Rainfall Amount Prediction with Climate and Geospatial Data

JR Lee, Ayumu Yamagishi

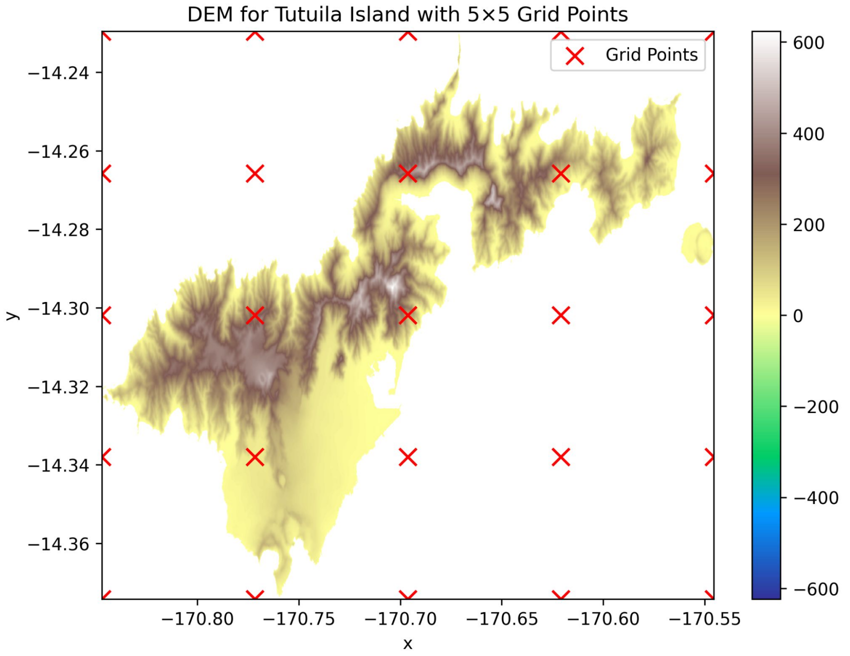
Presentation Overview

- Introduction to the project
- Approaches & Methodology
- Hyperparameter Tuning
- Single Best Model
- Ensemble Cross-Validation Model
- Comparison and Analysis
- Conclusion

Introduction

3 Key Components for Prediction

- Climate reanalysis data
- Local and regional elevation patches
- Temporal Information



Rainfall Prediction Models

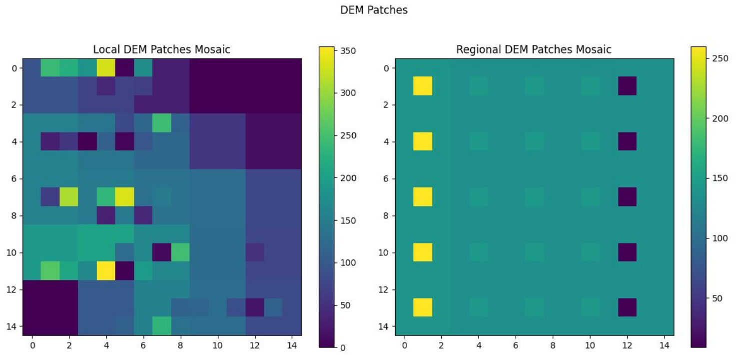
- **Single Best Model**
  - A single hyperparameter-optimized LAND (Location-Agnostic Neural Downscaling) model
  - Uses the best hyperparameters from hypertuning
- **Ensemble Cross-Validation Model**
  - A robust ensemble of 5 of the single best model
  - Trained at 5 different random seeds with 5-fold cross-validation
  - Ensemble of 5x5 = 25 total models

