SOIL MOISTURE RETRIEVAL USING THE GROUND WATER DATASET USING MACHINE LEARNING

K Rambabu[1] , Adapa Saibabu[2]

1. Assistant Professor (HOD) MCA, DEPT, Dantuluri Narayana Raju College, Bhimavaram, Andharapradesh

Email id: - kattarambabudnr@gmail.com

2. PG Student of MCA, Dantuluri Narayana Raju College , Bhimavaram, Andhara pradesh

Email id: - adapasaibabu23@gmail.com

ABSTRACT

In this research, a data-driven modeling approach using machine learning algorithms was compared to artificial neural networks (ANNs) for predicting transient groundwater levels in a complex groundwater system under variable pumping and weather conditions. Various prediction horizons were considered, including daily, weekly, biweekly, monthly, and bimonthly timeframes. It was discovered that, overall, both approaches exhibited comparable modeling performance in terms of prediction accuracy and generalization. However, ANN demonstrated superior performance, especially for longer prediction horizons, where limited data events were available for model development. This suggests that ANN has the potential to serve as a valuable and practical tool, particularly in cases with a scarcity of measured data for future predictions. Additionally, the study revealed a high level of consistency between the training and testing phases of modeling when employing ANN, as opposed to the machine learning algorithm.

1. INTRODUCTION

Water below the land surface appears in two zones - saturated and the unsaturated zone. When rainfall occurs, a part of it infiltrates into the ground. Some amount of this infiltrated rain is held up by the upper layer of soil in its pore spaces. This layer is immediately below the land surface and contains both air and water and is known as the unsaturated zone. When all the soil pores are completely filled with water, then water seeps further down through the fractures in the rock. After a certain depth all pores in the soil are completely filled with water, this part forms the saturated zone. The top of saturated zone is known as the water table and water in this zone is called the groundwater.

2. LITERATURE SURVEY AND RELATED WORK

1. **Title: "Soil Moisture Retrieval Using Remote Sensing and Machine Learning Techniques"**

- **Authors: Dorigo, W., Wagner, W., Hohensinn, R., et al.**

- **Publication Year: 2010**

- **Summary: This paper discusses the use of satellite remote sensing data and machine learning algorithms for soil moisture retrieval. It provides a good introduction to the topic and the challenges associated with it.**

2. **Title: "Estimation of Soil Moisture from Remote Sensing and Ground-Based Observations Using Data Assimilation with a Nonlinear Filter"**

- **Authors: Reichle, R. H., & Koster, R. D.**

- **Publication Year: 2004**

- **Summary: This paper explores the assimilation of remote sensing and ground-based data for soil moisture estimation. It introduces the concept of data assimilation, which can be relevant for your project.**

3. **Title: "A Review of Soil Moisture Measurement"**

- **Authors: Robinson, D. A., Campbell, C. S., & Hopmans, J. W.**

- **Publication Year: 2008**

- **Summary: While not specific to machine learning, this review paper provides an excellent overview of various soil moisture measurement techniques, which can be useful for selecting the appropriate dataset for your project.**

4. **Title: "Soil Moisture Estimation through Satellite Remote Sensing: A Machine Learning Perspective"**

- **Authors: Crow, W. T., & van der Schalie, R.**

- **Publication Year: 2015**

- **Summary: This paper focuses on the application of machine learning techniques, such as neural networks and support vector machines, for soil moisture estimation from satellite remote sensing data.**

3. EXISTING SYSTEM

Groundwater level is an indicator of groundwater availability, groundwater flow, and the physical characteristics of an aquifer or groundwater system. Due to increased population and decreased groundwater recharge, the demand increases and it may not be feasible to check the draft of groundwater resources. The only available option is to increase the recharge rate to the aquifer by suitable means. Therefore it is necessary to quantify the present rate of groundwater recharge, monitor the change in water table depth and then predict the future trend of water table depth before any intervention.

Disadvantage:

- Any phenomenon, which produces pressure change within an aquifer, results into the change of ground water level.
- These changes in ground water level can be a result of changes in storage, amount of discharge and recharge, variation of stream stages and evaporation.

4. PROPOSED SYSTEM

This is mainly in the form of estimation of the magnitude of a hydrological parameters. The factors that influence and control the groundwater level fluctuation were determined to develop a forecasting model and examine its potential in predicting groundwater level. Models for prediction of water table depth were developed based on Artificial Neural Networks (ANN) with different combinations of hydrological parameters. The best combination was confirmed with factor analysis. The input parameters for groundwater level forecasting were derived using Time Series Analysis (TSA).

Advantage:

- Most of the researches used ANN alone to predict groundwater level.
- But the present study incorporated factor analysis along with time series forecasting to increase the accuracy and usefulness of prediction.

