Master Thesis GEO511

"Interrater reliability of two qualitative soil moisture assessment schemes – Case studies from Tanzania and Switzerland"

Müller Daniela
daniela.mueller@geo.uzh.ch
09-711-243

Faculty representative:
Prof. Dr. Jan Seibert
jan.seibert@geo.uzh.ch

Supervised by:
Prof. Dr. Jan Seibert
jan.seibert@geo.uzh.ch

Dr. Michael Rinderer
michael.rinderer@geo.uzh.ch

Physical Geography H2K

Department of Geography University of Zurich - Irchel
Winterthurerstrasse 190
CH-8057 Zürich

Universität Zürich[UZH]

Submission: June 2015

Abstract

Soil moisture is an important component in the global, regional and local water cycle. The knowledge of the spatial distribution of soil moisture is of interest for various reasons. To capture these variability in soil moisture content there exist numerous measurement methods based on different techniques. What these methods generally all have in common is that they are relatively time-consuming and are in need of expensive measurement devices. Therefore a low cost method with low demand, cost, and time to estimate soil moisture qualitatively under humid environmental conditions was developed. To investigate the influence of training and to extend the scheme to drier conditions, within the framework of this thesis field campaigns in Switzerland and Tanzania were conducted.

In Tanzania the existing qualitative scheme was adapted to local conditions. Interrater reliability and the correspondence between qualitative classes and quantitative differences in soil moisture content were tested with students and farmers in two independent tests. To evaluate the influence of information or training, participants were divided into two different subgroups. The same division was made for the test in Zürich however in this case the qualitative scheme for humid environmental conditions was used. To investigate if the qualitative wetness classification scheme is applicable with a minimum of information in written form and on a voluntary basis (crowd sourced) a long-time field campaign was carried out in Brunni, Alptahl in Switzerland.

Results show that participants of the different tests showed a good agreement among individual raters. The new scheme for semi-arid environmental conditions is capable of capturing shallow soil moisture differences and qualitatively defined wetness classes represent quantitative differences in soil moisture content. It became evident that a more detailed introduction as well as training do enable an improvement of interrater reliability of the qualitative wetness classification. Results from the long-time field campaign in Brunni, Alpthal suggest that an oral introduction leads to better results than a minimum in written form. The total number of participants appears rather small compared to the number of people passing by but it needs to be assessed if adjustments of the test layout and facilities could possibly improve test results.

Acknowledgment

I would like to thank my supervisor Dr. Michael Rinderer for his great support throughout the development of this thesis as well as my supervisor and faculty representative Prof. Dr. Jan Seibert for his helpful advice at various stages of this thesis. For the diverse field campaigns last year I was dependent upon assistance of many people. I thank the staff and students of Nelson Mandela African Institution of Science and Technology and farmers of the Mungushi, Kichangani and Kigongoni furrow who participated in the soil moisture assessment in April and June 2014, respectively. I also thank the geography students of the University of Zurich who participated in the soil moisture assessment in October in Zürich. I acknowledge the support of our local partners of the Pangani Basin Water Board, the Upper Kikuletwa Water Users Association (Tito Kitomari) and local village and furrow leaders. I would like to thank Tobias Siegfried (iMoMo project coordinator), Hosea Sanga (local iMoMo-project manager), Hans Komakech (local lead project partner), Pascal Oechslin, Beat Lüthi, Sebastian Stoll and Manfred Kaufmann (field assistance in Tanzania), Hanna Pevna, Sandra Pool and Ling Wang (field assistance in Zürich), Renée Müller (field assistance in Brunni), Guido Wiesenberg and Anett Hofmann (lab assistance and soil characterization), Philip Jörg (geo- and satellite data), Matthieu Bolay (anthropo-technologic issues), Alfayo Miseyeki (translation). Many thanks to the Genossame Trachslau that made available their land in Brunni, Alpthal for the field campaign there. Last I thank the Swiss Agency for Development and Cooperation (SDC) for financial support of the project: "Qualitative Soil Moisture Assessment in Semi-arid Conditions (Tanzania/Africa)" as part of the Global iMoMo Initiative (www.imomohub.org).

Contents

Figures

Tables

1. Introduction

Soil moisture or soil water content is described as the amount of water that is stored in soil pores or bound to the solid matter and can be removed by drying with a temperature of 105°C (Dyck and Peschke, 1995). Willhelm (1993) however defines soil water as the part of underground water that is stored in the unsaturated zone. Underground water storage exists in two different zones: the unsaturated zone that contains water and air and in general occurs immediately below the land surface, and the saturated zone in which all interconnected openings are completely filled with water (Karamouz et al, 2012). Even though that only about 0.005 % of the global water resources are stored in this unsaturated zone (Strahler und Strahler, 1999), soil with its water budget is an important component in the global, regional and local water cycle. It controls on exchanges of energy and matter and physical processes, especially, the partitioning of the available energy at the Earth's surface into latent and sensible heat exchange with the atmosphere (Petropoulos et al. 2014). Soil moisture content influences feedback between land surface and climate as well as the exchanges of trace gases on land. This, in turn, has an impact on the dynamics of the atmosphere boundary layer and hence on weather and global climate (Patel et al. 2009). The preliminary degree of saturation of a soil impacts the response of a river catchment to rainfall or snowmelt and subsequent flood generation (Penna et al. 2011). With wetter soil conditions surface runoff starts earlier and is more intense. Additionally, the knowledge of the spatial distribution of soil moisture is helpful to determine the potential of the water input, whether from rain or snowmelt, for infiltration, overland flow, floods and erosion as well as the consequential impacts on streams, reservoirs, infrastructure and human life (Petropoulos et al. 2014).

Knowledge about soil moisture is an important information for sustainable water resources management, plant water requirements, plant growth and productivity, irrigation management and to decide on the right moment to carry out cultivation procedures. This is the case particularly in ecosystems of arid and semiarid regions where water is a limited resource. (Petropoulos et al. 2014)

The variability of soil moisture is controlled by various factors, among others the soil properties. Soils vary in texture, content of organic matter, structure, colour and the existence of macroporosity. This affects the fluid transmission, the retention properties as well as the rate of evaporation. Another important factor is topography. The distribution of soil moisture is for example influenced by variations in slope, curvature and relative elevation. Infiltration, drainage and runoff depend on slope angle, whereas aspect influences evapotranspiration via solar irradiance. The curvature influences the convergence of lateral flow. Also vegetation (type and density) can influence soil moisture variability in multiple ways. The canopy pattern affects throughfall and the shading effect controls the rate of evaporation. Furthermore the soil hydraulic conductivity is affected through root activity, water for transpiration can be absorbed from the soil and there is an addition of organic matter to the soil surface. Other factors are mean moisture content, depth to water table and meteorological factors as solar radiation and precipitation depth. (Famiglietti et al., 1998)

To capture soil moisture there exist numerous measurement methods. We can distinguish between ground based and remote sensing methods. For ground based techniques to measure soil moisture the instrument is in direct contact with the soil particles. These "in-situ" measurement techniques have the advantage of providing more precise data, soil moisture can be measured in different depths and the instruments can be accurately calibrated. Well known methods are the gravimetric method (cf. chapter 2.3.), nuclear methods like neutron scattering or gamma attenuation and electromagnetic methods like time domain reflectometry (cf. chapter 2.1.), frequency domain reflectometry (cf. chapter 2.2.) or for example tensiometer techniques (Petropoulos et al. 2014). Remote sensing methods are more suitable for larger study areas that need to be covered uniformly. The techniques are more complex and costly and only the upper most part of the soil can be tested. Active or passive microwave sensors as well as thermal sensors are commonly used. (Dobriyal et al. 2012)

As diverse as these methods may be, they are all relatively time-consuming and expensive when using a large number of measurement points (Famiglietti et al., 1998). There was a growing interest in developing a low cost method with low demand and cost and time to estimate soil moisture. In related disciplines, as in soil science or avalanche risk assessment, qualitative field methods for practical field application are well established (Rinderer et al., 2012). The content of sand, silt and clay in a soil can be estimated with some simple methods such as forming ribbons with different diameters or forming a ring (Schlichting et al. 1995). Some attempts to use qualitative criteria to capture shallow soil moisture differences were made in the past. Dunne and Black (1970) and Dunne et al. (1975) mapped saturated areas qualitatively for their studies about runoff production. A main aim of Ambroise et al. (1996) was to assess and model the spatial and temporal variability of the water cycle components and of the main factors controlling them. The extent of the saturated areas of their research catchment were assessed and mapped by visually identifying them. Inamdar and Mitchell (2007) also identified the surface-saturated areas through visual identification and mapping of surface-saturation and soil coring. The Scotland & Northern Ireland Forum for Environmental Research (SNIFFER) prepared a field survey manual to provide information about different types of wetlands and how to identify them in the field in 2009. The division between the different types are based on easily recognisable (visual) characteristics, such as hydrological features, vegetation and the landscape setting. Also Kulasova et al. (2014) used vegetation to map saturation. Together with common quantitative methods, changes in the patterns of a simplified classification of plant communities based on the requirements of plants for water were used. In 1998 the Natural Resources Conservation Service of the United States Department of Agriculture (USDA) published a photo guideline for estimating soil moisture by feel and appearance. For each photo an available water capacity range and a short description are given. Unfortunately no explanation is given how these values were determined and how accurate this method is when used by non-experts. Blazkova et al. (2002) defined five different qualitative categories of surface wetness for their model evaluation. However, they used just the wettest classes as they were interested in saturated areas and no systematic test of the methods performance was made. It remains unclear if this method is reliable when used by different persons and if the qualitative classes with their indicators correlate with real quantitative differences in soil water content. In 2012 Rinderer et al. developed a new wetness classification scheme for field application based on qualitative wetness criteria for wet environmental conditions (Sensing with boots and trousers). Interrater reliability as well as the correlation of the qualitative indicators with quantitative differences in soil water content were tested in a field campaign with 20 geography students from the University of Zurich. They found that their results were encouraging and that non-experts can reliably assess soil wetness in the field. When classifying the wettest and driest classes raters showed a high level of agreement, whereas intermediate wetness classes showed less agreement. As this method is relatively new and not yet tested sufficiently further tests and studies were required. Rinderer et al. (2012) pointed out that training beforehand might help to obtain a higher degree of agreement and would potentially minimize systematic bias. This potential training effect is under investigation as part of this thesis. A distinction is made between a short verbal introduction, a more detailed introduction with a few examples in small groups, and an introduction coupled with training and in an extra test an introduction in written form (minimum of required information summarized on one page). A request of iMoMo to support their project in Tanzania necessitated an adaptation of the scheme to arid environmental conditions and consequently conduct tests.

iMoMo stands for Innovative Technologies for Monitoring, Modeling and Managing Water and their aim is to foster the development and deployment of new water-related information and communication technologies (ICT) that are geared to collect crowdsourced data, synthesis of online data and the analysis as well as the distribution / exchange of the resulting knowledge for improved decision-making in the water sector. The idea is that a stakeholder taking part in this project is able to collect data and gets information in return that for example help him to achieve a better crop yield. As the pressure on existing water resources is continuously growing the available water hast to be used with care. The project in Tanzania is realised at the Themi River catchment which experiences dramatic shifts towards extreme water scarcity as consumptive allocation and pollution increase in line with population growth. Therefore especially downstream communities no longer receive the runoff quantities that they were used to in the past. This diminished water resources cause dispute between the different users as unjust allocation is a consistent issue. Hence a more widely accepted tool to support decision making is needed. The adapted soil wetness classification scheme for drier conditions should be tested for suitability as a decision making tool.

Next to a reduction of demand, cost and time to estimate hydrological variables by qualitative methods it possibly also allows to include a greater amount of people into data collection. If people do not require expensive devices and a specific education in this field theoretically everyone could help. This idea of crowd sourced collection of hydrological data was already investigated by some authors in the past. In Turner and Richter's (2011) paper volunteers mapped the spatial extent of wet reaches on the San Pedro Riparian National Conservation Area (SPRNCA). With the incorporation of volunteers, monitoring could be conducted over a large, continuous geographic domain. Lowry and Fienen (2013) present a research on building a crowdsourced database of inexpensive distributed stream stage measurements. On staff gauges installed signs encourage citizen scientist to voluntarily send hydrologic measurements via text message to a server. From their analysis they conclude that crowdsourced data collection is a supplemental method for collecting hydrologic data and a promising method of public engagement. In Buytaert et al. (2014) a summary of the current state of understanding of citizen science in hydrology and water resources can be found. In this thesis the potential of the wetness classification scheme by Rinderer et al. (2012) as a method for crowd sourced data collection was investigated.

For the purposes of this master thesis the author planned and conducted hydrological field experiments independently, collaborated at hydrological field experiments of the iMoMo project, and made herself familiar with the procedure and the applied methods. Data to answer the following research topics by means of the field experiments were collected and then processed. Specifically, the following research questions were addressed:

- Is the qualitative soil wetness classification scheme capable of reliably capturing differences in shallow soil moisture content in an arid environment?
- Are the qualitatively defined wetness classes representative of quantitative differences in soil moisture content?
- Can interrater reliability of the qualitative wetness classification be improved by training?

- Is the qualitative wetness classification scheme applicable with a minimum of information and on a voluntary basis (crowd sourced)?

First some background information on time domain reflectometry, frequency domain reflectometry as well as the gravimetric method and the sedigraph technique is given. In chapter 3 the different study areas in Tanzania and Switzerland are introduced individually. Chapter 4 gives an overview of the different qualitative and quantitative methods to estimate and measure soil moisture as well as site survey and laboratory techniques to determine the soils grain-size distribution and its type. Results obtained in this study are summarized in chapter 5. Qualitative and quantitative soil moisture data, the comparison of qualitative and quantitative soil moisture data as well as results of the soil analysis are discussed in different subchapters for the four study areas. The before presented results are discussed in chapter 6 and is followed by the conclusion (chapter 7).

2. Background

This chapter gives a brief introduction to the three used methods to measure soil moisture as well as the sedigraph method to determine grain-size distribution. This information is in particular meant for readers with no background knowledge in this field.

2.1. Time Domain Reflectometry

The time domain reflectometry (TDR) technique is based on the relation between the measured soil dielectric constant and the water content of a soil (Rajkai and Rydén 1992; Pepin et al. 1992). Water has a high static relative dielectric permittivity and is many times higher than values for air and for commonly found soil solids (Ferré et al. 2000). For TDR measurements electrodes inserted into the soil send a high frequency transverse electromagnetic wave along a transmission line. The time elapsed between sending the pulse and receiving the reflected wave is being measured and allows the determination of the dielectric constant. (Formula: Ferré et al. 2000, Jones et al., 2002)

$$\varepsilon_{ra} = \left(\frac{ct}{2L}\right)^2$$

For a wide range of soils the empirical relationship between relative dielectric permittivity ε_{ra} and volumetric water content was determined in the laboratory. Soil bulk density, ambient temperature and salt content showed to have no essential influence on this relationship. (Petropoulos et al. 2014; Ferré et al. 2000)

The methods advantages are surely its non-destructive character, that the used instrument is portable, easy to install and safe to operate. In addition, it is less labour intensive than other soil moisture measurement techniques and in general no soil specific calibration is required (Dobriyal et al. 2012).

The TDR device used for the study in Mungushi, Tanzania was the device “TRIME-PICO 64” of the company IMKO. This TDR sensor is equipped with two isolated parallel rods (16 cm length and ø6.0 mm) that can be inserted into the soil. The sensor has a measurement volume of 10x16mm (1250ml).To control the sensor and to save measurement data the sensor was combined with the PICO-BT Bluetooth module and a PDA. The sensor can be switched on by the PICO-BT and when the sensor is inserted to the soil, measurements can be started with the program PICO-TALK on the PDA. After a few seconds the values for the moisture, temperature, and electrical conductivity are available for output. All results can be viewed directly on the PDA and can be loaded on the computer as a data file.

For a measurement the time of arrival of specific predefined voltage levels is measured by the sensor. The time elapsed between sending the pulse and receiving the reflected wave is being measured and allows the determination of the dielectric constant. The sensor does not measure soil moisture content directly but rather uses the measured dielectric constant to determine it. The sensor receives a basic calibration by IMKO prior to shipment. This basic calibration serves to gauge for the cable length and probe mechanics as thickness of the rod coating or rod length. Two measurements, one in dry and one in water-saturated glass beads, are conducted and the calculated calibration data is stored in the sensor. With every change of sensor rods calibration has to be performed again. As a standard the system operates with a universal calibration for mineral soils. As this universal calibration is limited by parameters like clay content, organic content or bulk density, soils exceeding the defined limits, a material-specific calibration can be conducted. (IMKO, 2013)

2.2. Frequency domain reflectometry

Frequency Domain Reflectometry (FDR) is similar to TDR except that it provides an estimate of the soil moisture content based on changes in the frequency of signals due to the dielectric soil properties (Robock et al. 2000). These changes in resonant frequency and therefore variations in soil moister content can be measured with an electrical circuit using a capacitor and an oscillator (Dobriyal et al. 2012).

2.3. Gravimetric Method

The gravimetric method is a widely used method to determine soil moisture. For this technique a soil sample from the field gets extracted with a core cutter with a known volume. The weight of the wet sample is measured before being dried in an oven at a temperature of 105°C for at least 24 hours. To receive the concentration of water the samples weight before and after drying is compared. The density of the soil is used to convert the total amount of soil moisture into volumetric water content. (Reynolds 1970; Petropoulos et al. 2014)

The method is relatively easily applicable and gives a high level of accuracy. Certainly, one of its limits is its destructive character and therefore the impossibility of repeated sampling. Because of the necessary drying process the time for processing a sample is relatively high compared to other methods for measuring soil moisture in the field. Another limitation is that it is time and labour intensive and hence not appropriate for large scale field campaigns (tests). (Petropoulos et al. 2014)

2.4. Sedigraph

There are several methods available to determine the grain-size distribution which is crucial to determine the soil group. The sedigraph method presented here is based on Stokes Law. The law states that particles in a liquid reach a final (terminal) settling velocity that is proportional to the size of the particle (Quelle bis jetzt UCL geolab, egli et al. 2013). The following equation describes the terminal settling velocity v for a particle (Simons and Sentürk, 1992):

$$v\mathrm{p} = \frac{2\, r^2 g\, (\rho p - \rho f)}{9\eta}$$

Here, r is the radius of the particles, ρp the mass density of the particles, ρf the mass density of the surrounding fluid, g the gravitation acceleration and η the dynamic viscosity. The sedigraph makes use of this principle and determines the concentration of particles remaining at decreasing sedimentation depths in a cell filled with suspension (Syvitski, 1991). A x-ray beam is passed through the cell, the intensity of the incident radiation is plotted against the time and grain-size distribution is calculated from the intensity profile (Egli et al. 2013). The x-ray transmission of the cell filled with suspensions compared to the transmission of the cell filled with only sedimentation liquid gives concentration values (Syvitski, 1991).

3. Study Areas

In this chapter the different study areas are presented in separate subchapters. First the two study areas in Tanzania then the study area in Zürich, Irchel and Brunni, Alpthal.

3.1. iMoMo study areas in Tanzania

For the iMoMo project, two different study areas were chosen for the field campaigns in April and June, one site at the village Mungushi and one site at the nearby village Kitshangani. Both are situated in the north-east of Tanzania south of the city Arusha and Mount Meru (Fig. 1). The areas are located in the (arid) downstream region of the Themi River catchment about 1000 m above mean sea level. These two locations seemed to be suitable as both communities are affected by the dwindling resource water due to their crop cultivation.

