Dynamic Structure Learning through Graph Neural Network for Forecasting Soil Moisture in Precision Agriculture

Anoushka Vyas [1] Sambaran Bandyopadhyay [2]

[1]Department of Computer Science, Virginia Tech, Arlington, VA, USA [2]Amazon

Introduction

Soil moisture is defined as the amount of water level present in the top layers of the soil that interacts with the atmosphere through evaporation and transpiration.

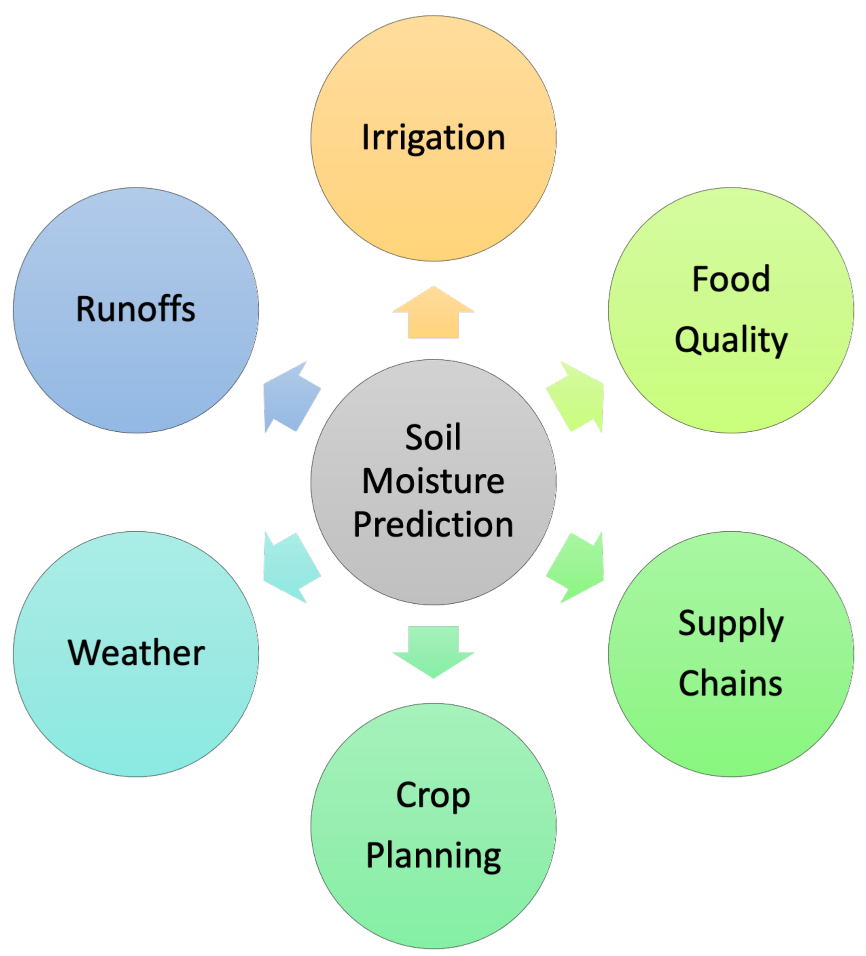

Figure 1:Applications of Soil Moisture Prediction

Even though soil moisture is quite small in amount in a specific region, it significantly affects all kinds of hydrological, biological, and biogeochemical processes, weather patterns, runoffs, and erosion.

