Interdisciplinary Project in Data Science

Pfennigbauer Johannes, 11902046

August 11, 2024

Project Summary

Title: The Impact of Subsurface Scattering on Microwave-Derived Soil Moisture Retrievals

Supervisor: Dr Martin Schobben, Dept. of Geodesy and Geoinformation, TU Wien

Co-Supervisor: tba

Start of Project: July 2024

Domain-Specific Lecture: 120.110 Data Retrieval in Earth Observation

1 Research Question

Since 2010, soil moisture has been considered an essential climate variable in the Global Climate Observing System. It is known to be a reliable predictor of floods and droughts, an essential parameter for agricultural management, and a meaningful estimate of other climate variables. [1]
A common technique to measure soil moisture content on a large scale is microwave remote sensing, which uses microwave backscatter signals as an estimate of soil moisture. This technique is based on the premise that the backscatter signal increases with the soil moisture content due to the change of the dielectric contrast at the soil-air interface in wet soils. However, under certain circumstances such as frozen soils, this technique does not deliver reliable results. Similarly, an anomalous behaviour was found when microwaves are reflected from dry soil, especially in arid regions. This effect, called subsurface scattering, occurs when the microwaves pass through the first few centimetres of dry soil and get reflected by hard surfaces beneath, as illustrated in Figure 1. This phenomenon results in an inverse relationship between soil moisture and the backscatter signal compared to the typical assumption. [2]
As a consequence the data obtained using the traditional assumption would show high soil moisture content in these arid regions, where in situ measurements prove that it is actually

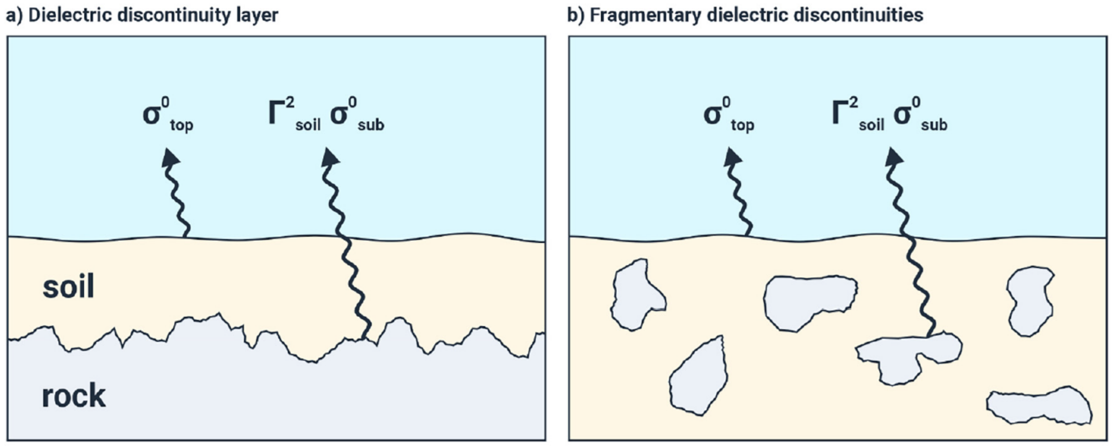

Figure 1: Illustration of subsurface scattering from [2]. Rock surfaces beneath the top soil layer (a) and dispersed stones (b) are the main causes of the phenomenon.

very dry. Therefore, the remote sensing group of Vienna University of Technology has developed a method to detect such anomalies on radar images using a reference soil moisture dataset. The analysis shows that the soil type and the coarse fraction of the soils could be reliable variables for the detection of this type of anomaly. [3] [4]
This project aims to investigate the described anomaly using auxiliary data such as terrain elevation, soil group and soil composition on the example of Australia. Therefore the research question is formulated as follows:

> What are the main drivers of the subsurface backscattering anomaly in dry, arid regions under the presence of rocks or distinct horizons?

2 Methodology

To answer the research question I will build a regression model based on the research from the remote sensing group at the Vienna University of Technology and their findings and enhance it with additional data to further investigate and explain the occurrence of such anomalies. The goal is to identify the main influences from a range of variables such as terrain elevation, soil type, soil composition and other potential data attributes. The study area is Australia.

2.1 Input Data

The occurrence of subsurface scattering is obtained by the correlation of Sentinel-1 backscatter timeseries and reference soil moisture data from ERA5Land provided by the European Center for Medium-Range Weather Forecasts (ECMWF) at 9 km resolution. The data is provided by the remote sensing group of the Vienna University of Technology.

