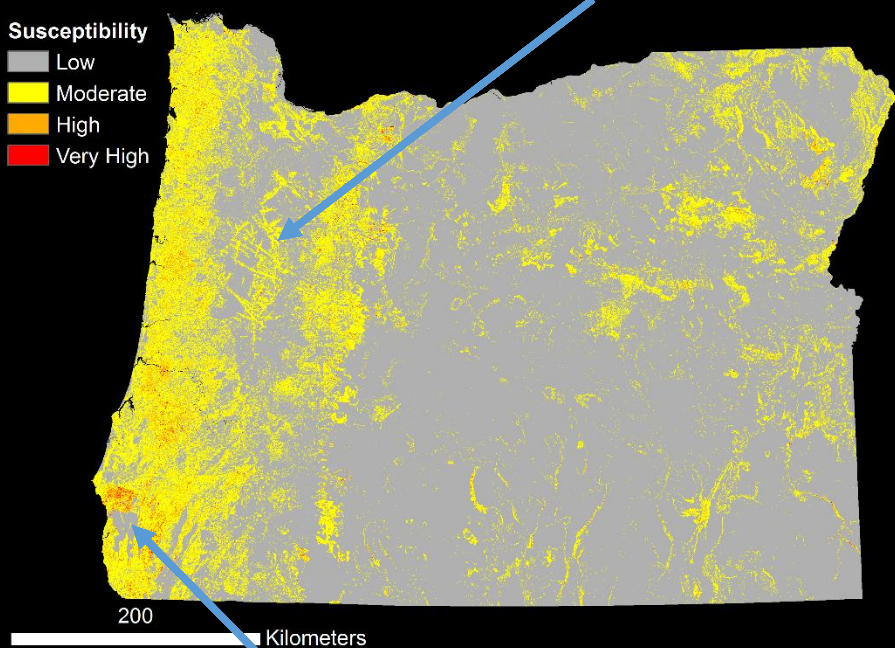
U.S. Geological Survey, 2016, The National Map: 3DEP products and services: The National Map, 3D Elevation Program Web page.