Problem Challenges

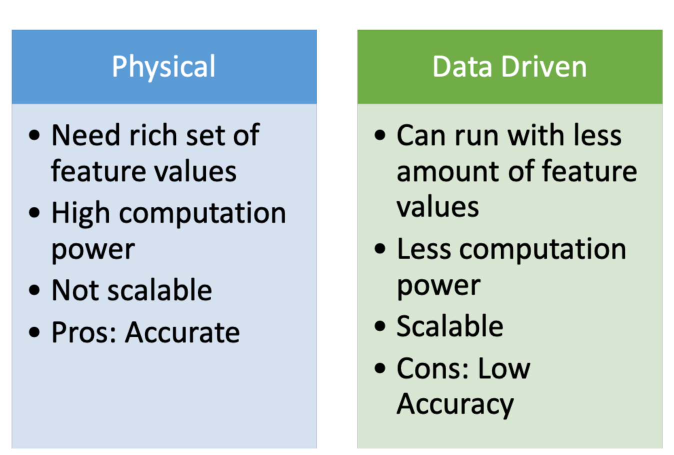

Figure 2:Pros and cons of methods

- Existing data driven approaches inherently assume those soil moisture values to be independent over different locations like simple SVM regression.
- There is spatial dependency in soil moisture values:
  - They depend on weather, crop cycles, soil quality of nearby locations.
  - It is difficult to learn in conventional ML algorithms because of huge data requirements.
  - Not all nearby locations can be correlated in soil moisture.
- Soil moisture at any location changes over time:
  - It can be due to the changes of various features of land and weather.
  - This requires a temporal (or dynamic) model for soil moisture forecasting.
- There is a need for semi-supervised training approach:
  - There are points with available feature values but missing soil moisture values (device and communication failures).
  - Obtaining ground truth soil moisture is expensive.
- There is no ground truth graph structure given as an input for the problem of spatio-temporal data modeling required for soil moisture forecasting.

Problem Statement

- We are given with a set of N locations over a geographic region.
- There are T time steps (weekly / monthly resolution).
- For each time step, ground truth soil moisture (SM) values are given for a subset of locations For each location and time step, there are some input features (such as temperature, relative humidity, precipitation, NDVI, etc.) available.
- Our goal is to predict (forecast) the soil moisture at future time steps such that the forecasting model considers both past soil moisture and input feature values, and also able to learn and exploit the relation between different locations in their soil moisture values.

