University of Tartu
Faculty of Science and Technology
Institute of Ecology and Earth Sciences
Department of Geography

Master's thesis in Geoinformatics for Urbanised Society (30 ECTS)

Soil moisture and temperature in wetland soils derived from remote sensing

Hleb Lazovik

Supervisors:
Jaan Pärn, PhD,
Sandeep Thayamkottu, MSc

Tartu, 2023

Abstract

Soil moisture and temperature in wetland soils derived from remote sensing

The study presents an approach to predict soil moisture and temperature in wetland soils from remotely sensed data. The focus is given to wetland soils using Landsat 8 remotely sensed imagery and a global ground truth dataset. To predict soil temperature, the Surface Temperature from Landsat 8 Level 2 Science Products is used. NDMI, MSI, and NDWI are picked as predictors of soil moisture. The ΔT*NDVI method is used as a soil moisture predictor as an attempt to adapt the trapezoid model approach to a global dataset. The main means of estimating performance of various predictors is regression analysis. The results demonstrated good (R^2 = 0.70) agreement between the ground and remotely sensed soil temperature measurements. For soil moisture, an agreement between the ground and remotely sensed measurements was found after using ΔT*NDVI ($p < 0.001$; R^2 = 0.39). I further discuss limitations in the use of ΔT*NDVI, namely poor applicability on boreal climates and bare soil samples.

Keywords: Landsat, NDVI, remote sensing, soil moisture, soil temperature, wetland

CERCS code: P510 – Physical geography, geomorphology, pedology, cartography, climatology

Annotatsioon

Märgalade mullaniiskuse ja mulla temperatuuri prognoos kaugseireandmete põhjal

Uurimistöö eesmärk on tuletada kaugseireandmetest maakera soomuldade niiskust ja temperatuuri. Töös kasutati Landsat 8 kaugseirepilte, mis verifitseeriti, ning ülemaailmsetel välitöödel mõõdetud andmestikku. Kasutati Landsat 8 *Level 2 Science Product* maapinna temperatuuri andmestikku. Mullaniiskuse tuletamiseks valiti kõigepealt kolm näitajat: NDMI (ingl *normalized difference moisture index*), MSI (ingl *moisture stress index*) ja NDWI (ingl *normalized difference water index*). Lisaks rakendati ΔT*NDVI (ingl *delta temperature multiplied by normalized difference vegetation index*) meetodit. Näitajate võimekust hinnati peamiselt regressioonanalüüsi abil. Tulemused näitasid tugevat positiivset seost (R^2 = 0,70) väliandmestiku ning kaugseire abil mõõdetud mullatemperatuuride vahel. Mullaniiskuse puhul andis olulise seose ΔT*NDVI rakendamine ($p < 0,001$; R^2 = 0,39). Uurimistöös arutletakse ΔT*NDVI kasutamist piiravate tegurite üle ning eelkõige selle piiratud rakendatavuse osas boreaalses vööndis ning taimkatteta muldadel.

Märksõnad: Landsat, NDVI, remote sensing, soil moisture, soil temperature, wetland

CERCS kood: P510 – Füüsiline geograafia, geomorfoloogia, mullateadus, kartograafia, klimatoloogia

Table of Contents

List of Figures

List of Tables

1. Introduction

Wetlands are one of the world's most important ecosystems. They are valuable as sources, sinks, and transformers of a multitude of chemical, biological, and genetic materials. They are sometimes described as kidneys of the landscape because of their functions: they are downstream receivers of water and waste from both natural and human sources. Wetlands mitigate both floods and drought by stabilizing water supplies. They cleanse polluted waters, protect shorelines, and recharge groundwater aquifers. The extensive food chain and rich biodiversity that they support have also earned them the nickname of nature's supermarkets. In the landscape, they provide unique habitats for a wide variety of flora and fauna. On a global scale, wetlands are important carbon sinks and climate stabilizers (Mitch & Gosselink, 2015).

Today, satellite-derived remote sensing data is an essential source of information for the purposes of study and monitoring of the wetlands. Soil moisture and temperature are some of the multiple parameters which could be potentially studied using remote sensing data. The crucial importance of soil moisture and temperature as parameters determining the functioning of wetland ecosystems is evidenced by the fact that their study is necessary for multiple purposes, including N_2O management (Pärn et al. 2018, Mitch & Gosselink, 2015), natural hazards prediction (Petropoulos et al. 2013), and racking the impacts of climate change (Brekke et al. 2009). Soil temperature can determine wetland biological activity. The primary controlling mechanisms in methane emissions are soil moisture, indicated by the water table position, and soil temperature (Mitch & Gosselink, 2015).

Landsat products are used extensively for the purposes of detecting soil moisture and temperature. In this context, it is common to refer specifically to the land surface temperature (LST). Multiple indices, such as the normalized difference vegetation index (NDVI), normalized difference water index (NDWI), normalized difference moisture index (NDMI), and temperature-vegetation dryness index (TVDI) are used.

Existing studies tend to focus on some particular study area, rarely aiming for global coverage, and only doing so for soil temperature. The research gap this scientific paper aims to fill lies in between, as the possibilities of studying soil moisture and temperature in a particular biome globally are assumed to be unexplored. In this research, attention is given to wetland soils globally, with the aim of presenting and evaluating a model, which would

explain the temperature and soil moisture there with optimal accuracy. For this purpose, analyses are performed, involving the ground data with multiple parameters, including the soil temperature at different depths and volumetric soil water content, and Landsat-derived indices and calculations.

The study aims to answer the following research questions:

1. How well do Landsat-derived soil temperature and moisture measurements agree with ground data?
2. Is it possible to create a global model for wetland soils that derives soil moisture and temperature from Landsat data within reasonable error margin?

The study consists of five chapters: 1. Introduction, 2. Theoretical overview, 3. Data and methodology, 4. Results, and 5. Discussion and conclusions. In the first chapter, an introduction to the research is given, and research questions are asked. In the second chapter, the assessment of literature related to the subject is presented, and further methodological decisions are given theoretical justification. In the third chapter, the datasets are described, and the methods used in the research are listed and explained. In the fourth chapter, the results of the research are presented. In the fifth and final chapter, the results are described and given an explanation, suggestions are made for further research in this direction, and conclusions are made reflecting on the research questions asked in the introduction.

2. Theoretical overview

Giving wetlands an exact, objective, and universal definition is challenging. Multiple definitions exist, which depend on the objective and field of interest of the user. For geology, soil science, hydrology, biology, ecology, sociology, economy, political science, public health science, and law, the wetlands definitions may be different, and even contradictory. It is common, however, to include the following components when defining wetlands:

1. Presence of water at the surface or within the root zone.
2. Uniquely wet soil, different from adjacent uplands.
3. Vegetation adapted to wet conditions, and absence of flood-intolerant biota (Mitsch & Gosselink, 2015).

An international definition of wetlands was adopted in the Ramsar Convention by the International Union for the Conservation of Nature and Natural Resources (IUCN):

> For the purposes of this Convention wetlands are areas of marsh, fen, peatland or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish, or salt including areas of marine water, the depth of which at low tide does not exceed six meters (Goodwin, 2017).
>
> [Wetlands] may incorporate riparian and coastal zones adjacent to the wetlands, and islands or bodies of marine water deeper than six metres at low tide lying within the wetlands (Goodwin, 2017).

In the field of derivation of soil moisture and temperature from remote sensing imagery, a significant amount of research has already been made. Various methods and concepts were introduced, tested, and evaluated under different conditions and circumstances. Study areas in those papers vary in both size and natural conditions, ranging from particular wetlands to global coverage.

2.1. Remote sensing of soil temperature

In the literature related to the derivation of soil moisture from remote sensing data, soil temperature (specifically LST) is usually featured as a necessary part of the analysis, required for the calculation of soil moisture. For this reason, soil temperature derivation precedes soil moisture derivation, and the theory behind soil temperature derivation is addressed first. Landsat data is commonly used in studies related to the derivation of LST from remote sensing data, both in and outside the context of wetlands.

The study by Kaplan & Avdan (2018) investigates the monthly dynamics of wetlands and finds correlations between several multi-sensor parameters. LST and NDVI were extracted from Landsat 8 using a tool developed in ERDAS Imagine, and dual polarization backscatter values (VH-VV) were extracted from Sentinel-1. Balikdami wetland in Turkey was used as a study area. The results showed a strong correlation between the LST and the NDVI values (0.94), and a strong correlation between the microwave (VH) and both thermal and optical parameters (0.81).

The research by Parastatidis et al. (2017) aims for global coverage rather than focusing on a particular study area. It explores the estimation of LST for the globe from Landsat 5, 7, and 8 thermal infrared sensors. A single channel algorithm was used to preserve consistency between different Landsat missions' datasets. Three sources were used for the emissivity data: ASTER, MODIS, and NDVI. An overall root mean square error of 1.52 °C was observed and it was confirmed that the accuracy of the LST product is dependent on the emissivity. A web application was developed by the authors of the study, offering the calculation of LST from Landsat images and a choice of three different emissivity sources.

In the research by Taloor et al. (2021), LST, NDMI, and normalized difference water index (NDWI) were retrieved over the Ravi basin in India and Pakistan. Landsat 8 data was used. NDWI and NDMI were calculated from Bands 5 (NIR), 3 (Green), and 6 (SWIR 1). To calculate LST, NDVI and fractional vegetation cover (FVC) were calculated first, and land surface emissivity (LSE) was derived from them. Surface radiance was calculated from LSE and at-sensor temperature, derived from Band 10 (TIRS), and was used to calculate LST. The results were validated by six field observations. Their comparison has shown a similar pattern and correlation described by the authors as "excellent". The average difference between LST and observed temperature is 1,6 °C, with the calculated LST being systematically more

different from 0 °C. The difference could be explained by a wide range of air temperatures in the study area, with the highest being +30 °C and the lowest being -31 °C, making the model less specific and accurate. The imperfection of the methods used can be cited as a reason as well since the achieved accuracy in this study was worse than that in the study by Parastatidis et al. (2017), which had global coverage, but a more precise methodology.

2.2. Remote sensing of soil moisture

Due to its complexity and lack of globally applied approaches, I gave more attention to the theory behind predicting soil moisture values from remote sensing data. I examined the trapezoid model approach separately due to its importance for the methodology of my research.

2.2.1. Overview

The research by Zeng et al. (2004) deals with the assessment of soil moisture in the semiarid environment in the northeastern Tibetan Plateau in China. It examines the relationship between NDVI, surface temperature, and soil moisture, shows the "dry edge" and the "wet edge" of the given parameters, representing low and high soil moisture respectively (described in more detail in *2.2.2. Trapezoid model approach*), and introduces the Soil Moisture Index (SMI), derived from the surface temperature and NDVI. The results were compared with the spatial distribution of desertification, and a conclusion was made that low SMI values do have a positive relation to the severity of desertification.

Featuring the dry edge and the wet edge in the remote sensing-derived soil moisture studies is not exclusive to Zeng et al. (2004). In the study by Uniyal et al. (2017), soil moisture was estimated from Landsat 7 and 8 imagery in the context of evaluating the Soil and Water Assessment Tool (SWAT) in two sub-catchments of the Ilmenau River basin in Northern Germany. Temperature-Vegetation Dryness Index (TVDI) was calculated using NDVI and the brightness temperature at the top of the atmosphere and was based on the empirical parametrization of the dry edge and the wet edge.

The study by Klinke et al. (2018) addresses the issue of acquiring soil moisture in wetlands from remote sensing data, employs several methods (including their combined usage), and evaluates their accuracy. The study area is a flood plain of the river Peene near Demmin, Mecklenburg-Western Pomerania, Germany, observed between 2000 and 2016. A large amount of satellite data was used, including Landsat 4 and 5, Landsat 7, Landsat 8, Sentinel-1, and Sentinel-2. Ground data was gathered through repeated measurements, as well as from the TERENO (Terrestrial Environmental Observatories; Bogena et al. 2012) test site DEMMIN (Durable Environmental Multidisciplinary Monitoring Information Network; Borg et al. 2014). The first method applied is the Ellenberg system (Ellenberg & Leuschner 2010), in which soil moisture is derived from bio-indicators, and rather than representing

actual soil moisture content, mean soil moisture conditions present at the specific site are represented. The second method used in the research is the universal triangle method (Price 1990), which is based on the idea that the surface temperature is directly dependent on soil moisture, and which uses the Soil Moisture Index (Zeng et al. 2004), NDVI, and LST. The third method involved estimating soil moisture using Support Vector Regression (SVR, Pasolli et al. 2015) techniques. As a result, the highest accuracy was achieved using the three methods combined (coefficient of determination of 0.93).

The case of the study by Klinke et al. shows that the existing studies are being used in the new research, with new methodologies being developed based on the older ones. In this study, we can see the Soil Moisture Index, introduced by Zeng et al., being used in the universal triangle method, and ultimately in the development of the model describing soil moisture in wetlands from remote sensing data with better accuracy.

In the studies by Kaplan & Avdan (2018) and Uniyal et al. (2017), it is possible to see the scientific value of studying soil moisture and temperature based on remote sensing data for other fields of science. In both cases, the techniques and methods to acquire soil moisture and temperature values are not the main focus of the study (in the case of Kaplan & Avdan it is wetlands monitoring, and in the case of Uniyal et al. it is evaluation of SWAT), but necessary components of the research. This could be regarded as an argument in favor of the relevancy of and scientific justification for studying the derivation of soil moisture and temperature from remote sensing data.