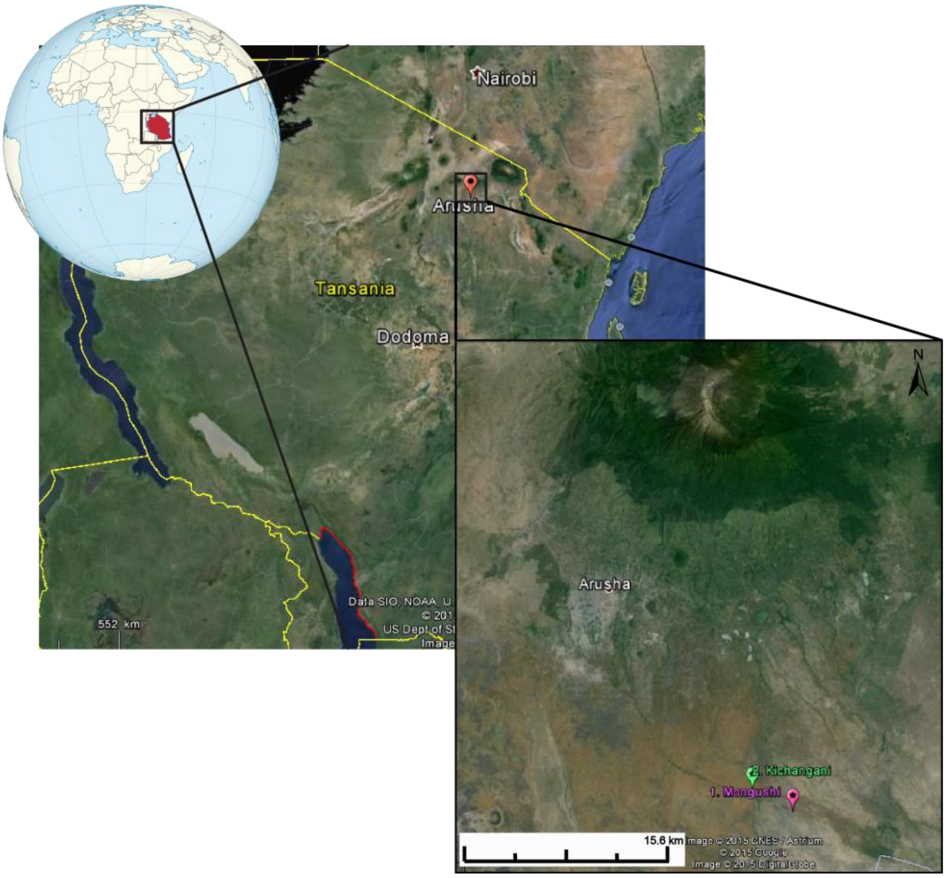


Figure 1: Map extracts of Africa (left), Tanzania (centre) and Arusha with surroundings; map basis: Google Inc. (2013), Wikipedia (2015).

The Themi River catchment is a small subcatchment of the larger Pangani River basin in Eastern Africa and covers an area of approximately 363 km^2 ranging from 4522m amsl to 939m amsl (Komakech and Van der Zaag, 2011). Themi River has several tributary streams which all originate at (from) the south-western slope of Mount Meru and flows into the Shambarai swamp downstream (Komakech and Van der Zaag, 2011). From March until May and October till December the area experiences two rainy seasons while the inter-seasonal rainfall is relatively low. The average rainfall varies from about 1400 mm/yr in the highlands (above 1700m amsl) to about 500 mm/yr in the lowland (below 1300m amsl). The seasonality of rainfall in Tanzania is driven mainly by the migration of the Inter-Tropical Convergence Zone (ITCZ) (McSweeney et al., 2010). This belt of very low pressure and heavy precipitation forms near the equator and changes its position over the course of the year which causes these two distinct wet periods in the north and east (McSweeney et al., 2010).

In the lowlands (where Mongushi and Kichangani are situated) land use is mainly dominated by subsistence farming and livestock rearing (Komakech and Van der Zaag, 2011). Common cultivation products are corn, beans, watermelon etcetera. Because of limited rainfall Farmers in this area rely heavily on irrigation. The study area in Mongushi contains unmade road, and mainly unplanted fields, and soils without vegetation cover due to the arid climate (Fig. 2). The area in Kichangani is on planted fields entirely (Fig. 2).

Figure 2: Pictures of parts of the study field in Kichangani (left) and Mongushi (right).

3.2. Study area Irchel Park

One of the study areas is located at Irchel Park and adjoining sections of forest in the city of Zurich in the Swiss Central Plateau (Fig. 3). Irchelpark is a near to nature landscape park and serves as recreation area for members of the University of Zurich and the general public (Stadt Zürich, 2015). With about 500m amsl Irchel Park lies slightly higher than the city centre.

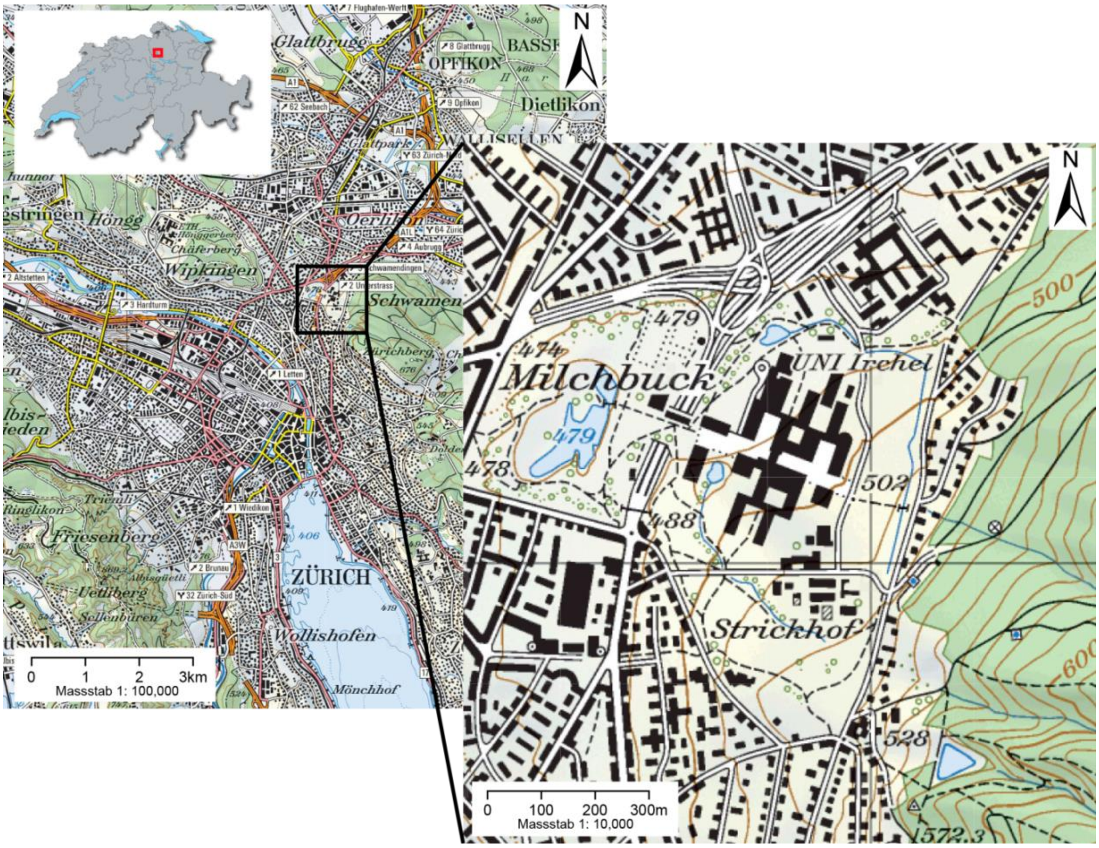


Figure 3: Map extracts of Zurich (left) and Irchel (right); map basis: Swisstopo (2007).

Zurich is within the temperate zone and has a mean annual (ground-surface) temperature of 9.3°C. There is precipitation all year round with a maximum during spring and summer (ca.1100mm/yr). The site for the tests in the park and forest is situated east of the University buildings. The used park area has meadow and a section of a little stream that is slightly affected but relatively natural. The forested area is part of the Zürichberg and categorized as a woodruff beech forest (Amt für Raumentwicklung, 2015). The used forest ground was without vegetation and partly covered with litter. The area was chosen because of its proximity to the university. In this way, a larger amount of students could be mobilized to take part in the test.

3.3. Study area Brunni, Alpthal

The fourth study area is located in the Alpthal, a valley south of Einsiedeln in the Swiss Prealps (Fig. 4). The exact site is above the village Brunni at the foot of the mountain Mythen between about 1180m and 1280m amsl. The area is along the main hiking trail and contains meadow, a crossing stream and forest.

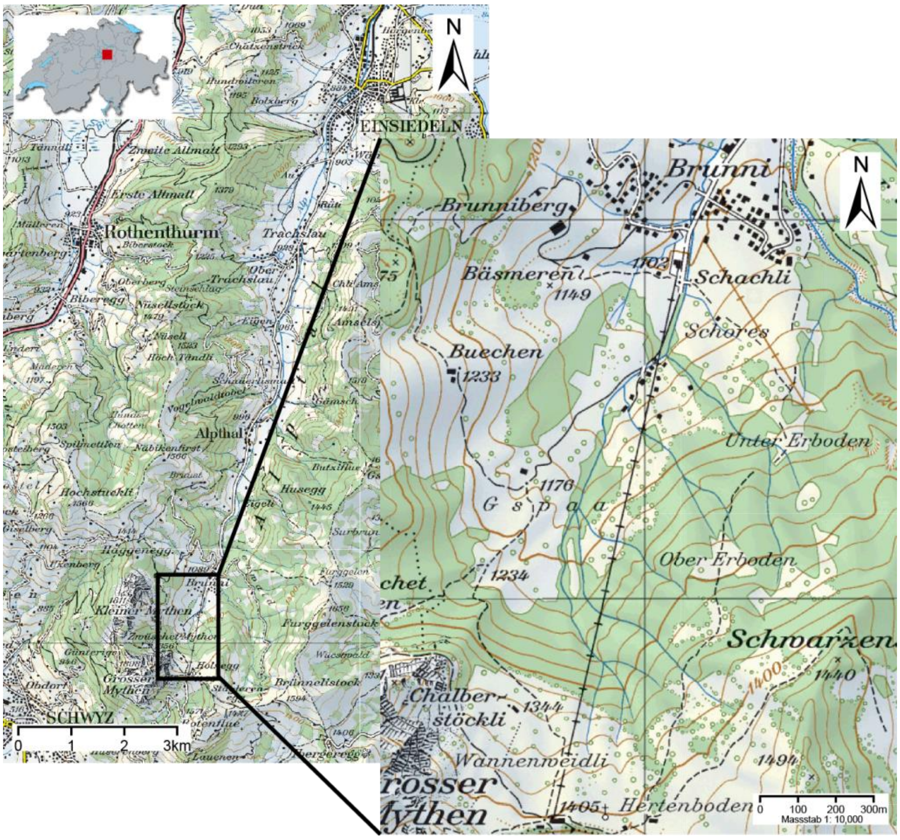


Figure 4: Map extracts of Alpthal (SZ) (left) and Brunni / Mythen (right); map basis: Swisstopo (2007).

Like Zurich the Alpthal is within the temperate zone and at the nearest meteorological gauging station (Einsiedeln, 910m amsl) a mean annual (ground-surface) temperature of 6.7°C was measured. Since Brunni is about 200m superior a mean annual (ground-surface) temperature of approximately 5°C can be assumed when using a temperature gradient of 0.6°C/100m. Compared to the Swiss average of about 1500mm/yr Einsiedeln and Alpthal receive a higher amount of precipitation. Alpthal has an annual precipitation often exceeding 2000mm. The basin location intensifies the up the valley increasing tendency of annual precipitation as well as of heavy precipitation (Burch et al., 1996). The area was chosen because of the large number of hikers that pass by and the proximity to the expert's place of residence. (MeteoSchweiz, 2015)

4. Methods

In the method chapter, first the methods used to measure soil moisture qualitatively are elucidated. These include the qualitative wetness classification scheme for wet environments as well as for the arid environments. Afterwards the different test layouts of the study areas and corresponding quantitative soil moisture measurement techniques as TDR, FDR and the gravimetric method are presented. At the end, the soil identification with its site survey and the soil texture analysis in the lab are explained.

4.1. Qualitative soil moisture measurements

In different subchapters the two qualitative wetness classification schemes are presented. First the scheme for wet environmental conditions followed by the classification scheme for dry environmental conditions.

4.1.1. Scheme for humid regions

The qualitative wetness classification scheme used here was developed in 2012 by Rinderer et al.. The scheme is based on seven qualitatively defined classes ranging from the driest class (class 1) to the wettest class (class 7) (Fig. 5). Each class is described by qualitative indicator criteria that is either recognizable through sight, sound or touch. To visually support those descriptions, icons for each class were added to the existing scheme. Class 1 would allow a person to sit down on the ground and the trousers would stay dry. If the trousers of a person sitting on the ground get moist after some minutes it is class 2 and if they get wet in the same time span the wetness class 3 would be assigned. Class 4 would not allow a person to sit on the ground without the trousers getting wet immediately. If a squelchy noise can be heard when stepping on the ground but no water is visible the site would be classified as class 5, if however water squeezes out of the topsoil when stepping on it with a boot it would be class 6. Is water seen on the soil surface, then the site would be classified as class 7. The authors points out though that it is not necessary that people using the classification scheme actually sit down on the ground each time they estimate soil moisture of a site. It is rather supposed that people do have some experience or can imagine whether they would get wet or stay dry when sitting on the ground. Further, vegetation and litter layer should be removed beforehand, as it can influence the classification. Additionally, raters have to be aware of possible influences as for example rainfall during the test or morning dew that might lead to misleading wetness class assignments.

Class		Qualitative indicator criteria
1		The trousers of a person sitting on the ground would stay dry
2		The trousers of a person sitting on the ground would get moist after some minutes
3		The trousers of a person sitting on the ground would get wet after some minutes
4		The trousers of a person sitting on the ground would get wet immediately
5		Squelchy noise can be heard when stepping on the ground but no water is visible
6		Water squeezes out of the topsoil when stepping on it with a boot
7		Water can be seen on the soil surface

Figure 5: Wetness classification scheme for wet environmental conditions (Rinderer et. al., 2012).

4.1.2. Scheme for semiarid regions

The qualitative wetness classification scheme for arid environmental conditions was especially developed for the iMoMo project, as the existing scheme was not applicable for drier conditions. Through interviews with local farmers in Tanzania and testing a wide range of soil water conditions in the field the scheme was adapted to local needs. The scheme is also based on seven qualitatively defined classes ranging from the driest class (class 1), that is described as dust dry, to the wettest class (class 7) water ponding on the soil surface (Fig. 6). Class 2 is defined to be dry, but with some moist look while drier than optimal for seeding would be classified as class 3. If soil water conditions are optimal for seeding crops class 4 would be assigned and class 5 if it is wetter than optimal and it is possible to form a solid brick. To classify a soil surface as class 6 water has to liquefy when someone steps on the soil. To ensure that the scheme is also understood by the local population all the terms and definition were as well translated into Swahili (e.g. kavu sana for very dry; Swahili scheme in the appendix Fig. 39). Special care was taken that the used Swahili terms and definitions were effectively used in this context by local farmers. As in the preceding scheme it is not meant that farmers form a brick all the time but rather imagine these conditions from their every-day life. Vegetation cover or litter layers should be removed beforehand and one has to keep in mind that soil surface conditions can be affected by recent rainfalls, dew or strong evaporation.

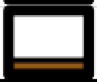
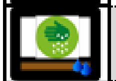
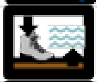
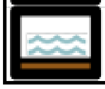
Icon	Class	Classname	Description
	1	very dry	"dust dry"
	2	dry	dry, but with some moist look
	3	below optimal	drier than optimal
	4	optimal	optimal for seeding crops
	5	above optimal	wetter than optimal - one can form a solid brick
	6	wet	when you step on the soil, water liquifies
	7	very wet	water ponding on the soil surface

Figure 6: Soil wetness classification scheme for dry environmental conditions (Rinderer et al., 2015).

4.2. Test Layout

In different subchapters the different measurement concepts of the four study areas are presented. First the two study areas in Tanzania followed by the study areas in Switzerland.

4.2.1. Tanzania April

To examine the qualitative wetness classification scheme for arid environments two datasets were compiled. The first test took place on the 10th of April in 2014 in the village Mungushi and a second in July 2014 in a nearby village (cf. chapter 4.1.2.). To analyse the inter-rater reliability, a possible effect of training together with the general performance of the qualitative method, the test was carried out as follows: For each qualitative class to be represented roughly equally often, the measurement points were assigned in advanced. Furthermore, care was taken that both prevailing soil types (“red soil” and “white soil”) were taken into account. A group of experts tested a great number of spots with the qualitative and the quantitative method to define suitable measurement points on the day prior to the test. The quantitative values used to determine a range of soil moisture values for a certain wetness class were derived previously based on this survey. This class ranges can subsequently be used as a reference for the inter-rater reliability analysis. On the day of the test, the selected 40 sampling points were marked with numbered flags along a 1.4 km parcours (Fig. 8). 14 master, 9 PhD students and 3 Professors of the Nelson Mandela African Institution of Science and Technology (NM-AIST) together with 14 local farmers, in total 40 people, conducted the test. PhD students as well as people with a higher education (professors) are gathered in the group PhD+. When referring to all of the farmers, master students and PhD+ the expressions F*all*, M*all* and PhD+ are used. (The students could voluntary attend a workshop at their university held by Michael Rinderer and iMoMo giving a lecture on the in 2012 developed "Boots and trousers" method.). To examine a possible training effect students and farmers were divided into two groups, one receiving a short introduction on method and form sheet either in English or Swahili while the other group got a more detailed introduction with representative sites of wetness classes 1, 4 and 7. When referring to the farmers and students with a short introduction the expressions F*not* and M*not* are used and when referring to farmers and students with more detailed introduction the expressions F*inf* and M*inf* are used. When referring to all people with a short or more detailed introduction, the expressions Inf*all* and Not*all* are used. Then, all participants passed the course individually and filled their estimations (one class per side) in the provided form. At the same time at least one expert passed the course as well. To avoid certain effects that alter soil wetness conditions on the soil surface people were asked to remove the upper most few centimetres. No rainfall occurred during the day of the test and to reduce a possible drying effect due to evaporation as well as interaction participants had to start shortly after another.