Approaches and Methodology

Data Resources

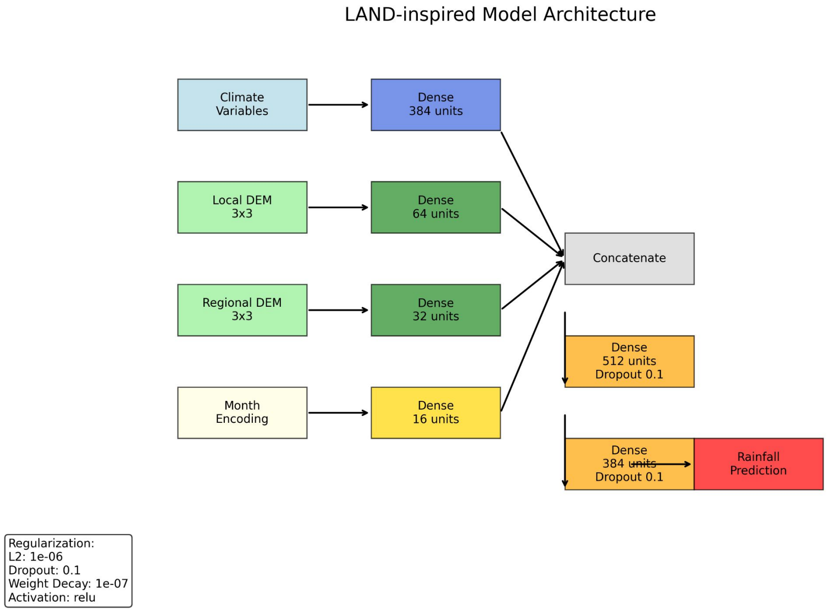
- **Climate Reanalysis Data**
  - 16 atmospheric variables from climate reanalysis datasets
  - E.g., air temperature difference between pressure levels, geospatial heights, humidities, moistures, etc.
- **Digital Elevation Model (DEM) Data:** Topographical information at two scaled formats:
  - **Local DEM:** 3×3 patches with 4 km per cell (12km total coverage)
  - **Regional DEM:** 3×3 patches with 20 km per cell (60km total coverage)
- **Temporal Information:** One-hot encoded months to observe seasonality


```
climate_air_2m,climate_air_temp_diff_1000_500,climate_hgt_1000,climate_hgt_500,climate_merid_moist_700,climate_
275.73466874871866,39.70901766470388,102.24777960152323,5898.380802091502,181.78138355199079,1018.9342196665802
275.7397319096871,39.707133367350934,102.24777960152733,5898.373255491722,181.7437203525616,1018.7214928250726,-
275.74479507065223,39.70524906999617,102.24777960152369,5898.365708891884,181.7060571531324,1018.5087659835644,-
275.74985823162035,39.70336477264277,102.24777960152505,5898.358162291959,181.66839395370323,1018.2960391420568
275.754921392587,39.701480475289145,102.24777960152414,5898.350615692296,181.63073075427405,1018.0833123005489,-
275.60164060969186,39.73727393300828,102.1465388636193,5900.482831752073,183.2521468043148,1016.3018975656315,-
275.60673370181075,39.73538527440701,102.14653886361839,5900.474849120568,183.214171738385,1016.0897338210121,-
275.61182679392823,39.733496615807326,102.14653886362112,5900.466866489383,183.17619667245557,1015.877570076392
275.61691988604304,39.73160795720628,102.14653886361839,5900.4588838579075,183.13822160652597,1015.665406331773
275.62201297815824,39.729719298606824,102.14653886362112,5900.450901226723,183.1002465405964,1015.4532425871547
275.468612470663,39.765530201312004,102.04529812571218,5902.584861412441,184.72291005663874,1013.669575464682,-
275.4737354939287,39.763637181464446,102.04529812571445,5902.576442749647,184.68462312420877,1013.457974816952,-
```


```
month_0,month_1,month_2,month_3,month_4,month_5…
0.0,0.0,0.0,0.0,0.0,1.0,...
```

Neural Network Architecture



- **Climate Branch**
  - 16 variables, dense layer with batch normalization
- **Local DEM Branch**
  - Flatten 3x3 local DEM patch, dense layer with batch normalization
- **Regional DEM Branch**
  - Flatten 3x3 regional DEM patch, dense layer with batch normalization
- **Month Branch**
  - One-hot encoded month, dense layer with batch normalization

