Hybrid Machine Learning Soil Moisture Forecasting

Eduart Murcia [1]
Sandra M. Guzmán, Assistant professor[1]

[1]Smart Irrigation and Hydrology- Agricultural and Biological Engineering Department, Indian River Research and Education Center, Fort Pierce, FL

Rationale

Accurate estimations of soil water content are needed to create decision-support systems for **irrigation management** in agriculture. **Hybrid** approaches combining **signal preprocessing** techniques and **ensemble learning** algorithms could provide the necessary accuracy for a reliable **machine learning (ML) forecast** system.

Background

In our lab we had **first-hand experience** working with **soil moisture sensors data** for irrigation management and from there we identified the following **challenges:**

- **Complexity** created from non-linear and non-stationary time series.
- **Single algorithms** are **limited** to performing well only **on specific** parts of the **data**.

Methods

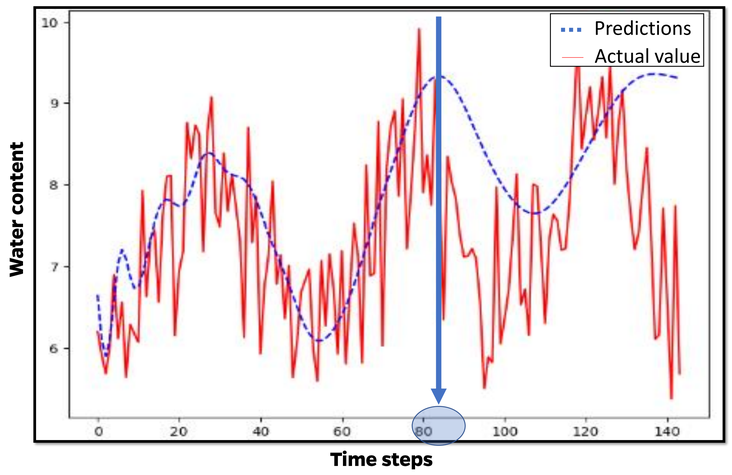
1. We implemented two **signal decomposition** techniques to **reduce the complexity** of the time series data.
2. We implemented an **ensemble learning** algorithm known as a **Mixture of Experts (MoE)**, which creates **expert** models for **specific** sections of the time series using a gating network.
3. For the **MoE** we used **sequential** training of the expert algorithms and a **Gaussian Gating Network** that were trained using an Expectation-Maximization (EM) algorithm to produce point estimation.

Results MoE

1. With the **MoE** were able to **extend the time** in which the model performs relatively well **from 82 to 2000-time steps.**
2. During the **training and validation**, the models performed **exceptionally well**, however, the performance was reduced in the **testing**, which suggests some signs of overfitting.

3. The EM optimization algorithm made the convergence of the model to global minima dependent on the initialization values.

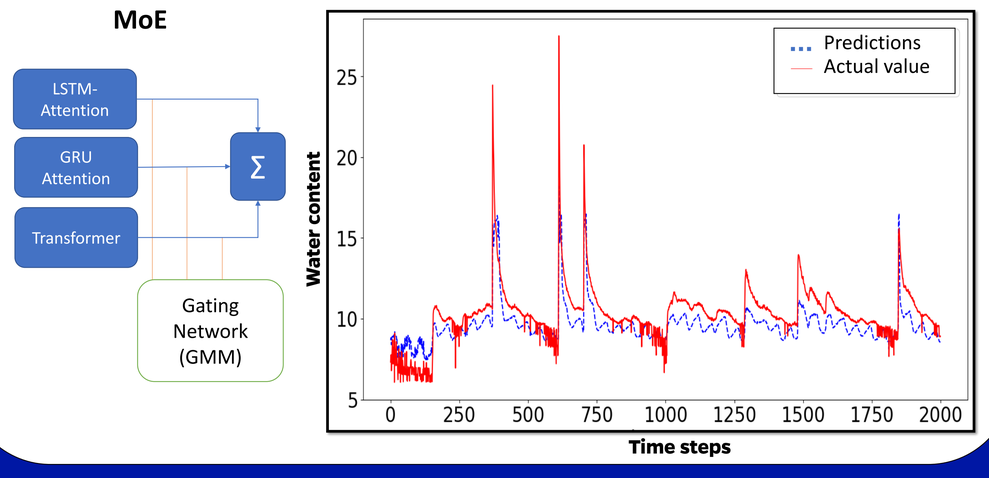
Stage 1 : Signal preprocessing



- Incorporating **signal decomposition** techniques caused an average reduction in the **mean absolute error (MAE) of 44%** and **increased** the **data interpretability.**

- After around **84-time steps (42 hours)** the model **deteriorates.**

Stage 2: Ensemble Learning MoE



- With the **MoE** were able to **extend the time** in which the model performs relatively well **from 82 (42 hours) to 2000-time steps (42 days).**

Results MoE

Table 1. Loss (Mean Absolute Error) on training and validation

	Sensor 1	Sensor 2	Sensor 3
Loss training	0.002	0.012	0.029
Loss Validation	0.002	0.040	0.140

Table 2. Mean Absolute Error & Mean Squared Error on test dataset

	Sensor 1	Sensor 2	Sensor 3
MSE	1.370	3.055	3.181
MAE	0.953	1.365	1.542

Conclusions

- A **combination of signal preprocessing and ensemble learning** algorithms can **increase** the **accuracy** of **ML-based soil moisture forecasting** systems for up to **several weeks**, which meets the irrigation management needs of most agricultural systems.
- **•Signal preprocessing techniques contribute mainly to reducing the complexity of times series, while ensemble learning approaches contribute the most to increasing the forecast model's legitimacy over time.**

Future work

1. Developing a gating network that allows incorporating multiple heterogeneous models to extend the applicability of the model to a wider range of time series characteristics.

Acknowledgment

- **This work is supported by the University of Florida IFAS Archer Early Career Seed Grant.**
- **National Institute of Food and Agriculture, U.S. Department of Agriculture, Hatch project under accession number 1021250**
- **Smart Irrigation and Hydrology Lab members.**
- **Dr. Catia Silva.**

Interested?
Scan me!

emurciabotache@ufl.edu - sandra.guzmangut@ufl.edu