In the papers mentioned above, particular natural objects, regions, or otherwise places of interest are used as study areas. The study by Parastatidis et al. is an exception, as in this case LST is being calculated for the entire globe. When it comes to estimating soil moisture, every mentioned study deals with a local and relatively small area. It allows assuming that a scientific gap exists for deriving soil moisture and temperature globally, but specifically for wetland soils. Placing the focus of the study in this way would allow for a necessary balance between specificity and generality, resulting in a global and reasonably accurate model.

2.2.2. Trapezoid model approach

The trapezoid model approach is well-described by Sandholt et al. (2000). In the study, it is linked to the Temperature-Vegetation Dryness Index (TVDI), which follows the trapezoid concept, and is used to predict soil moisture.

The idea behind TVDI is to obtain information about soil moisture through the relationship between LST and NDVI. Their combination provides information on vegetation and moisture conditions at the surface. A scatterplot of remotely sensed LST and a vegetation index often results in a triangular shape (Price 1990, Carlson et al. 1994), or a trapezoid shape (Moran et al. 1994) if a full range of fractional vegetation cover and soil moisture contents is represented in the data. A simplified LST/NDVI plot is shown in Figure 1. According to Friedl & Davis (1994), at least for well-watered surfaces, the relationship between LST and NDVI is more directly related to surface soil moisture, through the increase of the thermal inertia of the soil with surface soil moisture rather than as a limiting control on latent heat.

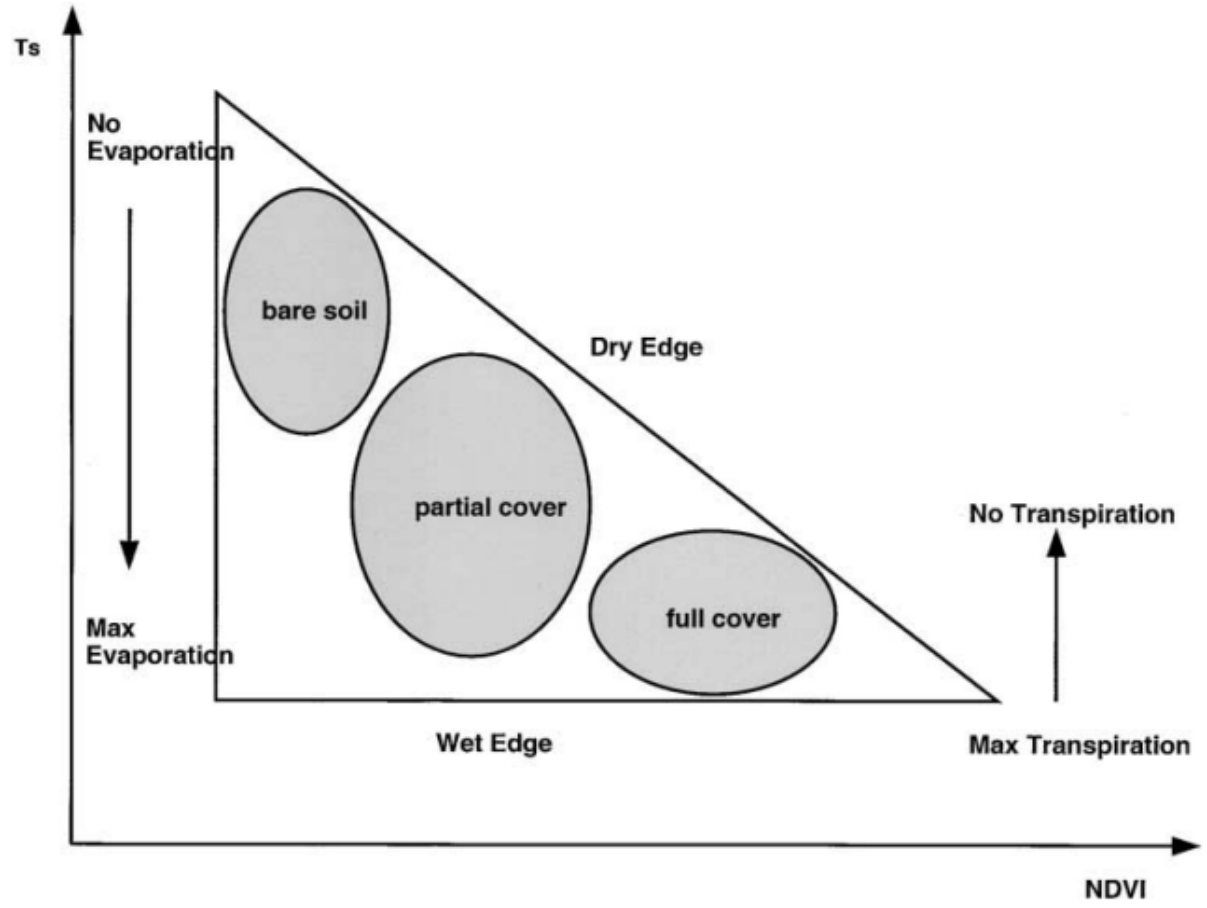

Figure 1. Simplified LST/NDVI plot (after Lambin & Ehrlich, 1996); Ts = LST.
Source: Sandholt et al. (2000)

In Figure 1, the left edge represents bare soil, ranging from top-down dry to wet. With an increase in NDVI, maximum LST decreases. The upper ("dry") edge represents dry conditions, and the lower ("wet") edge represents wet conditions.

Based on findings in the literature, Sandholt et al. (2000) suggested the following mechanisms as those determining a sample location in the LST/NDVI plot: fractional vegetation cover, evapotranspiration, thermal properties of the surface, net radiation, and atmospheric forcing and surface roughness. The factors are summarized in Figure 2. It is stated that, although no direct relation between surface temperature and surface soil moisture is evident, soil moisture is a critical factor in the mechanisms involved.

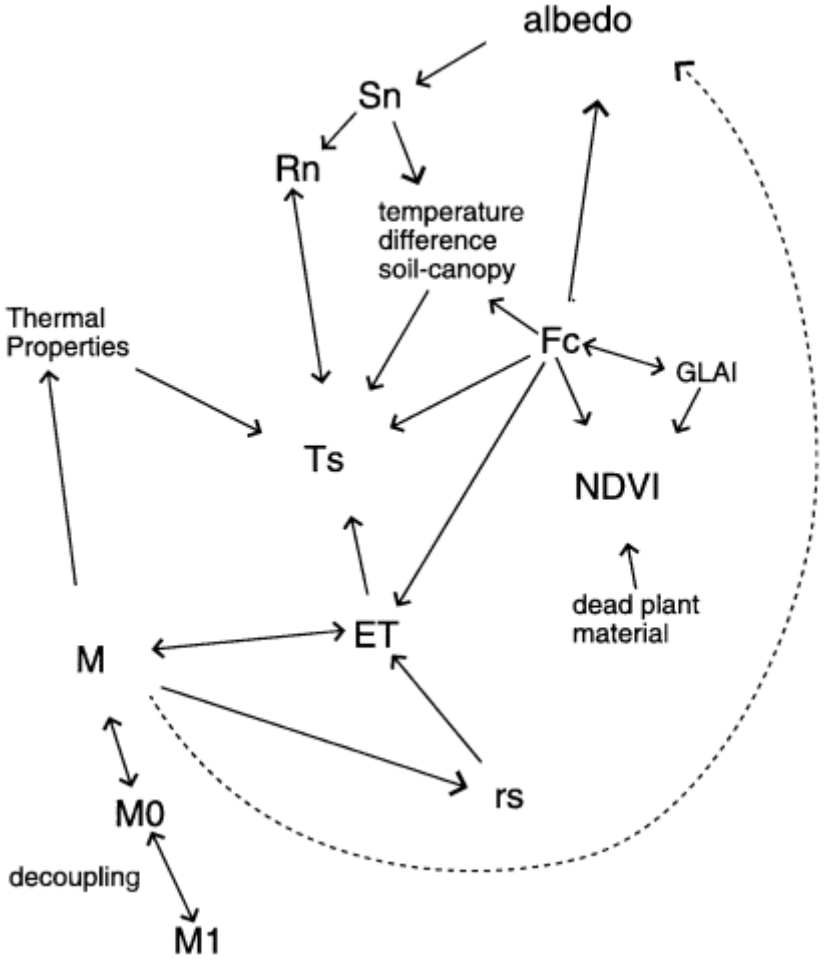

Figure 2. Factors determining LST. Ts = LST; Sn = shortwave net radiation balance; Rn = net radiation balance; GLAI = green leaf area index; Fc = fractional vegetation cover; ET = evapotranspiration; rs = stomatal resistance; M1 = soil moisture content (root zone); M0 = top soil moisture content.
Source: Sandholt et al. (2000)

Based on those concepts, Sandholt et al. (2000) proposed TVDI, defined as:

$$TVDI = \frac{Ts - Tsmin}{a + bNDVI - Tsmin}$$

where *Tsmin* is the minimum LST in the triangle, *Ts* is the observed LST in the sample, *NDVI* is observed in the sample, *a* and *b* are defined to fit the dry edge (*Tsmax* = *a* + *bNDVI*), and *Tsmax* is the maximum LST for a given NDVI. TVDI ranges from 0 (wet edge) to 1 (dry edge). It is assumed that the main source of LST variation is soil moisture, which implies that the method should be applied locally, with climatic and orographic conditions within the research area being homogenous. In the case of Sandholt et al. (2000), the study area is 140x140 km^2 square with generally flat relief in northern Senegal. The researchers concluded that TVDI was closely related to surface soil moisture simulated with the MIKE SHE model (R^2 = 0.70), and described the results as "promising", despite finding the estimation of TVDI parameters in the dry season problematic.

The trapezoid model has its limitations. In the study by Burdun et al. (2020), an LST-based trapezoid model (used in the research under the name Thermal-Optical TRApezoid Model, TOTRAM) was tested on two Estonian bogs, Linnusaare and Männikjärve. The study discovered a lack of correlation between TOTRAM and water table depth, which is strongly linked to soil moisture. An explanation was suggested, that evapotranspiration and vegetation growth in boreal climates are much more limited by energy rather than by moisture (McVicar et al. 2012). This limitation of TOTRAM was demonstrated by Garcia et al. (2014), and it was suggested for TOTRAM to be applied only to the regions where water is a limiting factor rather than energy.

Hitherto, a trapezoid model approach was only used for local studies. Hence, an attempt at covering global data would mean, idiomatically speaking, navigating uncharted waters, and would thus have a scientific value.

3. Data and methodology

3.1. Data

The research implies establishing relations between wetland soil temperature and water content data on the ground and satellite imagery. For this reason, the input data can be divided into ground data and satellite data.

3.1.1. Ground data

The ground data is an aggregate dataset, composed for the paper by Pärn et al. (2018). The dataset features 1544 samples of wetland soils from 58 different sites worldwide (Figure 3) and includes 41 parameters, including soil temperature at different depths and soil water content. The sites are located in tropical (A), temperate (C), and continental (D) climatic zones. They correspond to various landscapes: high mountains, lowlands, mountain valleys, and river valleys. The sites include lands in active agricultural use (arable, plantations, pastures), natural landscapes, and peat extraction sites. Data was collected by climate monitoring stations between 2011 and 2017 during vegetation periods. The list of sites is shown in Annex 1.

The following parameters are particularly important to the study:

1) *Coordinates.* Although one site may contain more than 40 samples that were collected from different spots of the same site, all samples within each <100m site transect share the same coordinates, and distance between the samples is assumed to be within the accuracy of satellite positioning.
2) *Volumetric water content (VWC, m^3/m^3).* The main measurement of soil moisture in this study.
3) *Temperature, °C, measured at 10 cm depth.* The main measurement of soil temperature in this study.
4) *Air temperature at the nearest weather station in the warmest month of the year, °C.* Although not ground data per se and acquired from KNMI Climate Explorer (http://climexp.knmi.nl), it is present in the ground dataset for normalizing the soil temperature to local climatic maxima (Pärn et al. 2018).