Susceptibility

4 Validate results

Biased

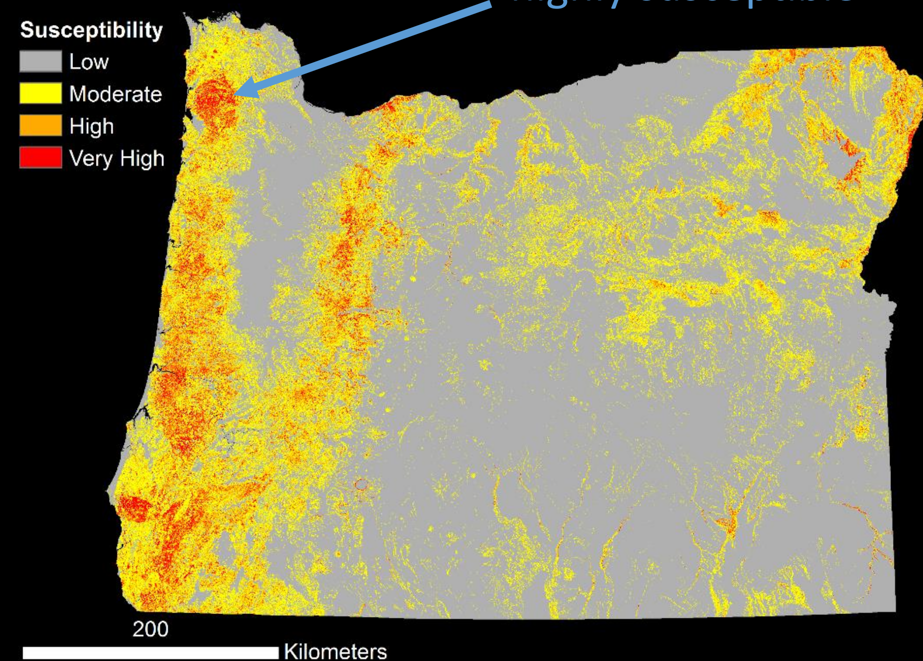
Overemphasis on distance to fault



Underestimate in specific geologic units

Bias-mitigated

Coast Range identified as highly susceptible



Discussion

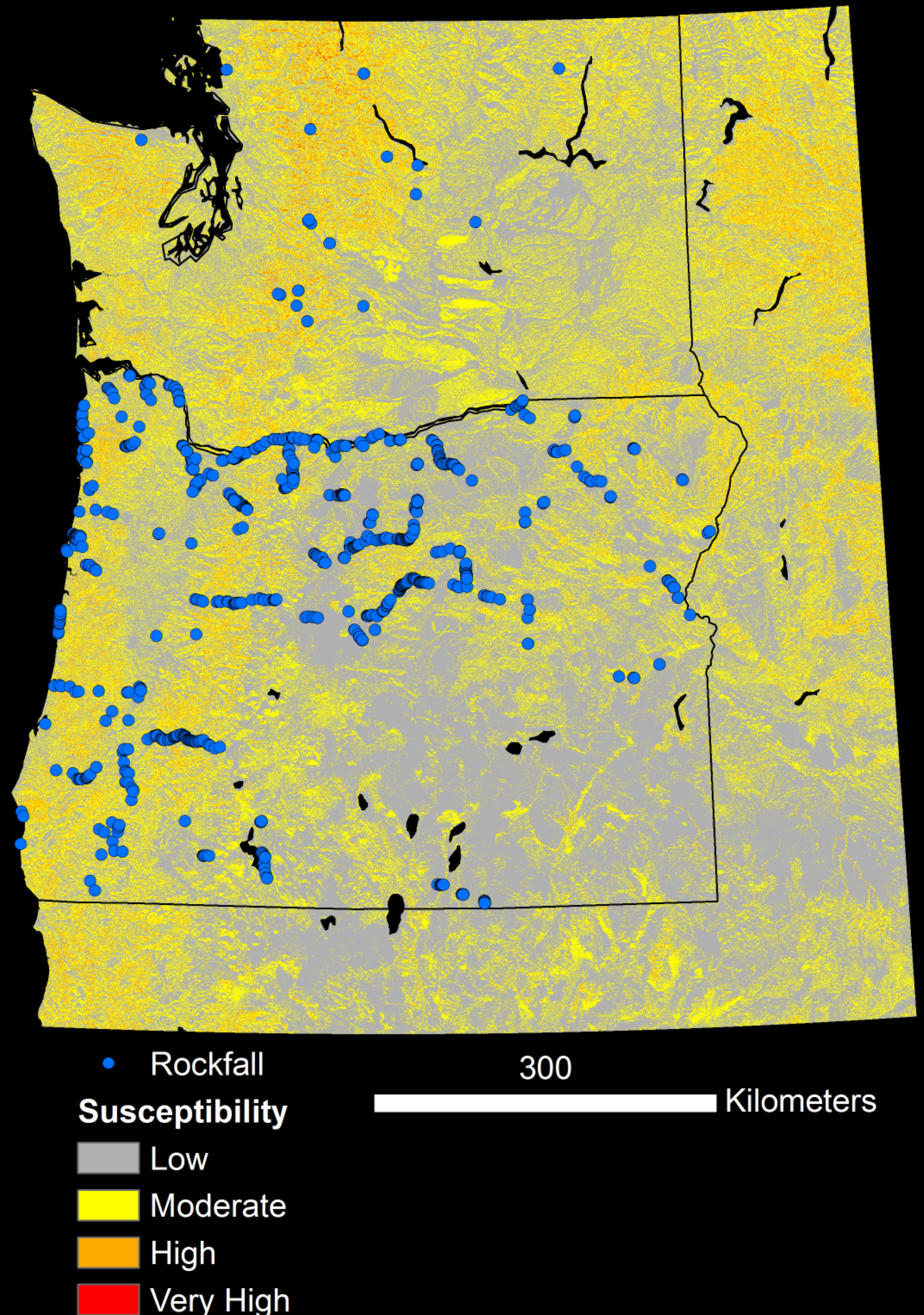
- While we believe that the PNLI is concentrated near highways and cities due to reporting bias, some of this effect is probably due to anthropogenic disturbance.
- Therefore, susceptibility may have been overestimated in the areas away from highways (validation zone). The estimates near highways should have represented anthropogenic effects correctly.

Conclusions

- Reporting bias can have a strong effect on the fitting of empirical landslide models.
- Although many strategies for bias mitigation could be employed, the simplest approach delivers generally plausible results that are most reliable in the most critical locations: along major highways and rail lines.