Simultaneously, TDR measurements with the "TRIME-PICO 64" device (Fig. 7) of the company IMKO were conducted. The universal calibration for mineral soils was chosen as the important parameters are almost all within the set limits. Clay content varies considerably over the different measurement points (21 out of 40 points were further examined in the lab and clay contents between 4 % and 53 % were found) but only two points showed values above the set limit of 50 % by IMKO. Organic content and bulk density of the top soil of the study are within the set limits. Without soil specific calibration, measurements are less accurate and the volumetric water content values hold some uncertainty. When measuring actual volumetric water content values sites with high soil water contents can cause difficulties. IMKO indicates that their device has a lower accuracy for conditions with more than 40% volumetric water content. Three TDR measurements were taken at each sampling point and the surrounding 15cm^2. With a larger number of measurements collected, the repeatability of the results and instrument errors can be ascertained (Penna et al., 2009). It has to be noted that surface soil moisture can be highly variable over space (Western et al., 2004) and depending on the exact measurement spot different water contents can result. The 3 values were averaged and for each sampling point volumetric water content was available for comparison to qualitative classes or gravimetric measurements. To control the sensor and to save measurement data the sensor was combined with the PICO-BT Bluetooth module and a PDA. To protect the rods from breaking, the soil was pre-drilled prior to probe insertion. To ensure the contact between sensor rods and the soil the driller tool has slightly thinner and shorter rods. In this way measurement errors due to air gaps can be minimised. Additionally the sensor should not be repeatedly inserted and removed at the same spot (Jones et al., 2002). The soil around the rods gets slightly compacted when the sensor or the pre-driller is stuck into the ground. This can result in a reduction of the measured water contents (Rothe et al., 1997). Of the 40 measurement points of the field test only 16 could be measured by TDR because the device broke and the remaining points were only determined gravimetrically. Additionally to the TDR measurements, gravimetric sampling was accomplished for each sampling point as control for the averaged TDR values.With a steel cylinder (5 cm ⌀) at approximately 10 cm depth below the soil surface 100cm3 soil samples were extracted for every measurement point. The soils samples where then filled into sealable plastic bags and transported to the lab. The samples were taken as so called disturbed soil samples, as they do not keep the pristine soil structure in the collection process. Each sample was filled into a separately numbered aluminium bucket and then dried with 105° C in the lab oven for at least 24 hours. Buckets with the soil samples were weighted wet and dry and with, this loss of water and the known volume of the sample volumetric water content could be calculated. (IMKO, 2013)

Figure 7: ML3 ThetaProbe sensor with Moisture Meter HH2 (left), TRIME-PICO 64 sensor with PICO-BT Bluetooth module and PDA (middle) and GS3 Greenhouse sensor with cable for the logger connection (right)

4.2.2. Tanzania June

The second test took place in June in 2014 in the village Kichangani. Results of the first test in Mungushi were not as good as anticipated. A repetition of the test with some adaptions was supposed bring the desired improvements. Alike the first test in April, the measurement points were assigned in advanced by an expert by qualitative testing. In this way it could be ensured that each qualitative class is represented often enough. Since in July the area is experiencing a dry phase, measurement points of the wetter classes (that can be assigned to wetter classes) could only be found on recently irrigated fields. All 42 sampling points lay within planted fields and were marked with numbered flags. The test was performed by 18 farmers and 7 experts (members of the iMoMo project with a certain experience with the method are categorized as experts) (Fig. 9). When referring to all of the farmers and experts the expressions F*all* and E*all* are used. To test whether an improvement compared to the first test in April can be achieved with a better and longer introduction or training, farmers were divided in small subgroups. When referring to farmers with a basic introduction the expression F*not* is used and when referring to farmers with training the expression F*inf* is used. For each measurement point one wetness class had to be chosen. During the day of the test no rainfall occurred and evaporation was considered to be small as all participants finished the test within 1 hour.

Simultaneously FDR measurements were made and samples for the gravimetric method collected. At each spot and the surrounding 15cm^{2,} two FDR measurements were conducted and one gravimetric sample obtained. Uncertainty for single measurements could be reduced with multiple measurements. For the measurements two individual devices from the same company and model type were used. In this way it could be ensured that none of the devices had an offset and time for measuring at each sampling point could be reduced. Measurements of s sampling point that had offset of ≥8 Vol% were repeated and the four measurements were averaged for further analysis. Analogous to chapter 4.2.1 volumetric water contents for the gravimetric samples were determined. For the FDR measurements the ML3 ThetaProbe of the company Delta-T Devices Ltd (Fig. 7) was used. This FDR sensor is equipped with four parallel stainless steel rods (60 mm length and ø 3 mm) that can be inserted into the soil. To operate the sensor the Moisture Meter type HH2 was utilized. All the settings can be carried out by the Moisture Meter and a measurement is performed within seconds. The results are stored on the Moisture Meter but only the last one is available. Water content and the sensor output in milliVolts are saved for every measurement point with the time signature and consecutive sample numbers and can be loaded on the computer as a CSV data file. The ThetaProbe uses a sample volume of approximately 60 x 30mm and measures the difference between the output wave and the return wave frequency to determine soil moisture (cf. chapter 2.2). With a linearization table and soil-specific parameters the mV readings get converted into soil moisture. For sensors, and organic and mineral soil parameters linearization tables are pre-installed. However, it is possible to enter own soil type parameters to achieve a greater accuracy. Here the default parameter values for mineral soils were chosen. Soil-specific calibration in order to determine the parameter values a0 and a1 experimentally was not possible, as the time required for a laboratory calibration was not available. Measurements are less accurate without soil specific calibration and the volumetric water content values hold some uncertainty. Measuring actual volumetric water content sites with high soil water contents can cause difficulties. The manufacturer indicates that their devices have lower accuracies for conditions with more than 50% volumetric water content. The sensor was either inserted directly into the soil or in case of extremely hard soils after the use of the pre-driller. The advantages of this device compared to the IMKO TDR device presented prior are the shorter rods and therefore measurements close to the soil surface. As the measurements serve as comparison for the qualitative values of the soil surface this is ideal. Besides, the Moisture Meter is battery-operated and has clear advantages over the TDR (has to be recharged) when used far away from any sockets. (Delta-T Devices, n.y.; Delta-T Devices, 2013a, Delta-T Devices, 2013b)

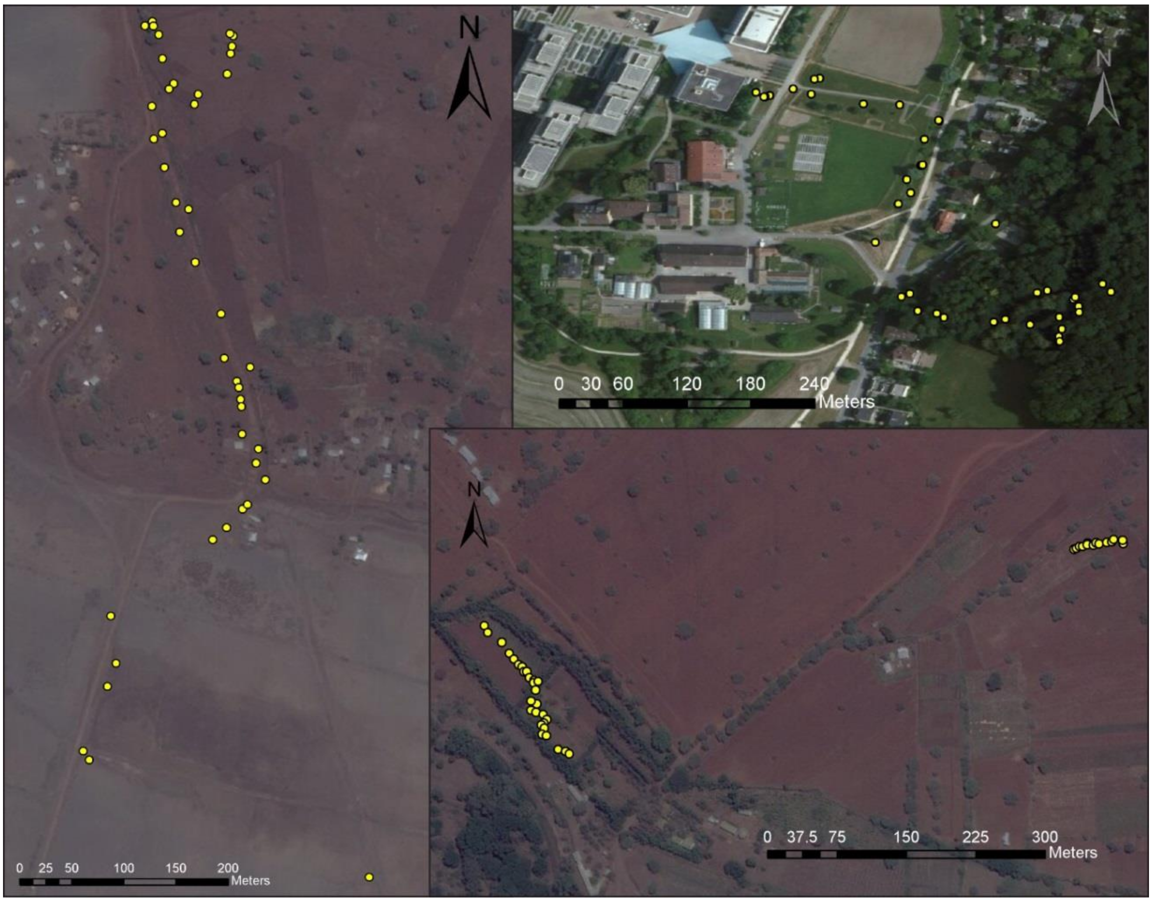


Figure 8: The different sampling parcours based on GPS data for Mungushi (left), Kichangani (right, bottom) and Irchel (right, top); map basis: ArcGIS World Imagery (2015).

4.2.3. Zürich Irchel

To also examine the importance of training using the qualitative wetness classification scheme for wet environmental conditions another dataset was compiled. The possible effect of training and the general performance of the qualitative method was of interest. The test was conducted on the 31. of October at Irchel, Zurich. Measurement points were assigned in advance by two experts by qualitative testing. As in the two previous tests in this way it can be ensured that each qualitative class is represented enough often. Points were distributed in forested area, on meadows and along a brook side. The natural soil moisture conditions in late autumn did not allow having enough sample sites with wetness class 1 even after several days without rainfall. To ensure that the driest class was represented in the test, bricks of soil were extracted the day before, dried in an oven and then placed back in the field. In total 40 sampling points along a 0.6km parcours (Fig. 8) were marked with numbered flags. 9 Bachelor students, 9 Master students and 3 PhD students conducted the test (Fig. 9). When referring to all of the bachelor and master students the expressions B*all*, M*all* are used. The 3 PhD students were not considered as an individual group as their number is not representative. They are only included if all raters are considered. To examine a possible effect of training the group was divided into two subgroups. One group received a short introduction on method and form sheet while the other group got a more detailed introduction with representative sites of three wetness classes. When referring to all people with a short or more detailed introduction, the expressions Inf*all* and Not*all* are used. For bachelor and master students with a short or more detailed introduction, the expressions B*inf* and B*not* respectively M*inf* and M*not* are used. All participants passed the test individually and finished within 45 minutes. For each measurement point only one wetness class could be assigned. No rainfall occurred during the test day and evaporation was considered to be relatively small, as the test was only very short and net radiation which influences evaporation (Karamouz et al., 2012), is considerably smaller than in summer. At the same time FDR measurements were made and samples for the gravimetric method collected. At each spot 4 FDR measurements were conducted, with each device two measurements, and one gravimetric sample obtained. Analogous to chapter 4.2.1 volumetric water contents for the gravimetric samples were determined. FDR measurements were conducted as described in chapter 4.2.2. Here as well the default parameter values for mineral soils were chosen. Because of the relatively diverse study area containing meadows and forest soils a uniform calibration was not possible and a laboratory calibration would have been too time-consuming. The measured quantity here is volumetric water content and gives only information of the volumetric % amount water of the whole considered soil body. As information about the pore volume is missing no statement about saturation can be made.

Figure 9: Test participants in Kitchangani (left), Alpthal (middle) and Zurich (right) while estimating soil moisture qualitatively.Photo (left): M. Rinderer.

4.2.4. Brunni, Alpthal

To examine if the qualitative classification method is applicable with limited information in written form and on a voluntary basis (crowd sourced) a test parcours was placed at the heavily used hiking trail between Brunni and Holzegg. It is estimated that on a good day during holidays or a weekend day with good weather an approximate number > 100-200 people passed the test area. The experiment ran from 28 September 2014 until 4 November 2014. Along a path 500 m long, 10 measurement points were selected and marked with an orange-yellow stake and the corresponding number (Fig. 10). Locations for the 10 measurement points were assigned by an expert and based on terrain characteristics or the presence of a water source. To cover as many different wetness classes as possible, points were distributed over meadow with sloped location, in troughs, on hilltops, along a small brookside, and in the forest. On both sides of the parcours (point 1 and point 10) information boards and material boxes were placed (Fig. 11). People interested (independent of his or her age or education) to participate in the test could find pre-printed forms (see page 74 in the appendix) and writing utensils in the material boxes. The pre-printed forms contained short instructions, the classification scheme, an empty table to fill in the estimated soil moisture values for every point, as well as questions about personal data (compulsory: date and time; voluntarily: name, age, profession / education). The test could be carried out in both walking directions and completed forms could be returned into the provided portfolio inside the material boxes on either side of the parcours. On five days (preferential weekend days or days with good weather) during the running experiment one expert conducted qualitative reference estimations and soil moisture was measured continuously with FDR sensors. At each of the 10 sampling points measurements were taken in a 15 minute interval. As mentioned earlier, every spot was marked with a yellow-orange stake with the corresponding number where test participants had to estimate the shallow soil moisture. To prevent people from influencing the measurements by compressing the soil, the sensor was buried a little behind the stake in a depth of approximately 8-10 cm. The logger was attached to a metal rod driven into the soil. For each point, daily volumetric water content values were generated by calculating the average of all values per point per day. This method does not consider daily variations in surface soil moisture content but as these variations were negligibly small this approach seems appropriate for the purpose. For the entire duration of the test rain gauge measurements near measurement point 1 were carried out. For days without reference of the expert, a soil moisture class had to be assigned with the help of FDR measurements, meteorological data and expert knowledge to every sampling point. The 5 days with reference estimation were taken as base. Volumetric water content variability of point 9 and 10 over the entire test period and precipitation data was viewed. In between these reference estimations, class assignments were changed relative to changes in conditions. Rainfall events for example caused an increase of soil moisture class while periods without precipitation may have a decrease in wetness class entailed. Special conditions of certain sampling points have been taken into account by expert knowledge. While for example trough sites remain wet for some time, hilltop positions dry faster. Accordingly, wetness class assignments decrease shorter after a rain event. It should be noted that the expert didn't know any of the results from quantitative measurements at any time during the ongoing field campaign and the qualitative estimations are therefore unaffected.

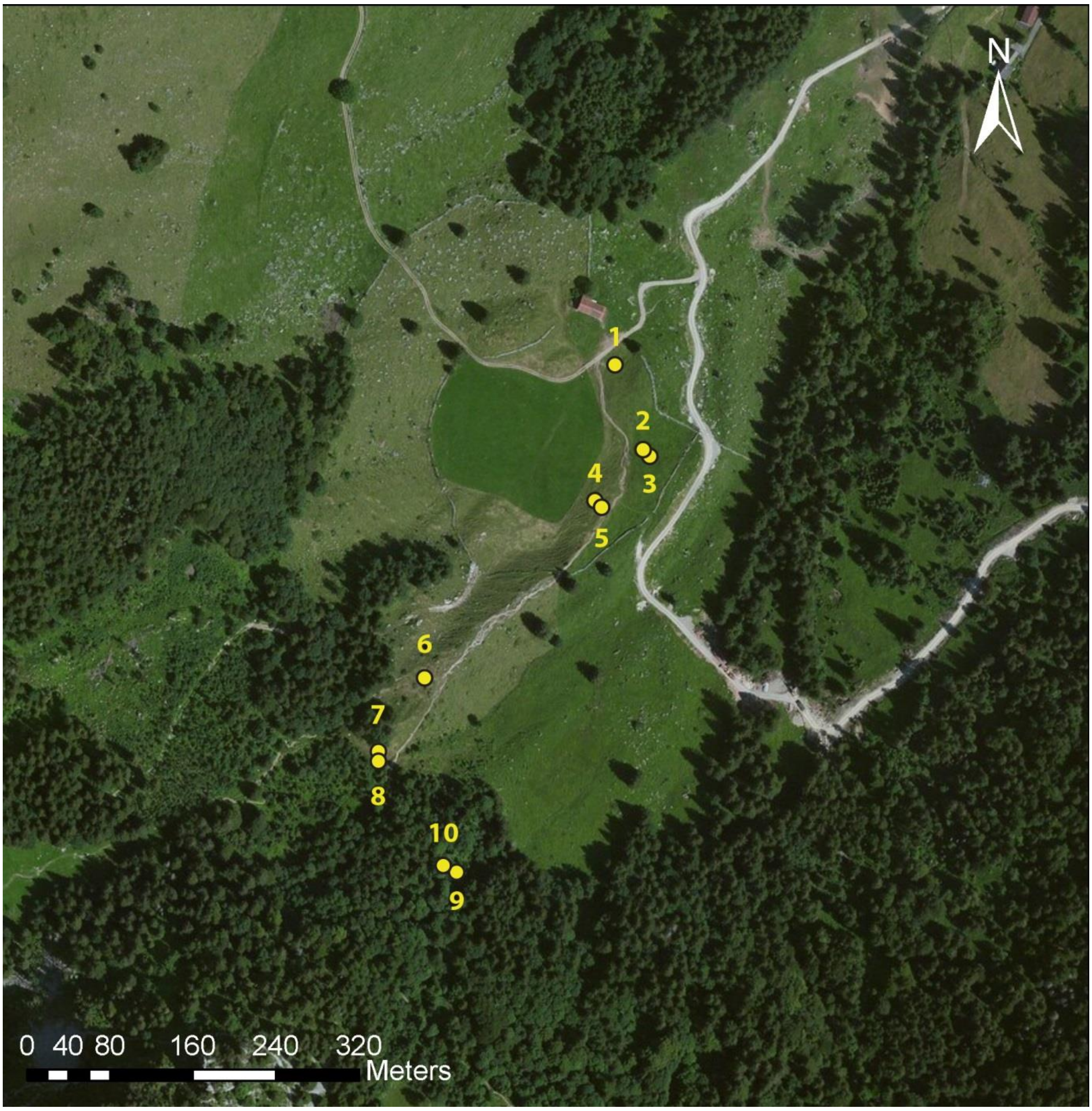


Figure 10: The sampling parcours based on GPS data for Brunni, Alpthal; map basis: ArcGIS World Imagery (2015).

For the FDR measurements the Greenhouse sensor (Fig. 7) of the company Decagon Devices was used. Together with the Em50G Remote Logger the sensor can be used for continuous measurements over a longer period. The sensor is equipped with three parallel stainless steel rods that can be inserted into the soil. Alternatively, the whole sensor can be buried into the soil. To operate the sensor it has to be connected to the battery-operated logger and settings carried out with the provided software “ECH_2O” via field laptop. Once the settings are defined, measurements are carried out automatically and data is stored on the logger. The stored data can either be downloaded directly, using the USB port and a field laptop or via internet from Decagon’s Internet server. The Em50G Remote Logger allows transmitting the stored data over the internet with GSM / GPRS (Global System for Mobile communications / General Packet Radio Service) cellular module. The GS3 sensor is able to measure volumetric water content, temperature and electrical conductivity. To determine the volumetric water content a 70MHz oscillating wave is sent to the sensor rods that charges according the dielectric of the material. The measured charge is expressed as dielectric permittivity and is then converted to water content by a calibration equation. It is possible to perform a media-specific calibration to enhance accuracy, for this test however, standard calibration was used. (Decagon Devices, 2014a; Decagon Devices, 2014b)

Figure 11: Information board (left) and orange-yellow stake with corresponding site number (right) at study area Alpthal.

For precipitation measurements the rain gauge model ECRN-100 of Decagon Devices was used. This is a high resolution rain gauge with two internal tipping spoons (one tip per 0.2 mm of rain). To operate the rain gauge it was connected to the Em50G remote logger and settings could be chosen via field laptop. The rain gauge was fitted about 1.70 m above ground and was attached to a metal rod driven into the soil. One precipitation dataset was generated with both the daily precipitation data of the meteorological weather station “Alpthal” operated by MeteoSchweiz (Precipitation, daily total from 05.40 AM till 05.40 AM the following day) and the measured precipitation by the installed rain gauge at the test site in Brunni (a 15 min measurement interval). Because the rain gauge was installed just at the parcours, very local rain events could be captured better and the high-resolution data enabled the control on the beginning of a rain event. Possible major changes of soil moisture during the day affecting qualitative soil moisture estimation could be detected and considered for the further analysis. Possible measurement errors due to heavy storms, the consequent disturbance of the measurement device or just defective measurements of the device itself, could be detected and eliminated with the precipitation data of MeteoSchweiz. In case of a large discrepancy between the measured precipitation data and the precipitation data of MeteoSchweiz latter was given preference. Otherwise data measured by the rain gauge was used for the final precipitation dataset.