Our Solution Approach: DGLR

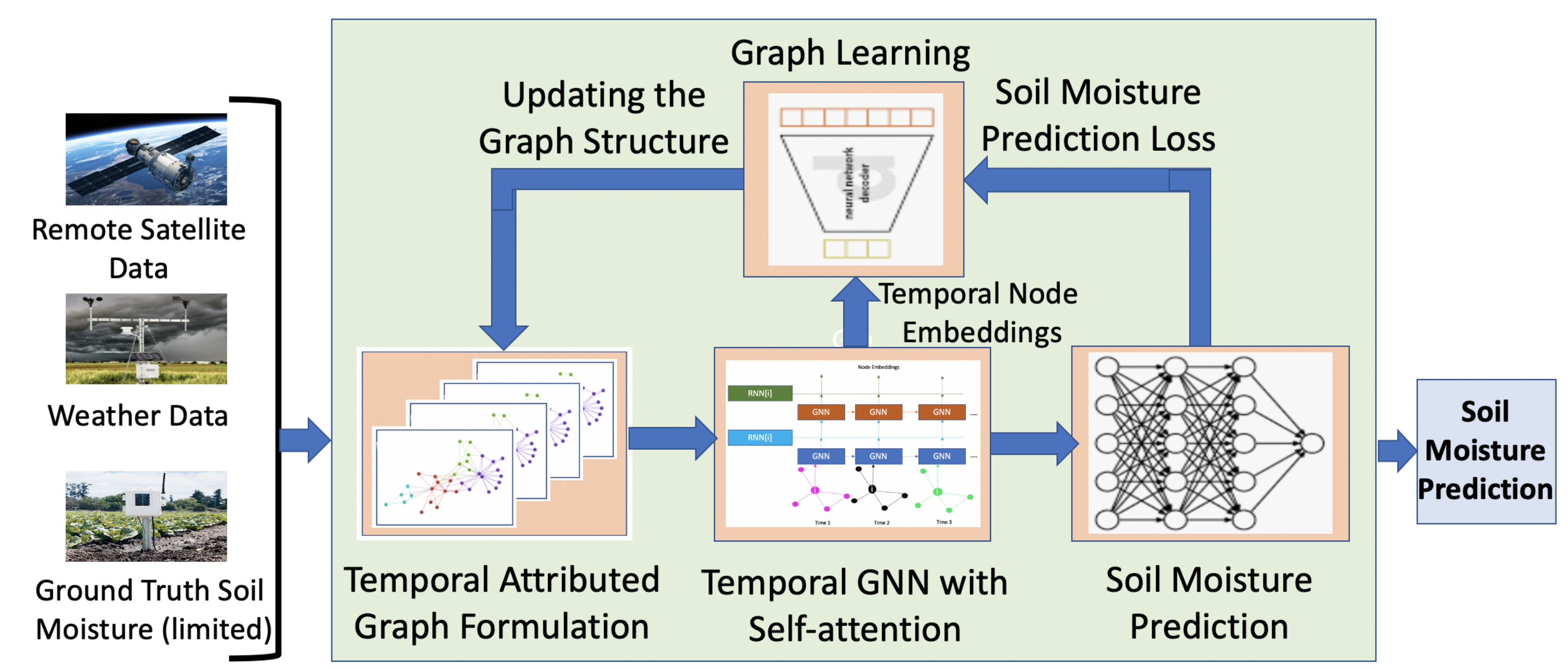

Figure 3:Our Solution Approach: DGLR

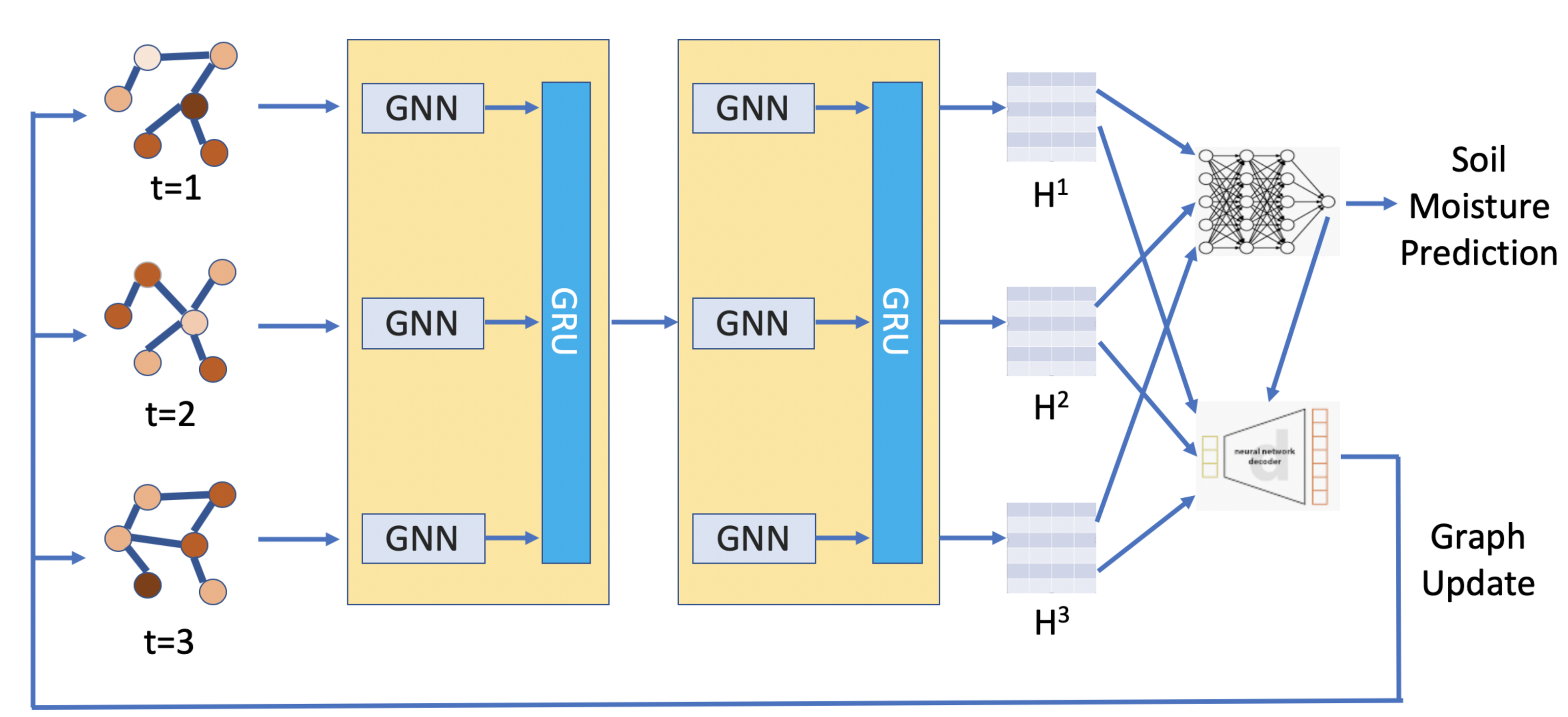

Figure 4:Our Algorithm Architecture: DGLR

Updating the Dynamic Graph Structures

$$\text{Graph Closeness Regularizer: } \mathcal{L}_{GC} = \sum_{t=1}^{T} \text{Distance}(A^t, \hat{A}^t) \tag{1}$$

$$\text{Feature Smoothness Regularizer: } \mathcal{L}_{FS} = \sum_{t=1}^{T} \sum_{\substack{i,j \in [N] \\ i \neq j}} \hat{A}^t_{ij} ||x^t_i - x^t_j||^2_2 \tag{2}$$

$$\text{Target Smoothness Regularizer: } \mathcal{L}_{TS} = \sum_{t=1}^{T} \sum_{\substack{i,j \in N^t_{tr} \\ i \neq j}} \hat{A}^t_{ij} (s^t_i - s^t_j)^2 \tag{3}$$

