

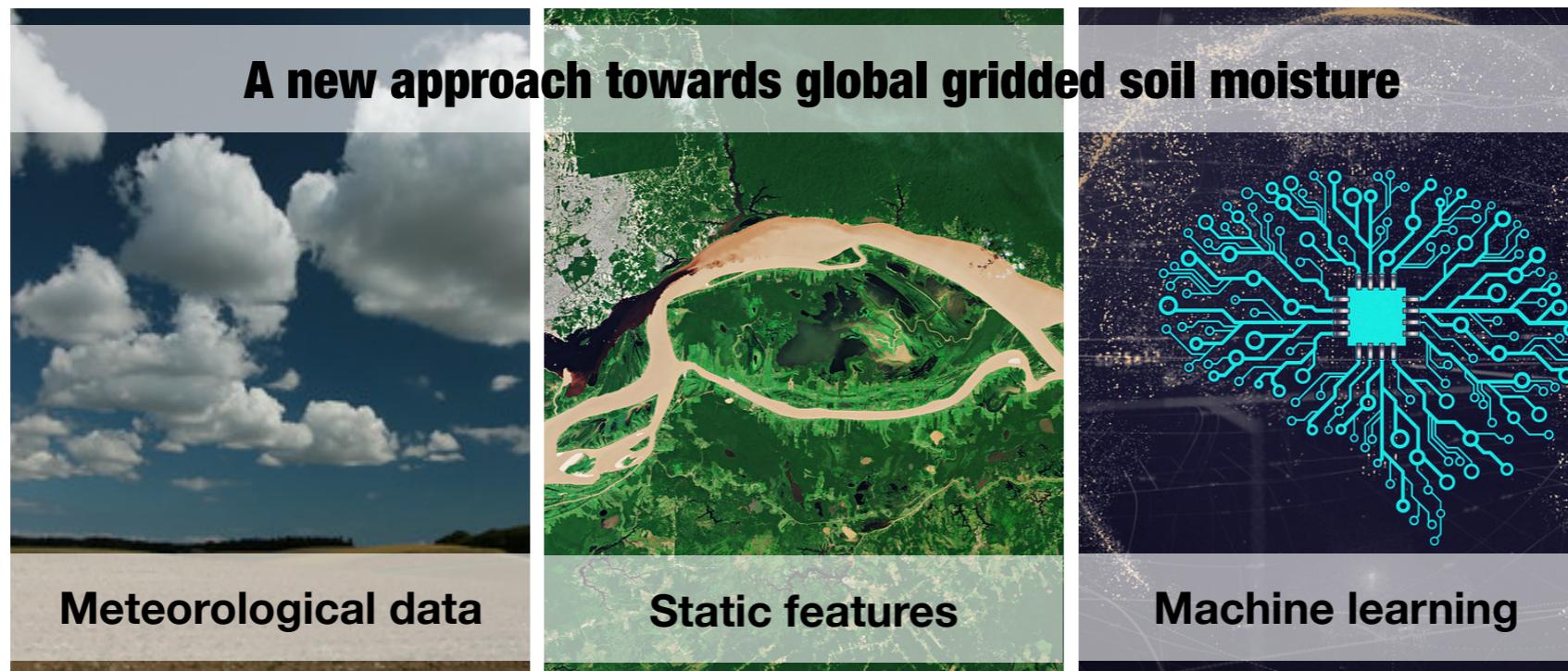
EGU2020-7428

Augmenting satellite-derived soil moisture with multiple data streams using machine learning

Sungmin O* and Rene Orth

Max Planck Institute for Biogeochemistry, Jena, Germany

(*email: sungmino@bgc-jena.mpg.de)

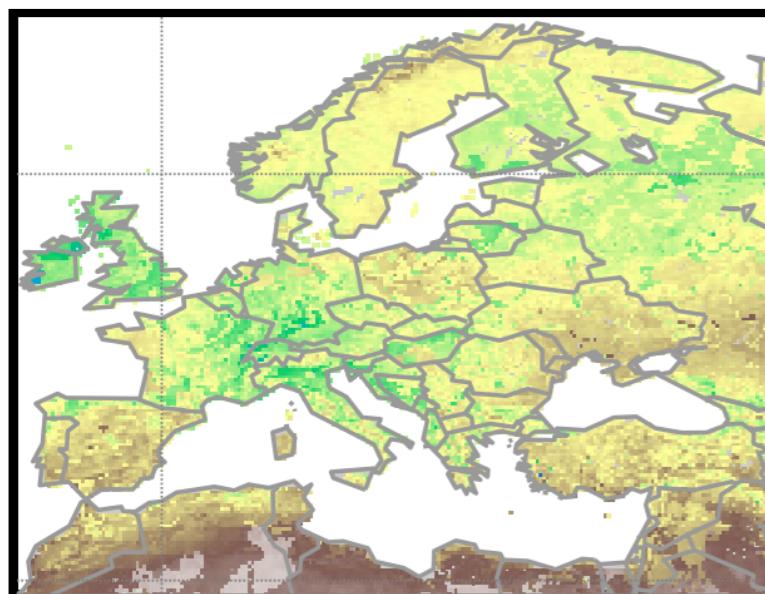


Photos are from <https://www.flickr.com/> ([Freddy Fehmarn](#)/[europeanspaceagency](#)/[mikemacmarketing](#); left to right) and edited

