Estimating Groundwater Recharge Variability in a Semi-Arid South African Catchment using Machine Learning

E. Bjerre, University of Copenhagen; J. Koch, GEUS; W.K. Dalum, University of Copenhagen now Fredensborg Municipality; N. Smith, University of Pretoria SA; K.G. Villholth, Water Cycle Innovation; S. Jessen, University of Copenhagen, K.H. Jensen, University of Copenhagen*

Introduction:
In semi-arid regions characterized by little and erratic precipitation and ephemeral river flow, groundwater is commonly the only perennial source of freshwater sustaining ecosystems and freshwater withdrawals for agricultural, domestic and industrial uses. Groundwater replenishment may occur either as (1) diffuse recharge, which refers to the recharge occurring across the landscape as precipitation infiltration or as (2) focused recharge, which represents groundwater replenishment via seepage from river channels during high flow occurrences. The latter component has been shown to contribute substantially to groundwater recharge locally. Estimation of both recharge components is highly uncertain and is constrained by the limited observational data both in time and space often prevailing in semi-arid regions. Moreover, their relative contributions at catchment scale remain underexplored.

Methods and data:
This study employs a data-driven machine learning approach to estimate annual groundwater recharge for the period 1970-2021 in the semi-arid Hout/Sand catchment (7,722 km2), Limpopo, South Africa. The Water Table Fluctuation (WTF) method is used to derive annual recharge estimates from 97 individual groundwater hydrographs spanning different time periods. Following, the recharge estimates are used to train a Light Gradient-Boosting Machine (LightGBM) model, generating fully distributed annual recharge maps at a 100 m resolution over the study period.

Results:
Results demonstrate considerable spatial variability in recharge, with high recharge values concentrated along river networks. Accordingly, proximity to rivers emerged as the most important co-variate in the LightGBM model.
Both focused and diffuse recharge are decreasing over the period 1970-2021, although focused recharge at a lower rate, ultimately causing a decline in the overall catchment-scale recharge. The relative contribution of focused recharge is increasing over time, due to stronger decline in diffuse recharge.

Discussion and take-home message:

- The machine learning approach is computationally efficient and robust, overcoming some limitations of conventional models caused by sparse data for calibration, uncertain parameters, and conceptual uncertainties in semi-arid catchments.
- The recharge training dataset based on the WFT method is subject to uncertainty due to the limited knowledge of specific yield used as input in the WTF calculations.
- The results suggest that although the relative importance of focused recharge increases over the 52-year period with rainfall intensification, is does not buffer the overall decrease in diffuse recharge, putting in question the hypothesis that focused recharge will act as a climate change buffer.


* *elbj@ign.ku.dk: University of Copenhagen, Department of Geosciences and Natural Resource Management, Øster Voldgade 10, 1350 København K, Denmark*