Experimental Evaluation

Missing	Spain		
SM %	RMSE (↓)	SMAPE % (↓)	Correlation (↑)
0%	0.031 ± 0.001	15.583 ± 1.008	0.764 ± 0.017
10%	0.035 ± 0.004	18.121 ± 1.430	0.750 ± 0.033
20%	0.039 ± 0.002	18.864 ± 1.516	0.738 ± 0.027
30%	0.039 ± 0.006	18.962 ± 1.112	0.661 ± 0.047

Table 1:Test set performance of DGLR with different proportions of missing soil moisture values in the training set.

Distance	Spain		
Threshold (km)	RMSE (↓)	SMAPE % (↓)	Correlation (↑)
8	0.034 ± 0.002	16.418 ± 0.810	0.752 ± 0.042
12	0.032 ± 0.003	16.441 ± 1.106	0.752 ± 0.028
16	0.031 ± 0.001	15.583 ± 1.008	0.764 ± 0.017
20	0.032 ± 0.004	16.556 ± 1.099	0.754 ±0.054

Table 2:Test set performance of DGLR with different distance thresholds in initialising the graph structure.

	Spain			USA		
Algorithms	RMSE (↓)	SMAPE % (↓)	Correlation (↑)	RMSE (↓)	SMAPE % (↓)	Correlation (↑)
SVR-Shared	0.061 ± 0.004	33.654 ± 1.236	0.020 ± 0.001	12.332 ± 0.848	39.013 ± 3.395	0.287 ± 0.031
SVR	0.052 ± 0.005	23.840 ± 2.453	0.083 ± 0.002	10.461 ± 0.718	27.377 ± 3.314	0.354 ± 0.031
Spatial-SVR	0.052 ± 0.007	23.752 ± 1.562	0.150 ± 0.013	10.245 ± 1.567	26.710 ± 4.049	0.351 ± 0.049
ARIMA	0.041 ± 0.008	19.002 ± 0.872	0.010 ± 0.001	9.290 ± 1.110	27.785 ± 1.306	0.159 ± 0.016
RNN-Shared	0.039 ± 0.003	23.443 ± 1.996	0.585 ± 0.048	7.814 ± 0.442	30.880 ± 2.355	0.191 ± 0.041
RNN	0.039 ± 0.003	21.722 ± 1.385	0.529 ± 0.049	7.338 ± 0.346	27.833 ± 1.812	0.273 ± 0.046
DCRNN	0.061 ± 0.003	31.832 ± 2.363	0.588 ± 0.062	8.161 ± 0.762	25.213 ± 1.762	0.393 ± 0.015
EvolveGCN	0.061 ± 0.004	31.789 ± 2.554	0.731 ± 0.022	8.526 ± 0.519	27.717 ± 1.919	0.415 ± 0.038
DGLR (Shared)	0.037 ± 0.003	17.533 ± 1.000	0.752 ± 0.020	7.526 ± 0.519	20.700 ± 1.210	0.493 ± 0.031
DGLR (w/o SL)	0.035 ± 0.003	15.599 ± 0.940	0.753 ± 0.017	6.970 ± 0.356	18.360 ± 0.987	0.517 ± 0.023
DGLR (w/o Sm)	0.033 ± 0.002	19.835 ± 1.005	0.759 ± 0.018	6.721 ± 0.176	17.991 ± 1.211	0.533 ± 0.013
DGLR (full model)	**0.031 ± 0.001**	**15.583 ± 1.008**	**0.764 ± 0.017**	**6.454 ± 0.111**	**17.002 ± 0.987**	**0.566 ± 0.008**

Table 3:Performance of soil moisture forecasting in test interval of different datasets.

Contact anoushkav@vt.edu and samb.bandyo@gmail.com for any questions.

Link to code: `https://github.com/AnoushkaVyas/DGLR`