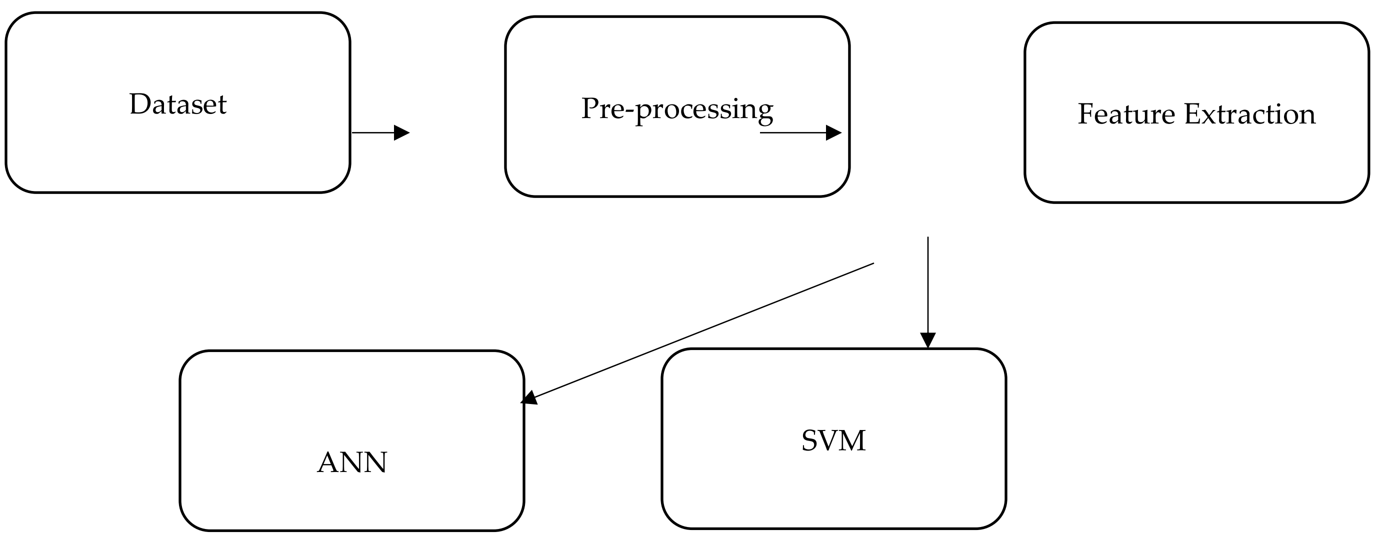


FIG 1 - System Architecture

5. METHODOLOGIES

MODULE

Field survey:

Field survey was carried out to establish the observation well locations suitable for the study area. The wells were selected in such a way that areas of different elevations are suitably covered. The spatial locations were identified by conducting GPS (Global positioning system) survey. The groundwater level was recorded periodically.

Journal of Engineering Sciences

Factor analysis:

In factor analysis the correlation between input parameters Potential evapotranspiration (PET), temperature, humidity and rainfall were analysed using Statistical Package for Social Sciences (SPSS) for monsoon and non-monsoon season. Any factor having component value less than 0.5 was extracted as it is less significant for the input combination.

Time series analysis (TSA):

In this phase the input parameters required for the prediction of groundwater level were forecasted. The values were forecasted based on previously observed data. In this study, time series analysis based on moving average method was adopted.

Prediction using ANN:

ANN is an information processing paradigm inspired by biological nervous systems, such as our brain. It consists of large number of highly interconnected processing elements, called neurons, working together. An ANN consists of input, hidden and output layers as shown in Fig 1 and each layer includes an array of processing elements. A Neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weights, and the activation function. The learning, training, performance and transfer functions used in this study are LEARNGDM, TRAINSCG, MSE AND TRANSIG respectively.

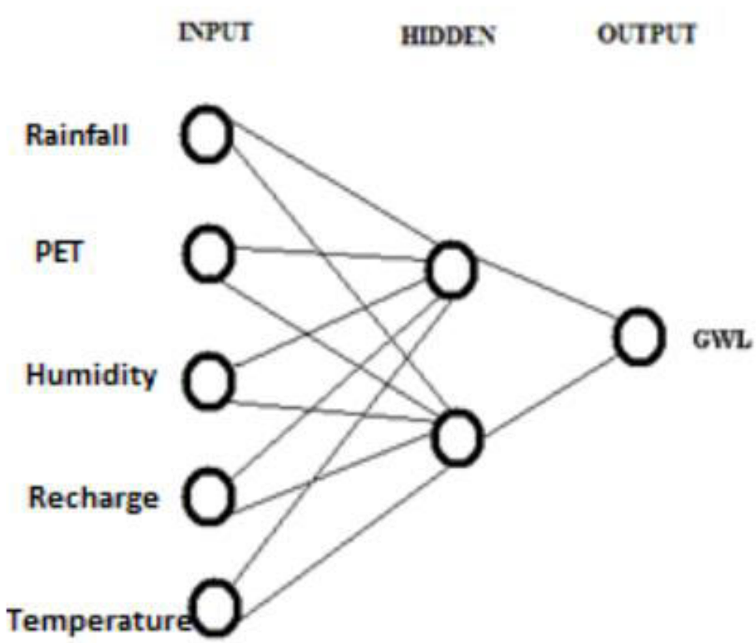


Fig 1 Prediction Layer

 Journal of Engineering Sciences

Modules:

1. Field survey
2. Factor analysis
3. Time series analysis (TSA)
4. Prediction using ANN

Field survey:

◎ Field survey was carried out to establish the observation well locations suitable for the study area. The wells were selected in such a way that areas of different elevations are suitably covered.

◎ The spatial locations were identified by conducting GPS (Global positioning system) survey.

◎ The groundwater level was recorded periodically

Factor analysis:

◎ In factor analysis the correlation between input parameters Potential evapotranspiration (PET), temperature, humidity and rainfall were analysed using Statistical Package for Social Sciences (SPSS) for monsoon and non-monsoon season.

◎ Any factor having component value less than 0.5 was extracted as it is less significant for the input combination.

Time series analysis (TSA):

◎ In this phase the input parameters required for the prediction of groundwater level were forecasted. The values were forecasted based on previously observed data.

◎ In this study, time series analysis based on moving average method was adopted.

Prediction using ANN:

◎ ANN is an information processing paradigm inspired by biological nervous systems, such as our brain.

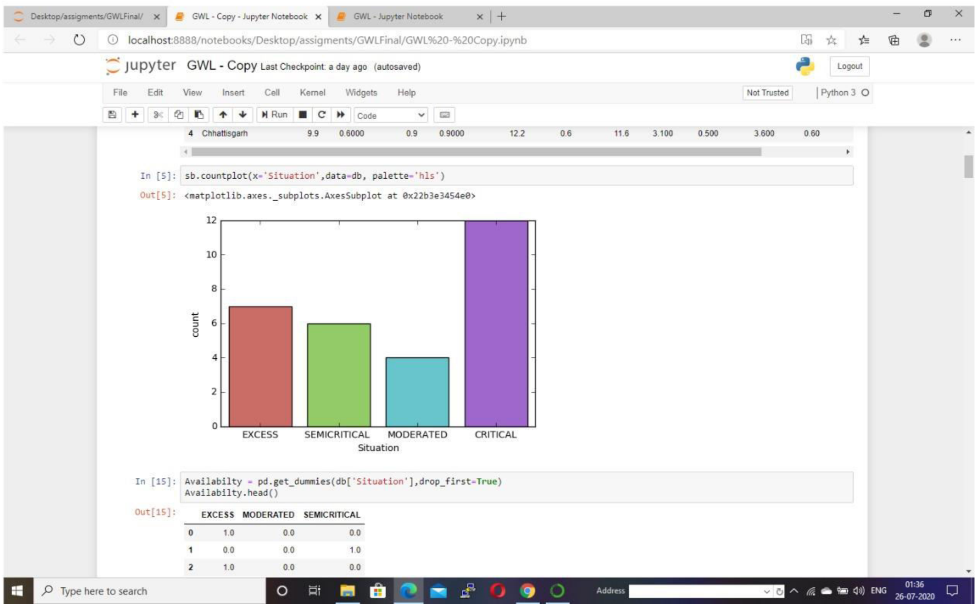
◎ It consists of large number of highly interconnected processing elements, called neurons, working together.

