**Using Deep Learning and Machine Learning models to improve the accuracy of Evapotranspiration Estimation using Penman-Monteith Equation**

**Abstraction**

Evapotranspiration is important component of hydrological cycle which includes combined loss of water through soil surface and transpiration from plant leaf.  
For better understanding and accuracy ET calculation it is the most critical component for effective water resource management, climate, agriculture needs.  
It has a great impact on crop production, irrigation planning and the direction of overall health ecosystem in the region. Looking at the global requirement of the water due to rapid population growth, climate change and urbanization ET calculation is essential for   
steady growth and environment protection. There are serval methods for ET calculation Penman-Monteith equation is one the most effective and scientific approach for it.  
It uses serval other metrological data wind speed, relative humidity, radiation, temp etc. It provides a sound comprehensive framework for ET calculation across a range. The rapid growth machine learning and deep learning has shown new avenue for complex, data driven discipline. The benefit of ML and DL for the estimation of ET is their ability to deal with different data sets and diverse outputs it can use satellite image, satellite based remote sensing data (NDVI and EVI) by working on these heterogenous data we can improve and enhance our understanding regarding evapotranspiration.

**Introduction**

The most crucial component is Surface Conduction (GS) which quantifies the movement of water vapor and carbon dioxide between vegetation and the atmosphere [1]. It significantly affects transpiration, photosynthesis, and energy flows, in addition to other ecological, hydrological, and atmospheric processes. Enhancing Gs accuracy becomes crucial since simulations of the global carbon cycles as well as the impacts that climate change has on land ecosystems would be resolved with time, thereby facilitating an improved understanding of the relationship between plants and the environment.

In the past, we have had to depend on empirical linkages and detailed field approaches to estimate Gs, however, as has become almost customary with methods, these have their informalities; they require a high-quality input data routine, a narrow time and area coverage, and assimilations that rarely capture the employment of many interconnected environments, even though they can sometimes be useful in a particular region.

In addition, the links that exist between atmospheric processes and plant physiology in physical models might not fully integrate the dynamic and lateral processes, hence, there is an urgent need for effective, future-oriented explanatory techniques that will allow for optimal utilization of the availability and variety of environmental data. Machine Learning has a wide range of capability that allows it perform complex functions which include nonlinear and dynamic integrations [5]. This makes it an ideal fit for solving biological problems in an ecosystem as it has intuitive aptitude to go through simple and complex databases and make sense out of them [6,7]. As far as Gs estimation is concerned, ML provides opportunities that allow it to handle the challenges that traditional approaches have been unable to handle.

previous study demonstrates the effectiveness of machine learning in ecological and environmental modelling [10]. ANNs or Artificial neural networks have been widely used because of their adaptability in approximating nonlinear functions, however, tree-based methods like CatBoost have also proven to be successful at processing categorical and numerical data with little preparation. To get more enhanced predictive accuracy we can use Ensemble techniques, which combine predictions from many models, by leveraging the benefits of individual algorithms. In both combining the meteorological elements with the remote sensing data and the meteorological factors, utilizing machine learning (ML) for Gs prediction remains a learning frontier. Through this research, we try to fill this void by developing a simple and robust machine learning based model to predict Gs using ensemble learning and comparing the predictions of an ANN, Cat Boost and a hybrid stacking ensemble model.

**Methodology**

**Dataset**:

The data for this experiment was recorded in Klinberg Germany between January 1 2004 and December 31 2014. The dataset contains both **meteorological** and **remote sensing** data, which are used to estimate **stomatal conductance (Gs)** that is the prime component needed to calculate **evapotranspiration (λE)** [1,2] using Prime-Medilyn equation.

Variable Selection:

The variables used in this analysis have been selected on the basis of their known relationship with previously conducted research work on stomatal conductance and evapotranspiration. As meteorological variables, temperature, precipitation, and solar radiation have also been associated with the energy and water balance processes that regulate stomatal behavior. Gross Primary Production or GPP is treated as one of the best choosing proxies of photosynthetic processes which in turn are also related to the processes involving the stomata. The Vapor Pressure Deficit VPD in this model is used as an index of atmospheric moisture stress which regulates the degree of opening and closing of the stomata. NDVI and NIRv, which are examples of remote sensing variables, were selected because they can evaluate vegetation greenness and photosynthetic activities of crops across large areas, thus, giving important spatial information. The integration of these two cases of meteorological and remote sensing data enables to estimate stomatal conductance in an integrated way.