4.3. Soil identification

To identify the two prevailing soil types (“white soil” and “red soil”) in Mungushi, Tanzania, two soil profiles were built. The following chapter is subdivided into a part about the site survey and the soil texture analysis in the laboratory.

4.3.1. Site survey

For each soil type a site which appeared to be suitable was chosen to build a soil profile. Profile 1 is located in the “red soil” and has a depth of 98 cm and profile 2 is located in the “white soil” and has a depth of 91 cm. (Fig. 12) Based on the guidelines for soil description (FAO, 2006) and the “Anleitung zur Standortaufnahme im Gelände” (Script GEO 241 based on Blume et al., 2011 and Stahr et al., 2000) a site survey had been made. Characteristics as for example depth, horizon boarders, abundance of rock fragments, soil type, structure, structure stability and colour were determined by use of easy field methods and only few simple aids. Soil samples of all horizons were collected and sent to the laboratory for further analysis. Notes about the soil characteristics were made in the field book and afterwards transferred to word and excel documents. A detailed documentation about the site survey and following soil textural analysis is gathered in one report.

Figure 12: Soil profile 1 (left) and 2 (right) with measuring sticks and markings symbolizing horizon boarders or small interlayers (depth of soil profile 1: 98cm and soil profile 2: 91cm).

4.3.2. Soil texture analysis (sieving, sedigraph)

Soil profile samples and selected gravimetric samples of the first field test in Tanzania were analysed. The samples of the two soil profiles had not been dried by then and were put into oven for about 24h at a temperature of 105°C. Each sample went through a sieving process with sieves with a mesh size of 2 mm and 250 µm and the proportion of the three grain-size ranges (>2 mm, <2 mm >250 µm and <250 µm) could be determined with a balance. Because the grain size distribution of three samples caused the sedigraph some trouble in the first round, additionally a sieve with a mesh size of 125 µm was used for these three.

For the sedigraph analysis 4 g of the soil sample <250 µm, respectively <125 µm, were weighed and put into suspension with 25 ml of tap water and 25 ml of sodium hexa metaphosphate (Egli et al., 2013). To obtain a good mixing the suspension was mixed with an ultra-sonic device for 5 minutes. The ultra-sonic device produces sound waves and thereby microscopic bubbles which when collapsing release energy by which the soil aggregates blast apart (Gee and Bauder, 1986). Afterwards, the suspension can be filled into the sedigraphs mixing chamber and the measurement can be started via associated computer program. Beforehand, the sedigraph was put into operation by placing and fixing all the hoses, rinsing the machine several times, and compiling a base line with a calibration solution.

5. Results

Various experiments and analysis were performed to answer the research questions. The results get presented subdivided into qualitative, qualitative vs. quantitative and quantitative soil moisture data chapters for all the different study areas.

5.1. Qualitative soil moisture data

In this chapter results of the qualitative soil moisture data for all 4 study areas are presented. Each field campaign is treated separately in its respective subchapter. For every study area, the deviation of wetness classification relative the median or a reference is presented. Additionally, the spread in wetness class assignment among the different wetness classes is shown for every test. The performance of individual raters was observed and two statistical tests were carried out to assess the agreement of different ratings within the dataset for the three similar tests of Tanzania and Zürich.

5.1.1. Tanzania: Test in April

To assess the overall performance of the qualitative classification scheme for arid environmental conditions the agreement between individual raters was investigated. The deviation of wetness classifications relative to the median of each sampling point was analysed and plotted as frequency distribution (Fig. 13). The median of all group classifications was chosen as reference as it is robust in terms of outliers and gives a good representation of the overall classification of a group of raters. In about 47% of all cases, the raters independently assigned the same wetness class and in about 87% of all cases, the raters assigned the same wetness class or were off by only one class (see Fig. 13). A classification difference of two classes (wetter and drier) was present in approximately 8% of all classification cases. An overestimation or underestimation of wetness by three or more wetness classes occurred in 5% of all cases.

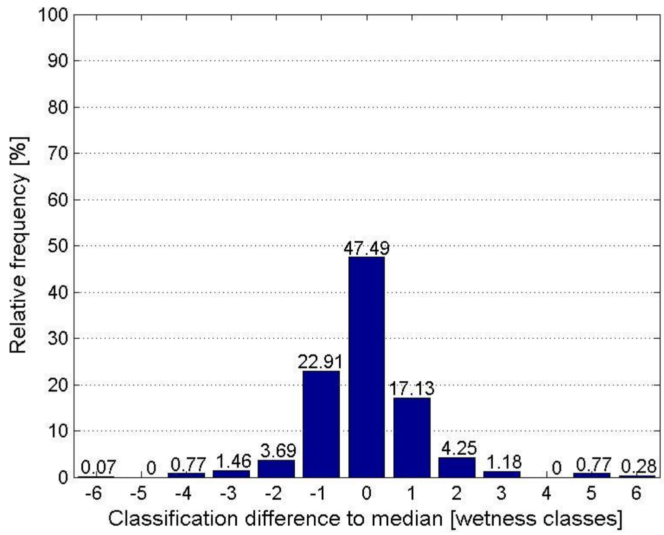


Figure 13: Deviation of wetness classifications relative to the median of each sampling point for all raters plotted as relative frequency distribution.

The separation of farmers, master students, and PhD+ showed that master and PhD+ students performed considerably better than farmers. In approximately 60% respectively 58% of all cases master and PhD students assigned the same wetness class and were off by one wetness class in 33% and 33% respectively (see appendix Fig. 40). Two groups were formed (Inf*all* and Not*all*) to assess if inter-rater reliability of the qualitative wetness classification can be improved by training. The agreement between individual raters of the two subgroups was investigated and is compared by relative frequency distributions of class assignments (Fig. 14). In approximately 50% of all cases, the raters of both groups (informed and not informed) independently assigned the same wetness class and in about 36% respectively 39% they were off by only one class. When looking only at the individual results of informed farmers or students compared to the not informed farmers or students, it is recognizable that assignments of farmers didn’t change much due to a better introduction. However informed master students achieved considerably better result than not informed master students (Fig. 15).

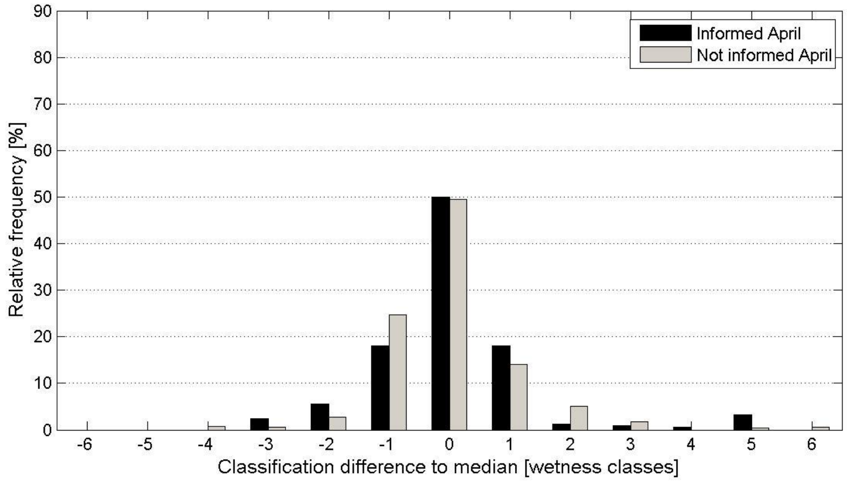


Figure 14: Deviation of wetness classifications relative to the median of each sampling point for the two subgroups "Inf*all*" and "Not*all*" plotted as relative frequency distribution.

The much narrower frequency distribution for the informed master students is especially noticeable. In over 99% of all cases, the raters of the group informed master students independently assigned the same wetness class or were off by only one class. In about 58% of all cases, raters of the group not informed master students assigned the same wetness class. At 77%, this value was nearly 20% higher for the group informed master students.

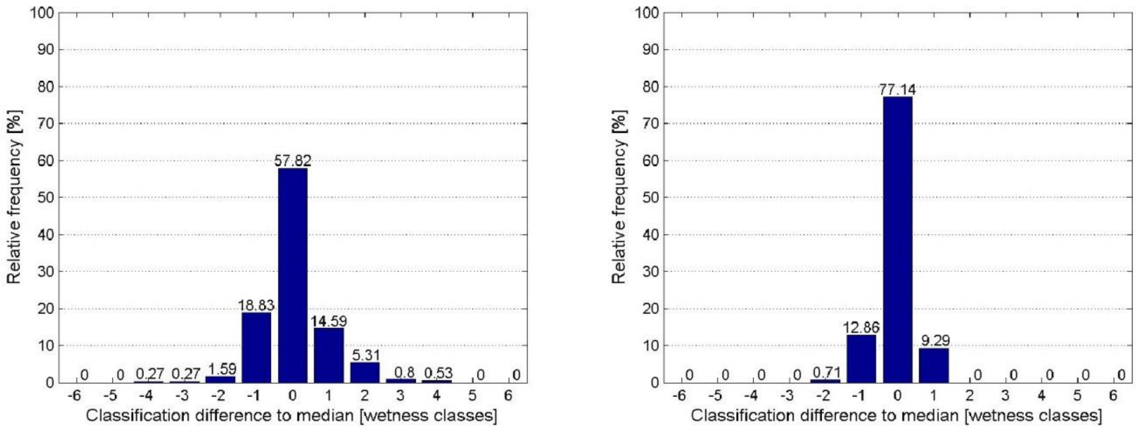


Figure 15: Deviation of wetness classifications relative to the median of each sampling point for "M*not*" (left) and "M*inf*" (right) plotted as relative frequency distribution.

In order to show if certain classes are more difficult to assign than others the spread in wetness class assignments among the seven different wetness classes for each sampling point was analysed (Fig. 16). In the plot points were sorted according to the median qualitative wetness classification of each sampling point (numbers 1 to 40). White circles show the median of the classifications, while gray shades indicate relative frequency of wetness class assignments.

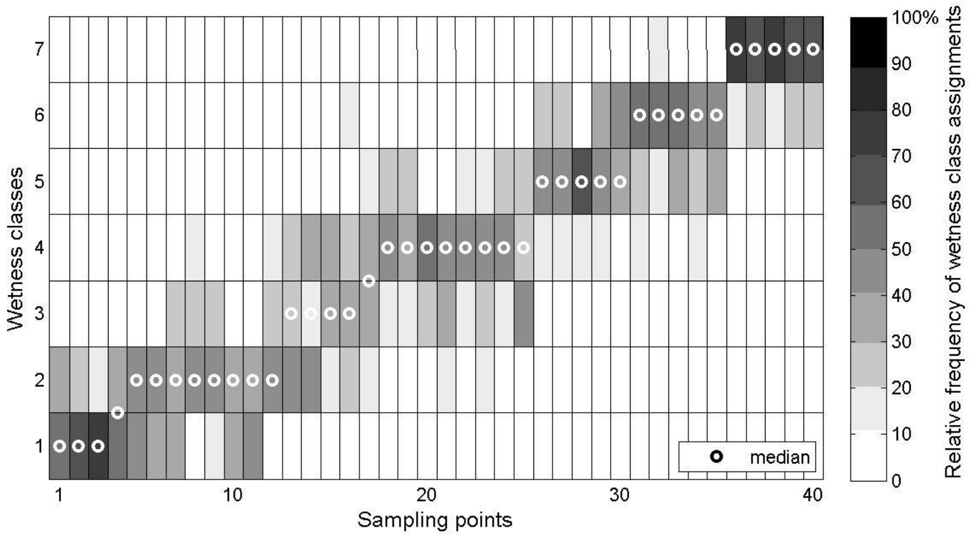


Figure 16: Wetness class assignments for each of the 40 sampling points numbered and sorted by the median of class assignments for each sampling point considering all raters. White circles show the median as reference and grey shades indicate relative frequency of wetness class assignments.

Class 7 and 1 show a high agreement among raters, this is represented by darker shades. All other classes show a stronger spread of wetness class assignments which is particularly large for wetness class 3. The number of sampling points per wetness class is not equally distributed as raters classification results can't be known in advanced. The better result of the informed master student group shown before can also be recognized in the spread in wetness class assignments for the individual groups (Fig. 41 in the appendix). Classes 5 and 6 show a much better agreement for raters of the informed master student group.

Another possibility to identify classes that are more difficult to assign and analyse if individual raters perform differently, is to analyse the mean difference of classification to the reference for all sampling points of a certain wetness class per rater (Fig. 17). Differences are either shown in blue or red, indicating too wet respectively too dry classifications compared to the reference. The mean differences of classification of the classes did not show a systematic pattern, except for class 7 and class 6 that was generally assigned too dry. It showed that over- and underestimations of the wetness conditions were rather random. Raters 1 and 5 showed the largest differences, classes 1-3 were classified too wet while classes 4-7 were classified too dry. In general, raters 1-14 had the tendency to classify classes 1-3 too wet and classes 4-7 too dry.

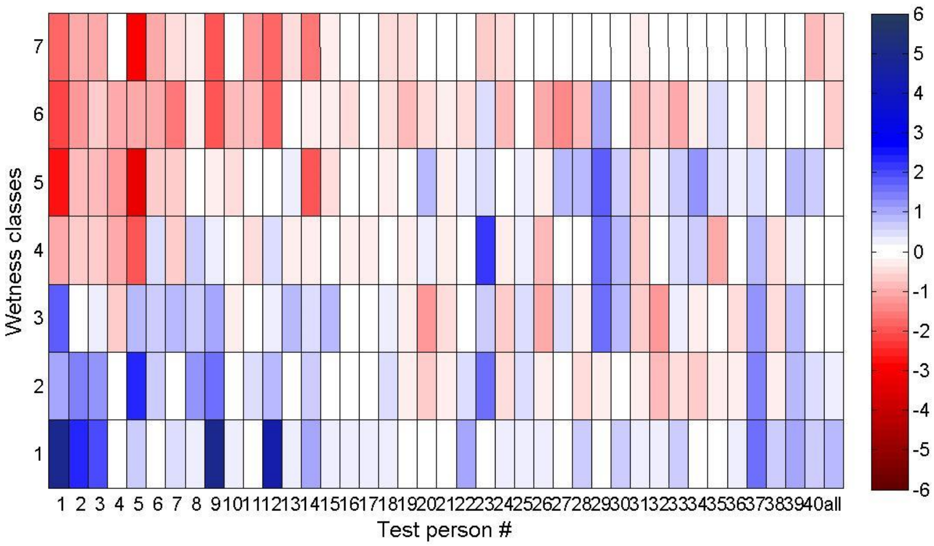


Figure 17: Mean difference of classification to the reference for all sampling points of a certain wetness class per rater considering all raters. Blue shades indicate positive classification differences (too wet) and red shades negative classification differences (too dry).

The dataset was divided into two subsets to examine if raters improve in the course of the test. The first 20 sampling points of the parcours were assigned as “firsthalf” (Fi*all*), and the second 20 sampling points as “secondhalf” (Se*all*). The deviations of wetness classification relative to the median of each sampling point of the first half and the second half for all test participants is plotted as relative frequency distribution (see Fig. 42 in the appendix). In about 48% of all cases for all raters of the first half of sampling points and about 46% for the second half of sampling points raters assigned the same wetness class. An overestimation or underestimation of wetness by one class occurred in approximately 38% (Fi*all*) and 39% (Se*all*) of all cases. For further information see Table (6) in the appendix.

Krippendorff's alpha (KA) and Cohen's kappa (CK) were calculated to quantify the degree of agreement among raters statistically. Cohen's Kappa is used to compare the agreement between one rater and a reference and Krippendorff's alpha is to assess the agreement among a group of raters. Krippendorff's Alpha is defined as follows:

$$\alpha = 1 - \frac{D_o}{D_e}$$

D_o is the observed disagreement among values assigned to units of analysis and D_e is the expected disagreement when assuming random assignments (Krippendorff, 2011). D_o is approaching zero and α theoretically tends towards 1 if there is a high degree of agreement among raters. If raters disagree completely and assignments were made randomly, D_o would equal D_e and therefore Krippendorff's alpha would theoretically be zero.

Cohen's kappa is a coefficient of agreement between two ratets and is defined as the proportion of agreement without chance agreement (Cohen, 1960). The equation for K is:

$$K = \frac{Po - Pe}{1 - Pe}$$

where Po is the relative observed agreement among two raters and Pe is the agreement if classes were assigned randomly. If the observed agreement equals the random agreement K theoretically equals zero and if the raters agree completely K theoretically equals 1. Instead of comparing two raters, here a rater was compared to a reference. Hypothetically the maximum possible value for Cohen's kappa is 1, but as raters did not assign classes equally frequent maximal attainable kappa values are normally below 1. Krippendorff's alpha and Cohen's kappa values for the test in April are summarized in Table 1. Both values are given in percentage, where 0% means no agreement and 100% a perfect agreement.

The Krippendorff's alpha value for all raters of the test in April shows an agreement of 66%. As already mentioned earlier, master and PhD+ students (KA of PhD+: 82%, KA of M*all*: 83%) performed considerably better than farmers (KA of F*all*: 42%). With 68% the Not*all* who only got a basic introduction reached a higher value than the group Inf*all* that received a more detailed introduction. For all participants of this first test in April, the median value for the maximal attainable kappa is 0.76. For this purpose the ration between the CK value and maximal attainable kappa was considered. The medians and the interquartile range (IQR) of CK/CKmax values for each group are presented in Table 1. CK/CKmax values obtained from the test in April ranged between 43% and 83% and the median CK/CKmax considering all participants was 51%.

Table 1: Krippendorff's alpha und Cohen's kappa values for the different groups. (F: farmers, M: master students, PhD+: PhD students and lecturers, Inf: informed (a more detailed introduction), Not: not informed (only a basic introduction), Fi: first half of sampling points, Se: second half of sampling points, *all*: all of this group have been considered, *inf*: informed (only the subgroup with the more detailed introduction has been considered), *not*:not informed (only the subgroup with the basic introduction has been considered)

Test	Groups	Krippendorff Alpha [%]	Median CK/CK_{max} [%] (IQR)
April	F*all*	42	43 (38-70)
	F*inf*	41	65 (53-79)
	F*not*	49	53 (46-60)
	M*all*	83	66 (52-72)
	M*inf*	91	83 (77-87)
	M*not*	81	67 (61-71)
	PhD+	82	66 (60-69)
	Inf*all*	51	57 (38-64)
	Not*all*	68	56 (36-60)
	Fi*all*	66	56 (42-70)
	Se*all*	64	57 (38-76)
	All	66	51 (37-62)

5.1.2. Tanzania: Test in June

Data analysis of the test in June was similar to the test in April. The deviation of wetness classifications relative to the median of each sampling point was analysed and plotted as frequency distribution (Fig. 18). In approximately 63% of all cases, the raters (*all*) independently assigned the same wetness class and in about 94% of all cases, the raters assigned the same wetness class or were off by only one class. In approximately 6% of all classification cases a classification difference of two classes or more (wetter and dryer) was observed.

When looking at farmers and experts individual results, it is noticeable that farmers reached a higher interrater agreement compared to the first test in April (Table 6 in the appendix). Farmers assigned the same wetness class in approximately 66% of all cases and they assigned the same class or were off by only one class in about 94%. Experts assigned the same wetness class in approximately 59% of all cases and they assigned the same class or were off by only one class in about 96%. Similar to the field campaign in April, the group of participants was divided into two subgroups (F*not* and F*inf*, only farmers as experts already have some experience). The agreement between individual participants of the two subgroups was investigated and is compared by relative frequency distributions (Fig. 19).