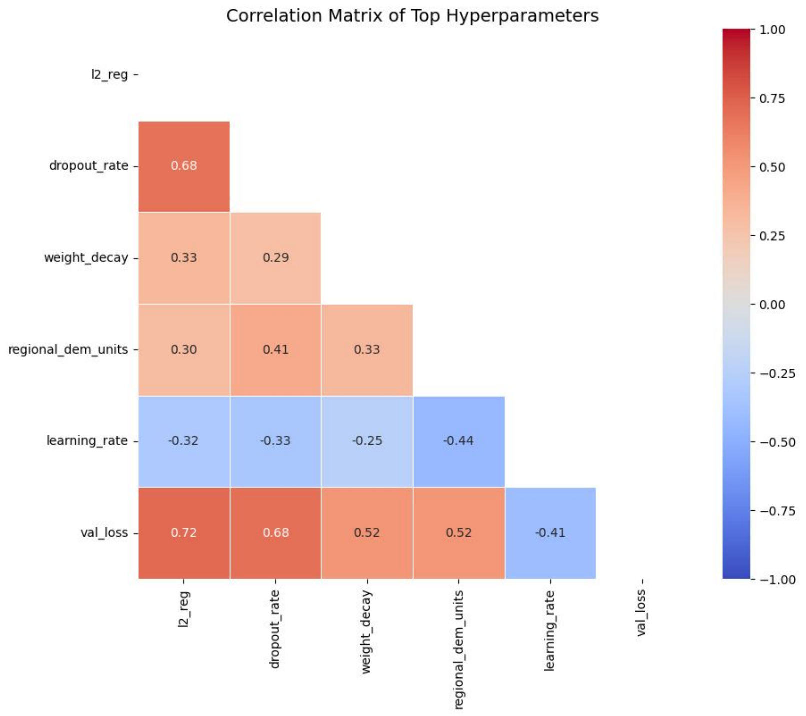
Hyperparameter Tuning

Tunable Parameters

Network Architecture: number of neurons in hidden layers (na, nb), units of each branch, activation function

Regularization: Dropout rate, learning rate for L2 regularization,

Optimization Parameters: Weights decay for AdamW, Learning rate schedule parameters



Tuning Process

na: 512	**nb**: 384
dropout_rate: 0.1	**l2_reg**: 1e-06
learning_rate: 0.01	**weight_decay**: 1e-07
local_dem_units: 64	**regional_dem_units**: 32
month_units: 16	**climate_units**: 384
use_residual: False	**activation**: relu

Algorithm: Bayesian Optimization with Hyperband early stopping

Number of Trials: 100

Epochs per Trial: Up to 50 with early stopping

Objective: Minimize validation loss (MSE)