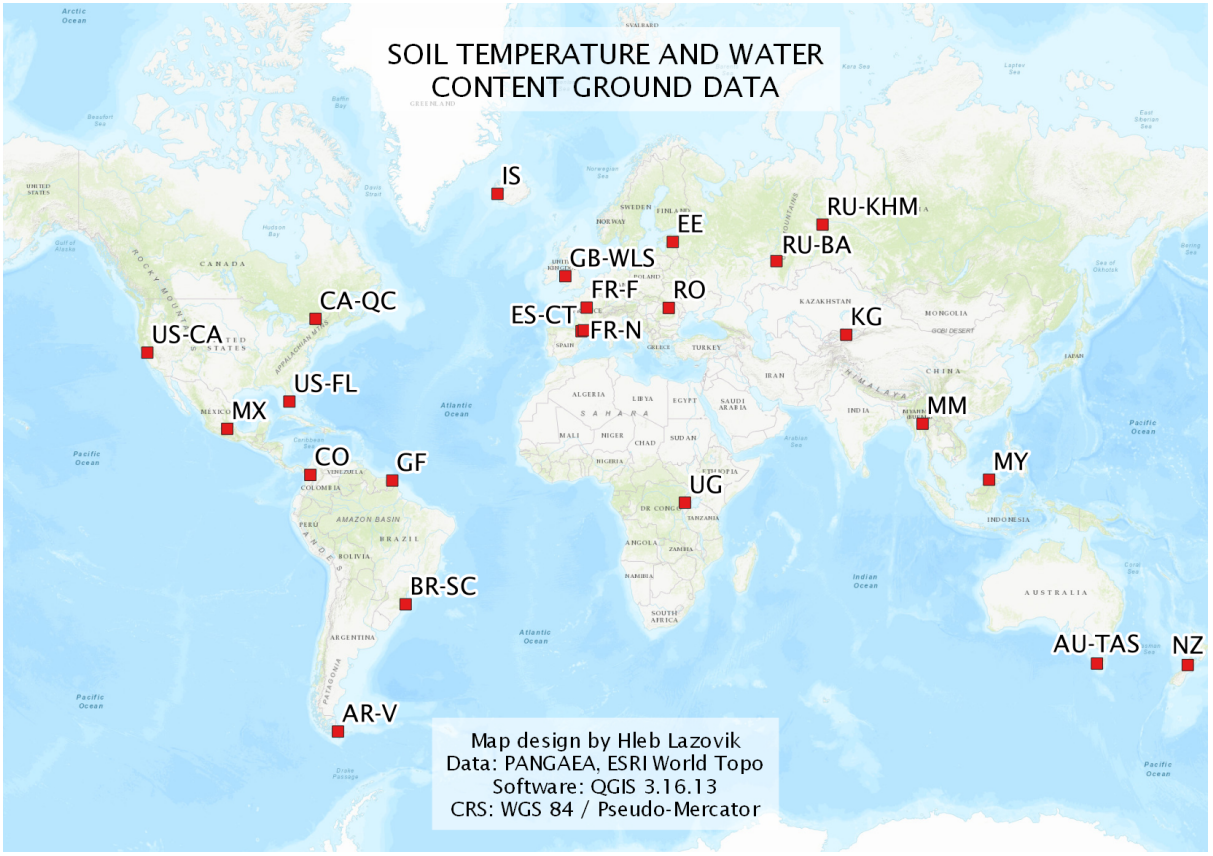

Figure 3. Location of the ground data.
Dataset source: Pärn (2018)

3.1.2. Satellite data

The satellite data used in the research is Landsat 8 imagery. It is acquired in moderate spatial resolution (30 m for OLI, 100 m for TIRS) in the visible, near-infrared, short wave, and thermal infrared wavelengths, organized into 11 bands (Figure 4). Landsat 8 imagery covers the Earth's terrestrial and polar regions. It is processed into 185 km × 180 km Level 1 terrain-corrected products that have a typical 950 MB compressed GeoTiff file size (Roy et al. 2014). In addition, Landsat 8 Level 2 Science Products contain Surface Reflectance, Surface Temperature, Intermediate bands used in calculation of the Surface Temperature products, and quality assessment masks (Landsat missions, 2022). Of particular importance to the research is Surface Temperature.

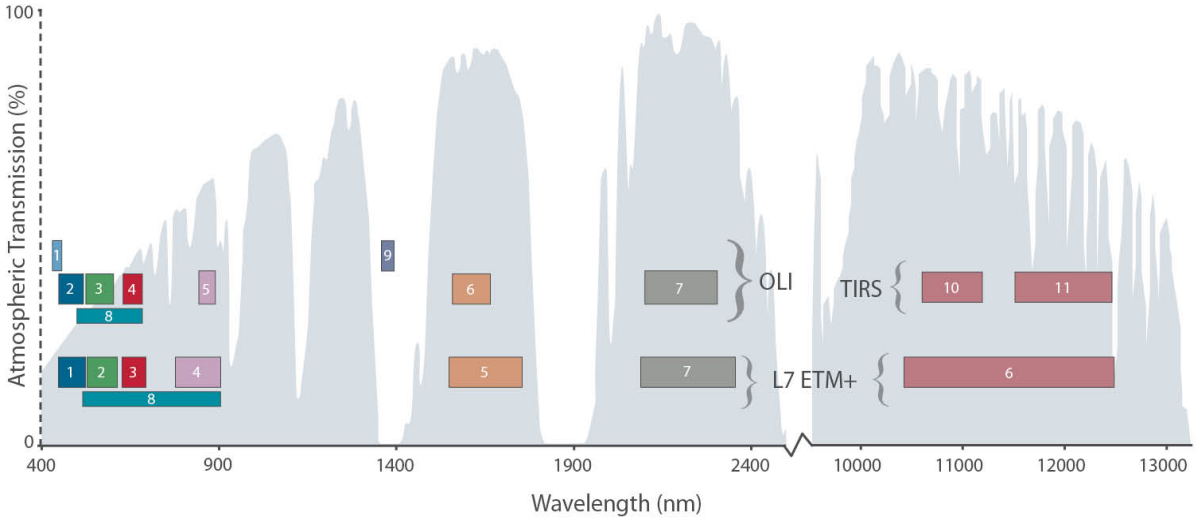

Figure 4. Landsat 8 spectral bands compared to Landsat 7.
Source: Landsat 8 | Landsat Science
https://landsat.gsfc.nasa.gov/satellites/landsat-8/

Images with certain properties were picked for the research. They were chosen individually for each site (although in most cases the same picture could be applied to multiple neighboring sites), taken in the same month of the year as the samples, although not necessarily in the same year (in most cases, the difference is up to two years). It was aimed for the cloud cover not to exceed 10% of the image, but for several sites, it was not possible. Every picture was manually checked so that the area of interest was not covered by clouds. In total, 29 images were used for 58 sites.

3.2. Methods

The simplified methodology of the research is shown with a flow chart (Figure 5). It represents the main steps taken during the research and includes the datasets, the processes, and the used software. The flow chart is created with Lucidchart for Google Workspace.

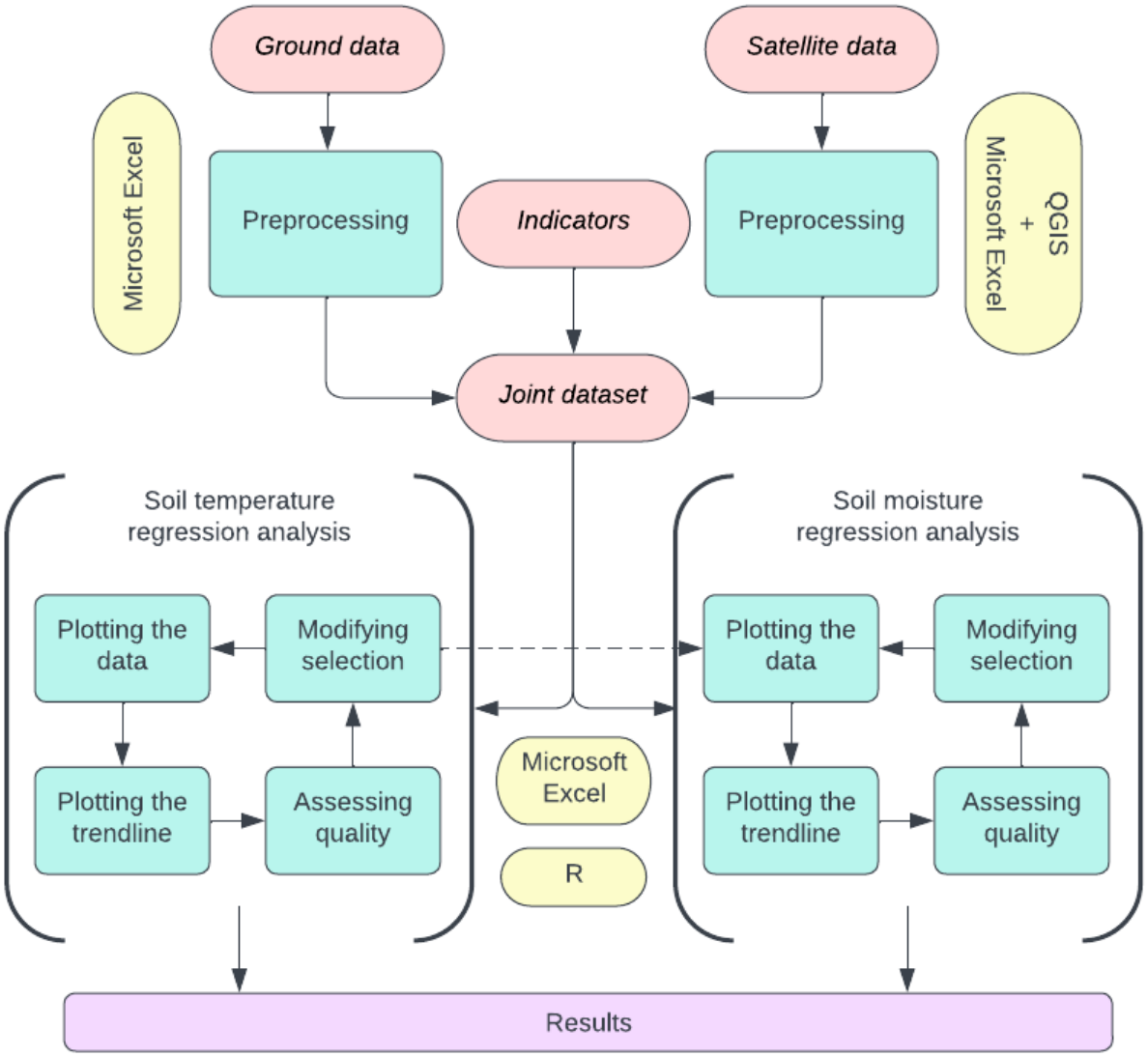

Figure 5. Flow chart of the methodology.

3.2.1. Data preprocessing

Due to the accuracy of positioning (section 3.1.1), the decision was made to operate with average parameter values per site, rather than with individual samples. For each site, an average value of every quantitative parameter was calculated for further use in the analysis. The calculations were made in Microsoft Excel for Microsoft 365 MSO.

The loss of ground data quality related to averaging samples with ranging values was estimated by calculating the standard deviation of values within a site. The average standard deviation of VWC is 0.13 m^3/m^3, and the average standard deviation of temperature at 10 cm depth is 1.43 °C. With average site values of VWC ranging from 0.17 m^3/m^3 to 0.96 m^3/m^3, and average site values of temperature at 10 cm depth from 3.38 °C to 29.37 °C, such values of average standard deviation within sites are expected to, at least to some extent, contribute to the model's inaccuracy.

In order to collect the band values from the satellite images, polygons were created, corresponding to the site locations. The creation of polygons was manual, with several rules being followed: the polygon should correspond to the site's location (although not necessarily exact coordinates), Vegetation, and Land use values (identified visually with images from Landsat and Esri World Imagery), and contain a visually homogenous area, in order to decrease the noise. The average polygon size is 0.09 km^2, with polygon sizes ranging from 0.001 km^2 to 1.08 km^2. The band values (Band 1 to 7), as well as Surface Temperature values, were extracted as average within a polygon. Both polygon creation and extraction of values were made in QGIS.

After the extraction, the band values were normalized according to the Landsat documentation (Landsat missions, 2022). For the band values, this meant applying the multiplicative scale factor 0.0000275 and the additive offset -0.2. For the Surface Temperature, the multiplicative scale factor was 0.00341802, and the additive offset was 149. The resulting Surface Temperature values represented absolute surface temperature in Kelvin, and prior to the analysis, they were converted to Celcius. The normalization was made in Microsoft Excel for Microsoft 365 MSO.

3.2.2. Remotely derived indicators

When assessing a physical parameter with the use of satellite imagery, it is common to organize the remotely sensed data into indicators of various complexity. Those could be the values of a single band or an index calculated from two or more bands, and in some cases additional data. Several indicators were used in this study.

Surface Temperature. Post-processed values derived from Landsat 8 Level 2 Science Products represented an estimate of surface temperature in Celcius, thus being fit for the regression analysis without further calculations, both as a source of LST values and as an

indicator of soil temperature measured at 10 cm depth (with a 10 cm difference between the physical meaning of the two parameters being an unwanted but ultimately accepted compromise).

NDMI. Normalized Difference Moisture Index, a seemingly logical choice to predict moisture, is highly correlated with canopy water content (Klemas & Smart 1983). For Landsat 8, calculated as (Band 5 – Band 6) / (Band 5 + Band 6).

MSI. Moisture Stress Index is used for canopy stress analysis, productivity prediction, and biophysical modeling, with higher values of the index indicating greater plant water stress and in inference, less soil moisture content (Welikhe et al. 2017). For Landsat 8, calculated as Band 6 / Band 5.

NDWI. Normalized Difference Water Index is a measure of liquid water molecules in vegetation canopies that interacted with the incoming solar radiation, and is sensitive to the total amounts of liquid water in the stacked leaves (Gao 1996). For Landsat 8, calculated as (Band 3 – Band 5) / (Band 3 + Band 5)

*ΔT*NDVI.* The index is proposed in this research as an attempt to apply the trapezoid model approach to a global dataset. It is calculated as (Tnorm - LST) * NDVI, with Tnorm being a local temperature normalizing LST and contextualizing it in the local climatic conditions. In the research, air temperature at the nearest weather station in the warmest month of the year is used as Tnorm, with the main justification being its use as a means to normalize the soil temperature to local climatic maxima by Pärn et al. (2018), as well as its convenient presence in the ground dataset.

The way the index uses the trapezoid approach is shown in Figure 6. Although the LST/NDVI trapezoid values are strictly local and exist in the context of the local climatic conditions, in most cases Tnorm is expected to be found between the dry edge and the wet edge. The difference between Tnorm and LST (ΔT) shows how far LST is from the "normal" value (and thus, how close it comes to the dry or the wet edge). Since the wet edge has a significant slope, NDVI is present in the formula to magnify ΔT. Thus, high NDVI values would mean that even relatively small ΔT would result in a high variance of soil moisture. Negative values would indicate relatively dry conditions (closer to the dry edge), and positive values would indicate relatively wet conditions (closer to the wet edge).

