**I-GUIDE MODEL CARD**

The I-GUIDE Model Card is an easy-to-use tool that will allow you to create documentation for each model that you create or use in a project.

Using this tool will help facilitate transparency and reproducibility about your project. It will also help you comply with relevant policies of journals, funding agencies, and universities.

The Model Card applies to:

* Pre-existing models acquired from other sources, e.g., produced by other researchers;
* Models you and your collaborators produced yourselves;
* Models you and your collaborators produced by integrating two or more other models (e.g., coupling).

**Model Card Attribution**

This Model Card template is an adapted version of the I-GUIDE Data Card template, which itself is based on Google’s *Data Cards Playbook* (https://pair-code.github.io/datacardsplaybook/).  
It has been restructured to address key considerations for geospatial model transparency, performance evaluation, and ethical deployment, in alignment with the I-GUIDE research lifecycle.

This work is shared under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).  
License details: <https://creativecommons.org/licenses/by-sa/4.0/>

A logo with text on it

AI-generated content may be incorrect.

* + 1. **BASIC INFORMATION**

|  |  |
| --- | --- |
| Model Card ID Number | *FHP-01* |
| Model Name | *Flood Depth Prediction across the Contiguous United States using Multimodal Earth Observation Data and Machine Learning* |
| Model Version | *0.01* |
| Persistent Identifier | *N/A* |
| Outputs Supported | *N/A* |
| Model Card Author | *Parnian Ghaneei, The University of Alabama, pghaneei@crimson.ua.edu* |

* + 1. **MODEL OVERVIEW**

|  |  |
| --- | --- |
| Model Type | *(Select all that apply)*  ☐ AI model: *Random Forest*  ☐ Statistical model: *(Specify type)*  ☐ Other: *(Specify)* |
| Purposes | *(Select all that apply)*  ☐ Classification  ☐ Decision support  ☐ Forecasting  ☐ Regression  ☐ Simulation  ☐ Spatial analysis  ☐ Other: *(Specify)* |
| Domains of Application | *(Select all that apply)*  ☐ Climate science  ☐ Economics  ☐ Environmental impact modeling  ☐ Geospatial analysis  ☐ Hydrology  ☐ Population modeling  ☐ Other social systems modeling: *(Specify)*  ☐ Other: *(Specify)* |
| Model Authors and Developers | Contributors:  Parnian Ghaneei  Xiao Chen  Jikun Liu  Temitope Akinboyewa  Rufai Omowunmi Balogun  Aleksander Berg  Wen Zhou, PhD. |
| Source and Acquisition Method | ☐ Acquired (from external source)  ☐ Developed internally  ☐ Integrated from multiple models (e.g., coupled) |
| User Licensing | ☐ Open source: *(Specify license type)*  ☐ Proprietary: *(Specify owner)*  ☐ Other restrictions on use: *(Specify restrictions)* |
| Storage Location | ☐ Repository: *(Name and link to dataset)*  ☐ Project-specific storage: *(Describe location)* |
| Access Control Policies | ☐ Open  ☐ Embargoed: *(Describe release timeline)*  ☐ Restricted: *(Describe access criteria)* |
| Use Case | *Train a Random Forest regressor on engineered spatial features to predict flood depth; evaluate, compute and save feature importances.* |

* + 1. **MODEL INPUTS AND TRAINING DATA**

|  |  |
| --- | --- |
| Model Inputs | * [*Digital Elevation Model at 10m and 1m*](https://www.usgs.gov/3d-elevation-program) * [*National Land Cover Dataset Impervious Surfaces*](https://www.usgs.gov/centers/eros/science/annual-national-land-cover-database) * [*CHIRPS Precipitation*](https://www.chc.ucsb.edu/data/chirps) * [*Sentinel 1 VV and VH*](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD) * [*Soil Moisture Active Passive*](https://developers.google.com/earth-engine/tutorials/community/smap-soil-moisture) * [*Weather Parameters*](https://www.weatherapi.com/)   + *Humidity*   + *Dewpoint Temperature*   + *Wind Direction (UV)*   + *Wind Gust (kPh)*   + *Feels Like Temperature*   + *Wind Chill*   + *Heat Index*   + *Chance of Rain*   + *Chance of Snow*   + *Visibiltiy (km)* |
| Input Data Types | *(Select all that apply)*  ☐ Raster  ☐ Tabular  ☐ Time Series  ☐ Vector  ☐ Other: *(Specify)* |
| Training Data Used | * *Target:* [*USGS High Water Mark Height*](https://www.usgs.gov/water-science-school/science/high-water-marks-and-flooding) * *Inputs: Same as ‘*Model Inputs’ |
| Training Dataset Representativeness | *The target data are limited and predominantly concentrated within certain land cover and land use categories. This restricted coverage reduces the representativeness of the dataset across broader landscape conditions.* |

