HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

DIGITAL

Catchment scale prediction of soil moisture trends from Cosmic Ray Neutron Rover Surveys using machine learning

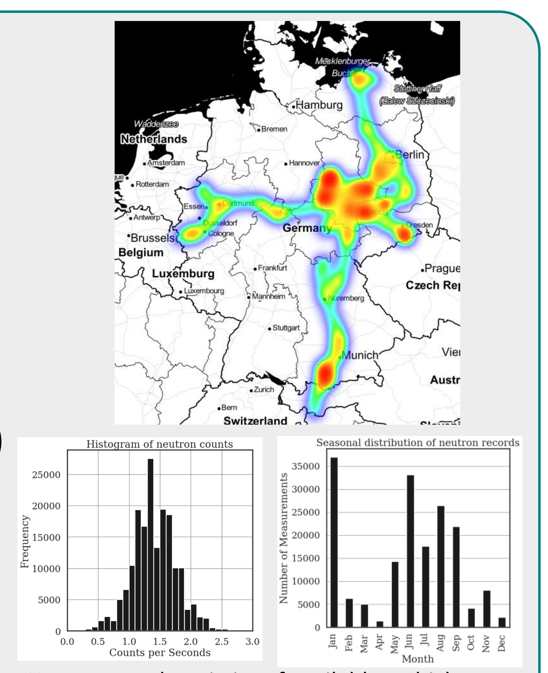
Erik Nixdorf[1], Marco Hannemann[1], Uta Ködel[1], Martin Schrön[1] and Thomas Kalbacher[1]

1 Helmholtz Centre for Environmental Research GmbH - UFZ

UFZ HELMHOLTZ Zentrum für Umweltforschung

Motivation

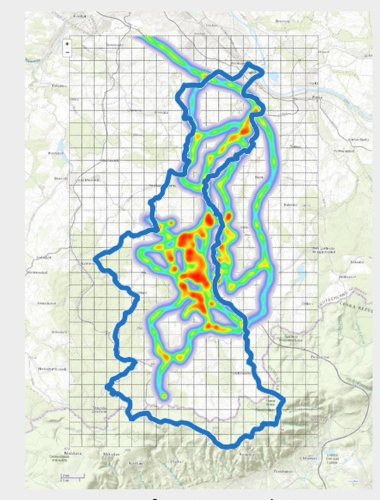
- Soil moisture is an essential critical state variable in land surface hydrology and a key component of microclimate
- Measuring principles exist on point scale (e.g TDR), mesoscale (e.g. vehicle-mounted CRNS) and large scale (satellite observation)
- We integrate CRNS records, features from ancillary datasets and machine learning algorithms towards an improvement of soil moisture predictions on the catchment scale



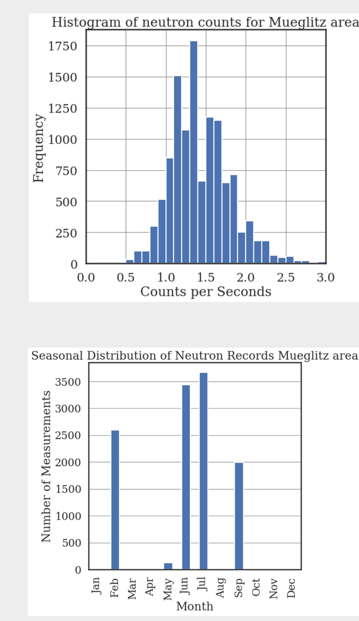
Heatmap and statistics of available vehicle-mounted (CRR) neutron records

Study Site

- Mueglitz River Basin south of Dresden (~209 km²)
- The basin is heavily impacted by heavy rain events and flash floods
- Ongoing event-driven intense watershed-scale research and monitoring efforts (bit.ly/2zNqKzi)