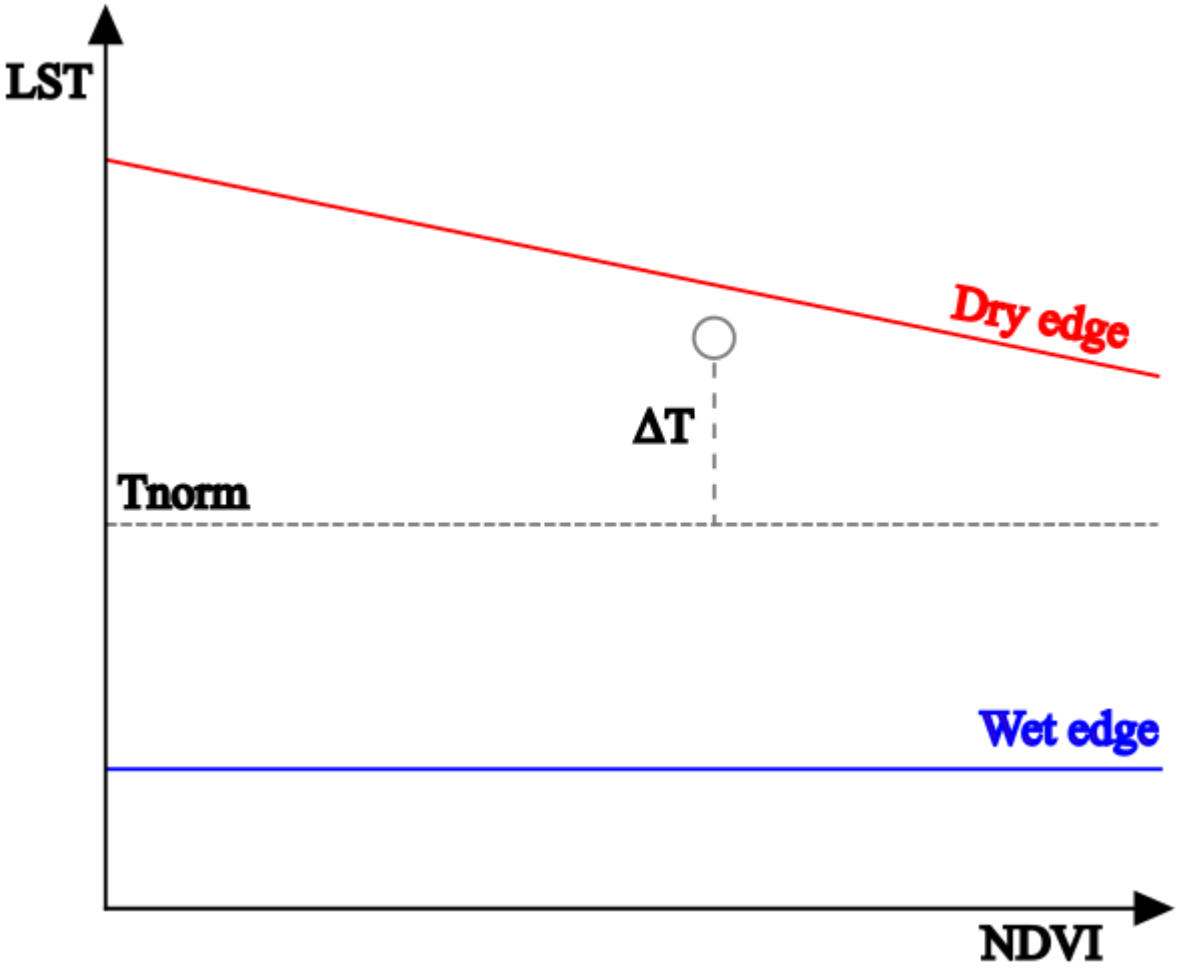

Figure 6. Visual concept of ΔT*NDVI.

3.2.3. Regression analysis

The regression analysis was performed in Microsoft Excel for Microsoft 365 MSO, as well as using R, separately for soil temperature and moisture. The steps included plotting and visualizing the data, plotting the trendline, assessing the capability of one set of variables to predict the other with the use of the R^2 coefficient, and managing the data with a goal to increase the R^2 values in a non-arbitrary way (most commonly through data filtering). The analysis had several iterations, with solutions considered optimal being presented in this paper.

During the course of the analysis as well as theoretical research, it became clear that predicting soil temperature shows significantly more accurate results compared to predicting soil moisture. Additionally, filtering out data coming from two locations (Kyrgyzstan and Romania, 7 sites in total) resulted in a striking improvement of R^2 values for predicting soil temperature (from 0.32 to 0.70), from which a conclusion was made that, for one reason or

another (for example, the date of the samples' collection and the satellite imagery being non-identical, or potential errors in locating the sample polygons) the connection between those samples and remotely obtained data was faulty. Thus, soil temperature analysis was used as a filtering guide for soil moisture analysis, and the sites were excluded from both calculations.

For soil moisture regression analyses, involving the trapezoid approach, the climate was used as a filtering parameter, although in a separate iteration. To address the theoretically poor performance of LST-based approaches to predict soil moisture in energy-limited climates, a set of calculations was made that excluded samples from boreal regions, in which climates were classified according to the Köppen climate classification as D (continental). Another iteration of calculations was made, excluding samples with bare soil: given the importance of NDVI in the proposed formula, as well as its physical meaning, it is assumed that the presence of vegetation in the sample is necessary for assessing soil moisture. Wetland soils are not generally expected to be bare, thus, this exclusion is assumed not to contribute negatively to the formula's applicability.

4. Results

4.1. Soil temperature analysis

For the soil temperature analysis, remotely derived and pre-processed LST values were used as a dependent variable, and soil temperature values measured at various depths were used as independent variables. Soil temperature measured at 10 cm depth was given special importance, while temperatures measured at lower depths were put into analysis to check an expected outcome, in which the ability of remotely-derived LST to predict soil temperature decreases with increasing depth.

In the first, unfiltered iteration of the soil temperature at 10 cm depth linear regression analysis, the R^2 value amounted to 0.32 (Figure 7). In the second iteration, two locations, in which the connection between the ground samples and remotely obtained data was presumed faulty (seven sites in total), were filtered out of the calculation. Much better results were obtained, with the R^2 value of 0.70 (Figure 8). Filtering parameters were kept for further analyses of the soil temperature at 20 cm, 30 cm, and 40 cm depths (Figures 9, 10, 11).

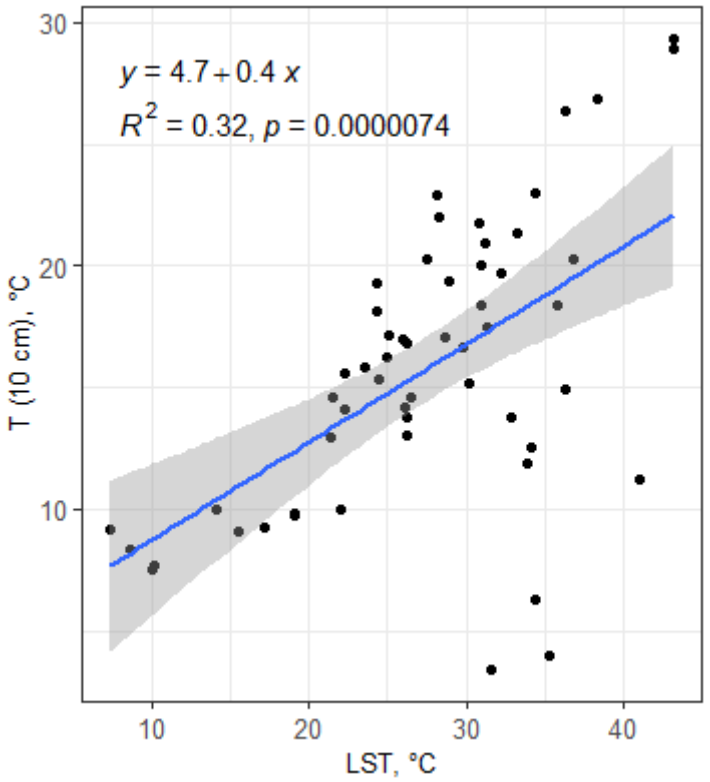

Figure 7. Soil temperature at 10 cm depth linear regression analysis, first iteration.

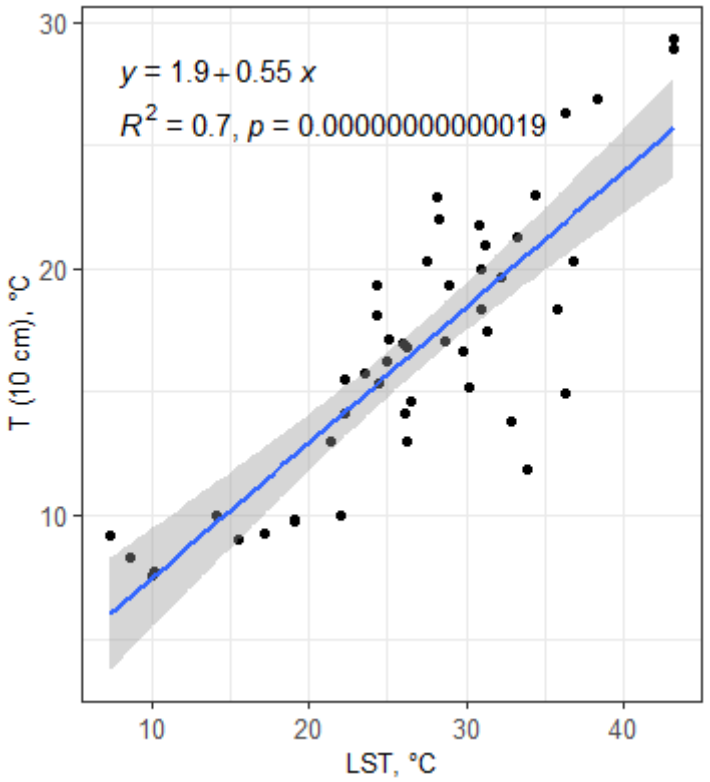

Figure 8. Soil temperature at 10 cm depth linear regression analysis, filtered.

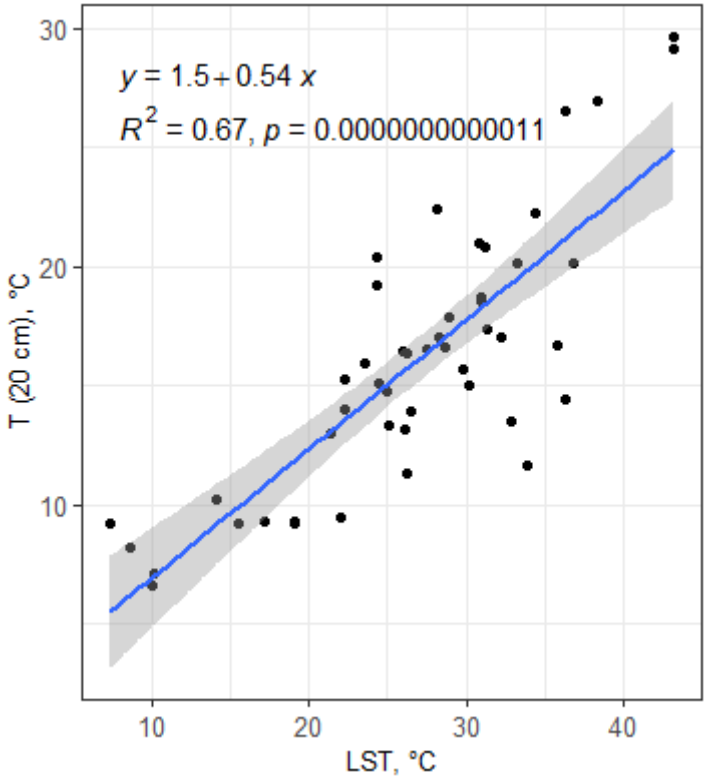

Figure 9. Soil temperature at 20 cm depth linear regression analysis, filtered.

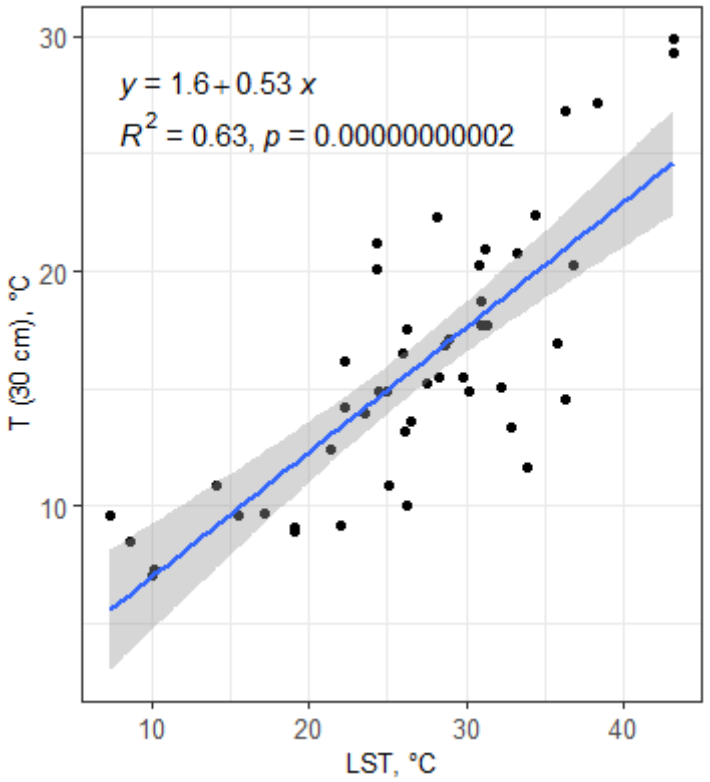

Figure 10. Soil temperature at 30 cm depth linear regression analysis, filtered.