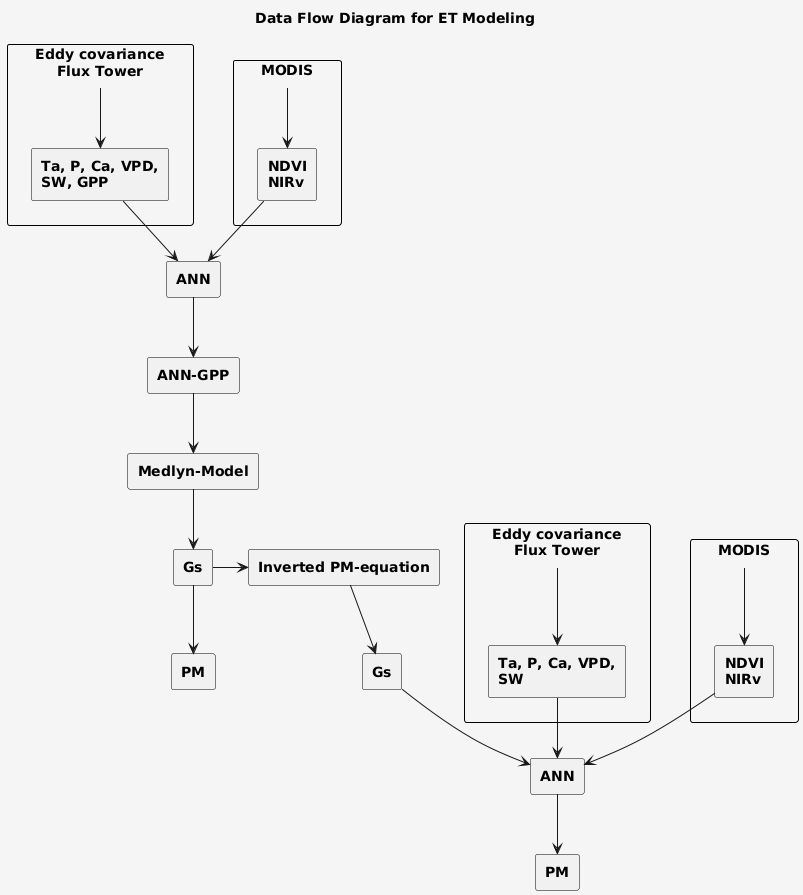


Figure1: P is precipitation, Ta is temperature, GPP is gross primary production, VPD is vapor pressure deficit, SW is solar radiation, Ca is carbon dioxide concentration, NDVI is normalized difference vegetation index, NIRv is near-infrared reflectance of vegetation

The meteorological data is obtained from FLUXNET (<https://fluxnet.org/>) while remote sensing features has to be calculated using the values of band 1 and band 2 from Modis data MOD09A1 (<https://modis.ornl.gov/>).

NDVI and NIRv are calculated using the following formulas:

where :

r1​: red band (visible light).

r2​: near-infrared (NIR) band [8].

The target class is Gs (stomatal conductance) which works as the input for Prime-Medilyn equation which is used to calculate λE (evapotranspiration):

where Rn is net radiation, G is soil heat flux, ∆ is the gradient of the saturation vapor pressure versus atmospheric temperature, ρ is air density, Cp is the specific heat at constant pressure of air, D is the vapor pressure deficit of the air, Ga is the aerodynamic conductance, and γ is the psychometric constant.

To calculate Gs, the following equation is used:

where Ca is CO2 concentration of the air, g1 and g0 are undetermined coefficients derived from regression analysis, GPP is gross primary production and D is the vapor pressure deficit of the air [4].

**Model Training:**

The ANN model is generally repeatedly trained to reduce overfitting and generate better results. Hence, we need to find the best structure of layers for our model for which we will try different number of neurons in layers and use the most optimised one based on our AIC score which is calculated using the formula:

The **Akaike Information Criterion (AIC)** is used to identify the best model by selecting the one with the **lowest AIC score**. This approach ensures that the selected model achieves the best **trade-off between goodness of fit** and **model complexity**, thus preventing overfitting. We also try and use a RNN model using similar approach to achieve better results.