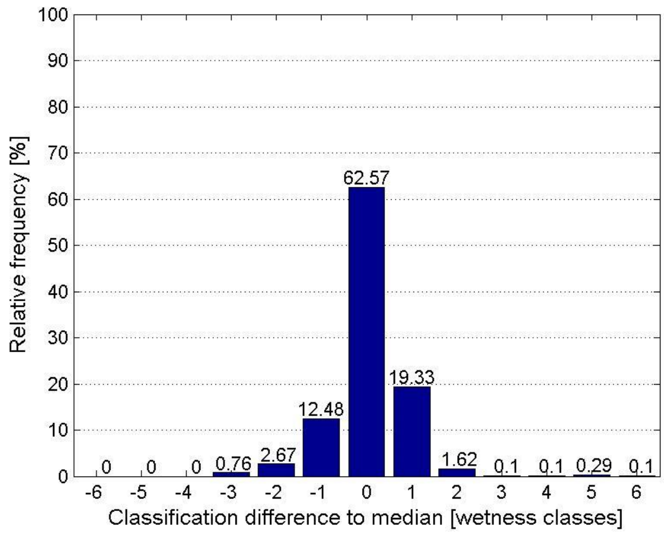


Figure 18: Deviation of wetness classifications relative to the median of each sampling point for all raters plotted as relative frequency distribution.

In about 69% of all cases for F*inf* and in about 66% for F*not* the raters independently assigned the same wetness class. In approximately 97% (F*inf*) respectively 89% (F*not*) of all cases, raters assigned the same wetness class or were off by only one class. A classification difference of two classes (wetter and drier) was present in approximately 2% of all classification cases for the group "Inf*all*". The group "Not*all*" had a higher value with about 7%. Classification differences of more than two classes occurred in less than 1% (F*inf*) but the group F*not* reached a value of about 3%.

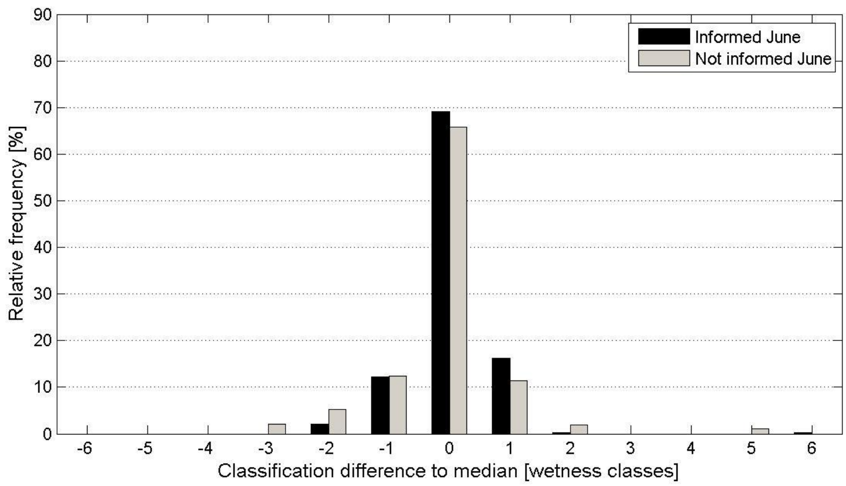


Figure 19: Deviation of wetness classifications relative to the median of each sampling point for the two subgroups "F*inf*" and "F*not*" plotted as relative frequency distribution.

For the second test in June the spread in wetness class assignments among the seven different wetness classes for each sampling point was examined as well (Fig. 20). Of all classes, class 3 shows the largest spread of class assignments while class 1 shoes the best agreement. The spread of class assignments of class 5 and 6 look very similar and show a relatively narrow frequency distribution when compared to classes 2 and 3. For none of the points the median wetness class was class 7. The number of sampling points per wetness class was not equally distributed as raters classifications were not known in advanced.

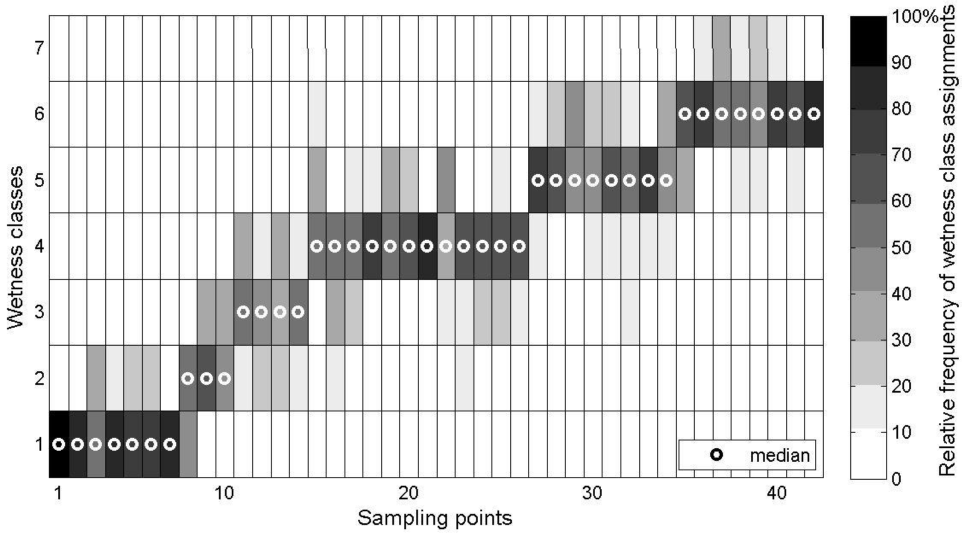


Figure 20: Wetness class assignments for each of the 42 sampling points numbered and sorted by the median of class assignments for each sampling point. White circles show the median as reference and grey shades indicate relative frequency of wetness class assignments.

The mean difference of classification to the reference for all sampling points of a certain wetness class per rater was plotted to see if individual raters performed differently or if certain classes were systematically over- or underrated (Fig. 21). In general no systematic pattern is recognizable and the mean classification differences to the reference exceed 2 classes in only 1% of all cases. All pads of class 7 remain white as no sampling point was assigned to the wettest class by the median. Some raters show the tendency to classify either too wet or too dry over nearly all classes while others alternate. As previously done during the test in April, the data set of the second test in June was divided into the two subsets "Fi*all*" and "Se*all*". The first subset contains sampling points 1-21 and the second sampling points 22-42. The deviations of wetness classification relative to the median of each sampling point of the first half and the second half for all raters is plotted as relative frequency distribution (Fig. 43 in the appendix).

For all raters in the first half of sampling points raters assigned the same wetness class in approximately 61% of all cases. They assigned the same wetness class or were off by only one class in about 95%. For all raters in the second half of sampling points raters assigned the same wetness class in approximately 64% of all cases and in about 94% they assigned the same wetness class or were off by only one class (see Table 6 in the appendix).

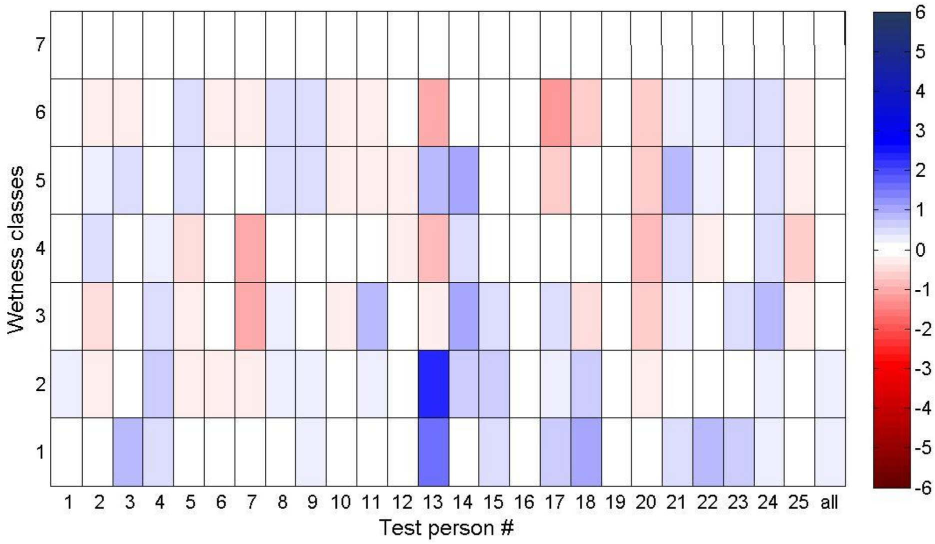


Figure 21: Mean difference of classification to the reference for all sampling points of a certain wetness class per rater. Blue shades indicate positive classification differences (too wet) and red shades negative classification differences (too dry).

Krippendorff's alpha (KA) and Cohen's kappa (CK) were calculated to also statistically quantify the degree of agreement. Krippendorff's alpha and Cohen's kappa values for the test in June are summarized in Table 2. The Krippendorff's alpha value for all participants of the test in June was 78% and the median CK/CKmax value (67%) is also higher compared to the one calculated for the test in April (see Table 1). KA of F*all* was similar to the one of E*all* but the median and the IQR of CK/CKmax of F*all* was 75% compared to 59%.F*inf* show a higher agreement among each other than farmers who just received a basic introduction (F*not*) (see Fig. 19). This circumstance is also evident from the KA values. KA of F*inf* was 87%, while KA of F*not* was 65%. For all raters ("All") of the test in June the values for the maximal attainable Cohen's kappa vary between 0.68 and 0.91.

Table 2: Krippendorff's alpha und Cohen's kappa values for the different groups. (F: farmers, E: experts , Fi: first half of sampling points, Se: second half of sampling points, *all*: all of this group have been considered, *inf*: informed (only the subgroup with the more detailed introduction has been considered), *not*:not informed (only the subgroup with the basic introduction has been considered)

Test	Groups	Krippendorff Alpha [%]	Median CK/CKmax [%] (IQR)
June	F*all*	76	75 (63-81)
	F*inf*	87	79 (77-85)
	F*not*	65	75 (70-83)
	E*all*	84	59 (56-70)
	Fi*all*	78	71 (55-86)
	Se*all*	79	75 (58-86)
	All	78	67 (59-73)

5.1.3. Zürich, Irchel

Data analysis for the data collected in Zurich was performed as described above. Figure 22 shows the deviation of wetness classifications relative to the median of each sampling point plotted as frequency distribution. In approximately 73% of all cases, the raters independently assigned the same wetness class and nearly 99% raters assigned the same wetness class or were off by only one class. An over- or underestimation of wetness by two classes occurred in less than 2% of all classification cases.

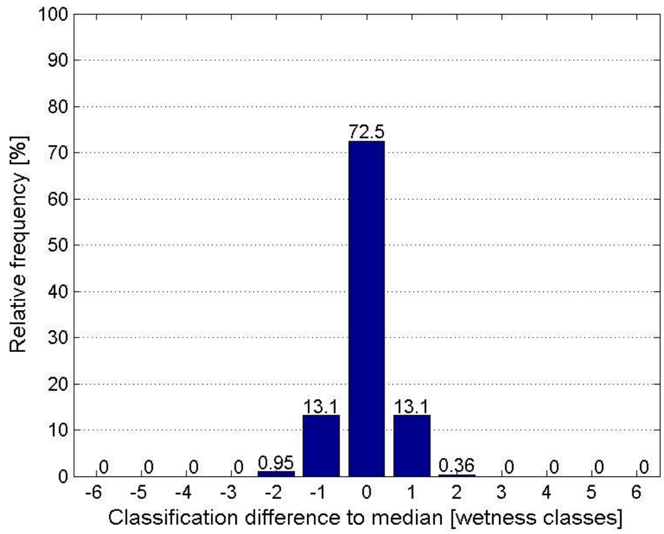


Figure 22: Deviation of wetness classifications relative to the median of each sampling point for all raters plotted as relative frequency distribution.

When looking at the individual results of master and bachelor students (Table 6 in the appendix) it is noticeable that bachelor and master students have relatively similar results. In approximately 71% of all cases, bachelor students assigned the same wetness class and in about 99% they assigned the same class or were off by only class. Master students assigned the same wetness class in approximately 74% and reached 98% with classification differences not larger than 1. The participants were divided into the two subgroups "Inf*all*" and "Not*all*". To ensure that results are not just influenced by one group, PhD, master and bachelor students were divided into the two different subgroups. The agreement between individual participants of the two subgroups was investigated and is compared by relative frequency distributions (Fig. 23). Results for these two subgroups looked very similar and in approximately 74% ("Inf*all*") and 75% ("Not*all*") of all cases, raters assigned the same wetness class. In about 99.5% of all cases for both groups, raters assigned the same wetness class or were off by only one class. Classification differences of more than two classes are below 1% for both groups.

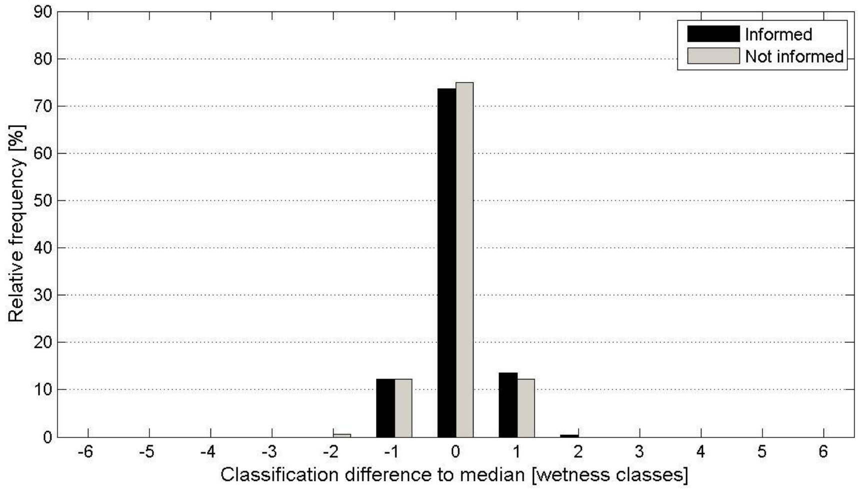


Figure 23: Deviation of wetness classifications relative to the median of each sampling point for the two subgroups "Inf*all*" and "Not*all*" plotted as relative frequency distribution.

Spread in wetness class assignments among the seven different wetness classes for each sampling point was examined for the data collected in Zurich as well. (Fig. 24). The three wettest classes show the best agreement. Raters had most difficulties to assign Class 2 and 3 as the largest spread of class assignments could be found for these two classes. Class 1 and class 4 showed a relatively narrow frequency distribution indicating a high degree of agreement accordingly. Even though raters classification results were not known in advanced, all classes were represented by at least 4 sampling points.

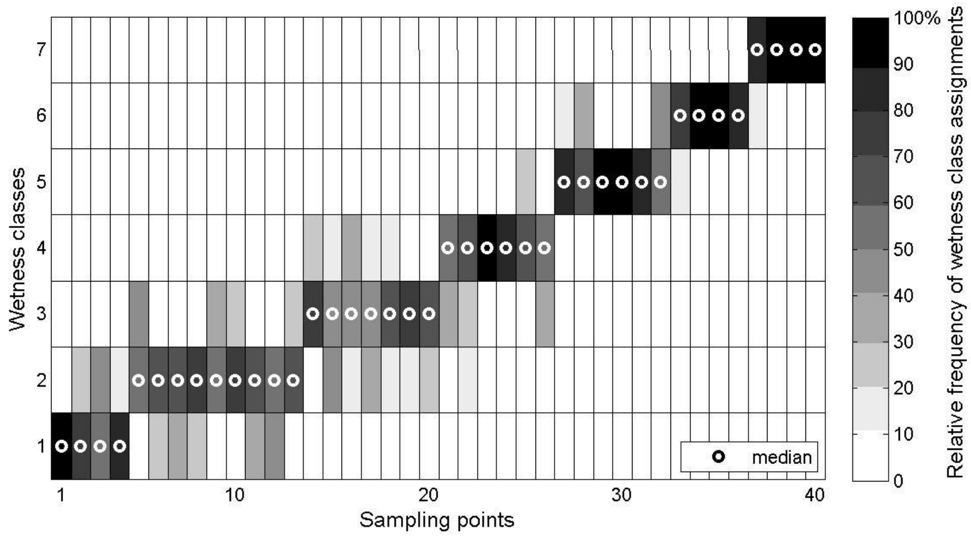


Figure 24: Wetness class assignments for each of the 40 sampling points numbered and sorted by the median of class assignments for each sampling point considering all raters. White circles show the median as reference and grey shades indicate relative frequency of wetness class assignments.

The mean difference of classification to the reference for all sampling points of a certain wetness class per rater was plotted to see if individual raters performed differently or if certain classes were systematically over- or underrated in the test in Zurich (Fig. 25).

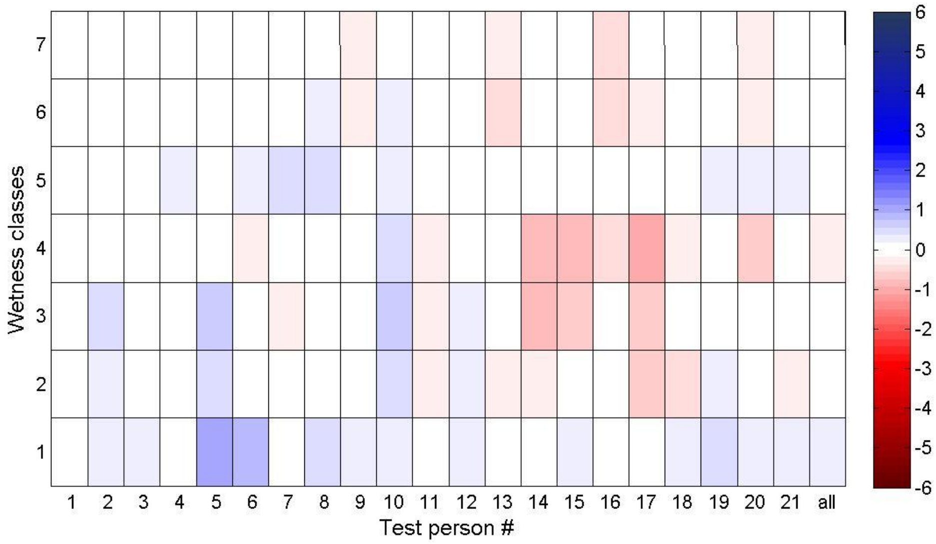


Figure 25: Mean difference of classification to the reference for all sampling points of a certain wetness class per rater. Blue shades indicate positive classification differences (too wet) and red shades negative classification differences (too dry).

It is apparent that the 3 wettest classes have the smallest classification difference (white or only slightly dyed / shaded pads). Class 4 shows the tendency of being classified slightly too dry while class 1 was classified slightly too wet. No clear pattern is visible for the other classes. As in the previous two tests, the data was divided into the two subsets "Fi*all*" and "Se*all*". The first subset contains sampling points 1-20 and the second sampling points 21-40. For the first half as well as for the second half of the data the deviation of wetness classification relative to the median of each sampling point for all test participants is plotted as relative frequency distribution (see Fig. 44 in the appendix). In approximately 70% of all cases, for all raters in the first half of sampling points raters assigned the same wetness class and in about 98% they assigned the same class or were off by only one class. In approximately 75% of all cases, for all rater in the second half of sampling points raters assigned the same wetness class and in more than 99% they assigned the same class or were off by only one class (see Table 6 in the appendix).