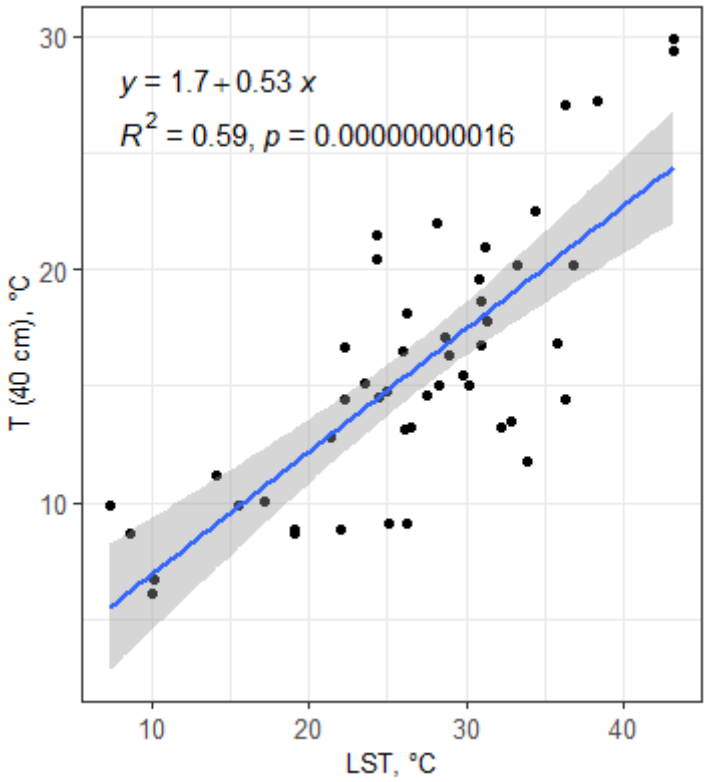

Figure 11. Soil temperature at 40 cm depth linear regression analysis, filtered.

Evidently, an expected outcome, in which the ability of remotely-derived LST to predict soil temperature decreases with increasing depth, is observed. The calculated R^2 values at various depths are shown in Table 1.

Table 1. Soil temperature linear regression analysis results at various depths.

Depth, cm	R^2
10	0.70
20	0.67
30	0.63
40	0.59

Overall, observing the R^2 and p-values, it could be said that the connection between the remotely derived data and the ground data is reasonably strong.

4.2. Soil moisture analysis

For the soil moisture analysis, remotely derived indices (NDMI, MSI, NDWI, and proposed ΔT*NDVI) were used as dependent variables, and volumetric water content (VWC, m3/m3) was used as an independent variable. The same filtering options as in the soil temperature analysis were applied. In addition, separate filterings were made to exclude the samples from boreal regions, as well as samples with bare soil.

4.2.1. NDMI, MSI, and NDWI as indicators

In the first iteration of the soil moisture analysis, already existing indices were picked, with an assumption of such a choice being a somewhat safe option.

Linear regression analysis with NDMI as a dependent variable has shown poor results, with the R^2 value of nearly 0.00 (Figure 12). Linear regression analysis with MSI as a dependent variable has shown results only marginally better compared to the previous one, with the R^2 value still near 0.00 (Figure 13). Linear regression analysis with NDWI as a dependent variable has performed poorly as well, with the R^2 value of nearly 0.00 (Figure 14).

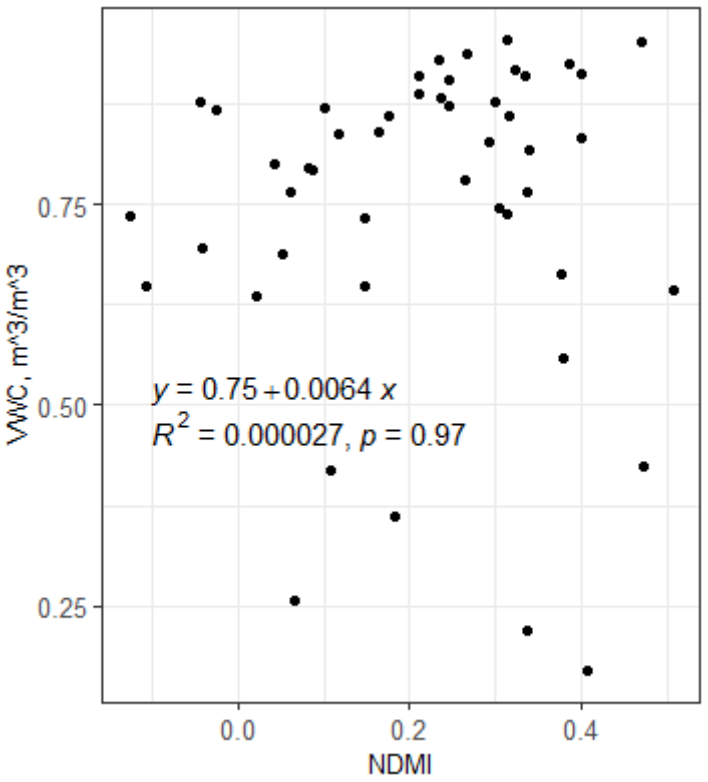

Figure 12. VWC linear regression analysis, NDMI as a dependent variable.

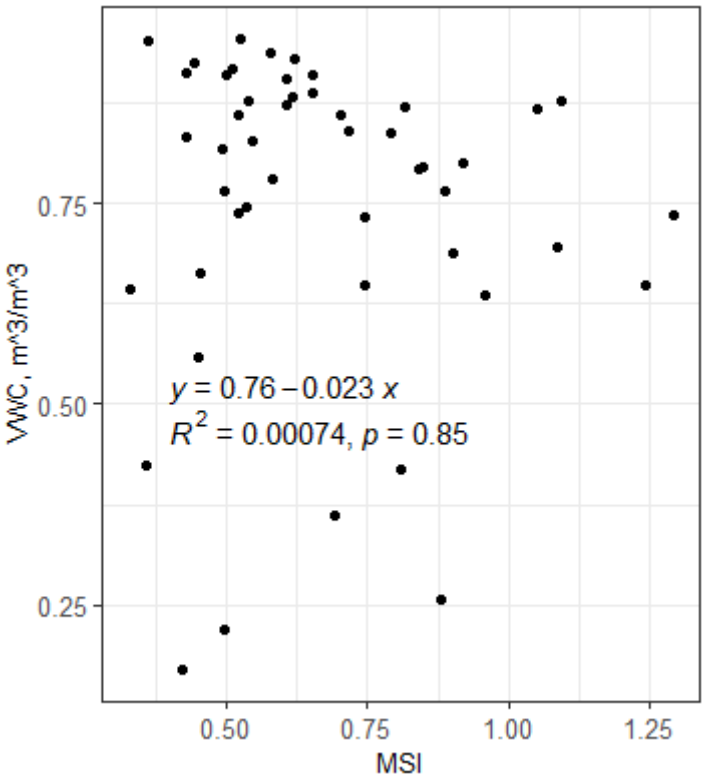

Figure 13. VWC linear regression analysis, MSI as a dependent variable.

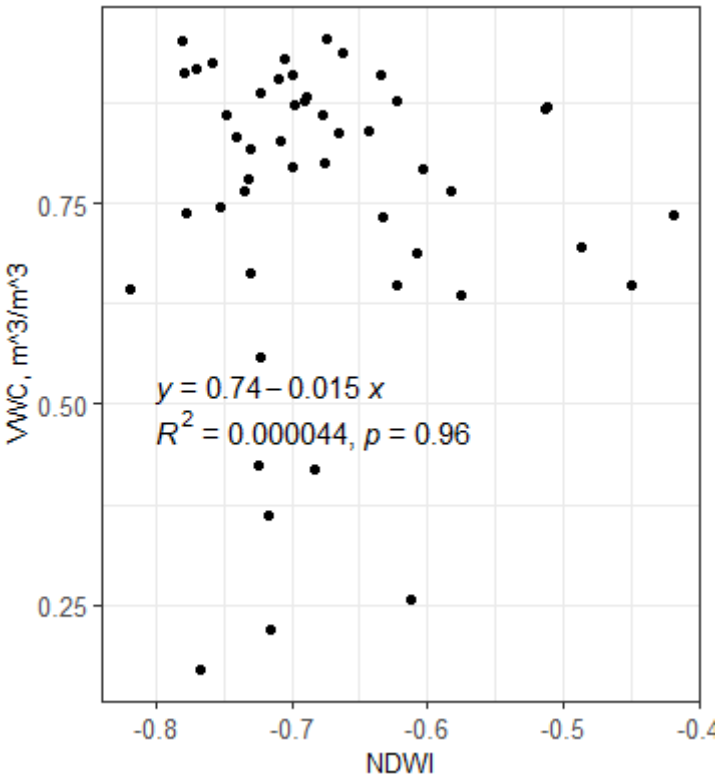

Figure 14. VWC linear regression analysis, NDWI as a dependent variable.

Conducting the soil moisture analysis with NDMI, MSI, and NDWI as dependent variables has demonstrated the poor suitability of those indices as soil moisture predictors within the context of the research. With p-values of over 0.8, a conclusion can be made that the null hypothesis is correct. Besides bringing a great deal of anxiety to the analyst, it served as an incentive to look for new and unorthodox approaches to the task at hand.

4.2.2. ΔT*NDVI as indicator

Linear regression analysis with ΔT*NDVI as a dependent variable, conducted with the same filtering parameters as the previously mentioned analyses utilizing NDMI, MSI, and NDWI, has shown significantly better results (Figure 15). The R^2 value of 0.16 is an evident improvement over 0.00, although still not an indicator of a good connection between the two variables.

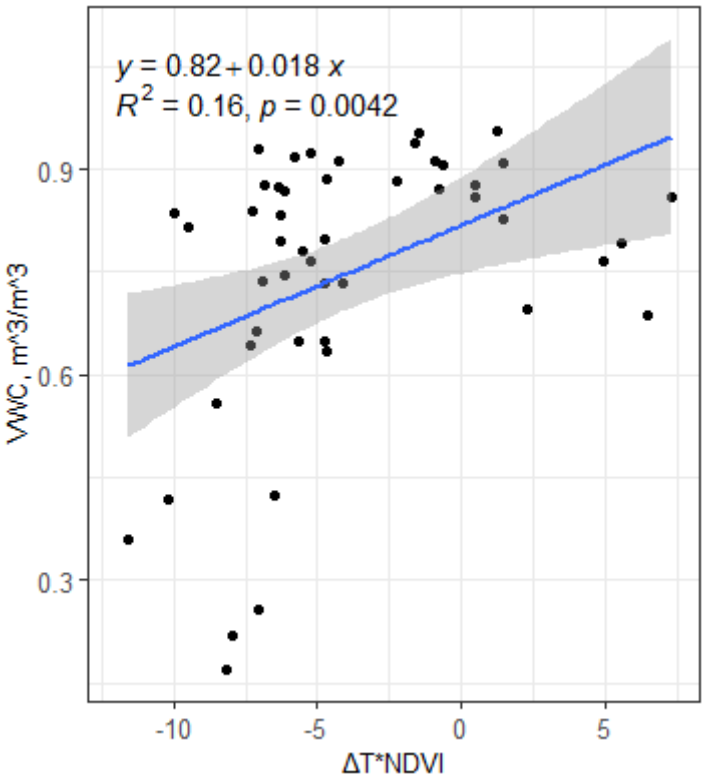

Figure 15. VWC linear regression analysis, ΔT*NDVI as a dependent variable.

Judging from the spread of samples on the graph above, a decision was made to attempt polynomial regression instead (Figure 16). The results improved significantly, with the R^2 value amounting to 0.31.

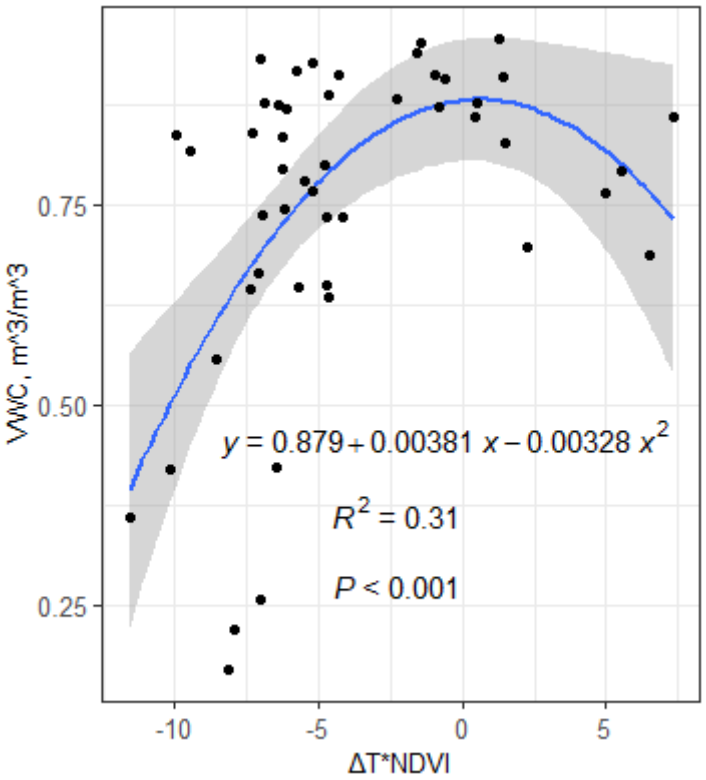

Figure 16. VWC polynomial regression analysis, ΔT*NDVI as a dependent variable.