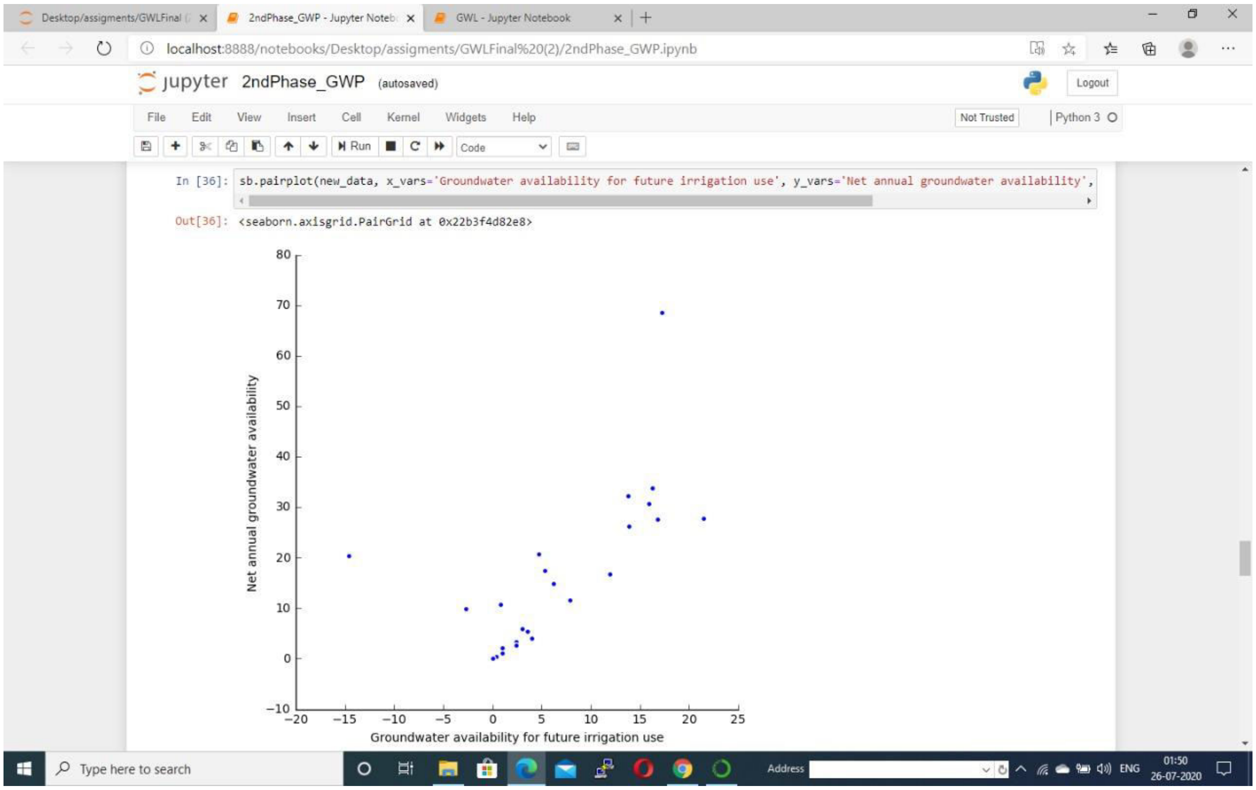
◎ An ANN consists of input, hidden and output layers as shown in Fig 1 and each layer includes an array of processing elements.

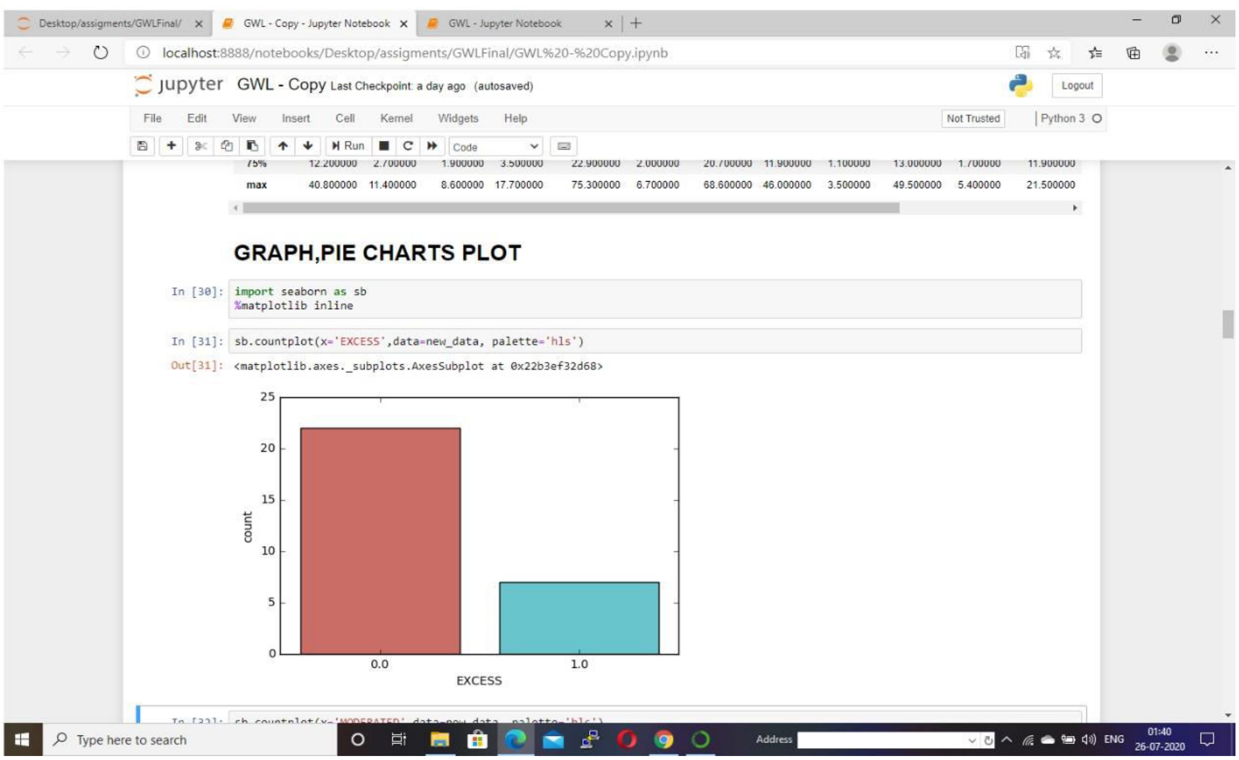
◎ A Neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weights, and the activation function.