Next steps

- Map susceptibility across the Pacific Northwest, with separate models for each landslide type
- Identify trends in landslide-triggering precipitation
- Apply lessons learned across the USA



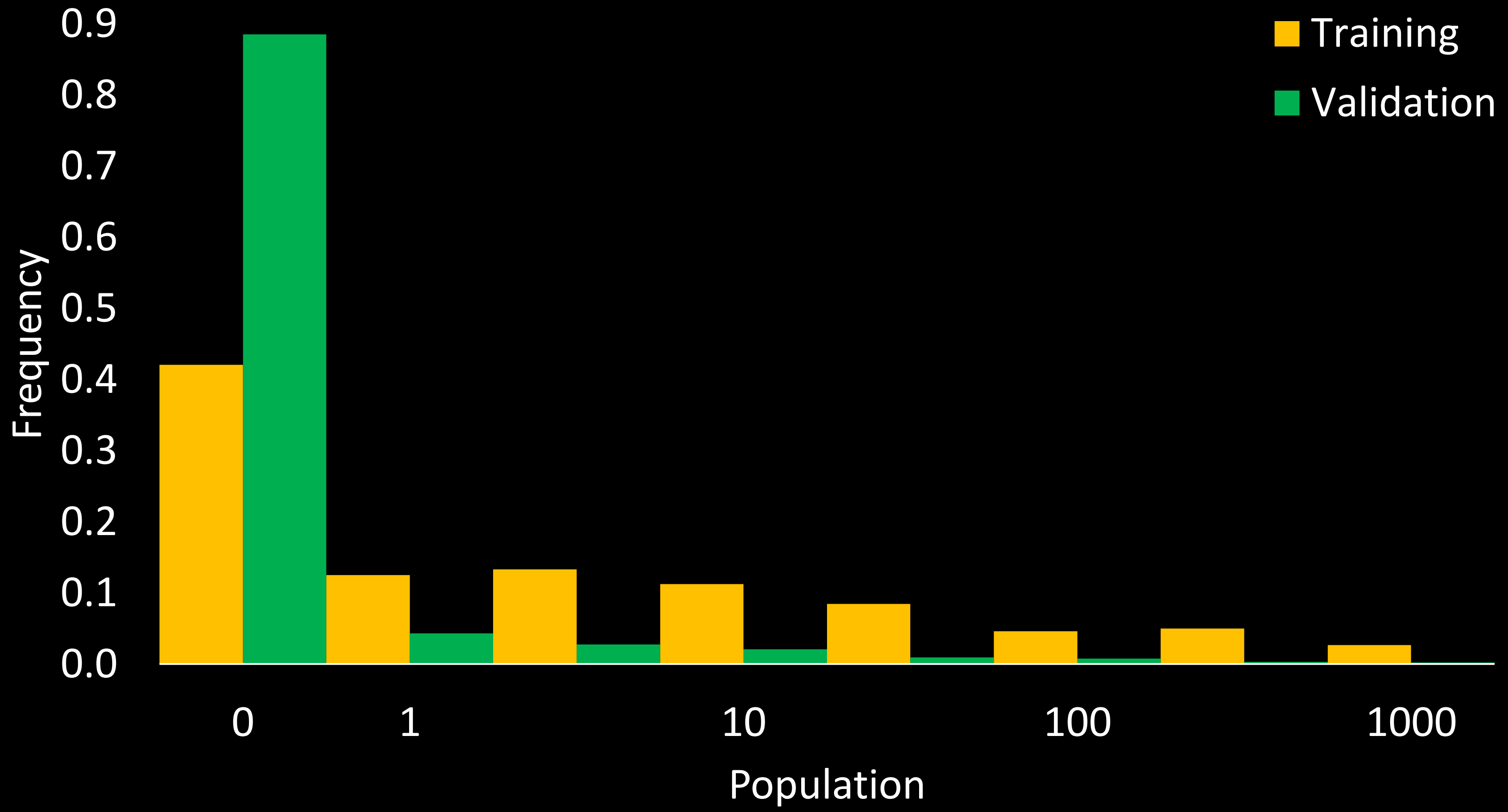
Thanks! Questions and suggestions welcome at: thomas.a.stanley@nasa.gov