Given the evident change of trend on the graph at the ΔT*NDVI values of around 0, and difficulties in justifying a polynomial connection between the two variables, it was proposed to apply a piecewise function instead. Due to the low amount of ΔT*NDVI values above 0, the connection in this segment of the graph was assumed to be insignificant. Thus, only the "negative" part of the graph was used in further analysis involving ΔT*NDVI (Figure 17). The R^2 value amounted to 0.30, only marginally worse compared to polynomial regression.

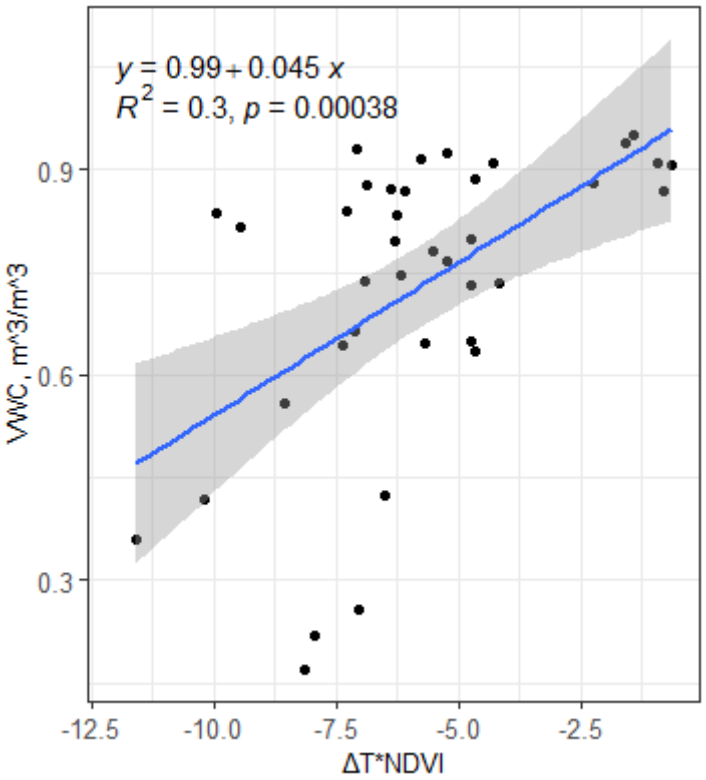

Figure 17. VWC linear regression analysis, ΔT*NDVI as a dependent variable, negative ΔT*NDVI values only.

4.2.3. Boreal regions excluded

A similar set of regression analyses with ΔT*NDVI as a dependent variable was conducted, with samples coming from climates classified according to the Köppen climate classification as D (continental) filtered out.

Linear regression analysis with boreal regions excluded has demonstrated a significant improvement compared to the one without boreal regions exclusion, evidenced by the R^2 value of 0.30 (Figure 18). Polynomial regression analysis has also shown significant improvement after filtering out boreal regions, with the R^2 value amounting to 0.40 (Figure 19).

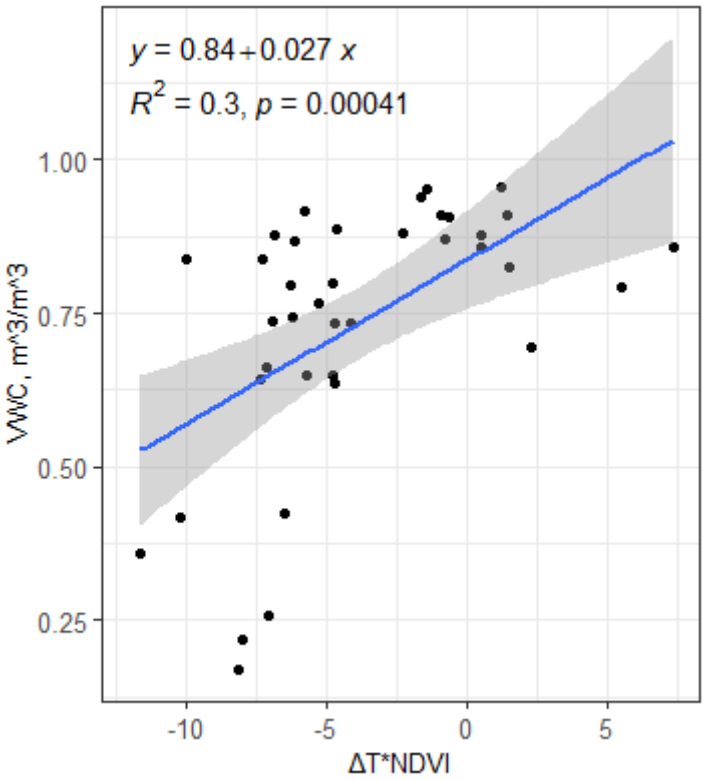

Figure 18. VWC linear regression analysis, ΔT*NDVI as a dependent variable, boreal regions excluded.

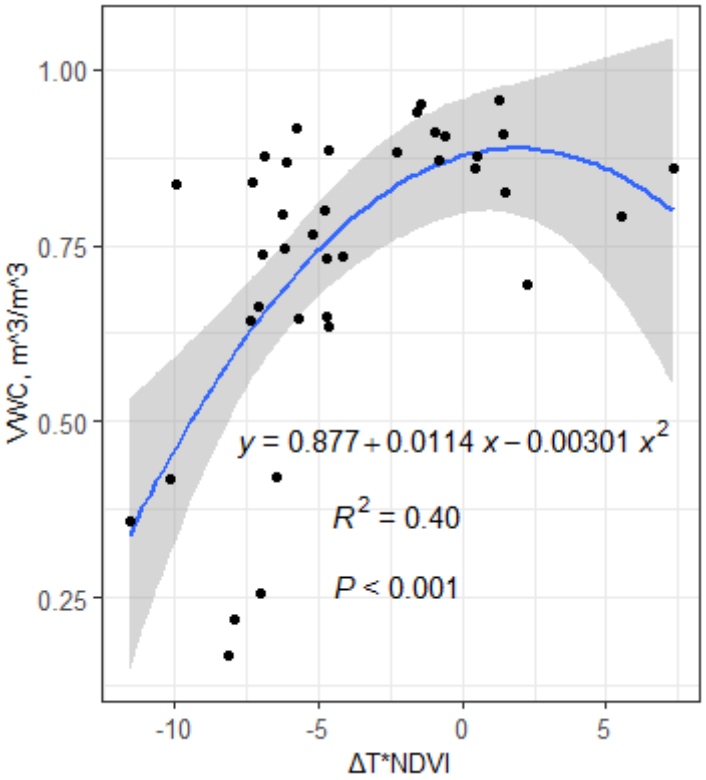

Figure 19. VWC polynomial regression analysis, ΔT*NDVI as a dependent variable, boreal regions excluded.

The pattern continues when conducting a linear regression analysis only in the "negative" part of the graph, with the R^2 value amounting to 0.38 (Figure 20).

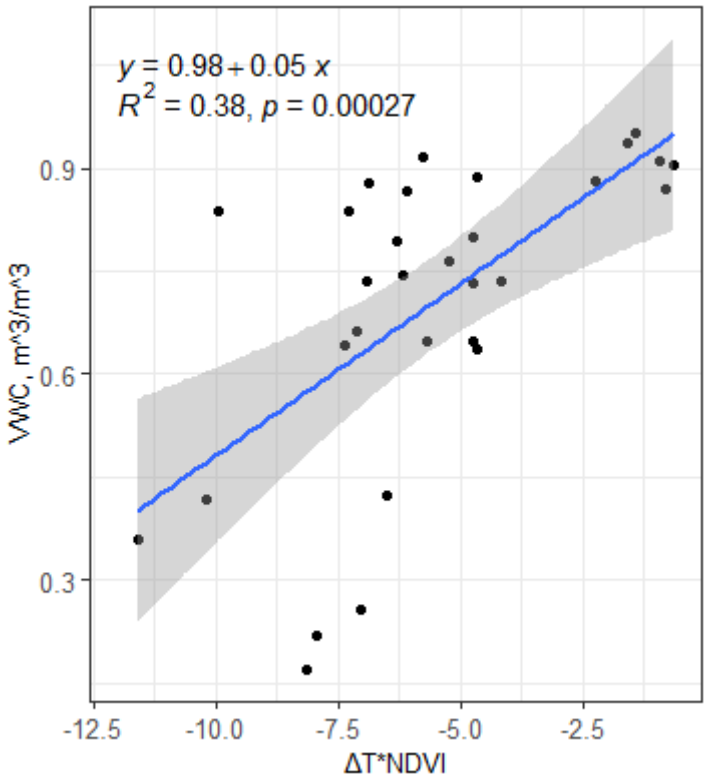

Figure 20. VWC linear regression analysis, ΔT*NDVI as a dependent variable, boreal regions excluded, negative ΔT*NDVI values only.

4.2.4. Bare soil excluded

Another iteration of analyses using ΔT*NDVI as a predictor was made, with samples with bare soil excluded from calculations.

Linear regression analysis with bare soil excluded has demonstrated results almost the same compared to the iteration without filtering, with the R^2 value of 0.18 (Figure 21). Polynomial regression analysis with bare soil excluded has shown some improvement compared to the iteration without filtering, although not as significant as the exclusion of boreal regions. The R^2 value amounted to 0.35 (Figure 22).

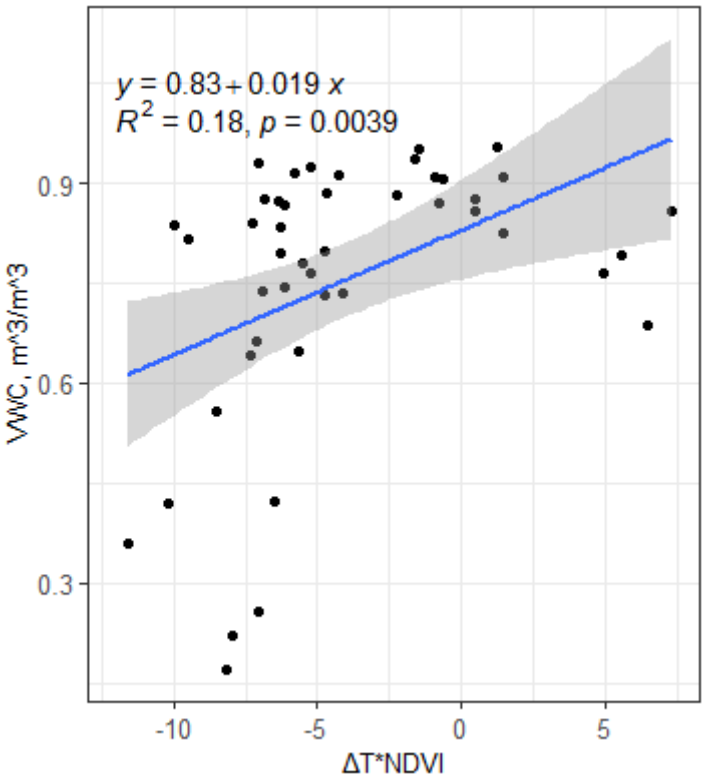

Figure 21. VWC linear regression analysis, ΔT*NDVI as a dependent variable, bare soil excluded.

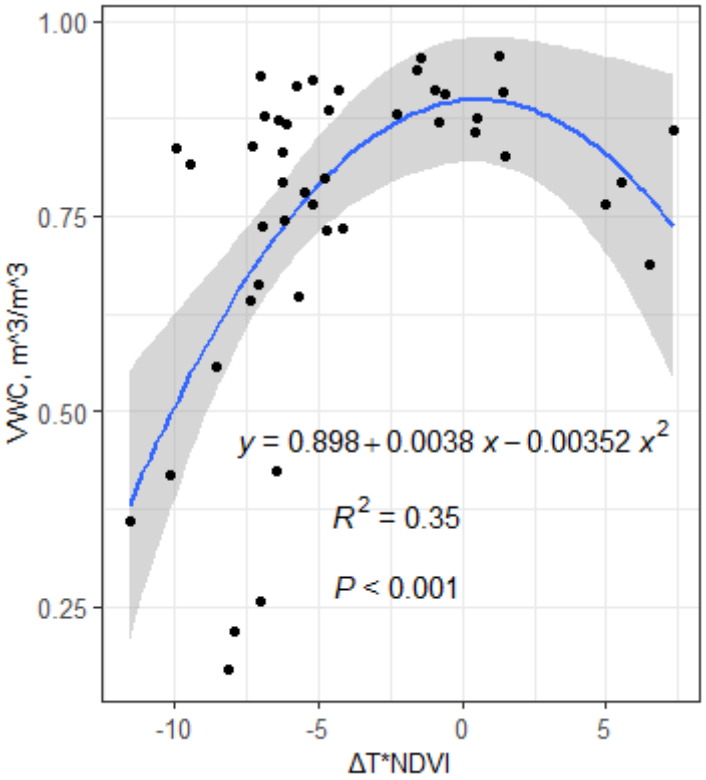

Figure 22. VWC polynomial regression analysis, ΔT*NDVI as a dependent variable, bare soil excluded.

The same pattern can be observed in the linear regression analysis only in the "negative" part of the graph. There the R^2 value amounted to 0.32 (Figure 23).