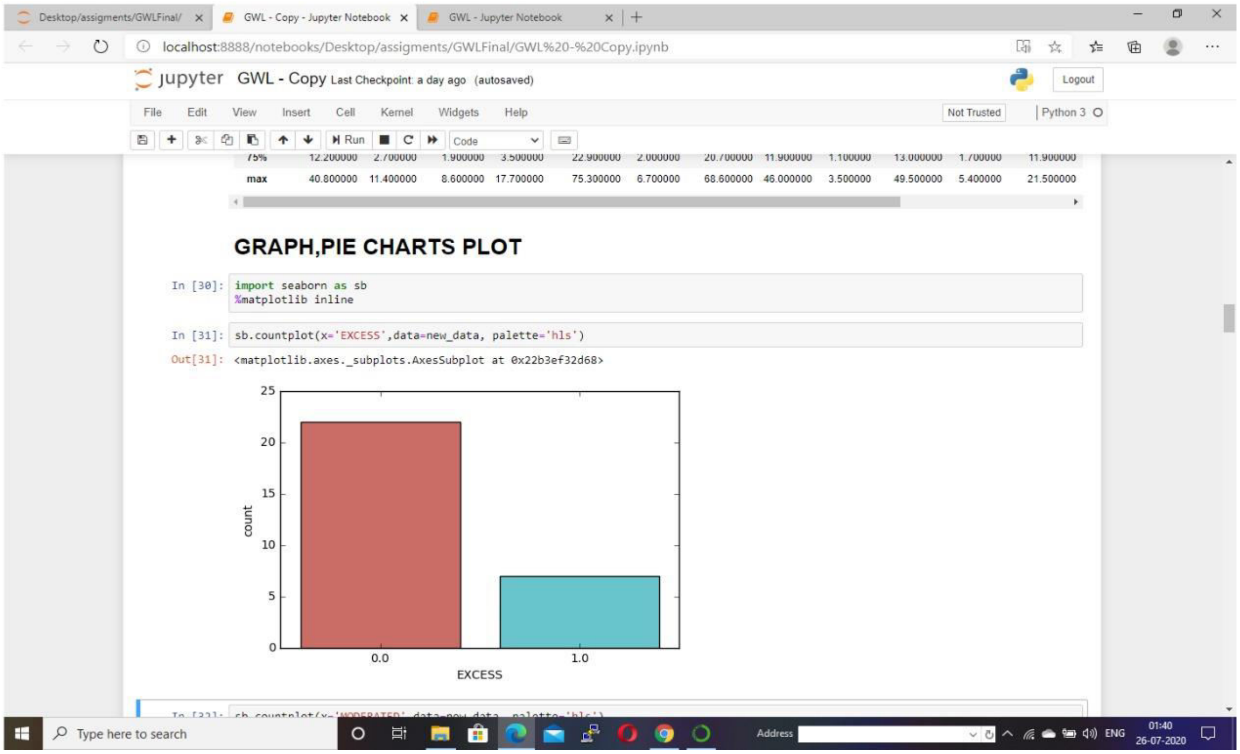
6. RESULTS AND DISCUSSION SCREEN SHOTS

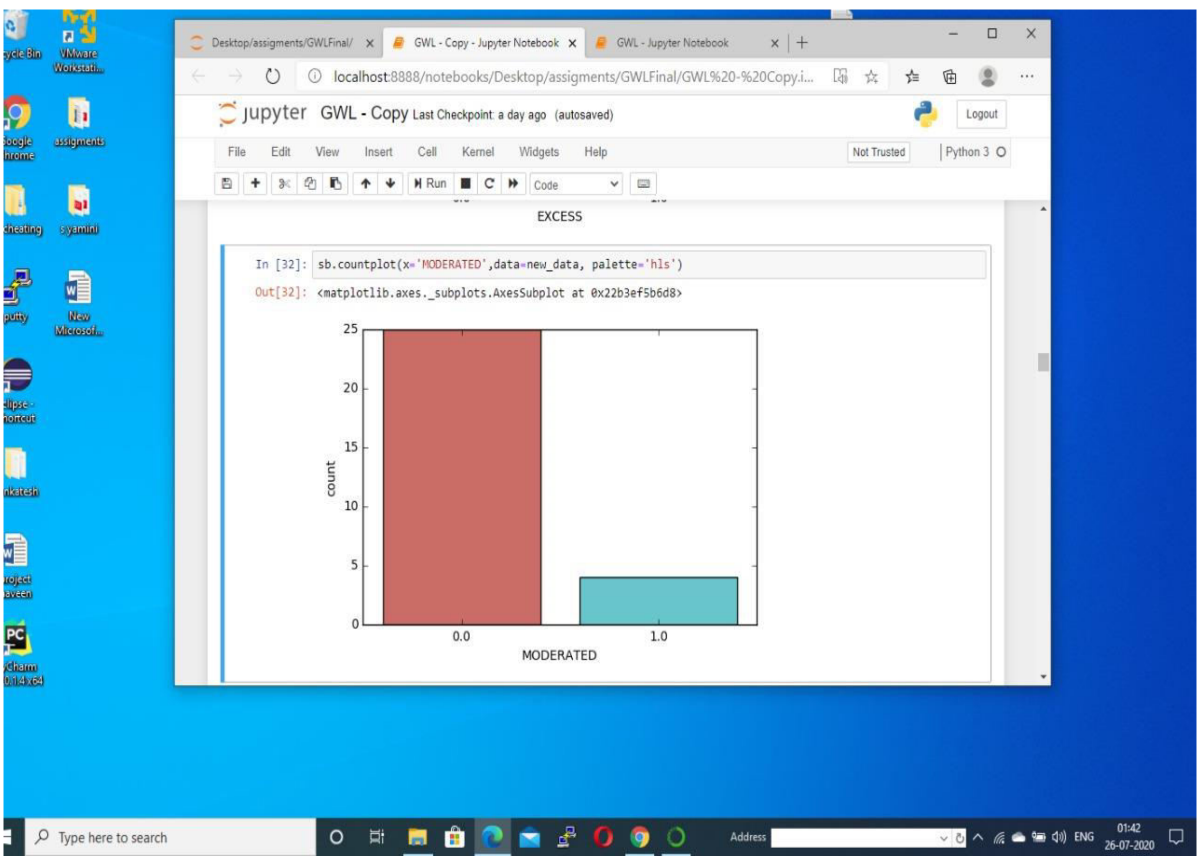
SCREENSHOTS:

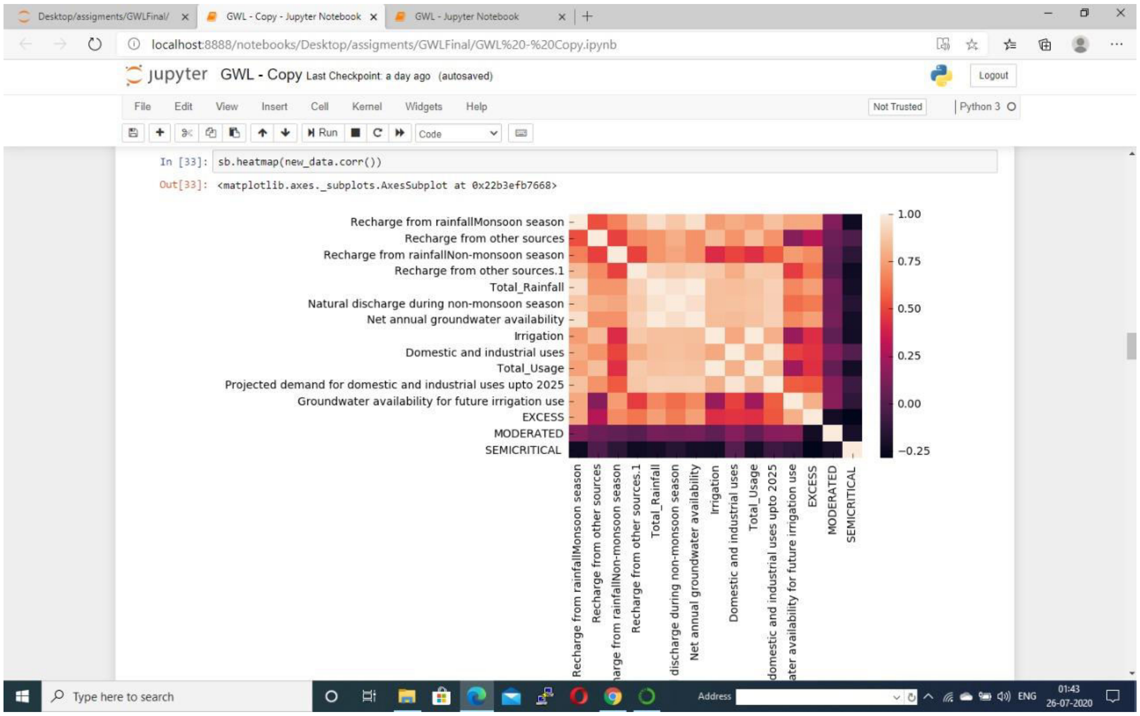



localhost:8888/notebooks/Desktop/assigments/GWLFinal%20(2)/2ndPhase_GWP.ipynb
jupyter 2ndPhase_GWP (autosaved)
In [36]: sb.pairplot(new_data, x_vars='Groundwater availability for future irrigation use', y_vars='Net annual groundwater availability',
Out[36]: <seaborn.axisgrid.PairGrid at 0x22b3f4d82e8>
Net annual groundwater availability
Groundwater availability for future irrigation use

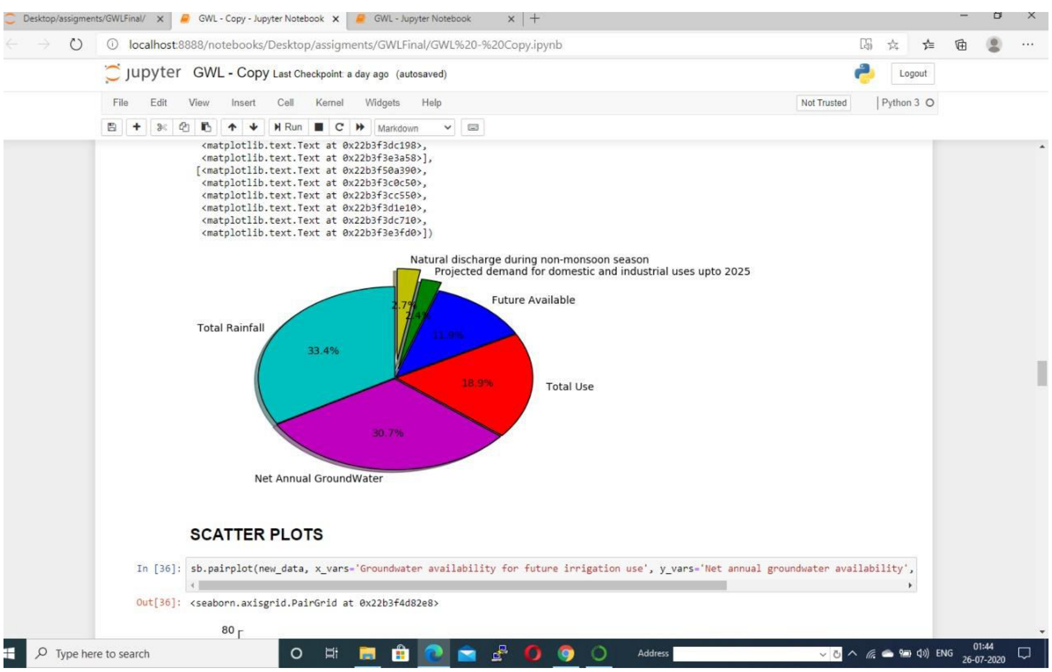

localhost:8888/notebooks/Desktop/assigments/GWLFinal/GWL%20-%20Copy.ipynb
jupyter GWL - Copy Last Checkpoint: a day ago (autosaved)
max 40.800000 11.400000 8.600000 17.700000 75.300000 6.700000 68.600000 46.000000 3.500000 49.500000 5.400000 21.500000
GRAPH,PIE CHARTS PLOT
In [30]: import seaborn as sb
%matplotlib inline
In [31]: sb.countplot(x='EXCESS',data=new_data, palette='hls')
Out[31]: <matplotlib.axes._subplots.AxesSubplot at 0x22b3ef32d68>
count
EXCESS