Table 3: Krippendorff's alpha und Cohen's kappa values for the different groups.(B: bachelor students, M: master students,, Inf: informed (a more detailed introduction), Not: not informed (only a basic introduction), Fi: first half of sampling points, Se: second half of sampling points, *all*: all of this group have been considered, *inf*: informed (only the subgroup with the more detailed introduction has been considered), *not*:not informed (only the subgroup with the basic introduction has been considered)

Test	Groups	Krippendorff Alpha [%]	Median CK/Ckmax [%] (IQR)
ZRH	B*all*	92	77 (72-80)
	B*inf*	91	80 (72-83)
	B*not*	93	86 (83-89)
	M*all*	91	82 (76-93)
	M*inf*	93	88 (81-89)
	M*not*	92	90 (86-95)
	Inf*all*	92	78 (70-91)
	Not*all*	92	82 (77-93)
	Fi*all*	91	84 (69-93)
	Se*all*	93	86 (77-100)
	All	92	77 (73-90)

As for the previous tests in Tanzania, Krippendorff's alpha and Cohen's kappa were calculated for the test in Zürich. Krippendorff's alpha und Cohen's kapp values for the test in Zürich are summarized in Table 3. The KA value for all participants of the test shows a very high agreement of 92% and also the median and the IQR of CK/CKmax is considerably higher than for the two tests in Tanzania (77, 73-90). There was no significant difference between Inf*all* and Not*all* (see Fig. 23 and Table 3). Inf*all* and Not*all* both have a KA value of 92% and the medians of CK/CKmax of these groups are similar

(Inf*all*: 78%, Not*all*: 82%). The median value of CK/CK max of master students (M*all*: 82%) is higher than the one and the one of bachelor students (B*all*: 77%).

5.1.4. Brunni, Alpthal

The test was on a voluntary basis and every person passing by, independent of his or her age or education had the possibility to take part at any time. For this reason the data was examined for plausibility. In total 6 out of 102 people had to be excluded for further analysis as their results indicate that they did not understand the task properly and haphazardly assigned wetness classes. This makes about 5.9% of the total amount of filled in forms that had to be eliminated. For about 67% of all 96 forms, wetness class estimations for all 10 sampling points were made. Estimations for 8-9 sampling points were made in 19% of all cases and 3-7 sampling points reached a value of 5%. 9% of all forms contained only two estimations of wetness classes.

To assess if the qualitative wetness classification scheme for humid environmental conditions is applicable with a minimum of information and on a voluntary basis, the agreement between individual raters and a reference was investigated. The deviation of wetness classification relative to the reference of each sampling point at a certain day was analysed and plotted as frequency distribution (Fig. 26). The median could not be taken into account for this analysis as not enough data is available for every day.

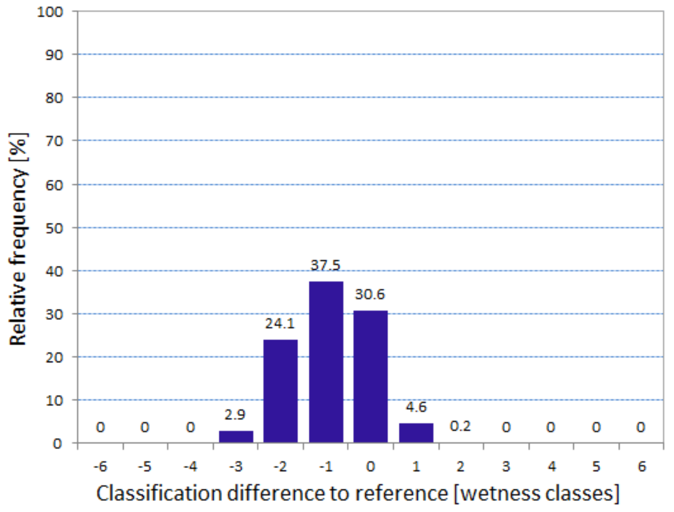


Figure 26: Deviation of wetness classification relative to the reference of each sampling point at a certain day for all raters plotted as relative frequency distribution.

The raters who participated in the test in Brunni tended to classify the sampling points too dry relative to the reference (see Figure 26). In approximately 31% of all cases, the raters independently assigned the same wetness class and in about 73% they assigned the same class or were off by only one class relative to the reference. While the classification difference to reference of -1 has a relative frequency value of about 38%, +1 only has a relative frequency value of about 5%. An overestimation or underestimation of wetness by two classes occurred in about 24% of all cases and in approximately 3% the raters were off by -3 wetness classes.

The wetness class assignments of the participants were plotted against the reference wetness class assignments to visualise the spread in wetness class assignments among the seven wetness classes for all sampling points (Fig. 27). The size of the circle is proportional to the quantity of assignments divided through the absolute number for the regarding wetness class assignments combination. In the optimal case all assignments would lie on the diagonal between point x=1, y=1 and x=7, y=7 indicated by the red line.

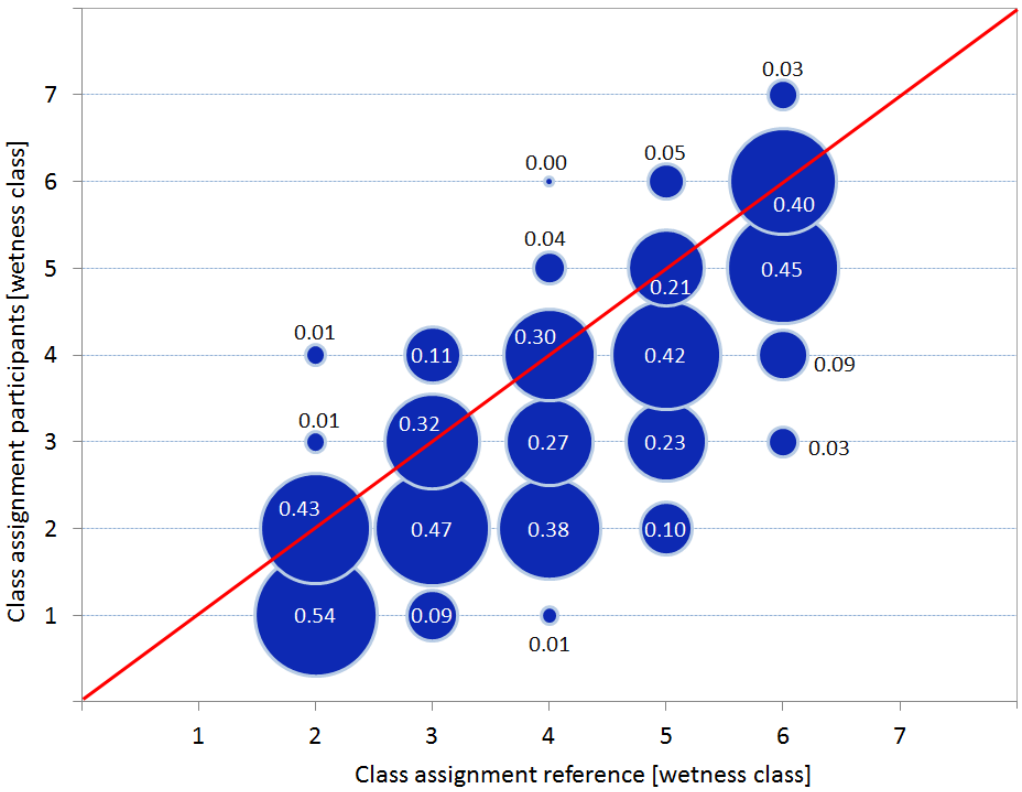


Figure 27: Spread in wetness class assignments among the seven wetness classes for all sampling points and all raters. The circle size is proportional to the quantity of assignments for the regarding wetness class assignments combination. The red line symbolizes the perfect assignment to the reference (1:1 line). Numbers on or next to circles refer to the proportion of numbers of assignments within this wetness class (reference).

This plot also shows that raters assignments tended to be drier than the reference assignment. For the reference class assignments no values for classes 1 and 7 are present as the reference person never assigned these two classes. In none of the cases the most frequent assignment was the same class as the reference. Class 4 showed the largest spread in class assignment (assignments spread over 6 classes) and the most frequent assignment of raters was off by two classes. Class 1 shows the the smallest spread of wetness class assignment ranging over 4 classes but most of them concentrated in only 2 classes.

5.2. Quantitative soil moisture data

In this chapter measured volumetric water content values of the different study areas are presented. They were either measured with TDR, FDR or gravimetric method. In case two different methods were used simultaneously at a test side, comparison between these two datasets can be made. In the following graph (Fig. 28) FDR and gravimetric volumetric water content values of the field test in Zürich are compared. The volumetric water content determined with FDR is generally higher across all measurement points but no clear pattern is recognizable.

It is quite different though for the measurements in Kichangani, where especially for wet sites volumetric water content determined by gravimetric sampling is considerable below the values determined by FDR (Fig. 29). Again no clear pattern in the differences between the volumetric water content determined by TDR and gravimetric sampling is recognizable for the 16 sampling points in Mungushi. All measured soil moisture content values for the different study areas are summarized in tables in the appendix (Tables 7, 8 and 9).

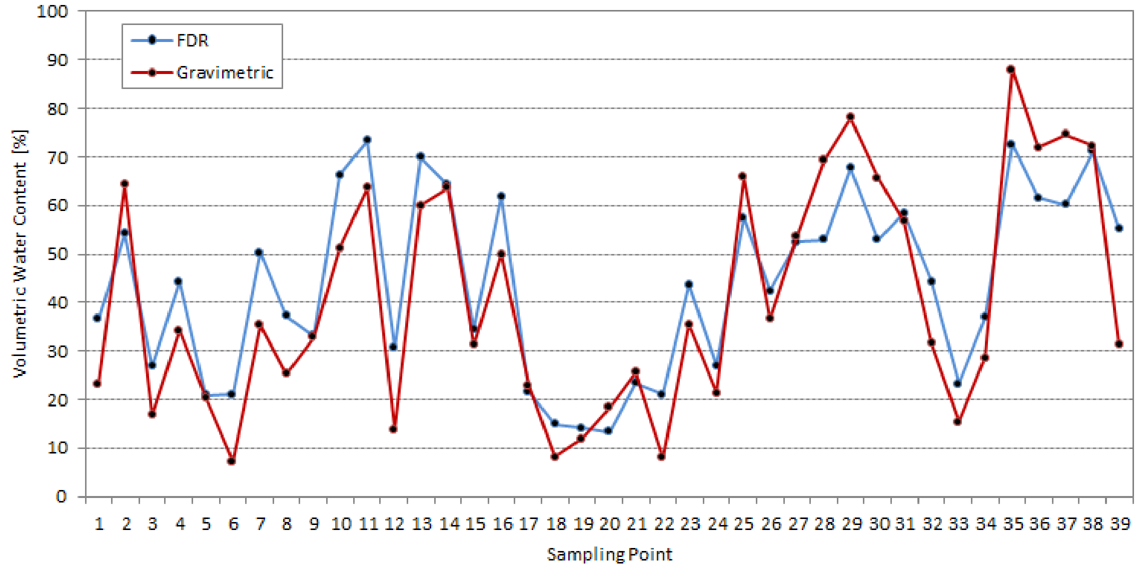


Figure 28: FDR and gravimetric measurements for all sampling points of the field test in Zurich, Irchel compared to each other.

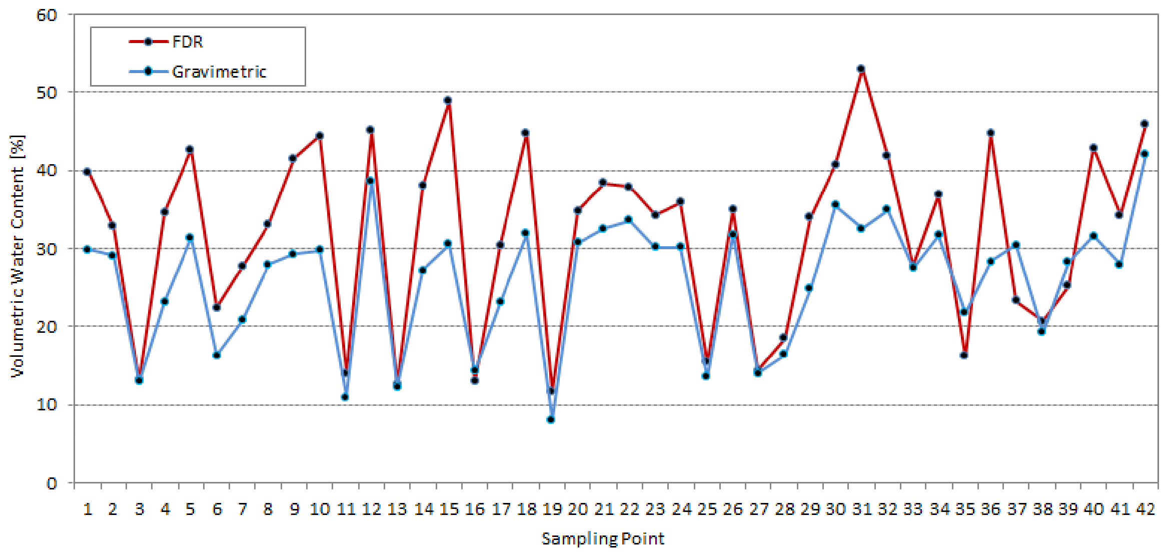


Figure 29: FDR and gravimetric measurements for all sampling points of the field test in Kichangani compared to each other.

The volumetric water content values measured by TDR or FDR device were compared to the gravimetrically determined volumetric water content values by means of regression analysis for both tests in Tanzania and for the test in Zürich. The analysis was performed with the statistics program SPSS. Volumetric water content values showed a good correlation with gravimetrically measured water contents for all tests. The following Pearson's correlation coefficients were observed: 0.61 for the test in Tanzania in April, 0.77 for the test in Tanzania in June and 0.82 for the test in Zürich. The calculated regression equations are as follows:

$$(1)\ \ TDR_{av} = 3.287 + 0.869 * gravim.\ (R^2 = 0.614)$$
$$(2)\ \ FDR_{av} = -0.965 + 1.261 * gravim.\ (R^2 = 0.773)$$
$$(3)\ \ FDR_{av} = 14.469 - 0.725 * gravim.\ (R^2 = 0.820)$$

TDRav respectively FDRav stand for the averaged volumetric water content values, measured with the TDR or FDR devices. The term gravim stands for the volumetric water content determined by the gravimetric method. For all 3 regression equations the explained variability (R^2) is listed. Explained variabilities of 61.4%, 77.3% and 82% were found for the calculated regression equations (1), (2) and (3).

Figure 30: Sampling points 9 and 10 of the study area in Brunni, Alpthal.

For the study area in Brunni, Alpthal only FDR measurements were conducted. These measurement series for all points are compared to precipitation data (Fig. 31). Except for sampling point 9 and 10

(Fig. 30) volumetric water content variations due to precipitation events are small. Surface soil moisture variations for the sampling points 1-8 could not be detected in a depth of approximately 8-10cm and therefore only measurements of sampling point 9 and 10 were considered for further analysis (raters vs. volumetric water content and the spread of volumetric water content for soil samples of each wetness class).

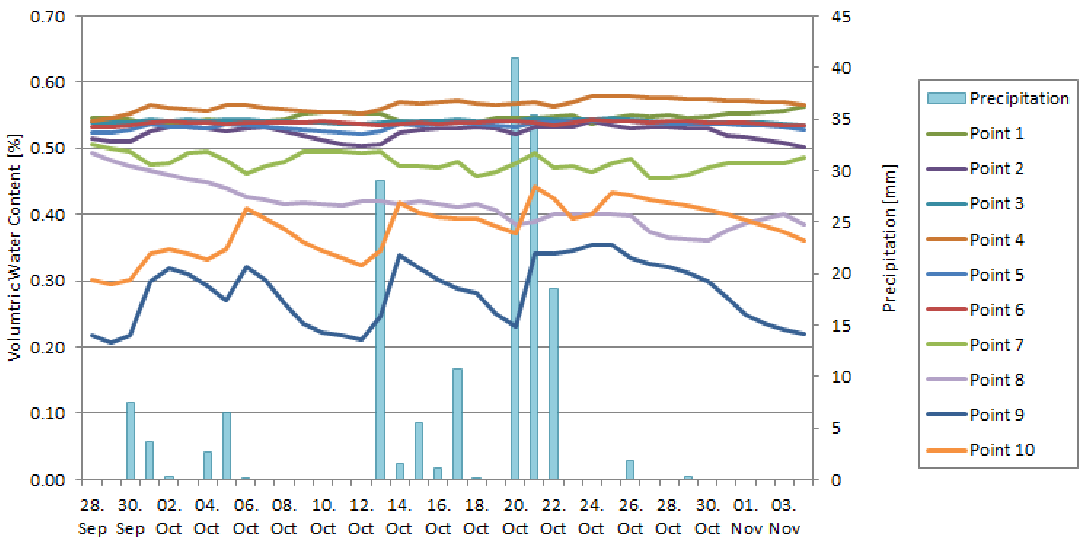


Figure 31: FDR measurements for all sampling points and precipitation measurements of the field test in Brunni, Alpthal.

5.3. Quantitative vs. qualitative determined/measured soil moisture

Measured volumetric water contents are compared to the qualitatively defined wetness classes for all field campaigns to asses if the qualitatively defined wetness classes represent quantitative differences in soil moisture content. First the results from the study areas in Tanzania are presented, followed by the study areas in Switzerland.

5.3.1. Tanzania: Test in April

To show the correspondence between qualitative wetness classes and quantitative measurements, the spread of volumetric water content for soil samples of each wetness class determined by the gravimetric method was plotted against the associate qualitative wetness class (Fig. 32). The qualitative wetness class assignments of the experts were used as reference. The median of all classifications was not considered as there was some time difference between the test and the withdrawal of the gravimetric samples. There is a difference in water content recognizable for the qualitative wetness classes. For soil samples collected in April the median volumetric water content ranged from 16% to 39%. Between classes 2 and 6, an increase of median water content and its 25- and 75% quantiles could be observed. Class 1 and 2 had a similar median water content and 25- and 75% quantiles. When looking at class 6 and 7, it is discernible that the median water content of class 7 is below the median water content of class 6 and also the class 7 25% quantile is below the quantile of class 6.

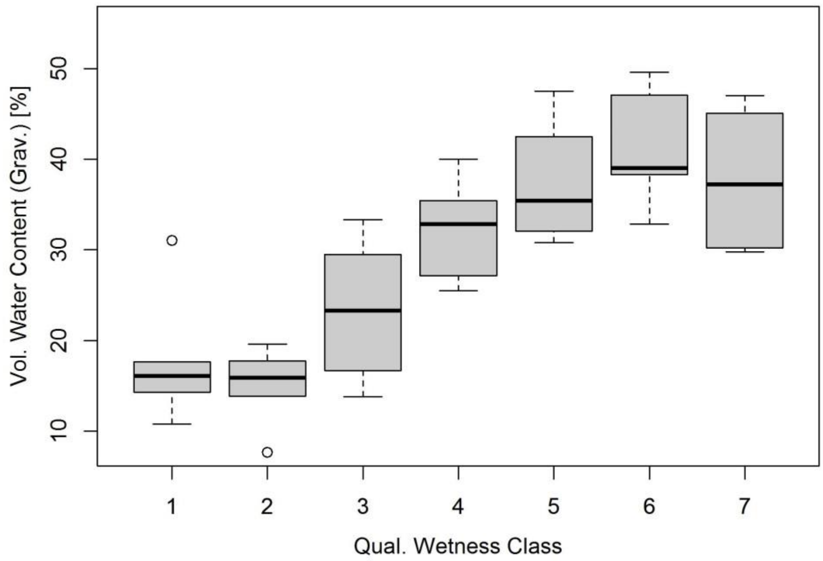


Figure 32: Volumetric water content for soil samples of each wetness class determined by the gravimetric method during the first test in Tanzania in April.