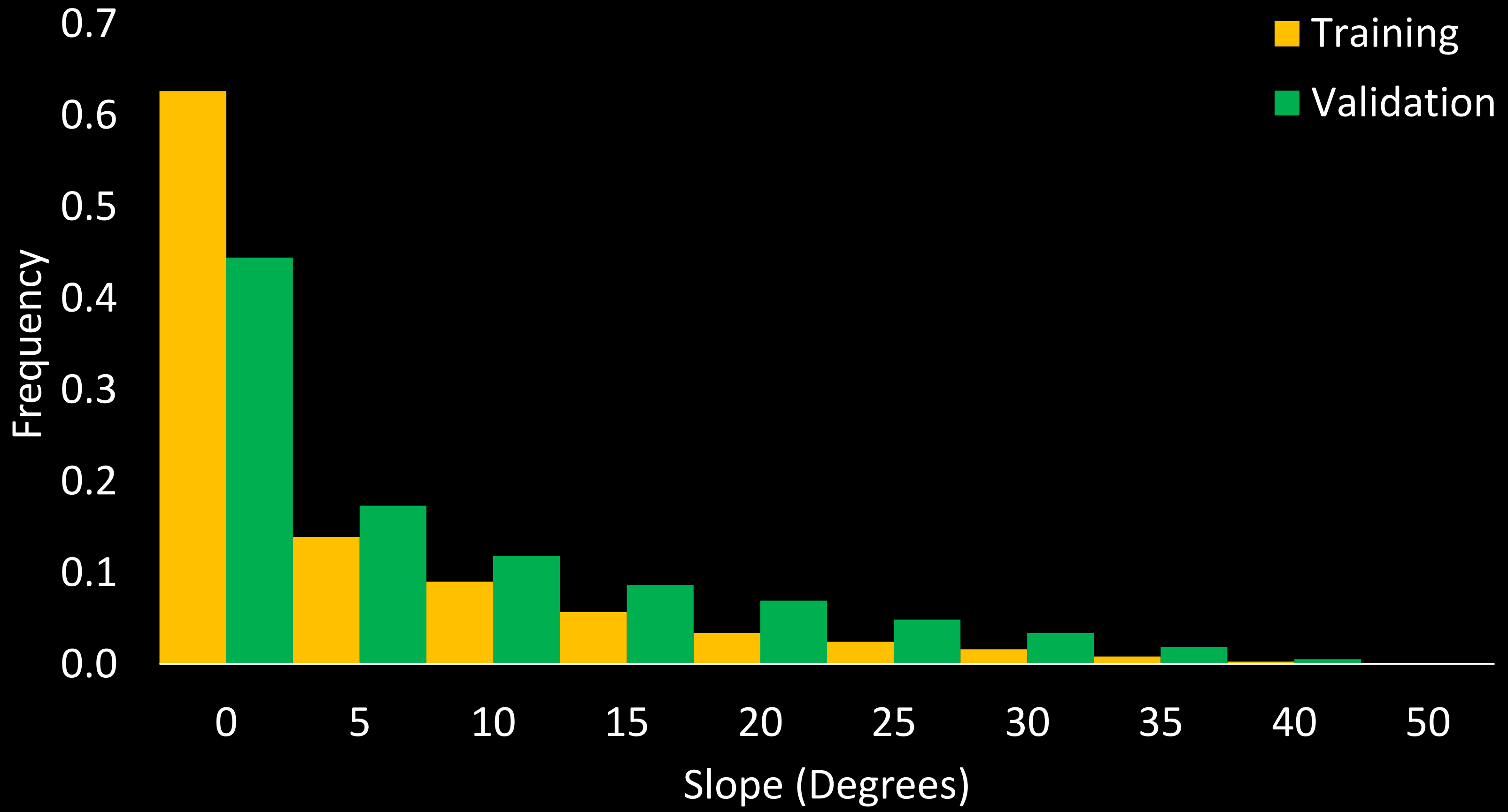
Acknowledgements

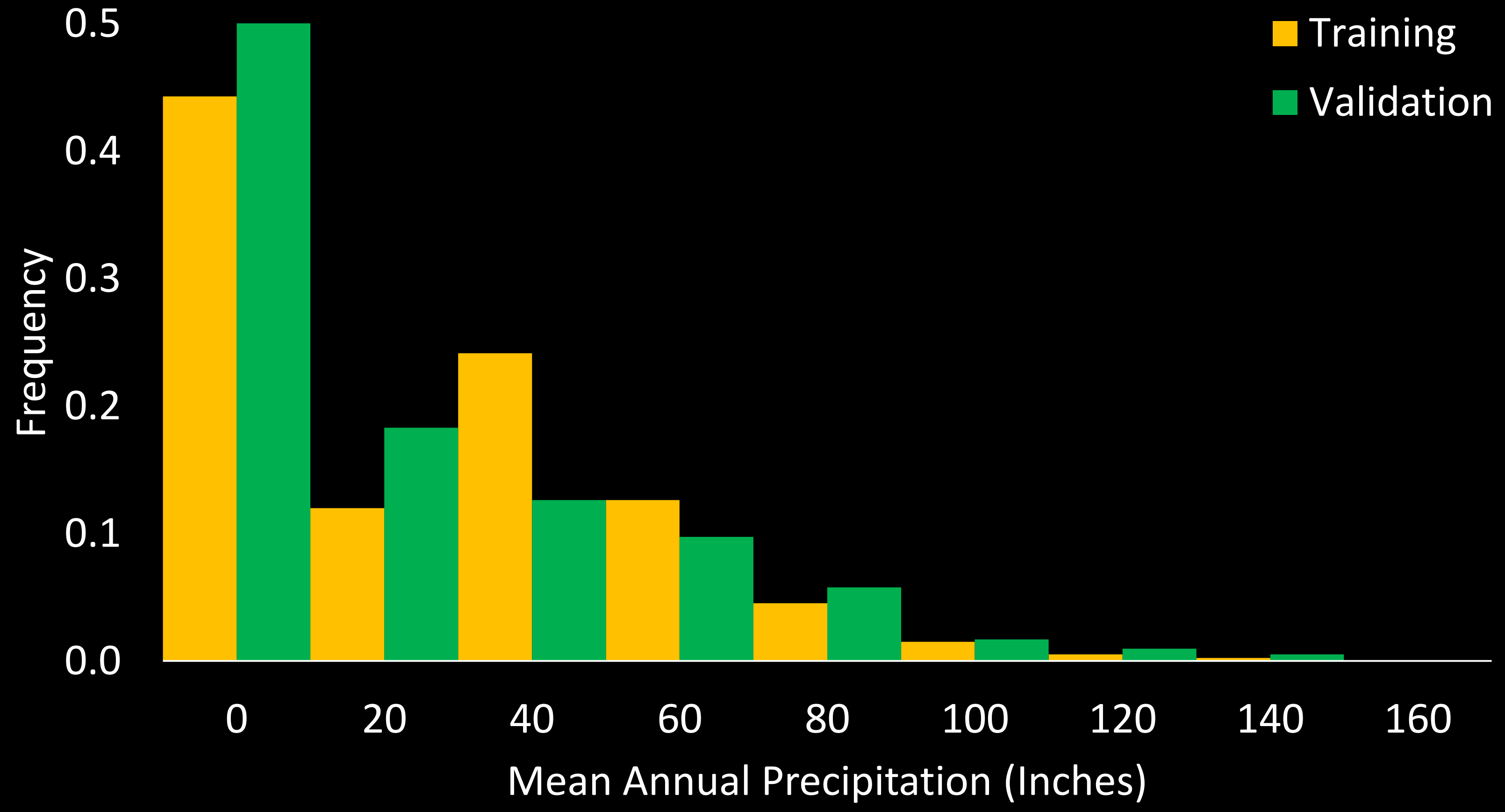
This research was funded by the NASA NNH14ZDA001N-INCA: Climate Indicators and Data Products for Future National Climate Assessments. This work would not have been possible without the data provided by DOGAMI, ODOT, and many others. We also thank Jordan Psaltakis for assembling the PNLI.

Recommended reading

- Montgomery, D.R., 1994, Road surface drainage, channel initiation, and slope instability: Water Resources Research, v. 30, no. 6, p. 1925–1932, doi: 10.1029/94WR00538.
- Steger, S., Bell, R., Petschko, H., and Glade, T., 2015, Evaluating the Effect of Modelling Methods and Landslide Inventories Used for Statistical Susceptibility Modelling, *in* Engineering Geology for Society and Territory - Volume 2, Springer International Publishing, Cham, p. 201–204.
- Steger, S., Brenning, A., Bell, R., Petschko, H., and Glade, T., 2016b, Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps.


Training
Validation
Frequency
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1
10
100
1000
Population


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Frequency
Training
Validation
0
5
10
15
20
25
30
35
40
50
Slope (Degrees)


Training
Validation
Frequency
0.5
0.4
0.3
0.2
0.1
0.0
0
20
40
60
80
100
120
140
160
Mean Annual Precipitation (Inches)