GRAPH,PIE CHARTS PLOT
In [30]: import seaborn as sb
%matplotlib inline
In [31]: sb.countplot(x='EXCESS',data=new_data, palette='hls')
Out[31]: <matplotlib.axes._subplots.AxesSubplot at 0x22b3ef32d68>
count
EXCESS
0.0
1.0

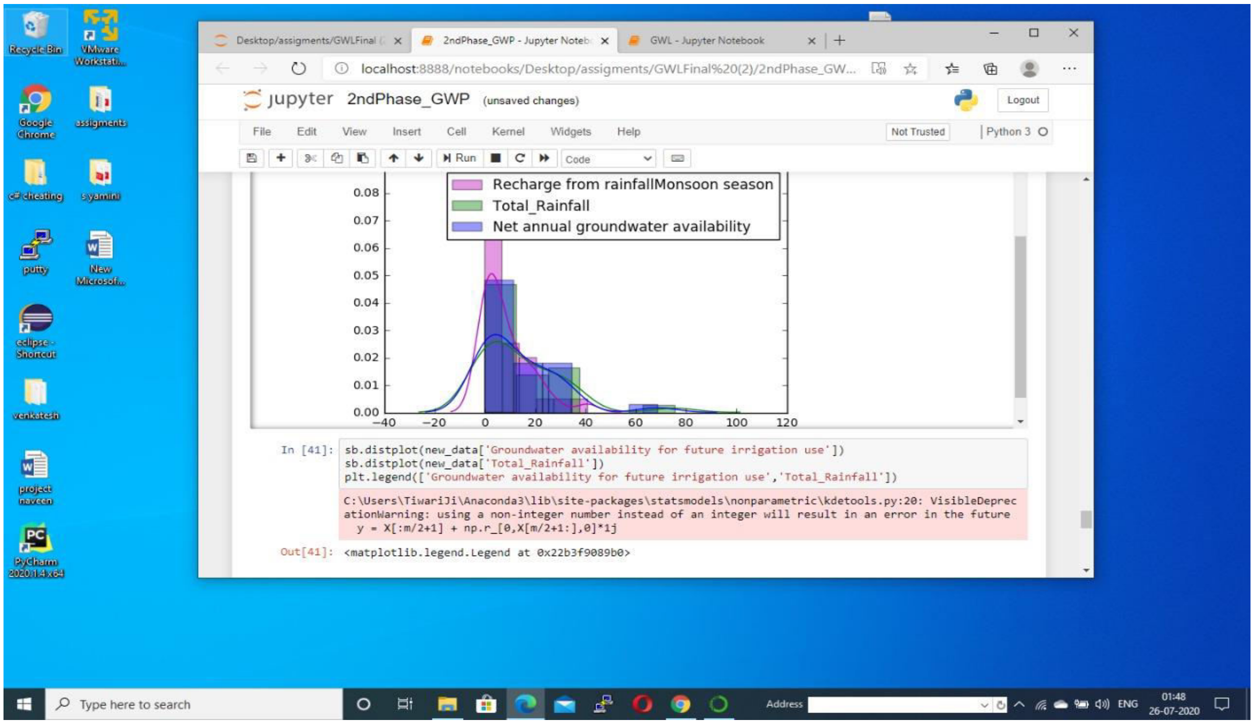

In [32]: sb.countplot(x='MODERATED',data=new_data, palette='hls')
Out[32]: <matplotlib.axes._subplots.AxesSubplot at 0x22b3ef5b6d8>
count
MODERATED
0.0
1.0