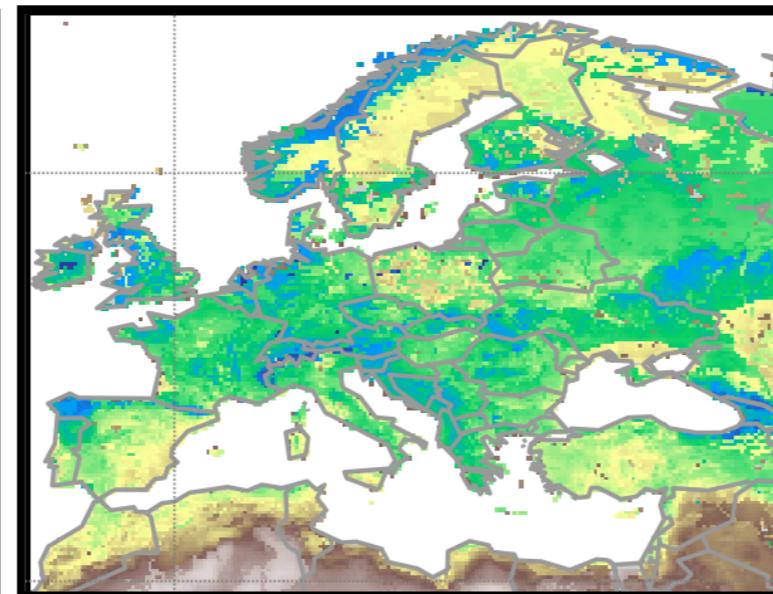
1. Main Results

Average surface soil moisture over Europe during 2015-2016

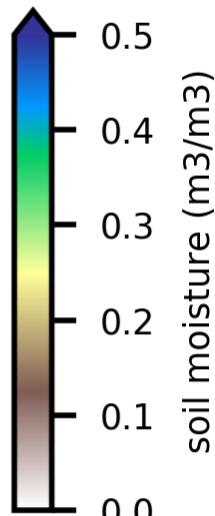
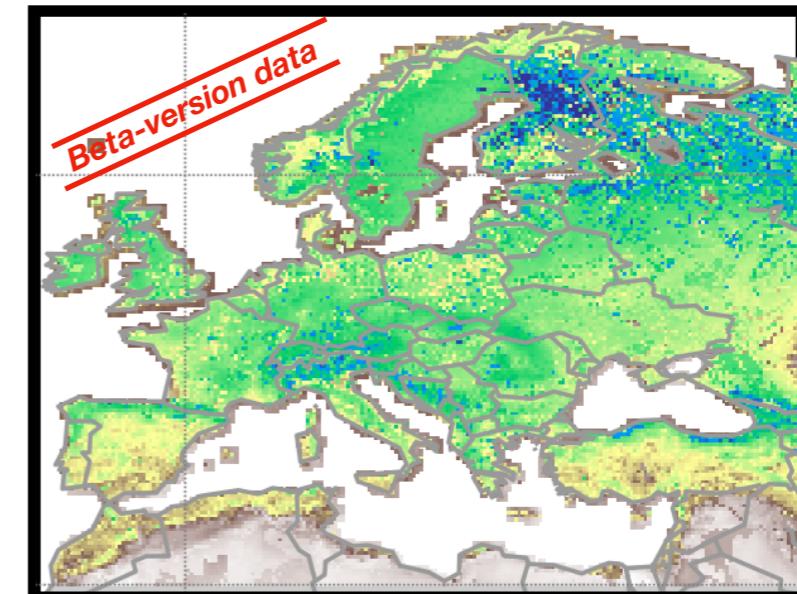
Satellite observation
(ESA CCI v04.4 COMBINED)



Modelled data
(GLEAM v3.3a)