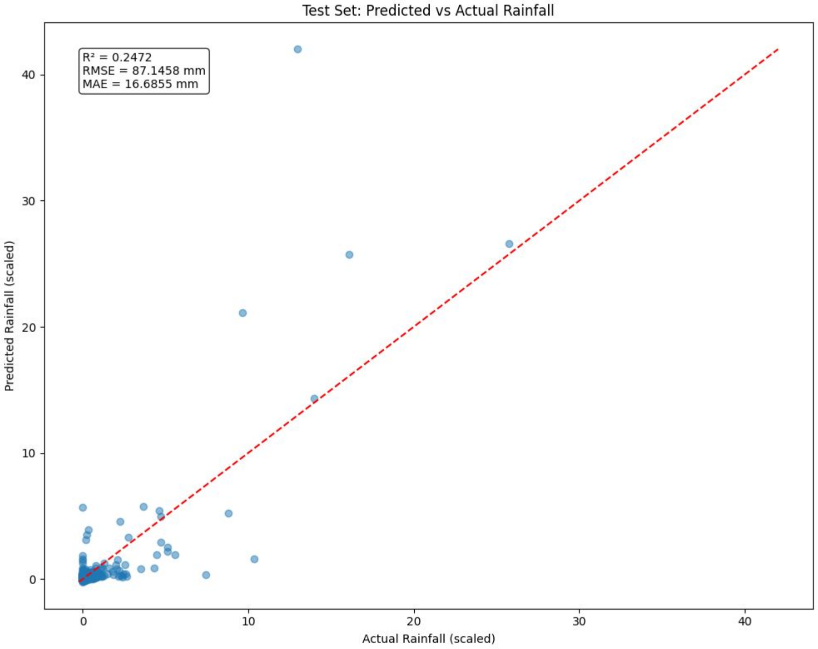
Early Stopping: Patience of 15 epochs

Single Best Model

Single Best Model Result

Validation
R^2: 0.7212
RMSE: .4913 in
MAE: .1324 in

Test
R^2: 0.2472
RMSE: .8715 in
MAE: .1669 in



Note: There is an error in the units. It should 1/100 in and not mm.

Ensemble Cross-Validation Model

Ensemble Cross-Validation Model Implementation

- 25-model structure
  - 5 cross-validation folds
  - 5 models per fold
- Training Process
  - Split Data
    - 90% for training data
      - Further split into 90% for training set and 10% for validation
    - 10% set aside for test data (final evaluation)
- Ensemble Aggregation: Average of 25 models in total

Ensemble Cross-Validation Model Result

Avg. across 5 folds

R^2: 0.625

RMSE: .6334 in

MAE: .1443 in

Best fold (Fold 5)

R^2: 0.701

RMSE: .5895 in

MAE: .1382 in

Held-out test set

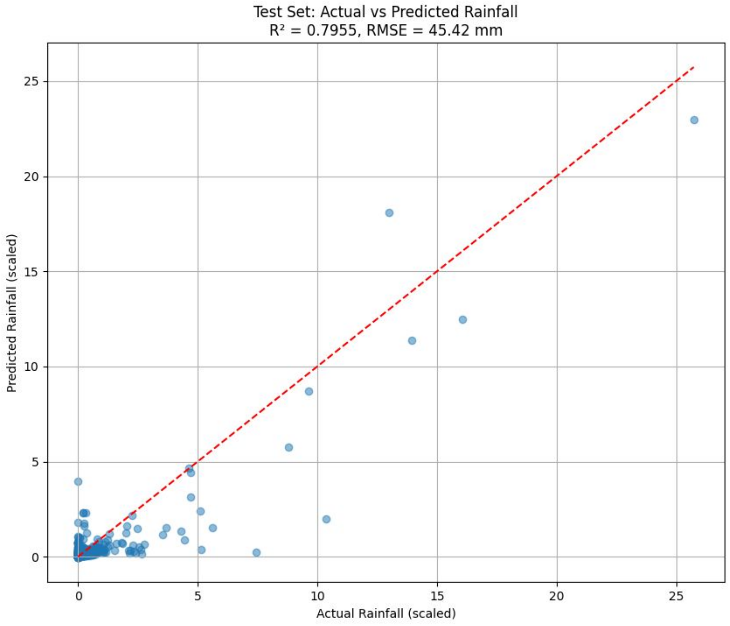
R^2: 0.796

RMSE: .4542 in

MAE: .1325 in

Ensemble Cross-Validation Model Advantages

- Reduced variance in prediction
- Improved generalization
- Robustness to initialization through different random seeds
- Less overfitting to validation data



Note: There is an error in the units. It should 1/100 in and not mm.

Conclusion

Conclusion

- We developed a rainfall prediction model using the LAND (Location-Agnostic Neural Downscaling) method with deep learning.
- Two models were compared:
    - Single Best Model: good validation performance, but weak test results.
    - Ensemble Cross-Validation Model: test R^2 of 0.7955
        - Reduced overfitting
        - Improved generalization
        - More reliable prediction results.
- This study shows that deep learning combined with ensemble strategies can significantly improve rainfall prediction performance (R^2 score).
    - Area for improvement: different metrics like Jaccard Index, etc.