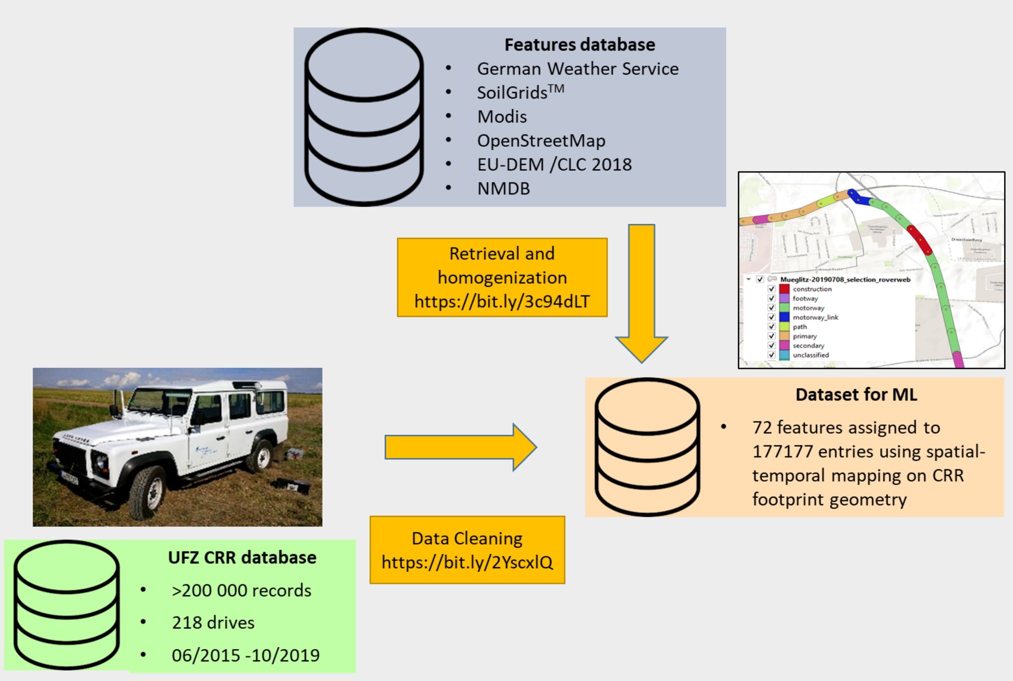


Heatmap of CRR records in Mueglitz River Basin



Methodology

´Data Preprocessing



Learner Application

scikit learn

- 3 learners selected for regression problem
- Trained with subset from entire CRNS dataset
- Performance tested with Mueglitz CRNS data

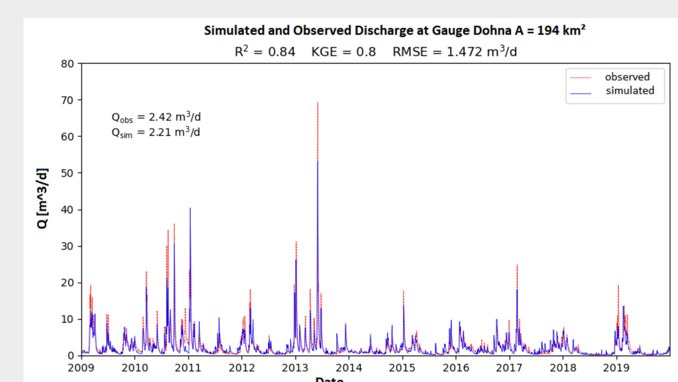
Learner	Parametrization	Runs
Linear Regression (LR)	RFE (True/False)* 5-72 Features (5 steps)* MinMax- and Standardscaler*	100
Random Forest (RF)	1-50 Trees (9 steps)	900
Artifical Neural Network (ANN)	1-4 hidden layers (4 steps) 1-19 hidden units (7 steps)	2800