* + 1. **MODEL STRUCTURES**

|  |  |
| --- | --- |
| Feature Selection | *All the input data have been utilized.* |
| Hyperparameters and Tuning | *Hyperparameters = {*  *'n\_estimators': [50, 100, 200, 300],*  *'max\_depth': [None, 10, 20, 30],*  *'min\_samples\_split': [2, 5, 10],*  *'min\_samples\_leaf': [1, 2, 4],*  *'max\_features': ['sqrt', 'log2', None],*  *'bootstrap': [True, False]*  *} Randomized search with 5-fold and scoring of mean squared error.* |
| Software and Dependencies | *Python;*  *Libraries:*  *numpy*  *pandas*  *matplotlib*  *sklearn* |

* + 1. **MODEL PERFORMANCE AND EVALUATION**

|  |  |
| --- | --- |
| Validation Approach | ☐ Cross-validation  ☐ Holdout set  ☐ Time series split  ☐ Other: *(Specify)* |
| Evaluation Results | *Final Test Set Performance:*  *R² Score: 0.4804*  *RMSE: 0.5130*  *MAE: 0.3549*  *MSE: 0.2631* |
| Testing or Validation Data Used | *Derived from the same combined features* |
| *(If model is integrated from multiple other models)*  Contribution of Constituent Models | *N/A* |

* + 1. **MODEL ADAPTATION AND CUSTOMIZATION (for acquired or integrated models only)**

|  |  |
| --- | --- |
| Source Models | *N/A* |
| Availability of Source Model Code | *N/A* |
| Modifications | *N/A* |
| Training Data Adjustments | *N/A* |

* + 1. **MODEL DEPLOYMENT AND USAGE**

|  |  |
| --- | --- |
| Computational Requirements | *(List hardware/software requirements)* |
| Geospatial Considerations | *The target data are limited and predominantly concentrated within certain land cover and land use categories. This restricted coverage reduce the representativeness of the dataset across broader landscape conditions.* |

* + 1. **TRANSPARENCY, EXPLAINABILITY, AND INTERPRETABILITY**

|  |  |
| --- | --- |
| Model Transparency | ☐ Fully transparent (rule-based, interpretable ML)  ☐ Partially transparent (some explainability features)  ☐ Black box (deep learning, complex ML models) |
| Explainability Features | ☐ Feature importance analysis  ☐ LIME  ☐ Sensitivity analysis  ☐ SHAP values  ☐ Other: *(Specify)* |
| Interpretability Challenges | *Nonlinear feature interactions can make local reasoning difficult; feature importance may be biased toward high-cardinality or high-variance features.* |
| Communication of Model Limitations | *Uncertainty quantification has not been included.* |

* + 1. **OTHER ETHICAL CONSIDERATIONS**

|  |  |
| --- | --- |
| Ethical Risks (Other Than Transparency, Explainability and Interpretability) | *(Select all that apply)*  ☐ Bias in training data: *(Specify)*  ☐ Intentional misuse risks: *(Specify)*  ☐ Privacy risks and surveillance: *(Specify)*  ☐ Security risks: *(Specify)*  ☐ Stigmatization of individuals or communities: *(Specify)*  ☐ Other: *(Specify)* |
| Measures Taken to Address Ethical Risks | *The dataset was randomly split, and the test and validation sets were standardized using the statistical properties of the training data to prevent data leakage* |
| Suitable Uses | *Decision support and exploratory analysis within data regime similar to training set.* |
| Unsuitable Uses | *Extrapolation to unseen regions.* |