A Kruskal-Wallis test was carried out to analyse if the differences in volumetric water content for the wetness classes are significant. It is a nonparametric statistical procedure for comparing more than two samples that are independent, or not related (Kruskal and Wallis, 1952; Corder and Dale, 2009). A Kruskal-Wallis test with an adjusted level of risk of 0.002 to take into account the number of comparisons (Corder and Dale, 2009) indicated that the volumetric water content of the different qualitative wetness classes was only statistically significant for classes 2, 7. A Kruskal-Wallis test with a level of risk of 0.05 indicated that the classes 1, 2, 3; classes 3, 4 and classes 4, 5, 6, 7 were not significantly different from each other.

5.3.2. Tanzania: Test in June

The median volumetric water content ranged from 14% to 32% for soil samples collected in Tanzania in June (Fig. 33). The qualitative wetness class assignments of the experts were used as reference. In general, there is a difference in water content recognizable for the qualitative wetness classes. Between classes 2 and 5 there is an increase of median water content and its 25- and 75% quantiles identifiable. Classes 1 and 2 have different median water contents but the IQR strongly overlaps. For classes 5, 6, 7 and 4, 5 there is nearly no difference in median volumetric water content nor interquartile range (IQR) observable.

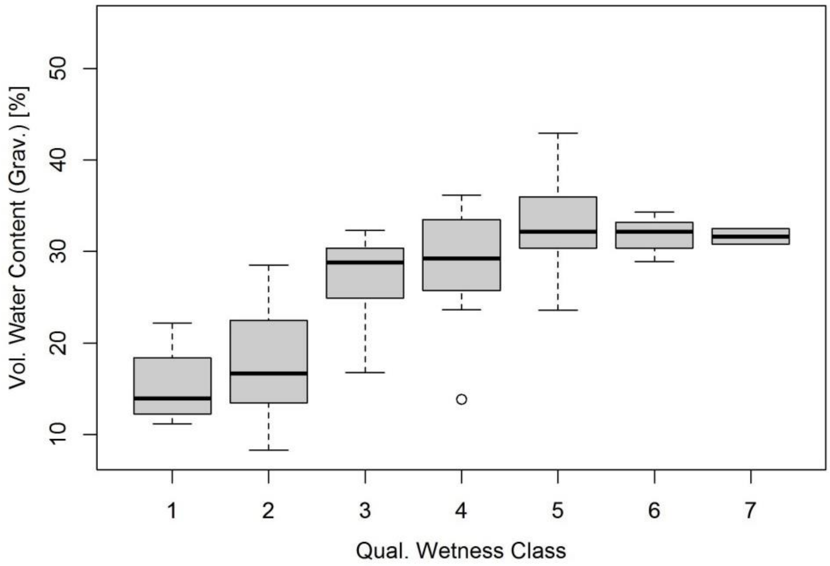


Figure 33: Volumetric water content for soil samples of each wetness class determined by the gravimetric method during the test in Tanzania in June.

A Kruskal-Wallis test with an adjusted level of risk of 0.002 indicated that the volumetric water content for the different qualitative classes was only statistically significant for classes 2, 5 and classes 2, 6. A Kruskal-Wallis test with a level of risk of 0.05 indicated that the classes 1, 2; classes 3, 4, 5, and classes 4, 5, 6 were not significantly different from each other. As only two samples for class 7 were available, this class could not be assessed.

5.3.3. Zürich, Irchel

In Figure 34 the spread of volumetric water content for soil samples of each wetness class determined by the gravimetric method was plotted against the associate qualitative wetness class. In this way the correspondence between the qualitative wetness classes and quantitative measurements could be investigated. The qualitative wetness class assignments of the experts were used as reference. For soil samples collected in Zürich the median volumetric water content ranges from 10% to 78%. For all 7 wetness classes the median and the IQR increased. With volumetric water content values between 23- and 66% class 4 shows the largest spread among the classes. Classes 1, 2, 3 and 6 all show a relatively narrow spread.

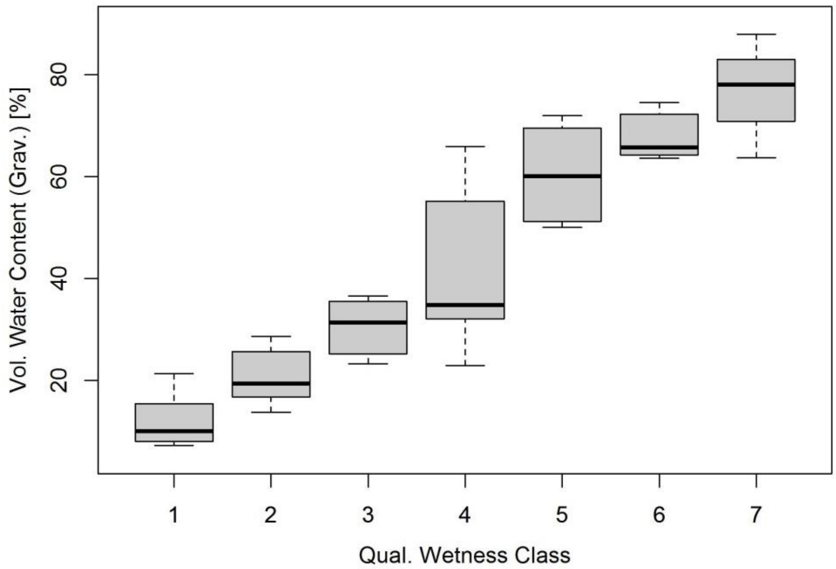


Figure 34: Volumetric water content for soil samples of each wetness class determined by the gravimetric method during the test in Zürich.

When using an adjusted level of risk to 0.002 the classes 1, 2; classes 1, 5, 6, 7 and classes 2, 3, 4, 5, 6, 7 are not significantly different from each other. A Kruskal-Wallis test with a level of risk of 0.05 indicated that the classes 3, 4; classes 4, 5 and classes 5, 6, 7 are not significantly different from each other.

5.3.4. Brunni, Alpthal

Comparisons between measured volumetric water content values and qualitative estimations of wetness classes for the field campaign in Brunni, Alpthal were only possible for 2 of the total 10 sampling points. As mentioned earlier in chapter 5.2, data from sampling points 1-8 had to be excluded as observed (qualitatively) surface variations of soil moisture could not be detected by the measurement devices. In Figure 35 the spread of volumetric water content measured with an FDR device for sampling point 9 and 10 were plotted against the associate qualitative wetness class (reference). It has to be mentioned that classifications used here were made by the test participants and not by a reference only. Hence, strong deviations of class assignments can cause an expansion of water content ranges. The median volumetric water contents range from 23% to 33% for point 9 and from 37% to 40% for point 10 over the whole measuring period. Soil moisture estimations for point 9 and 10 during the whole campaign never exceeded class 4, hence the spread of volumetric water content could only be shown for classes 1 to 4. There is an increase of median water content observable for both sampling points between classes 1 and 4. The IQRs of point 10 considerably overlap and the median volumetric water content were very similar. For point 9 this overlap is less strong and the median volumetric water contents are more distinct. The largest spread for both sampling points is found for class 2.

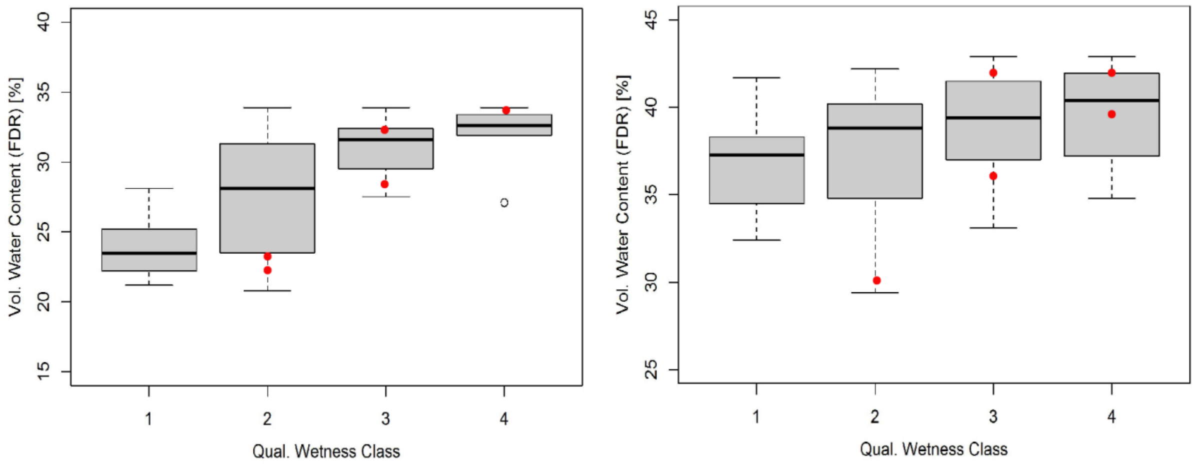


Figure 35: Measured volumetric water content (FDR) of each of the 4 occurring wetness classes during the test in Brunni, Alpthal. Sampling point 9 (left) and sampling point 10 (right). Red dots symbolize classifications of the reference person.

A Kruskal-Wallis test with an adjusted level of risk of 0.008 the volumetric water content for the different qualitative classes was only statistically significant for classes 1, 3 and classes 1, 4 (sampling point 9). Repeating the test with a level of risk of 0.05 indicated that classes 3, 4 of sampling point 9 were not significantly different from each other. For sampling point 10 the test indicated that all classes were not significantly different from each other.

In Figure 36 and 37 the time series of volumetric water content of sampling point 9 and 10 are presented together with the qualitative wetness classifications of the participants of each day. Classifications of participants are marked with circles and depending on the number of people assigning the same class at a certain day other colours and circle size were used. Darker colours as well as larger radius represent more people. FDR measurements are represented as time series. The tendency of assigning drier classes during periods of lower volumetric water content values and wetter classes during periods of higher volumetric water content values is visible for point 9 and 10. Even though there are class assignments that spread over 3 classes at certain days, it is recognizable that generally the participants' classifications correspond with the fluctuation of the volumetric water content.

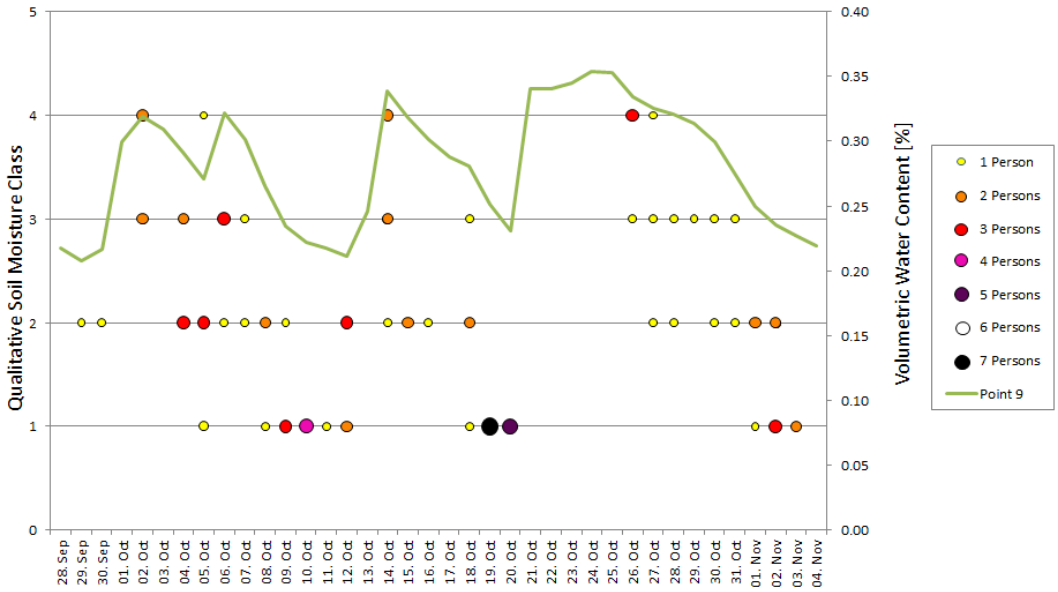


Figure 36: Time series of measured volumetric water content with the FDR device (green line) compared to raters' qualitative class assignments for sampling point 9 in Brunni, Alpthal. Larger radius and darker colours represent more people assigning the same class at a certain day.

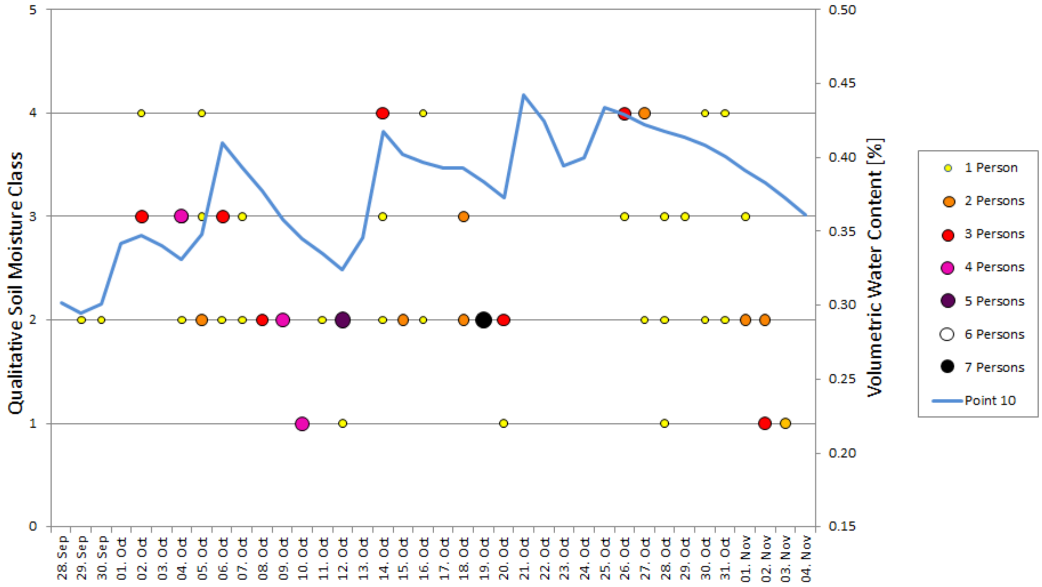


Figure 37: Time series of measured volumetric water content with the FDR device (blue line) compared to raters' qualitative class assignments for sampling point 10 in Brunni, Alpthal. Larger radius and darker colours represent more people assigning the same class at a certain day.

5.4. Soil characteristics

In this chapter the selected results of the soil profile analysis are presented. A detailed documentation about the site survey and following soil textural analysis is gathered in one report. Based on collected data and the World Reference Base for Soil Resources (2006), two reference soil groups and fitting prefixes and suffixes could be assigned. The WRB is the international standard taxonomic soil classification system that consists of 32 reference soil groups. Both soils were assigned as Cambisols, the first profile ("red soil") as Chromic Cambisol Colluvic Clayic and the second profile ("white soil") as Haplic Cambisol Siltic Ruptic. Results of the soil textural analysis of the two soil profiles of Tanzania are shown in table 4. With the help of Figure 38 and silt, clay and sand fractions obtained by the sedigraph method textural classes could be defined for the different soil horizons. In addition a dataset with the results of the soil textural analysis of 19 gravimetric samples of the field campaign in Mungushi taken at a depth of ca. 10 cm was compiled (Table 5). A selection of information about the soil profiles is summarized in table 10 in the appendix.

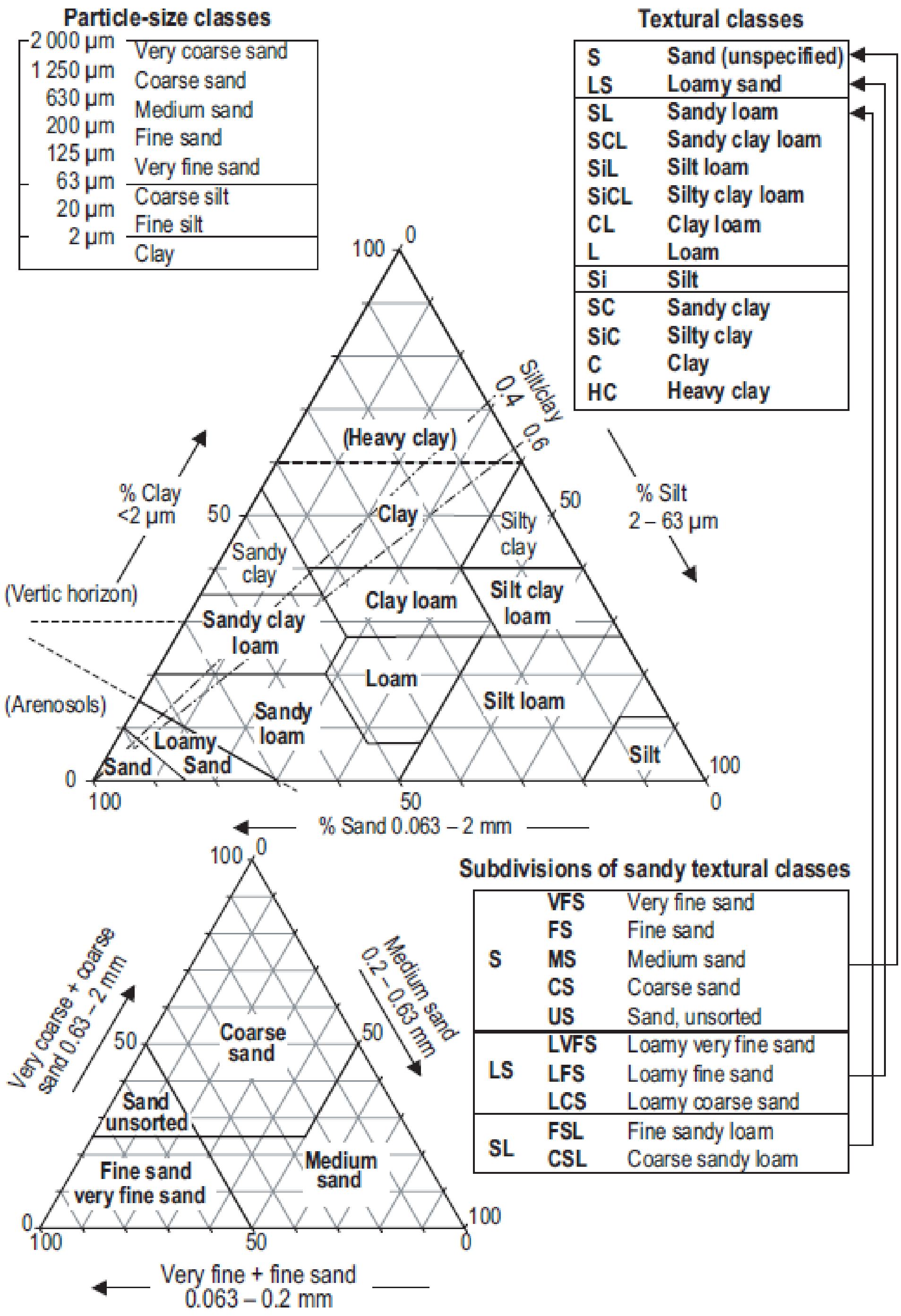


Figure 38: Relation of constituents of fine earth by size, defining textural classes and sand subclasses (FAO, 2006)

Table 4: Results of the soil textural analysis for soil horizons of the two soil profiles in Mungushi, Tanzania. Probes 22-27 refer to horizons 1-3 and two interlayers (25, 26) within the second horizon of the first soil profile. Probes 29-32 refer to horizons 1-4 of the second soil profile.