* Parameters of Linear Regression applied for all leaner

Hydrological Modelling

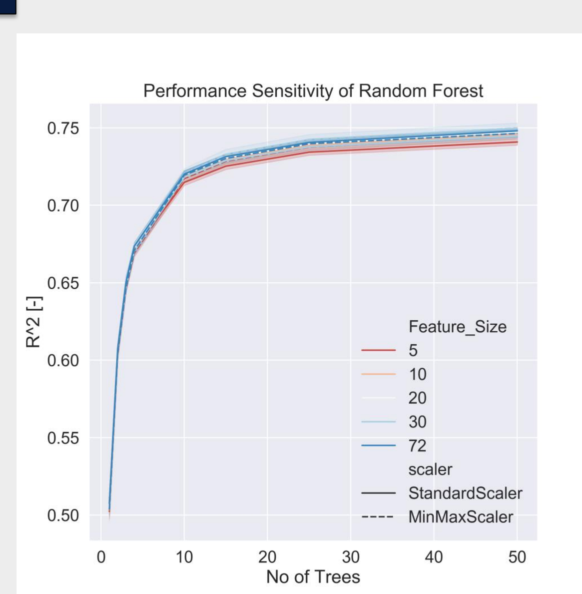
The mesoscale Hydrologic Model mHM

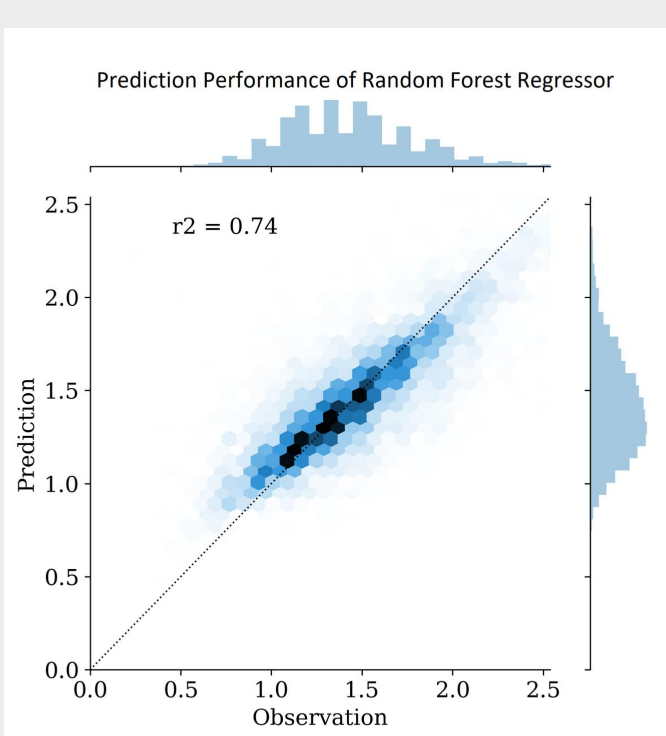
- Hydrological model of Mueglitz Basin set up using mHM 5.10 (Kumar et al, 2013)
- Model calibration using gauge data
- Provision of daily soil moisture estimates

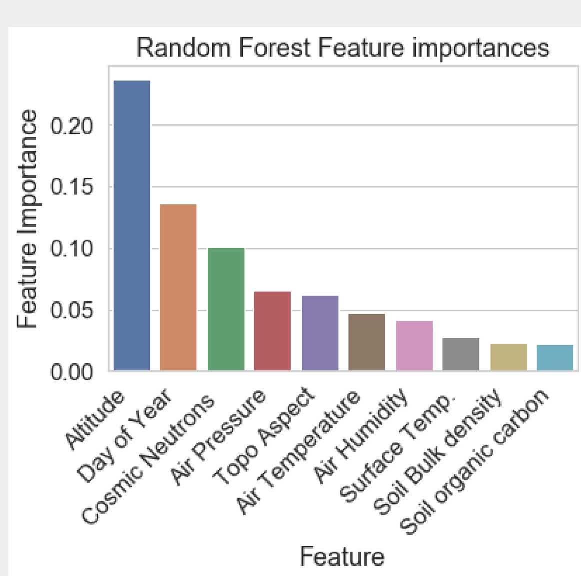


Model Performance

- RF (R^2_{max}=0.74) outperforms LR (R^2_{max}=0.28) and ANN (R^2_{max}=0.41)
- No of trees most sensitive model parameter
- Altitude, day of year and neutron input are most import features