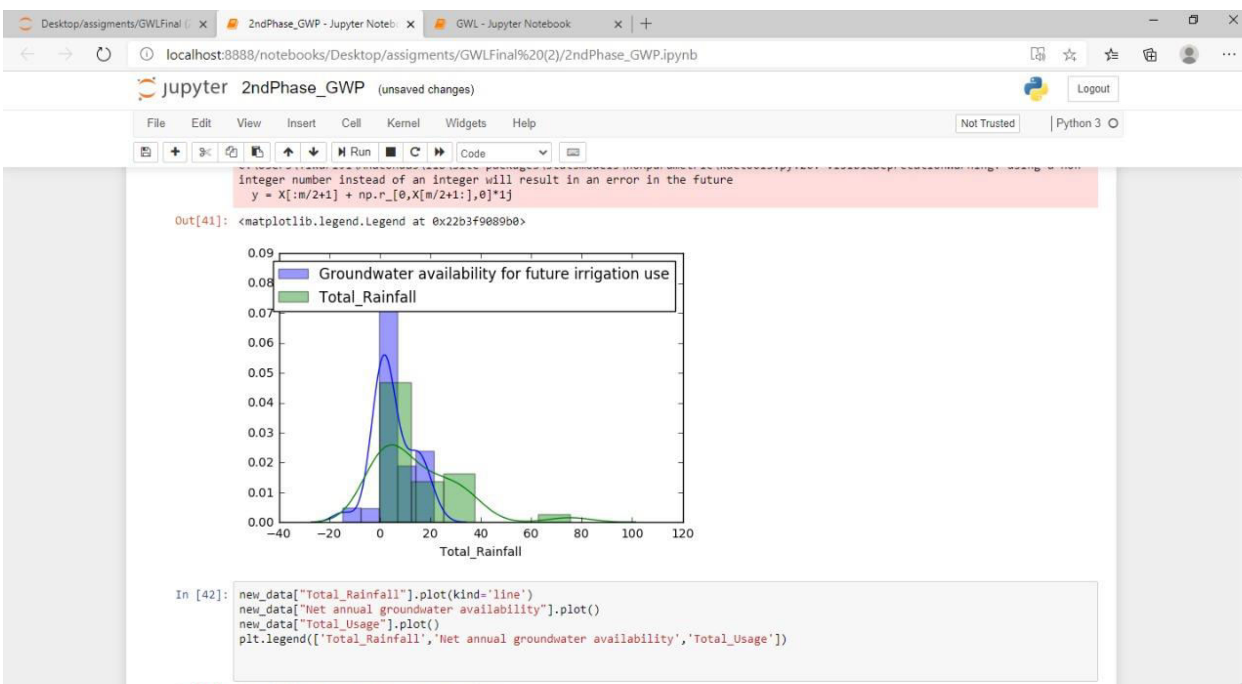


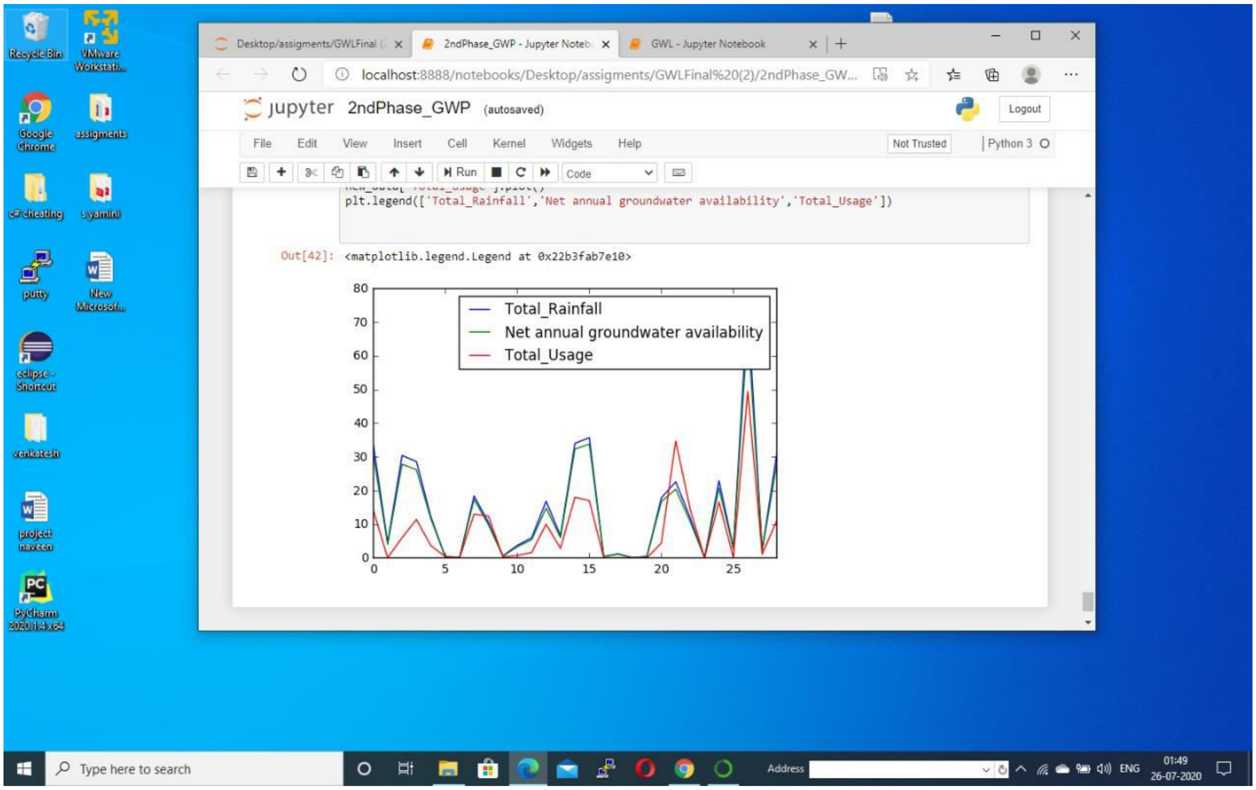
In [33]: sb.heatmap(new_data.corr())
Out[33]: <matplotlib.axes._subplots.AxesSubplot at 0x22b3efb7668>
Recharge from rainfallMonsoon season
Recharge from other sources
Recharge from rainfallNon-monsoon season
Recharge from other sources.1
Total_Rainfall
Natural discharge during non-monsoon season
Net annual groundwater availability
Irrigation
Domestic and industrial uses
Total_Usage
Projected demand for domestic and industrial uses upto 2025
Groundwater availability for future irrigation use
EXCESS
MODERATED
SEMICRITICAL
1.00
0.75
0.50
0.25
0.00
−0.25