In addition to deep neural networks, other machine learning models are explored, such as **XGBoost** and **CatBoost [9]** after which an **ensemble model** combining the best **Sequential** model and **CatBoost** model is also evaluated to further enhance performance [6]. To evaluate our model, Mean Square Error (MSE), Root Mean Square Error RMSE and R2 Score are used:

**Results Comparison and interpretation**

| **Model** | **R² Score** | **MSE** | **MAE** |
| --- | --- | --- | --- |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CatBoost** | 0.9217 | 0.1723 | 0.2122 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Stacked Ensemble** | 0.9338 | 0.1531 |  |  | 0.2021 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ANN** | 0.8881 | 0.2575 | 0.2657 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **RNN** | 0.8929 | 0.2477 | 0.2370 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **XGBoost** | 0.8415 | 0.3667 | 0.3889 |  |
|  |  |  |  |  |

**Artificial Neural Network(ANN)**

A graph of loss of a person

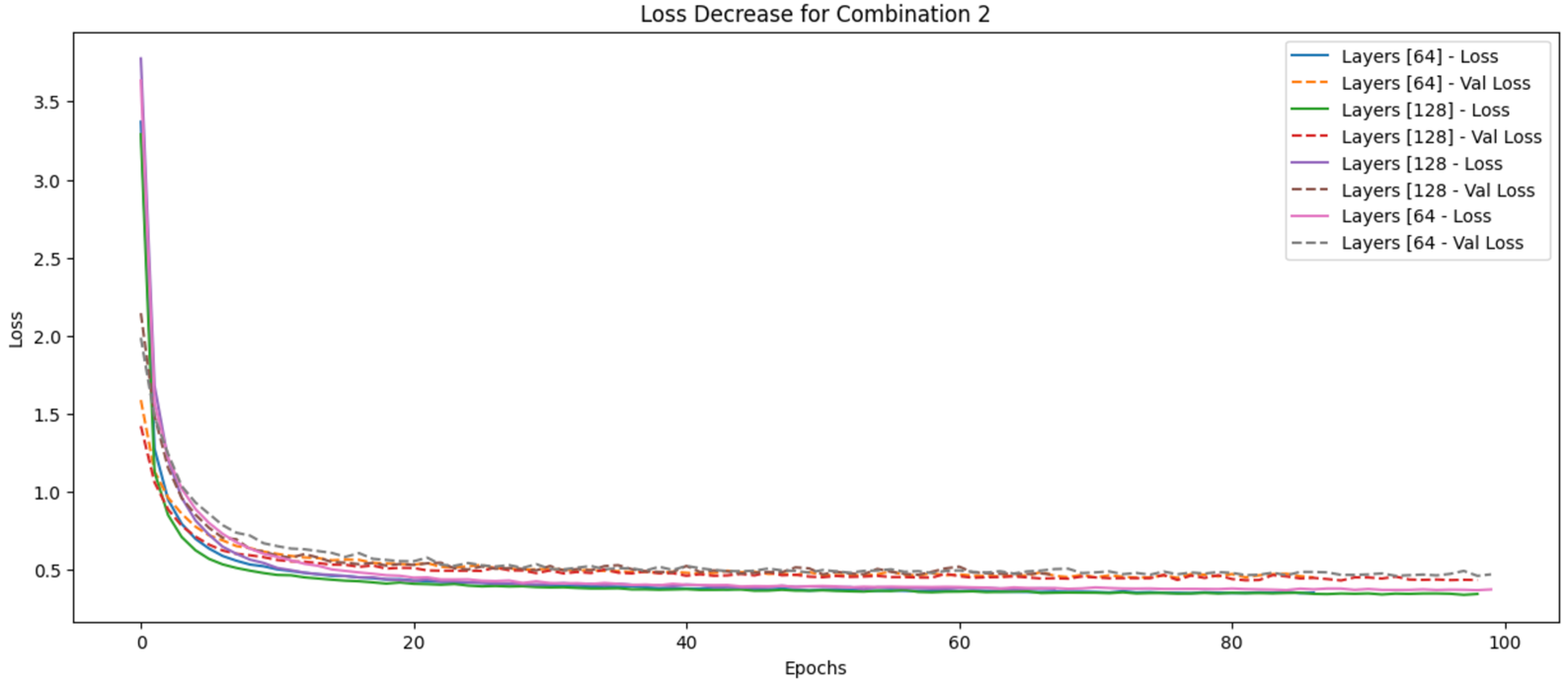
Description automatically generated with medium confidence

(The above plot shows decrease over epoch for different configurations of layer Configuration that have more layers[128] reflects similar loss trends which suggest comparable learning effect)

A graph of a graph

Description automatically generated with medium confidence

(R² score increases linearly over all configurations models having more layers may learn patterns effectively and lead to be better fitted model)



