Application of Random Forest Algorithm to estimate water use in a commercial forest plantation in South Africa (Oral)

Nkosinathi D. Kaptein[1]*, Ntsieni Mulovhedzi[2]

[1]University of the Free State, Bloemfontein, South Africa
[2]University of Mpumalanga, Nelspruit, South Africa
**Corresponding author: kapteinnd@ufs.ac.za*

ABSTRACT

Accurate estimation of tree transpiration is essential for sustainable forest water management, particularly in water-limited regions. The FAO-56 approach remains the most widely applied method for estimating evapotranspiration in agriculture; however, its application in commercial forestry is constrained by the lack of well-established crop coefficients and the complex physiological behaviour of forest stands. As a result, alternative modelling approaches are required. Recent advances in machine learning offer promising opportunities for capturing non-linear relationships between environmental drivers and plant water use.

In this study, a Random Forest (RF) multiple regression model was applied to estimate transpiration of six-year-old *Eucalyptus dunnii* trees at the Sabey experimental site near Nelspruit, Mpumalanga Province, South Africa. Continuous in situ measurements of tree transpiration were obtained using the heat pulse velocity method and used as reference data. Meteorological variables (air temperature, RH, gross rainfall, wind speed, solar radiation and soil water content) measurements were used as predictors in the RF model. The dataset was partitioned into calibration–validation (two-thirds) and independent testing (one-third) subsets to ensure robust model evaluation.

Variable importance analysis revealed that volumetric soil water content at 0.7 m depth was a key driver of transpiration, highlighting the strong dependence of tree water use on root-zone moisture availability. Modelled transpiration showed strong agreement with measured values, achieving a coefficient of determination (R^2) of 0.83, a root mean square error (RMSE) of 0.32, and a mean bias error (MBE) of 0.23.

These results demonstrate the strong potential of Random Forest modelling as a robust, accurate, and operationally efficient tool for estimating tree transpiration using readily measurable environmental variables. The approach provides a valuable alternative to conventional methods and offers significant potential for improving water-use assessments, plantation management, and decision-making in commercial forestry, particularly under conditions of increasing climatic variability.

Keywords: Machine learning, Water use, Calibration, Water management