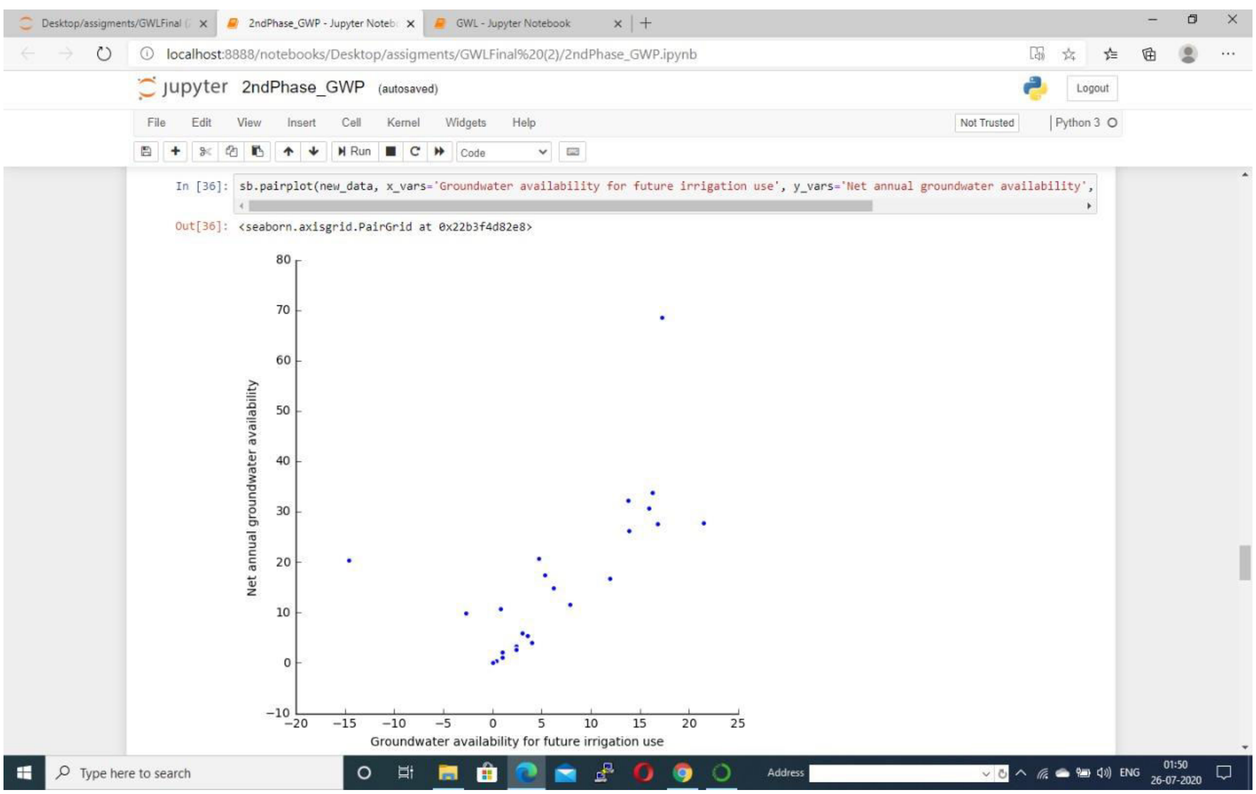


Natural discharge during non-monsoon season
Projected demand for domestic and industrial uses upto 2025
Future Available
Total Rainfall
Total Use
Net Annual GroundWater
33.4%
30.7%
18.9%
SCATTER PLOTS
In [36]: sb.pairplot(new_data, x_vars='Groundwater availability for future irrigation use', y_vars='Net annual groundwater availability',
Out[36]: <seaborn.axisgrid.PairGrid at 0x22b3f4d82e8>


Recharge from rainfallMonsoon season
Total_Rainfall
Net annual groundwater availability
In [41]: sb.distplot(new_data['Groundwater availability for future irrigation use'])
sb.distplot(new_data['Total_Rainfall'])
plt.legend(['Groundwater availability for future irrigation use','Total_Rainfall'])
C:\Users\TiwariJi\Anaconda3\lib\site-packages\statsmodels\nonparametric\kdetools.py:20: VisibleDeprecationWarning: using a non-integer number instead of an integer will result in an error in the future
y = X[:m/2+1] + np.r_[0,X[m/2+1:],0]*1j
Out[41]: <matplotlib.legend.Legend at 0x22b3f9089b0>


Out[41]: <matplotlib.legend.Legend at 0x22b3f9089b0>
Groundwater availability for future irrigation use
Total_Rainfall
Total_Rainfall
In [42]: new_data["Total_Rainfall"].plot(kind='line')
new_data["Net annual groundwater availability"].plot()
new_data["Total_Usage"].plot()
plt.legend(['Total_Rainfall','Net annual groundwater availability','Total_Usage'])


jupyter 2ndPhase_GWP (autosaved)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3
plt.legend(['Total_Rainfall','Net annual groundwater availability','Total_Usage'])
Out[42]: <matplotlib.legend.Legend at 0x22b3fab7e10>
Total_Rainfall
Net annual groundwater availability
Total_Usage
Type here to search
01:49
26-07-2020


localhost:8888/notebooks/Desktop/assigments/GWLFinal%20(2)/2ndPhase_GWP.ipynb
jupyter 2ndPhase_GWP (autosaved)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3
In [36]: sb.pairplot(new_data, x_vars='Groundwater availability for future irrigation use', y_vars='Net annual groundwater availability',
Out[36]: <seaborn.axisgrid.PairGrid at 0x22b3f4d82e8>
Net annual groundwater availability
Groundwater availability for future irrigation use
Type here to search
01:50
26-07-2020

Journal of Engineering Sciences

7. CONCLUSION AND FUTURE SCOPE

This paper introduces various machine learning algorithm before which we have collected weather information of both monsoon and non-monsoon then checked soil parameter after that for predicting transient groundwater levels in a complex groundwater system under variable pumping and weather conditions. Various prediction horizons were used, including daily, weekly, biweekly, monthly, and bimonthly prediction horizons. It was found that even though modelling performance (in terms of prediction accuracy and generalization) for both approaches was generally comparable.

8. REFERENCES

1.D. R. Legates, R. Mahmood, D. F. Levia, T. L. DeLiberty, S. M. Quiring, C. Houser, and F. E. Nelson, "Soil moisture: A central and unifying theme in physical geography," Progress in Physical Geography, vol. 35, no. 1, pp. 65–86, 2011.

[2] Y. H. Kerr, P. Waldteufel, J.-P. Wigneron, J. Martinuzzi, J. Font, and M. Berger, "Soil moisture retrieval from space: The soil moisture and ocean salinity (smos) mission," IEEE transactions on Geoscience and remote sensing, vol. 39, no. 8, pp. 1729–1735, 2001.

[3] E. G. Njoku, T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem, "Soil moisture retrieval from amsr-e," IEEE transactions on Geoscience and remote sensing, vol. 41, no. 2, pp. 215–229, 2003.

[4] R. H. Reichle, R. D. Koster, J. Dong, and A. A. Berg, "Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation," Journal of Hydrometeorology, vol. 5, no. 3, pp. 430–442, 2004.

[5] S. Lambot, E. C. Slob, I. van den Bosch, B. Stockbroeckx, and M. Vanclooster, "Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties," IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 11, pp. 2555–2568, 2004.

[6] M. S. Dawson, A. K. Fung, and M. T. Manry, "A robust statisticalbased estimator for soil moisture retrieval from radar measurements," IEEE transactions on geoscience and remote sensing, vol. 35, no. 1, pp. 57–67, 1997.

.