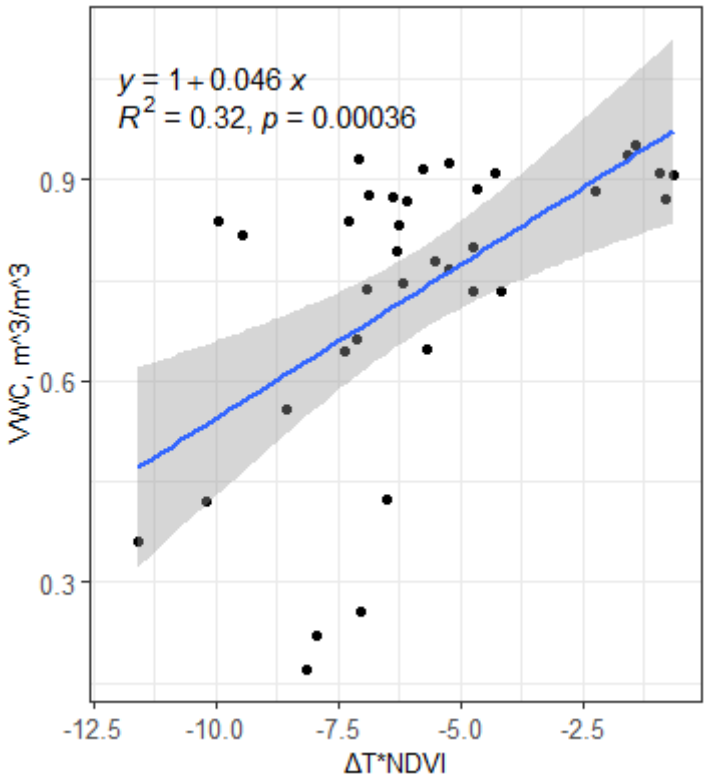

Figure 23. VWC linear regression analysis, ΔT*NDVI as a dependent variable, bare soil excluded, negative ΔT*NDVI values only.

Another iteration of a linear regression analysis was made, which excluded both boreal regions and bare soil samples, and where only the "negative" part of the graph was put into calculations. The results were the best among any linear regression analyses of soil moisture in this study, with the R^2 value of 0.39 (Figure 24).

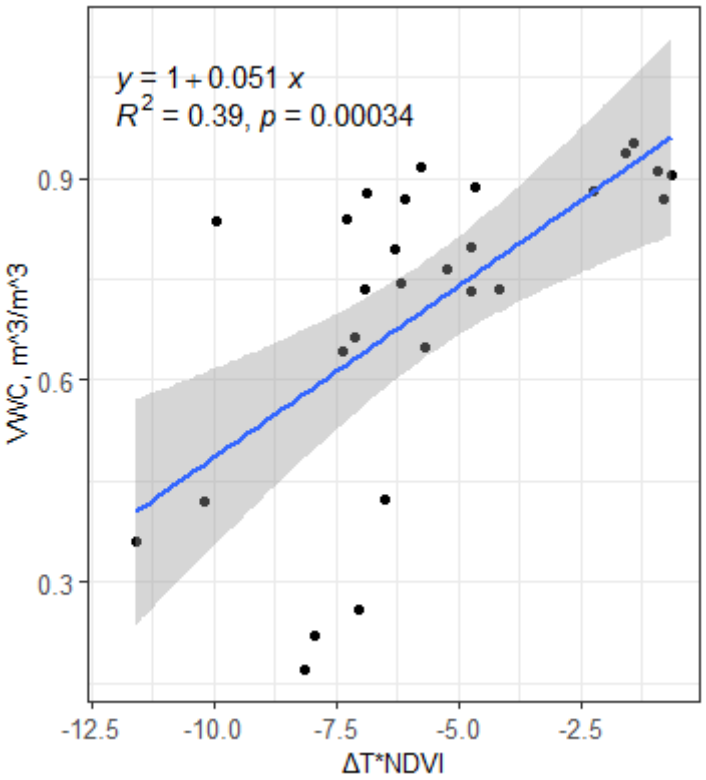

Figure 24. VWC linear regression analysis, ΔT*NDVI as a dependent variable, boreal regions excluded, bare soil excluded, negative ΔT*NDVI values only.

Evidently, the proposed ΔT*NDVI shows better results at predicting soil moisture compared to NDMI, MSI, and NDWI. The efficiency of ΔT*NDVI is better when the ΔT*NDVI values correspond to samples in non-boreal climates, and with vegetation. The calculated R^2 values of analyses with different predictors and filtering parameters are shown in Table 2. Excluding the same sites as in the soil temperature analysis is shown as a "Standard" filtering option, with other filtering options excluding the two "faulty" sites in addition to their respective filtering parameters.

Overall, observing the R^2 and p-values, it could safely be said that NDMI, MSI, and NDWI have shown exceptionally poor results as soil moisture predictors. A proposed ΔT*NDVI performs better, and in this case, p-values indicate very strong evidence against the null hypothesis. The best-observed R^2 result amounted to 0.40. While still not a particularly high value (for instance, when compared to the 0.70 value observed in the soil temperature analysis, using the same dataset and satellite images), a significant improvement over NDMI, MSI, and NDWI shows value in the proposed approach and can amount to a subject of discussion and further examination.

Table 2. Soil moisture regression analysis results.

Predictor	*Regression analysis type*	*Filtering options*	R^2
NDMI	Linear	Standard	0.00
MSI	Linear	Standard	0.00
NDWI	Linear	Standard	0.00
ΔT*NDVI	Linear	Standard	0.16
ΔT*NDVI	Polynomial	Standard	0.31
ΔT*NDVI	Linear	Positive ΔT*NDVI values excluded	0.30
ΔT*NDVI	Linear	Boreal sites excluded	0.30
ΔT*NDVI	Polynomial	Boreal sites excluded	0.40
ΔT*NDVI	Linear	Boreal sites excluded, positive ΔT*NDVI values excluded	0.38
ΔT*NDVI	Linear	Bare soil excluded	0.18
ΔT*NDVI	Polynomial	Bare soil excluded	0.35
ΔT*NDVI	Linear	Bare soil excluded, positive ΔT*NDVI values excluded	0.32
ΔT*NDVI	Linear	Bare soil excluded, boreal sites excluded, positive ΔT*NDVI values excluded	0.39

5. Discussion and conclusions

The main subject of discussion are the R^2 values, relative to each other as well as absolute numbers. Evident differences in the R^2 values are indicative of the different abilities of dependent variables to predict independent variables. One of the necessary goals of the discussion is thus to propose a physical explanation for those differences. The other goals are assessing the overall quality of the predictors, deriving conclusions from the proposed physical explanations, and proposing ways to improve the results.

5.1. Soil temperature analysis

The results of the soil temperature analysis are addressed first. The R^2 values there range from 0.59 (40 cm depth) to 0.70 (10 cm depth). Given that the predicted value is attributed to LST, and the ground data is attributed to the temperature at some depth, some lack of accuracy was expected. Given that the R^2 values decrease with depth, it can be said that the expectations were proven, and that, if there were LST values in the ground dataset, they would demonstrate a better connection with remotely obtained LST values. A simple linear interpolation suggests that, on the surface, the R^2 values would be 0.74, however, the relevance of using linear interpolation, in this case, is unproven, and thus 0.74 should be considered a rough and imprecise estimate. In any case, $R^2 = 0.70$ obtained from the soil temperature at 10 cm depth linear regression analysis shows a good connection between the dependent and independent variables.

5.2. Soil moisture analysis

The first point of discussion of the soil moisture analysis is the performance of NDMI, MSI, and NDWI. The R^2 values of 0.00 indicate the absence of connection between the VWC values in the dataset and those indicators, meaning that they are unusable in the context of the methodology of the research. The connection between one of those indices (MSI) and soil moisture was one of the subjects of the research by Harris & Bryant (2009). The correlation was significant; however, the responses were vegetation-type specific according to different water transport capacities. In the case of the current research, the species specificity condition was not met, with the indices being applied to samples with different vegetation (and, in some cases, without vegetation whatsoever). This explains the poor performance of MSI, and might possibly explain the poor performance of NDMI and NDWI as well.

The performance of the proposed ΔT*NDVI is a separate point of discussion. Judging from comparing the R^2 values, several conclusions can be made. ΔT*NDVI is evidently a better predictor of soil moisture than NDMI, MSI, and NDWI (within this geographical scope, climatic variety, and methodology). ΔT*NDVI predicts soil moisture less accurately than Landsat 8 Level 2 Surface Temperature predicts soil temperature. The performance of ΔT*NDVI as a soil moisture predictor is better when the studied surface has vegetation, and significantly better in non-boreal climates. Providing the polynomial relation between ΔT*NDVI and soil moisture with a physical explanation is challenging, and within this research, this relation is assumed to be dubious. Instead, the conclusion is made that the "positive" and "negative" parts of the ΔT*NDVI / VWC graph show separate patterns. Due to a lack of samples on the "positive" part, only the "negative" part was examined. A linear relationship between ΔT*NDVI and VWC was observed. It is also evident that the exclusion of boreal sites significantly improved the ΔT*NDVI performance, while exclusion of samples with bare soil only improved the performance incrementally.

All in all, ΔT*NDVI is a better choice than NDMI, MSI, and NDWI. In future research, it is recommended to explore further limitations of ΔT*NDVI (notably, in the "positive" part of the graph), as well as to experiment with other Tnorm values. In this research, air temperature at the nearest weather station in the warmest month of the year was used as Tnorm, and the lack of significant theoretical backing for this choice is evidence of its somewhat arbitrary nature. The global ΔT*NDVI approach has two slight limitations: the study areas should have vegetation and be located in a reasonably warm climate, in which the limiting factor in

evapotranspiration and vegetation growth is water, rather than energy. Needless to say, the ability to derive the aforementioned conclusions and suggest further directions of research indicate the presence of scientific value in this study.

5.3. Conclusions

To conclude the study, the research questions set in its introduction are to be addressed. Based on the answers, further conclusions can be made.

The study assessed the agreement between the ground measurements of soil temperature and soil moisture and Landsat-derived parameters. In the case of soil temperature, the connection was estimated to be reasonably strong ($R^2 = 0.70$), and in the case of soil moisture, it was fair ($R^2 = 0.39$; $p < 0.001$).

The global model deriving soil temperature from Landsat data already exists, and is part of the Landsat 8 Level 2 Science Products. In the context of being applied to a global wetlands dataset, the model performed well. The model proposed for deriving soil moisture from Landsat data has slight limitations (in bare soils and in boreal regions) which make its global application harder.

Judging from the R^2 and p-values, the soil temperature model can be assessed as somewhat accurate. The soil moisture model, however, can partly predict VWC values, however, if provided with further formula adjustments and more data to be calibrated on, the model might be improved to perform better and thus has scientific potential.

Reflecting on the answers to the research questions, it can be said that reasonably accurate Landsat-derived models were presented for predicting wetland soil temperature and moisture. The models are original, perform significantly better than NDMI, MSI, and NDWI, constitute a scientifically novel attempt to predict soil moisture on a wider scale and have thus potential for further use.

Märgalade mullaniiskuse ja mulla temperatuuri prognoos kaugseireandmete põhjal

Hleb Lazovik

Kokkuvõte

Sood on maakerale tähtsad ökosüsteemid, sidudes, muundades ja väljutades palju keemilisi, bioloogilisi ja geneetilisi materjale (Mitch & Gosselink, 2015). Märgalade uurimisel on tänapäeval oluline teabeallikas satelliidipõhised kaugseireandmed. Mullaniiskus ja -temperatuur on soo-ökosüsteemi fundamentaalsed näitajad, mida on võimalik uurida just kaugseire abil. Senised uuringud on keskendunud peamiselt mõnele konkreetsele uurimispiirkonnale ning harva on püütud uuringuga saavutada ülemaailmset esinduslikkust. Kui üldse, siis on eelkõige vaadatatud laiemalt vaid mulla temperatuuri. Käesoleva uurimistöö eesmärk oli täiendada ning anda hinnang mudelile, mille eesmärk on tuletada maailma soomuldade niiskust ja temperatuuri. Uurimistöös otsiti vastuseid kahele uurimisküsimusele:

1. Milline on seos välitööde käigus reaalselt mõõdetud andmete ning Landsati abil saadud näitajate vahel?
2. Kas on võimalik luua ülemaailmne mudel, mis piisava täpsusega tuletaks Landsati andmetest soomullaniiskust ja -temperatuuri?

Mullaniiskuse ja -temperatuuri prognoosimiseks kaugseirepiltide põhjal on juba tehtud märkimisväärne hulk uuringuid. Mitmete mudelite kontseptsioone ja meetodeid on erinevates tingimustes ja oludes seni mitmete uuringute raames selgitatud, analüüsitud ja täpsust hinnatud. Käesoleva magistritöö raames võeti peamiseks aluseks trapetsoidmudelil põhinev lähenemisviis. Trapetsoidil põhinevate indeksite idee on saada mulla niiskuse kohta teavet maapinna temperatuuri (ingl *land surface temperature,* LST) ja NDVI vegetatsiooniideksi (ingl *normalized difference vegetation index*) korrutise kaudu. Trapetsimudelil on siiski omad piirangud ning LST-põhist trapetsmudelit soovitatakse kohaldada ainult piirkondades, kus piiravaks teguriks on vesi, aga mitte soojusenergia.

Väliandmestik sisaldab 1544 mõõtmist 58l soomullaga alal üle maailma. Uurimistöös olid eeskätt olulised järgmised parameetrid: asukoha koordinaadid, mullaniiskus,

mullatemperatuur 10 cm sügavusel ja T_{norm} (aasta soojema kuu keskmine õhutemperatuur lähimas ilmajaamas).