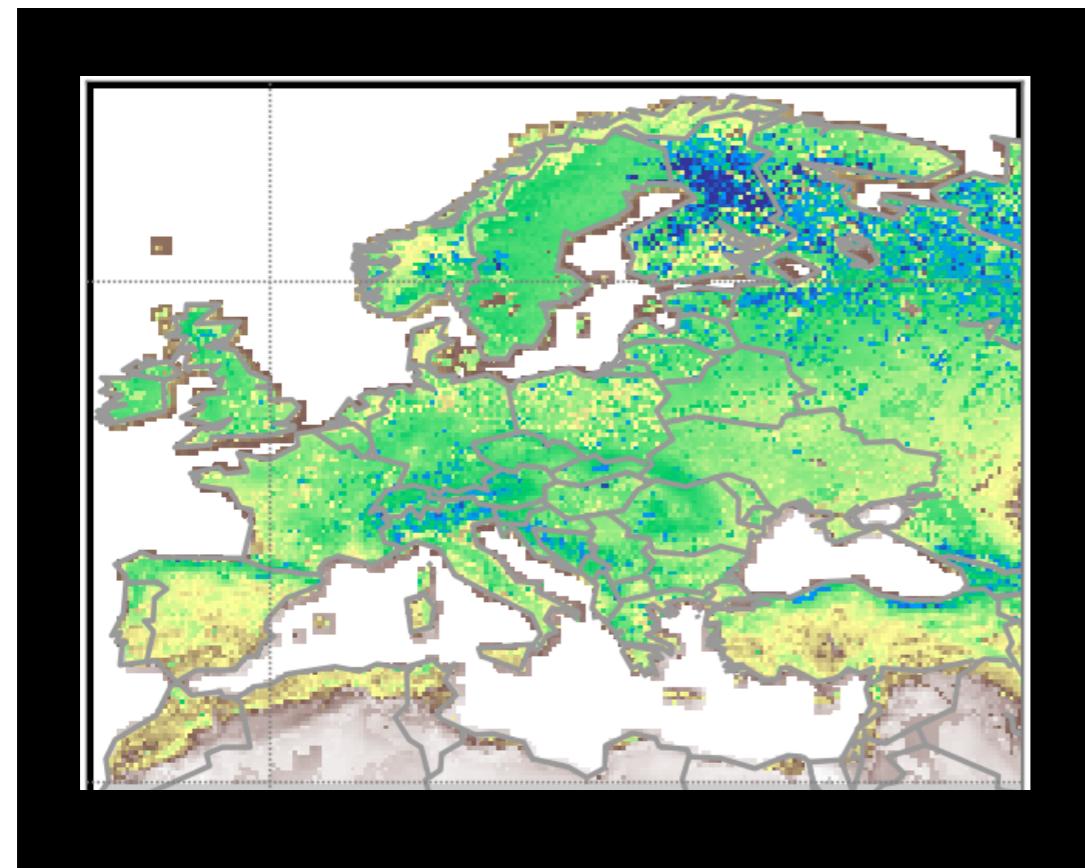


LSTM trained with in-situ data
(our new data)



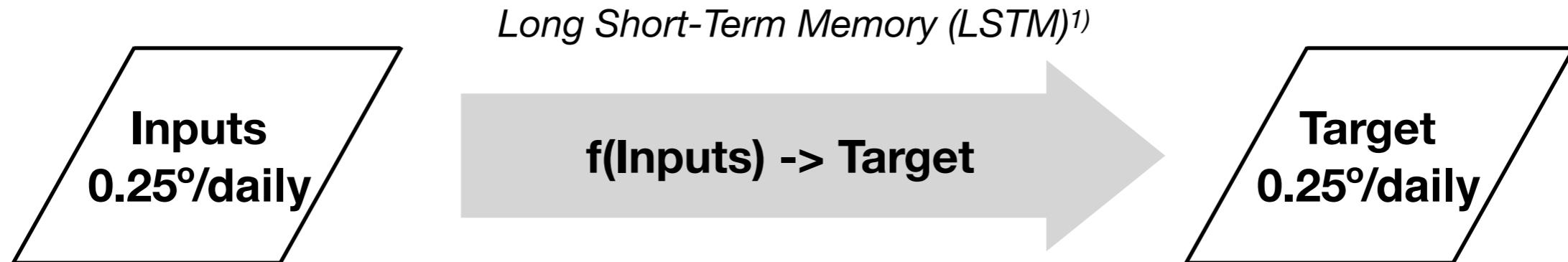
We show our new soil moisture data over Europe. **A Long Short-Term Memory (LSTM)-based neural network model** is trained to learn the relationship between multiple input variables and in-situ soil moisture measurements. As the input data are available all over Europe, the model can be applied to compute spatiotemporally seamless soil moisture data across the continent (see Method & Data). We aim to provide **surface and root-zone soil moisture data at a global scale for 2000-2018**.

2. Take-home message



- We present a novel machine-learning based approach for producing global gridded soil moisture data.
- LSTM can learn soil moisture dynamics from multiple input data streams and reproduce soil moisture at ungauged locations.
- We will provide a long-term, large-scale soil moisture dataset which is derived from in-situ measurements. Our new data can be complementary to existing satellite-based or modelled soil moisture data.

3. Method and Data



- ERA5 meteorological variables ↗
- Static features; elevation (GTOPO) ↗
soil type and landcover (HWSD) ↗
- Adjusted²⁾ ground measurements; in-situ data from International Soil Moisture Network (ISMN) ↗ over ~800 grid cells

1) Recurrent neural network designed to model time series and their long-term dependencies.

2) To estimate areal soil moisture, we adjusted the mean and standard deviation of point-level ISMN data to those of ERA5 gridded soil moisture. If more than one ISMN data are available within the same grid cell, their average was taken.