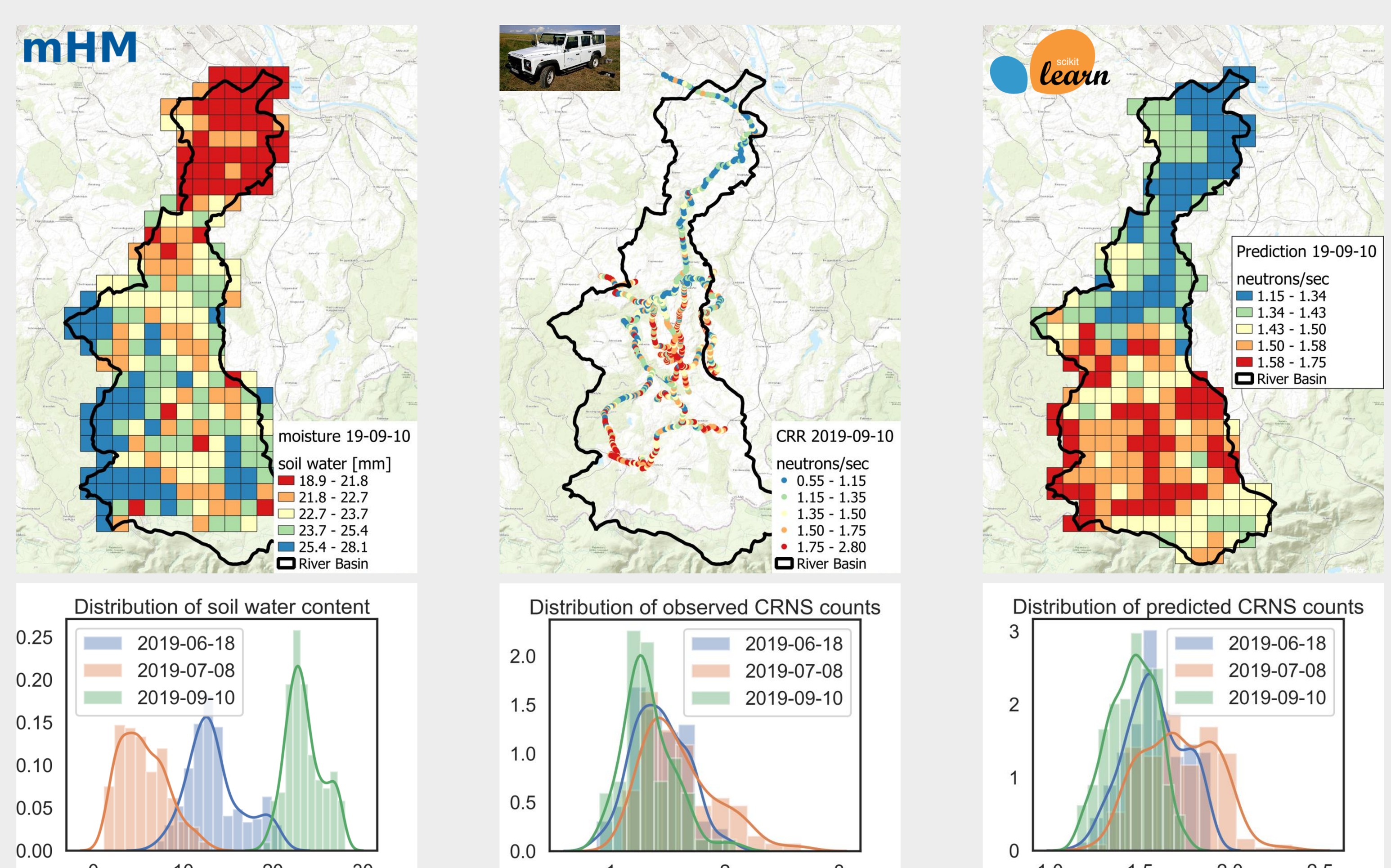






Basin-scale Prediction

- Considering the seasonal variations, both observed and predicted neutron counts resemble the inverse relationship to soil water content
- Spatial patterns are not represented properly→ improvement expected by using pre-corrected CRNS records



Conclusions

- Random Forest outperforms other approaches to predict CRNS records using exclusively features from free data sources
- Most relevant features are altitude, day of year, neutron input monitored at Jungfraujoch Neutron Monitor, air pressure and topographic aspect.
- Transferring the trained machine learning approach on the entire Mueglitz catchment, obtained CRNS estimates resemble the distribution of CRNS record data as well as the inverse relationship with modelled soil moisture on a seasonal scale
- Future developments will apply the developed machine learning workflow on corrected CRNS records and derived soil moisture estimates to improve soil moisture mapping on the catchment scale

References

- F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 85, pp. 2825–2830, 2011.
- M. Schrön *et al.*, "Improving calibration and validation of cosmic-ray neutron sensors in the light of spatial sensitivity," *Hydrol. Earth Syst. Sci.*, vol. 21, no. 10, pp. 5009–5030, 2017.
- R. Kumar, L. Samaniego, and S. Attinger, "Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations," *Water Resour. Res.*, vol. 49, no. 1, pp. 360–379, 2013.

DIGITAL EARTH

The project *Digital Earth* has been strategically initiated by all eight centres of the Helmholtz research field Earth and Environment (E&E). *Digital Earth* is funded by Helmholtz research field Earth and Environment (E&E). and co-financed from all eight E&E centres.

EGU General Assembly 2020

ESSI1.15
Towards SMART Monitoring and Integrated Data Exploration of the Earth System

EGU2020-3049