Kaugseireandmed pärinevad Landsat 8 satelliidilt. märgiga ala jaoks valiti välitööle ajaliselt lähim pilt. Andmete eeltöötlus hõlmas parameetrite väärtuste keskmistamist asukoha kohta, asukoha ümber polügoonide loomist ja sagedusribade normaliseerimist. Uurimistöös olid eelkõige olulised järgmised näitajad: maapinna temperatuur, NDMI, MSI, NDWI ja ΔT*NDVI. Trapetsoidmudeli väärtused on arvutatud järgmiselt: (T_{norm} – LST) * NDVI.

Näitajate sobivust hinnati peamiselt regressioonanalüüsi abil. Mitmel juhul kasutati paremate tulemuste saavutamiseks filtreerimist, et kõrvaldada eriliste omadustega alad. Näiteks jäeti välja boreaalsete piirkondade ja ilma taimkatteta alade ning positiivsete ΔT*NDVI väärtustega andmed. Saadud tulemused näitasid kaugseire abil mullatemperatuuri tuletamiseks keskmise tugevusega (R^2 = 0,70) seost väliandmestikuga. Väljas mõõdetud mullaniiskuse ning NDMI, MSI ja NDWI vahel olulist seost ei esinenud ($p < 0,8$). Statistiliselt oluline seos saadi ΔT*NDVI abil (R^2 = 0,39; $p < 0,001$). ΔT*NDVI kasutamisel toob uurimistöö välja mõned piirangud, näiteks on seda keerulisem rakendada boreaalses vööndis ja taimkatteta mullal.

Kokkuvõttes saadi Landsati andmeid kasutades piisava täpsusega mudelid soomulla niiskuse ja temperatuuri tuletamiseks. Mudelid on originaalsed, töötavad oluliselt paremini kui NDMI, MSI ja NDWI ning kujutavad endast teaduslikult uudset lähenemist mullaniiskuse prognoosimiseks kaugseireandmetest. Saab öelda, et mudelitel on potentsiaali edasiseks kasutamiseks.

Acknowledgements

I would like to thank my supervisors, Jaan Pärn and Sandeep Thayamkottu, for introducing me to the topic of this research, as well as for being helpful and providing valuable input and feedback throughout it. I would also like to give thanks to Iuliia Burdun for her valuable advice regarding the methodology and its limitations, as well as to Alexander Kmoch for his detailed feedback on my presentation. Last but not least, I would like to thank my parents, who supported and encouraged me throughout this journey.

References

Bogena, H., Kunkel, R., Puetz, T., Vereecken, H., Krueger, E., Zacharias, S., ... & Hajnsek, I., 2012. Tereno-long-term monitoring network for terrestrial environmental research. *Hydrologie und Wasserbewirtschaftung, 56*(3), 138-143. https://doi.org/10.17815/jlsrf-2-98

Borg, E., Fichtelmann, B., Schiller, C., Kuenlenz, S., Renke, F., Jahnke, D., & Wloczyk, C., 2014. DEMMIN-Test Site for Remote Sensing in Agricultural Application.

Brekke, L.D., Kiang, J.E., Olsen, J.R., Pulwarty, R.S., Raff, D.A., Turnipseed, D.P., Webb, R.S., & White, K.D., 2009. Climate change and water resources management—A federal perspective: US Geological Survey Circular 1331, 65 p. https://doi.org/10.3133/cir1331

Burdun, I., Bechtold, M., Sagris, V., Komisarenko, V., De Lannoy, G., & Mander, Ü., 2020. A comparison of three trapezoid models using optical and thermal satellite imagery for water table depth monitoring in Estonian bogs. *Remote Sensing, 12*(12), 1980. https://doi.org/10.3390/rs12121980

Carlson, T. N., Gillies, R. R., & Perry, E. M., 1994. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote sensing reviews, 9*(1-2), 161-173. https://doi.org/10.1080/02757259409532220

Ellenberg, H., & Leuschner, C., 2010. *Vegetation Mitteleuropas mit den Alpen: in ökologischer, dynamischer und historischer Sicht* (Vol. 8104). Utb.

Friedl, M. A., & Davis, F. W., 1994. Sources of variation in radiometric surface temperature over a tallgrass prairie. *Remote sensing of Environment, 48*(1), 1-17. https://doi.org/10.1016/0034-4257(94)90109-0

Gao, B. C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment, 58*(3), 257-266. https://doi.org/10.1016/s0034-4257(96)00067-3

Garcia, M., Fernández, N., Villagarcía, L., Domingo, F., Puigdefábregas, J., & Sandholt, I., 2014. Accuracy of the Temperature–Vegetation Dryness Index using MODIS under

water-limited vs. energy-limited evapotranspiration conditions. *Remote Sensing of Environment, 149*, 100-117. https://doi.org/10.1016/j.rse.2014.04.002

Goodwin, E. J., 2017. Convention on Wetlands of International Importance, Especially as Waterfowl Habitat 1971 (Ramsar). In *Elgar Encyclopedia of Environmental Law* (pp. 101-108). Edward Elgar Publishing. https://doi.org/10.4337/9781783477210.v.9

Harris, A., & Bryant, R. G., 2009. A multi-scale remote sensing approach for monitoring northern peatland hydrology: Present possibilities and future challenges. *Journal of environmental management, 90*(7), 2178-2188. https://doi.org/10.1016/j.jenvman.2007.06.025

Kaplan, G., & Avdan, U., 2018. Monthly analysis of wetlands dynamics using remote sensing data. *ISPRS International Journal of Geo-Information, 7*(10), 411. https://doi.org/10.3390/ijgi7100411

Klemas, V., & Smart, R., 1983. The influence of soil salinity, growth form, and leaf moisture on-the spectral radiance of. *Photogramm. Eng. Remote Sens, 49*, 77-83.

Klinke, R., Kuechly, H., Frick, A., Förster, M., Schmidt, T., Holtgrave, A. K., ... & Neumann, C., 2018. Indicator-based soil moisture monitoring of wetlands by utilizing Sentinel and Landsat remote sensing data. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 86*(2), 71-84. https://doi.org/10.1007/s41064-018-0044-5

Lambin, E. F., & Ehrlich, D., 1996. The surface temperature-vegetation index space for land cover and land-cover change analysis. *International journal of remote sensing, 17*(3), 463-487. https://doi.org/10.1080/01431169608949021

Landsat missions., 2022. *Landsat 8-9 Collection 2 Level 2 Science Product Guide.* Retrieved from https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/media/files/LSDS-1619_Landsat-8-9-C2-L2-ScienceProductGuide-v4.pdf

McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G., Thomas, A., ... & Dinpashoh, Y., 2012. Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation. *Journal of Hydrology, 416*, 182-205. https://doi.org/10.1016/j.jhydrol.2011.10.024

Mitsch, W. J., & Gosselink, J. G., 2015. Wetlands. John Wiley & Sons.

Moran, M. S., Clarke, T. R., Inoue, Y., & Vidal, A., 1994. Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. *Remote sensing of environment, 49*(3), 246-263. https://doi.org/10.1016/0034-4257(94)90020-5

Parastatidis, D., Mitraka, Z., Chrysoulakis, N., & Abrams, M., 2017. Online global land surface temperature estimation from Landsat. *Remote sensing, 9*(12), 1208. https://doi.org/10.3390/rs9121208

Pasolli, L., Notarnicola, C., Bertoldi, G., Bruzzone, L., Remelgado, R., Greifeneder, F., ... & Zebisch, M., 2015. Estimation of soil moisture in mountain areas using SVR technique applied to multiscale active radar images at C-band. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(1), 262-283. https://doi.org/10.1109/jstars.2014.2378795

Petropoulos, G. P., Griffiths, H. M., Dorigo, W., Xaver, A., & Gruber, A., 2013. Surface soil moisture estimation: Significance, controls and conventional measurement techniques. *Remote sensing of energy fluxes and soil moisture content*, 29-48. https://doi.org/10.1201/b15610-4

Price, J. C., 1990. Using spatial context in satellite data to infer regional scale evapotranspiration. *IEEE transactions on Geoscience and Remote Sensing, 28*(5), 940-948. https://doi.org/10.1109/36.58983

Pärn, J., 2018. Data for the paper Nitrogen-rich organic soils under warm, well-drained conditions are global nitrous oxide emission hotspots. *PANGAEA*. https://doi.org/10.1594/PANGAEA.885897

Pärn, J., Verhoeven, J. T., Butterbach-Bahl, K., Dise, N. B., Ullah, S., Aasa, A., ... & Mander, Ü., 2018. Nitrogen-rich organic soils under warm well-drained conditions are global nitrous oxide emission hotspots. *Nature communications, 9*(1), 1-8. https://doi.org/10.1038/s41467-018-04197-6

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., ... & Zhu, Z., 2014. Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment, 145*, 154-172. https://doi.org/10.1016/j.rse.2014.02.001

Sandholt, I., Rasmussen, K., & Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of environment, 79*(2-3), 213-224. https://doi.org/10.1016/s0034-4257(01)00274-7

Taloor, A. K., Manhas, D. S., & Kothyari, G. C., 2021. Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data. *Applied Computing and Geosciences, 9*, 100051. https://doi.org/10.1016/j.acags.2020.100051

Uniyal, B., Dietrich, J., Vasilakos, C., & Tzoraki, O., 2017. Evaluation of SWAT simulated soil moisture at catchment scale by field measurements and Landsat derived indices. *Agricultural Water Management, 193*, 55-70. https://doi.org/10.1016/j.agwat.2017.08.002

Welikhe, P., Quansah, J. E., Fall, S., & McElhenney, W., 2017. Estimation of soil moisture percentage using LANDSAT-based moisture stress index. *J. Remote Sens. GIS, 6*(2). https://doi.org/10.4172/2469-4134.1000200

Zeng, Y., Feng, Z., & Xiang, N., 2004, September. Assessment of soil moisture using Landsat ETM+ temperature/vegetation index in semiarid environment. In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium* (Vol. 6, pp. 4306-4309). Ieee. https://doi.org/10.1109/igarss.2004.1370089

Annex

Annex 1. List of ground dataset sites.

Site no.	*Location*	*Climate*	*Vegetation*
1	Russia, Bashkortostan	D	Typha latifolia
2	Russia, Bashkortostan	D	Grasses
3	Russia, Bashkortostan	D	Sedges
4	Russia, Bashkortostan	D	Sedges
5	Malaysia, Sabah	A	Oil palm
6	Malaysia, Sabah	A	Tropical evergreen forest
7	Brazil, Santa Catarina	C	Grasses
8	Brazil, Santa Catarina	C	Sedges
9	United States, California	C	Sedges
10	United States, California	C	Sedges
11	Spain, Catalonia	C	Sphagnum spp.
12	Colombia, Antioquia Department	C	Grasses
13	Colombia, Antioquia Department	C	Grasses
14	Estonia, Tartu County	D	Sedges
15	Estonia, Tartu County	D	Sedges
16	United States, Florida	A	Cupressaceae
17	United States, Florida	A	Cupressaceae
18	France, Centre-Val de Loire	C	Sedges
19	France, Centre-Val de Loire	C	Sedges
20	French Guiana, Cayenne	A	Eleocharis spp.
21	French Guiana, Cayenne	A	Eleocharis spp.
22	France, Occitania	C	Sedges
23	Iceland, Western Region	D	Sedges

24	Iceland, Western Region	D	Sedges
25	Kyrgyzstan, Naryn Region	D	Sedges
26	Kyrgyzstan, Naryn Region	D	Sedges
27	Kyrgyzstan, Naryn Region	D	Sedges
28	Mexico, Mexico City	C	Bare soil
29	Mexico, Mexico City	C	Grasses
30	Mexico, Mexico City	C	Typha domingensis
31	Myanmar, Shan State	A	Bare soil
32	Myanmar, Shan State	A	Grasses
33	Myanmar, Shan State	A	Grasses
34	New Zealand, Waikato	C	Sedges
35	New Zealand, Southland	C	Sedges
36	New Zealand, Waikato	C	Sedges
37	New Zealand, Southland	C	Sedges
38	Canada, Quebec	D	Herbs
39	Canada, Quebec	D	Sphagnum spp.
40	Romania, Suceava County	D	Sphagnum spp.
41	Romania, Suceava County	D	Sphagnum spp.
42	Romania, Suceava County	D	Sphagnum spp.
43	Romania, Suceava County	D	Sphagnum spp.
44	Russia, Khanty-Mansi Autonomous Okrug	D	Sphagnum spp.
45	Russia, Khanty-Mansi Autonomous Okrug	D	Sphagnum spp.
46	Russia, Khanty-Mansi Autonomous Okrug	D	Sedges
47	Australia, Tasmania	C	Gymnoschoenus sphaerocephalus
48	Australia, Tasmania	C	Grasses
49	Australia, Tasmania	C	Brassica rapa

50	Australia, Tasmania	C	Sphagnum spp.
51	Argentina, Tierra del Fuego Province	C	Sphagnum spp.
52	Argentina, Tierra del Fuego Province	C	Grasses
53	Uganda, Western Region	A	Bare soil
54	Uganda, Western Region	A	Cyperus spp.
55	Uganda, Western Region	A	Cyperus spp.
56	United Kingdom, Wales	C	Sphagnum spp and Ericaceae
57	United Kingdom, Wales	C	Sphagnum spp and Ericaceae
58	United Kingdom, Wales	C	Sedges

Non-exclusive licence to reproduce thesis and make thesis public

I, Hleb Lazovik

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, **“Soil moisture and temperature in wetland soils derived from remote sensing”** supervised by **Jaan Pärn, Sandeep Thayamkottu**.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0 , which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive license does not infringe other persons’ intellectual property rights or rights arising from the personal data protection legislation.

Hleb Lazovik

Tartu, 28.05.2023