(Training loss for all configuration decreases significantly in first 20 epoch which project model is improving its prediction)

A graph of a graph

Description automatically generated with medium confidence

(Model with 128 layers have higher R² score which suggest better performance , which may risk for overfitting)

A graph showing loss of a patient

Description automatically generated with medium confidence

(Graph reflects overfitting ,validation loss increases while training loss continues decreases significantly in first 20 epoch which project model may improves its prediction)

A graph of a line

Description automatically generated with medium confidence

(Model with 128 layers have higher R² score which suggest better performance as compared to model with 64 layers, which may risk for overfitting)

The best results for the ANN model were achieved using combination 2 with [64] layer model providing a R² score of 0.8881 while providing a low AIC value of 0.4947, barely above the lowest of 0.4763 while increasing the R² by over 4%.

**Recurrent Neural Networks (RNN)**

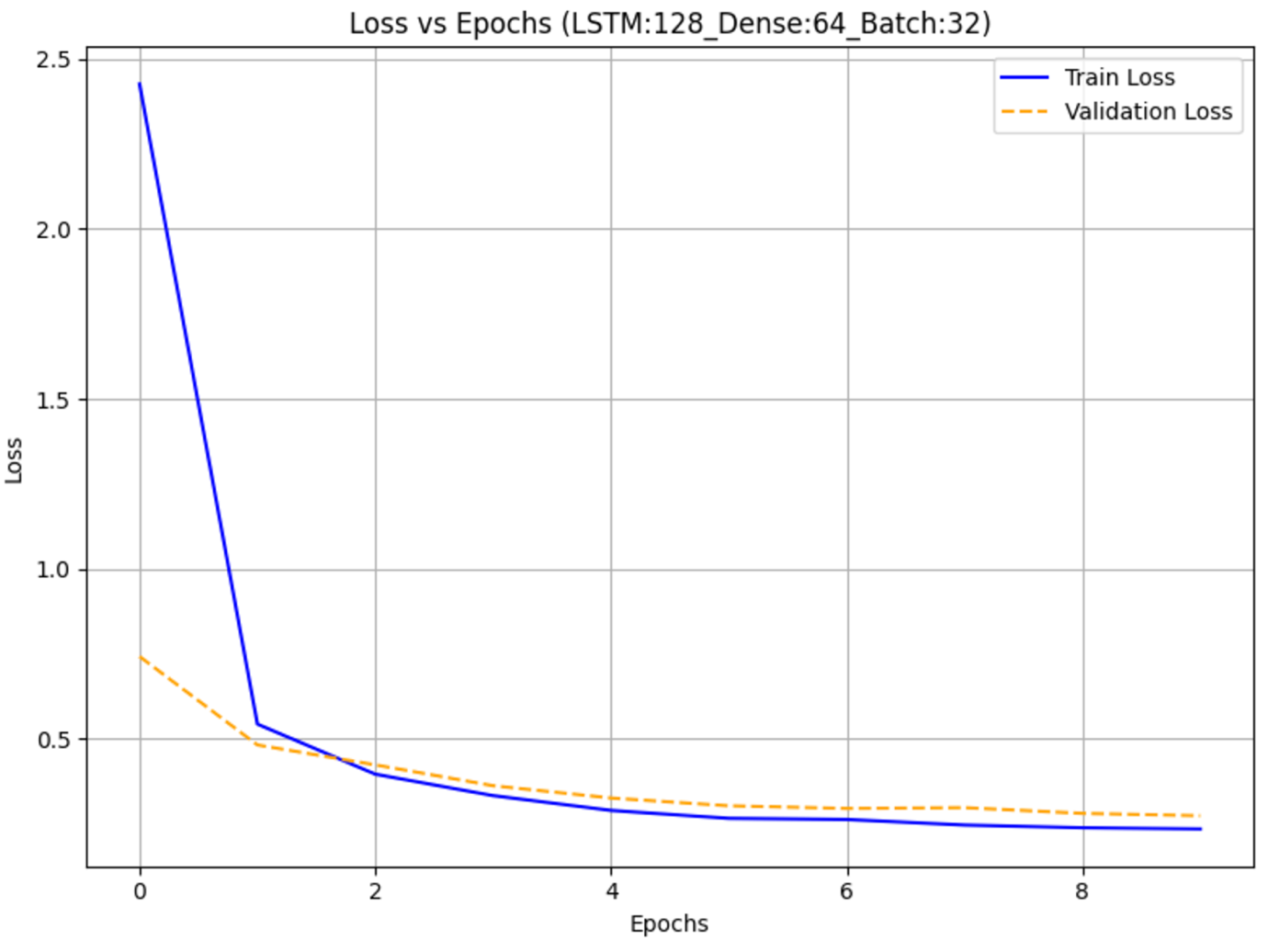
**A graph with blue and orange lines

Description automatically generatedA graph with a line

Description automatically generatedA graph with a line

Description automatically generatedA graph with blue and orange lines

Description automatically generatedA graph with blue and orange lines

Description automatically generated**

The above plots show slight overfitting after 2nd epoch which model is learning and training to perform better struggle to generalize unseen data. Overall, the AIC score for all the combinations was very high, showing that the models were complex and not a good fit for our dataset.

**XGBOOST**

The XGboost model provide a good result with R² score of 0.8415 after hypertuning but not the best results we could get.

**Catboost**

The catboost model was the best performer of our individual models providing a huge jump in R² score over ANN with 0.9217 in the metric, surpassing results for all other models while being relatively cost effective compared to Deep neural network models like ANN and RNN.

**Stacked Model**

In order achieve the best results possible, we used our ANN model with best AIC and performance metrics and stacked it with our hyper tuned catboost model using ensemble stacker to achieve the R² score of 0.9338. This study also highlights the effectiveness of hybrid models in ecological applications [6,10]

**Conclusion**

The above study emphasizes on the application of machine learning for the prediction of Stomatal Conductance (Gs) with the help of calculated evapotranspiration (ET). Among all the tested model Stacked Ensemble Model (ANN-CatBoost Hybrid) is the best performing model with highest R² score of [0.9338] and lowest Mean Squared Error(MSE) of 0.1531 which demonstrate its supremacy among all the other model for the prediction of Gs.