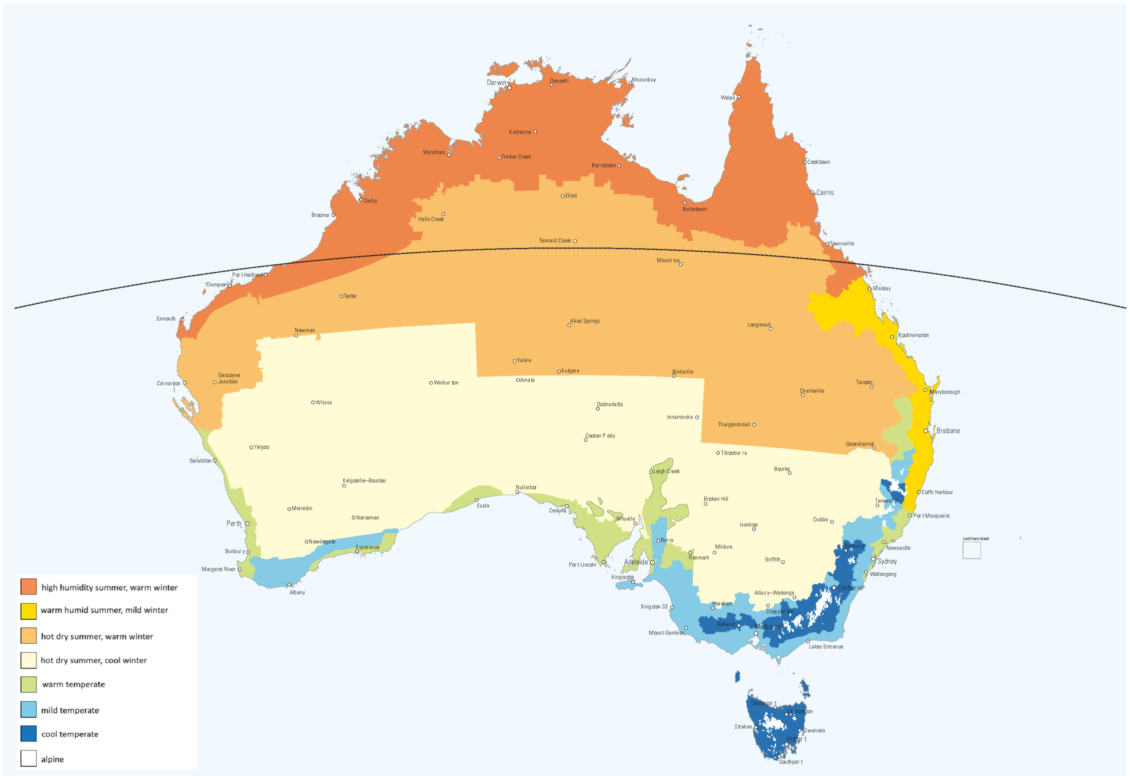

Figure 2: Climatic zones of Australia. Arid regions with dry summers in the inside (light orange and light yellow) and wet areas at the coast (dark orange, green, light and dark blue). Adapted from [5]

For a first model the mean soil moisture content over the last 30 years based on ERA5Land soil moisture data and the coarse fragment content of the soil from the Soil and Landscape Grid of Australia (SLGA) at a 90 meter resolution will be used.
All the data is collected for the study area, which spans along the 21 southern latitude from the western to the eastern end of Australia and covers various climatic zones from wet at the coast and dry in the interiors, as pictured in Figure 2.

2.2 Preprocessing

The data contains observations from pixels entirely over the sea and from deeper soil layers, which are not relevant for the scope of this study. Therefore, those datapoints will be removed. Furthermore, missing values in the data will be identified and removed if necessary. As the interaction of different data attributes could have a high impact on the target variable, I will also add such derived data attributes where applicable. All attributes will be standardized, so that the resulting coefficients are comparable.
Lastly, the data will be splitted temporarily into a training and test set in order to evaluate the performance of the model.

2.3 Modelling

The analysis on the anomaly will start with a simple multiple-regression model for high interpretability and understanding of the behavior.
To cover a more comprehensive list of potentially relevant attributes, the simple model will then be extended with derived or selected additional data attributes. At each step, the change in performance and behavior of the model will be noted.
Wagner et al. mention in their recent study [3] that the highly correlated nature of the environmental process, that leads to the subsurface scattering phenomena will be a challenging issue. To address this issue I want to use methods like Ridge regularization and Partial Least Squares (PLS) analysis, which are well suited for multi-collinear data. This approach could add valuable insights to identify the main drivers of the anomaly, which is the goal of this project.

2.4 Evaluation

For the regression models themselves the goodness of fit will be measured with common metrics like $RMSE$, R^2 and MAE, since they have proven to be robust and reliable.
In order to answer the research question, i.e. identifying the main drivers of the anomaly, a comparison of the resulting model coefficients from the multiple-regression part, the Ridge regularization and the PLS analysis will be used.
After inspecting the results and identifying the main drivers of the anomaly a deeper analysis on specific regions, that stand out during the comparison will be performed. This qualitative approach will be beneficial as it allows for the validation of model predictions in the context of local environmental or geological characteristics, providing insights that quantitative methods alone may overlook.

References

[1] K. Morrison and W. Wagner. Explaining anomalies in sar and scatterometer soil moisture retrievals from dry soils with subsurface scattering. *IEEE Transactions on Geoscience and Remote Sensing*, 58(3):2190–2197, Mar 2020.

[2] W. Wagner et al. Widespread occurrence of anomalous c-band backscatter signals in arid environments caused by subsurface scattering. *Remote Sensing of Environment*, 276:113025, Jul 2022.

[3] W. Wagner et al. Mapping subsurface scatterers from sar backscatter time series. In *EUSAR 2024*, 2024.

[4] Roland Lindorfer, Wolfgang Wagner, Sebastian Hahn, Hyunglok Kim, Mariette Vreugdenhil, Alexander Gruber, Milan Fischer, and Miroslav Trnka. Global scale maps of subsurface scattering signals impacting ascat soil moisture retrievals, Aug 2023.

[5] Australian Building Codes Board. Climate zone map, 2024. https://www.abcb.gov.au/resources/climate-zone-map [Accessed: 2024-08-11].