Master Thesis Call

Evaluation of Soil Moisture Data from ERA5 Using Sensor-Based Data in the DEMIN Region, Germany

Duration: 6 Months

Objective: To evaluate the accuracy and reliability of ERA5 soil moisture data by comparing it with ground-based sensor data in the DEMIN region, using advanced statistical models and time series analysis techniques.

Thesis Overview:
This thesis will focus on the comparison between soil moisture data from the ERA5 reanalysis dataset and sensor-based soil moisture data collected from the DEMIN region in Germany. The candidate will engage in extensive data preprocessing, exploratory data analysis, and the application of both traditional and state-of-the-art statistical models.

Key Areas of Work:

- Data collection and preprocessing: Aligning ERA5 and sensor-based soil moisture data for temporal and spatial consistency.
- Comparative analysis using traditional statistical methods (e.g., correlation, RMSE) to evaluate the performance of ERA5 data.
- Implementation of advanced statistical models, including ARIMA, Random Forest, and geostatistical techniques, for a deeper analysis of soil moisture variability.
- Application of the Granger causality test to explore potential causal relationships between ERA5 and sensor data.
- Detailed error analysis and model tuning to enhance the accuracy of predictions.
- Interpretation of results with a focus on soil moisture dynamics in the DEMIN region and comparison with traditional methods.

Expected Outcomes:

- A comprehensive evaluation of the ERA5 soil moisture data's accuracy in the DEMIN region.
- Insights into the potential causal relationships between ERA5 and sensor-based soil moisture data.
- Development of robust statistical models for soil moisture prediction and analysis.
- Recommendations for improving soil moisture monitoring in data-limited regions.

Who Should Apply:
This thesis is ideal for students with a background in environmental science, geoinformatics, hydrology, or related fields, and an interest in remote sensing, statistical modeling, and data analysis. Experience with programming languages such as Python or R, and familiarity with geospatial data, is highly desirable.

Contact Information: For more information, please contact Dr. Muhammad Usman (muhammad.usman@geo.uni-halle.de) at Department of Geoecology.

Detailed Plan

Evaluation of Soil Moisture Data from ERA5 using Sensor-Based Soil Moisture Data in the DEMIN Region of Germany with Advanced Statistical Modeling

Duration: 6 Months

Objective: To evaluate the accuracy and reliability of ERA5 soil moisture data by comparing it with ground-based sensor data in the DEMIN region, using advanced statistical models, including the Granger causality test.

Month 1: Literature Review and Data Collection

1.1 Literature Review

- **Week 1-2:**
  - Review literature on soil moisture estimation techniques and previous evaluations of ERA5 data.
  - Explore soil moisture dynamics in the DEMIN region.
  - Research advanced statistical models and the application of Granger causality tests in time series analysis.
- **Deliverable:** Comprehensive literature review document.

1.2 Data Collection

- **Week 3-4:**
  - Obtain sensor-based soil moisture data for the DEMIN region.
  - Download and preprocess ERA5 soil moisture data.
  - Ensure temporal and spatial alignment of the data.
- **Deliverable:** Organized and preprocessed datasets.

Month 2: Data Preprocessing and Initial Analysis

2.1 Data Preprocessing

- **Week 1-2:**
  - Handle missing data and align time series.
  - Prepare the data for time series analysis.
- **Deliverable:** Ready-to-analyze datasets.

2.2 Exploratory Data Analysis (EDA)

- **Week 3-4:**
  - Perform initial analysis and visualization of the datasets.
  - Identify trends and discrepancies between ERA5 and sensor data.
- **Deliverable:** EDA report with visualizations.

Month 3: Comparative Analysis

3.1 Statistical Comparison

- **Week 1-2:**
  - Apply traditional statistical measures (e.g., correlation, RMSE) to compare datasets.
  - Analyze seasonal and diurnal patterns.
- **Deliverable:** Statistical analysis report.

3.2 Introduction of State-of-the-Art Statistical Modeling

- **Week 3-4:**
  - Implement advanced models (e.g., ARIMA, Random Forest, Geostatistical models).
  - Apply Granger causality tests to assess potential causal relationships.
  - Cross-validate models to ensure robustness.
- **Deliverable:** Report on model implementation and Granger causality test results.

Month 4: Advanced Analysis with State-of-the-Art Models

4.1 Model Calibration and Validation

- **Week 1-2:**
  - Calibrate models using training data and validate on test data.
  - Ensure robustness with cross-validation and bootstrapping.

- **Deliverable:** Model calibration and validation report.

4.2 Error Analysis and Improvement

- **Week 3-4:**
  - Conduct detailed error analysis and model tuning.
  - Apply ensemble techniques if necessary.
- **Deliverable:** Enhanced model performance report.

Month 5: Interpretation and Discussion

5.1 Interpretation of Model and Granger Causality Test Results

- **Week 1-2:**
  - Interpret the results of the statistical models and the Granger causality tests.
  - Discuss the implications for soil moisture variability and causality in the DEMIN region.
- **Deliverable:** Draft of interpretation and discussion sections.

5.2 Comparison with Traditional Methods

- **Week 3-4:**
  - Compare results with traditional methods.
  - Evaluate the benefits of using advanced models and Granger causality tests.
- **Deliverable:** Comparative analysis report.

Month 6: Thesis Writing and Finalization

6.1 Thesis Writing

- **Week 1-2:**
  - Compile sections into a cohesive thesis document.
  - Write the introduction, methodology, and conclusion chapters.
- **Week 3-4:**
  - Review and revise based on feedback.
  - Prepare appendices and finalize the document.
- **Deliverable:** Complete thesis draft.

6.2 Finalization and Submission

- **Final Week:**
  - Proofread, format, and submit the thesis.
- **Deliverable:** Final submitted thesis.