Performance of the CatBoost is decent with R² score of 0.9217, which emphasis its robustness dealing with complex data sets.

ANN and RNN shows reasonable predictive capabilities but they are prone to Overfitting specially in deeper configuration [64 layers]. During the analysis of the Neural Network there is significant decrease in training loss in the first 20 epochs which shows effective learning of the model. The significant gap between validation and training loss reflects overfitting especially models with more numbers of layers. Model that has [64] layers achieved highest R² score which lead to high risk of overfitting after second epoch. To deal with overfitting regularization, simple architecture and early stopping are required for neural network. The result reflects the importance of metrological and remote sensing data which improve the model performance and robust prediction, hybrid model is suitable for the prediction of Gs and calculative ET estimation.

**References**

1. Monteith, J.L., 1965. Evaporation and environment. In: Fogg, G.E. (Ed.), *Symposia of the Society for Experimental Biology*, Vol. 19, Cambridge University Press, 205–234.
2. Jensen, M.E., Burman, R.D., Allen, R.G. (Eds.), 1990. Evapotranspiration and irrigation water requirements. *ASCE Manuals and Reports on Engineering Practice*, No. 70.
3. Rana, G., Katerji, N., 2000. Measurement and estimation of actual evapotranspiration in the field under Mediterranean climate: A review. *European Journal of Agronomy*, 13(2–3), 125–153. https://doi.org/10.1016/S1161-0301(00)00070-8.
4. Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: Guidelines for computing crop water requirements. *FAO Irrigation and Drainage Paper 56*. Rome: Food and Agriculture Organization of the United Nations.
5. Breiman, L., 2001. Random forests. *Machine Learning*, 45, 5–32. https://doi.org/10.1023/A:1010933404324.
6. Fang, L., Li, H., Chen, Y., 2020. Improving the accuracy of machine learning models in ecological modeling. *Ecological Indicators*, 113, 106197. https://doi.org/10.1016/j.ecolind.2020.106197.
7. Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.
8. Zhang, Y., Kong, F., Wu, Y., Gao, Z., 2020. The role of remote sensing data in evapotranspiration modeling: A comprehensive review. *Environmental Research*, 182, 109005. https://doi.org/10.1016/j.envres.2019.109005.
9. Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785.
10. Liu, Y., Zhang, S., Zhang, J., Tang, L., Bai, Y., 2021. Using artificial neural network algorithm and remote sensing vegetation index improves the accuracy of the Penman-Monteith equation to estimate cropland evapotranspiration. *Agricultural Water Management*, 250, 106826. <https://doi.org/10.1016/j.agwat.2021.106826>.