Probe	Weight of Clay [g]	Weight of Silt [g]	Weight of fine Sand (< 250 µ) [g]	Weight of coarse & medium Sand (<2mm > 250µm) [g]	Weight of Sand total (<2mm > 63µm) [g]	Total soil weight (<2mm) [g]	Clay [%]	Silt [%]	Sand [%]	Textural Class
22	29.49	17.24	0.57	2.50	3.07	49.80	59.21	34.63	6.17	C / HC
23	47.83	30.19	5.88	1.90	7.78	85.80	55.75	35.18	9.07	C
25	18.64	17.97	9.58	15.60	25.18	61.80	30.17	29.08	40.75	CL
26	4.98	4.69	2.13	31.20	33.33	43.00	11.58	10.91	77.51	SL / LS
27	11.78	10.14	1.17	100.70	101.87	123.80	9.52	8.19	82.29	LS
29	14.94	15.17	1.69	11.20	12.89	43.00	34.75	35.28	29.97	CL
30	2.73	24.26	9.21	1.40	10.61	37.60	7.26	64.53	28.21	SiL
31	2.89	32.21	0.90	0.00	0.90	36.00	8.02	89.48	2.51	Si
32	12.15	4.98	0.17	0.90	1.07	18.20	66.75	27.35	5.90	HC

Table 5: Results of the soil textural analysis for selected gravimetric soil samples. Probes 1-18 refer to corresponding sampling points of the parcours in Mungushi, Tanzania (within the "red soil"). Probes 19-21 refer to sampling points 38, 39 and 42 (within the "white soil")

Probe	Weight of Clay [g]	Weight of Silt [g]	Weight of fine Sand (< 250 µ) [g]	Weight of coarse & medium Sand (<2mm > 250µm) [g]	Weight of Sand total (<2mm > 63µm) [g]	Total soil weight (<2mm) [g]	Clay [%]	Silt [%]	Sand [%]	Textural Class
1	7.58	9.46	1.77	30.60	32.37	49.40	15.34	19.14	65.52	SL
2	19.54	11.19	0.47	9.30	9.77	40.50	48.25	27.62	24.13	C
3	23.81	11.37	0.92	73.50	74.42	109.60	21.72	10.37	67.90	SCL
4	7.05	4.87	0.68	13.70	14.38	26.30	26.82	18.52	54.66	SCL
6	13.15	8.20	1.55	14.50	16.05	37.40	35.17	21.92	42.91	CL
8	9.96	7.50	2.13	30.90	33.03	50.50	19.73	14.86	65.41	SL
9	20.08	12.64	2.08	5.00	7.08	39.80	50.45	31.77	17.78	C
10	15.79	11.11	1.99	7.10	9.09	36.00	43.87	30.87	25.26	C
11	7.90	6.30	2.21	10.20	12.41	26.60	29.69	23.67	46.64	SCL
12	5.12	5.46	1.72	37.20	38.92	49.50	10.34	11.03	78.63	LS / SL
13	9.56	9.41	2.82	28.10	30.92	49.90	19.16	18.87	61.97	SL
14	18.53	13.89	5.48	7.10	12.58	45.00	41.18	30.86	27.96	C
15	14.72	10.90	2.88	8.30	11.18	36.80	40.00	29.63	30.37	CL
16	16.29	10.43	2.27	1.60	3.87	30.60	53.25	34.10	12.65	C
17	8.97	6.96	1.28	11.50	12.78	28.70	31.24	24.25	44.51	CL
18	10.35	7.79	1.76	12.20	13.96	32.10	32.25	24.27	43.48	CL
19	1.53	21.04	8.03	3.40	11.43	34.00	4.49	61.89	33.62	SiL
20	12.30	5.02	0.29	19.80	20.09	37.40	32.88	13.41	53.70	SCL
21	11.28	24.68	5.24	7.80	13.04	49.00	23.03	50.36	26.61	L /SiL

6. Discussion

The presented subchapters of the discussion refer to the research questions raised at the beginning. First the performance of the qualitative soil wetness classification scheme for arid environmental conditions is discussed. Further it is evaluated if the qualitatively defined wetness classes represent quantitative differences in soil moisture content. At the end of the discussion the influence of different levels of introduction or training on the raters degree of agreement is assessed and the applicability of the qualitative method on a voluntary basis (crowd sourced) is discussed.

6.1. Qualitative scheme for semiarid regions

The aim was to evaluate if the qualitative soil wetness classification scheme is capable of reliably capturing differences in shallow soil moisture content in a semiarid environment. In both tests where this scheme was used (both study areas in Tanzania) a high agreement in wetness class assignments was found among master (M*all*) and PhD+ students during the test in April and farmers (F*all*) and experts (E*all*) during the test in June. The best results were achieved by F*all* and E*all* in June. The group of farmers (F*all*) performing the test in April was the only group showing a lower inter-rater reliability. Still, they agreed or were off by only one wetness class in 81% of all cases. These generally good results are encouraging and suggest the robustness of the qualitative method. This lower agreement of farmers in the test performed in April can be explained to a certain extent by the introduction that was carried out in too large groups. In this way it was not possible to ensure, that all participants had understood the method. Further, the structure of the form sheet caused some trouble. The wetness class for each sampling point had to be assigned by placing a tick mark into a corresponding box. As the matrix was slightly confusing, slipping out of position was not impossible. Even though it was only allowed to choose one class per site some participants placed more than one ticking mark. With some modifications of the form sheet, the site labelling, and the way of introduction for the second test in June, the inter-rater reliability of farmers could be improved considerably. During the second test introductions were given in small groups of 5 and instead of placing a tick mark in a pre-printed matrix-type of assessment form, the estimated soil moisture class had to be written down for each sampling point. Both the site number and the word “kituo” (English: “station”) were written on every flag. In both tests dry to intermediate wetness classes (2, 3 and 4) seemed to have been most difficult to assign and showed the largest spread (Fig. 16, 20). By altering group sizes and form sheet design the bias of individuals was also reduced in the second test.

Especially distinct misclassification over several classes could nearly be eliminated in June as slipping out of position while placing a tick mark was not possible anymore in this test.

When comparing these to compiled datasets from Tanzania to the tests with geography students in Switzerland (Rinderer et al., 2012 and Zürich, Irchel in 2014) it can be detected that agreements were lower (April) or similar (June). While students from Switzerland agreed in 70% respectively 73% of all cases and agreed or were off by only one wetness class in 95% respectively 99% of all cases, test participants from Tanzania agreed in approximately 47% (April) respectively 63% (June) of all cases. When considering deviations of up to +/- one wetness class, test participants from June reached nearly the same value as students in 2012 (Rinderer et al. 2012) and a little less than students in 2014 (Zürich, Irchel). Agreements of participants of the first test in April are lower than the one of students from Switzerland but higher than the one of participants of the test in Brunni. When considering distinct misclassifications, again results from June and Switzerland (2012, 2014) are very similar and for both less than 2 % were off by more than 2 classes and deviations of more than 3 classes did not occur at all. Distinct misclassification of more than 3 classes occurred in about 4.5% of all cases for the test in April. In general, intermediate classes seem to have been most difficult to assign (Tanzania in June and April; Switzerland 2012, October 2014 and Sep.-Nov 2014)

6.2. Quantitative differences in soil moisture content

In order to prove or disprove if qualitatively defined wetness classes are representative of quantitative differences in soil moisture content, measurements for both classification schemes (humid, semiarid) were considered (all four tests). Even though this was already investigated for the classification scheme for humid environmental conditions by Rinderer et al. (2012) it is going to be discussed in this thesis as well as the agreement between qualitative and quantitative soil wetness is strongly dependent on the soil type. It is to be expected that conditions are diverse as the tests were performed at two different locations (Test by Rinderer et al. in 2012 and Zürich, Irchel) with different raters. In all four tests qualitative wetness classes reflected actual differences in volumetric water content. But the median volumetric water content of the 3 wettest classes were very similar for both tests in Tanzania and in case of the second test (June) also classes 3 and 4 had nearly the same median volumetric water content (see Fig. 32 and 33). The median volumetric water content of the two driest classes were very similar for the test in June and nearly equal for the test in April. Rinderer et al. (2012) also observed considerable overlap in volumetric water content of individual qualitative wetness classes and discussed the number of classes. As median and spread of volumetric water content of certain classes are very similar, the merging of them could be contemplated. On the other hand a reduction of classes would have consequences on classification results: The resolution of soil moisture patterns would be smaller and hence involve a loss of spatial information; Even though a reduced number of classes would likely be easier for raters to assign it would limit the identification of wetness variability over time. (Rinderer et al., 2012)

As for the test in Brunni, Alpthal, only 2 of the 10 sampling points could be considered (cf. chapter 5.1) only the driest to intermediate classes (1-4) are included here. Although median volumetric water content increased with ascending wetness class, differences are rather small for all classes for point 10 and classes 3 and 4 for point 9 and the IQR overlapped. The largest spread of volumetric water content was found for class 2 for both sampling points. In general the range of volumetric water content covered by these 4 classes was relatively narrow when compared to the other tests in Tanzania, Zürich and Alpthal (Rinderer et al., 2012). One problem could be that the test was running from the end of September until the beginning of November and changes in water content about 10cm below soil surface were not as strong as they would be with higher temperatures and more intense solar irradiance. Distinguishable moisture changes on the soil surface could possibly remain undetected by measurement devices buried below the surface. The author therefore suggests to further test the qualitative scheme throughout the season to detect possible seasonal changes.

For the test campaign in Zürich with the classification scheme for humid environmental conditions median values increased with ascending class over all wetness classes. It is discernible (Fig. 34) that median volumetric water contents were more distinct than in the African dataset exception of classes 3 and 4. The largest spread of volumetric water content was found for class 4 and to a smaller extent for classes 5 and 7. When comparing to results of the other tests in Tanzania, classes of Zürich apparently reflect actual differences of volumetric water content the better. Similar to the range found by Rinderer et al. in 2012 (approximately 23-95% volumetric water content) a large range of volumetric water content from about 7- 88% was found for Zürich. Volumetric water content ranges obtained from the two field campaigns in Tanzania are much narrower. It needs to be mentioned that the classification scheme developed by Rinderer et al. (2012) was adapted especially for humid environmental conditions and was tested under such conditions. Therefore the median volumetric water content of a dry class of this scheme is more similar to a wetter class of the scheme developed for arid environmental conditions. As resembling qualitative indicators on the soil surface can be associated with different volumetric water contents, a calibration to the local soil types is necessary if absolute water contents are of interest (Rinderer et al., 2015). Another issue to keep in mind is that varying soil properties are expected to influence the relation between qualitative wetness classes and associated volumetric water content (Rinderer et al., 2012). This was noticeable for wet classes in Tanzania for soils with high clay contents and even though moisture condition of the soil surface changed, there was only a limited change in volumetric water content measured in shallow depths below the surface. It can be assumed that the clay content effected a damming of the water on the soil surface while the condition below remained alike.

6.3. Training effects and crowd sourcing

To evaluate if interrater reliability of the qualitative wetness classification can be improved by training and if interrater reliability changes due to different forms of introduction all 4 dataset were considered. For the Tanzanian master students of the first test in April and for the Tanzanian farmers in the second test in June there was an improvement observable for the "*inf*" group (M*inf* and F*inf*; Fig. 15 and Fig. 19). Classification differences of more than two classes could be prevented for the M*inf* students and in approximately 77% of all cases, raters assigned the same wetness class. Master students with only a basic introduction during the test in Tanzania in April here named "M*not*" however only agreed in about 58% of all cases and deviations of up to 4 classes were found. For the framers in June the differences between the two groups were less distinct but still recognizable.

For the test with students in Zurich no difference could be determined between Inf*all* and Not*all* (Fig. 23). In over 99% of all cases Inf*all* and Not*all* independently assigned the same wetness class. As the interrater-agreemeent of the "Not*all*" students is already very good, especially compared to the two tests in Tanzania, the need for a more detailed introduction seems not necessary. While in the case of Tanzanian framers communication difficulties could possibly have caused additional problems and consequently affected the results, students from Zürich could be introduced in their mother tongue by the author or an expert himself. Farmers of the first test in April on the other hand showed no improvement due to a more detailed introduction. The F*inf* as well as the F*not* showed large spreead of class assignments relative to the median of their group. As mentioned earlier in chapter 6.1 some enhancements of the form sheet and way of introduction had to be made for the second test in June. It is possible that errors due to difficulties with the form sheet are not evenly distributed over all groups of participants. Even though more detailed introductions were made, too large group sizes and possibly incomplete information as well as participants joining later could negatively influence the results. When comparing Tanzanian farmers of the first test with Tanzanian farmers of the second test

(Table 6 in the appendix) it became evident that a more detailed introduction and training can help to improve the interrate reliability of the qualitative wetness classification. Farmers (F*all*) in June agreed better than Farmers (F*all*) in April. In addition to this improvement by 20%, strong deviations are also less frequent. Regarding the test in Zürich and the test by Rinderer et al. (2012) Swiss students show very good results even without a more detailed introduction or training. The question arises whether the results for Swiss farmers would be similarly to the experience with farmers in Tanzania.

For the long-term field campaign in Brunni, Alpthal the information was kept on a minimum and in a written form. In general, a low agreement in wetness class assignments was found among the raters (Figure 26). In approximately 31% of all cases, raters independently assigned the same wetness class as the reference. The peak of agreement is shifted towards negative classification differences and it can be recognized that raters generally classified too dry. Compared to the results of the 3 other tests performed in the context of this theses and the results obtained by Rinderer et al. (2012) this value is low. Deviations of more than 3 classes did not occur but it should be mentioned here that 6 raters were excluded in advance as their classification results indicated that the concept of the method was not understood. Qualitative classifications by the expert are only available on 5 days of the whole test time. For days without reference of the expert a soil moisture class had to be assigned with the help of FDR measurements, meteorological data and expert knowledge to every sampling point. These reference assignments are based on only one person and even though it was an expert they can be biased. In case of a repetition of such an experiment it is suggested to increase the number of days with reference estimations or even generate a continuous reference dataset in order to reduce errors. The test was running from mid to late autumn and surface soil moisture variability was not measureable with the FDR devices buried about 10cm below the soil surface with the exception of two sampling points. A shift of the test time into summer would very likely improve the usefulness of these continuous measurement series.

Over a period of 37 days a total of 102 people voluntarily conducted the test in Brunni, Alpthal. During this time, there was good weather to a great extent and autumn holidays for a lot of cantons in Switzerland. On 20 out of 37 days there was no precipitation at all and on 5 days the amount of precipitation was below 1mm. On a good day during holidays or a weekend day with good weather an estimated number of >100-200 people passed the test area. The highest amount of participants per day was 7 people. Compared to the probably large number of hikers passing by, this is rather unsatisfying. On the one hand, this may be because people were not interested in taking part in such a test or even reading the information, on the other hand it may be that the size of the information board was not appropriate. Even though the board was very noticeable and positioned directly at the main hiking path it was necessary to approach it to read the information about the project. A larger board with a font size legible from further away could possibly increase the number of participants. Form sheets as well as writing utensils were always available in both material boxes and could not have influenced the amount of filled forms. Although the number of participants was not remarkably, the results of this first attempts are still an acceptable beginning. Adjustments to the test layout and facilities could possibly have a positive influence and improve test results. If findings of future investigations are satisfying a testing of a mobile application could be envisaged.

6.4. Limitations of the qualitative scheme

It is necessary to note that limitations of the qualitative classification scheme exist. As the method relies on wetness indicators derived from soil surface properties results might be influenced by certain factors. Vegetation as for example moss covering the soils surface may hold a considerable amount of water and thus be wetter compared to the soil surface. Conversely, litter layers covering the surface can be relatively dry compared to the soil below. If such overlays are not removed for testing the spot in question it possibly be over- or underestimated by raters. Other sources of error of misjudgement may be drizzle, dew or evapotranspiration that can alter the moisture situation of the soil surface. Another limitation is that, although soil moisture at the surface is expected to be related to that at depth for most soil types, only the soil surface properties are assessed (Rinderer et al., 2015). Especially with regard to the use of the scheme developed for semiarid environmental conditions this is of importance. The scheme should serve farmers (or generally the population) in countries with arid or semiarid conditions as a decision making tool for a fair allocation of the diminishing water resources for the irrigation of cropland. For many crops though it is significant what moisture condition is present at the root depth rather than at the soil surface. In regions with a sufficient supply of water as for example Switzerland the knowledge of soil moisture is however preferably of use in science. (Rinderer et al., 2012)

7. Conclusion

Findings of this thesis show the potential of soil wetness classification schemes based on qualitative indicators that are capable of capturing shallow soil moisture differences in semi-arid and humide temperate environments. Participants of the different tests showed a good consensus among individual raters. A good introduction and a readily understandable layout of the form sheet were however important requirements in the Tanzanian study. Higher agreements clearly indicate that qualitative indicators of the wettest and driest wetness classes seem to be easier to assign than intermediate classes. This study showed that qualitatively defined wetness classes represent quantitative differences in soil moisture content for both soil wetness classification schemes (scheme for humid environmental conditions and scheme for semiarid environmental conditions). Even though the volumetric water content of individual wetness classes was overlapping in the majority of cases the median volumetric water content increased with ascending wetness class. A possible reduction of class number could counteract overlapping but would involve a loss of spatial resolution and would limit the identification of wetness variability over time. On the basis of these tests it became evident that a more detailed introduction as well as training do enable an improvement of interrater reliability of the qualitative wetness classification. Farmers and master students in Tanzania achieved significantly better results when receiving a longer introduction with some examples or after doing a short training. Notably, strong deviations could be reduced or even eliminated. For Swiss students the group receiving a more detailed introduction did not show better inter-rater reliability. It would be interesting to investigate if Swiss farmers would perform similarly well. In comparison with the tests in Tanzania and in Zurich lower agreement in wetness class assignments was found among the raters in the long-time field campaign in Brunni, Alpthal. Raters tend to classify soils too dry when compared to the reference. Also total number of participants appears rather small in comparison with the number of people passing this hiking path every day. It needs to be assessed if adjustments of the test layout and facilities could possibly have positive influences and improve test results. Despite these findings results of this first attempts are still an acceptable beginning and further investigations are required.

In general it can be said that the classification schemes presented here provide robust and reliable results. The method is quick and easily applicable for non-experts and no additional measurement device is required. These circumstances make this method especially useful in remote areas without a sufficient supply of energy or areas with limited financial resources to buy expensive quantitative measuring devices (TDR, FDR). The project “iMoMo – Innovative Monitoring and Modeling of Water” is currently working on the idea of collecting crowd-based environmental data via SMS or

USSD. Soil moisture information could be obtained from local farmers classifying soil wetness with the scheme presented here and the collected data could be applied to model calibration and data assimilation. In return, information about soil water stress and suggestion on how to use and allocate the increasingly diminishing water resources could be provided to local farmers. In so, the simple method tested in this thesis could improve living conditions of the local society. An optimised distribution of water could enhance food production and minimise dispute between different users.

8. References

Ambroise, B., Freer, J., and Beven, K. (1996) Application of a generalized TOPMODEL to the small Ringelbach catchment, Vorges, France. Water Resources Research, 32, 2147-2159

Amt für Raumentwicklung Kanton Zürich, Abteilung Geoinformation (2015) Gewässer-Ökomorphologie. http://maps.zh.ch/ (Accessed 16 March 2015)

ArcGIS – World Imagery (2015) Sources: Esri, DigitalGlobe, Earthstar Geographics, CNES/Airbus DS, GeoEye, USDA FSA, USGS, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

Blazkova, S., Beven, K. J. and Kulasova, A. (2002) On constraining TOPMODEL hydrograph simulations using partial saturated area information. Hydrological Processes, 16, 441-458

Blume, H.P., Stahr, K. and Leinweber, P. (2011) Bodenkundliches Praktikum – Eine Einführung in pedologisches Arbeiten für Ökologen, insbesondere Land- und Forstwire, und für Geowissenschaflter. 3. Aufl., Spektrum, Heidelberg

Burch, F., Forster, F. and Schleppi, P. (1996) Zum Einfluss des Waldes auf die Hydrologie der Flysch-Einzugsgebiete des Alptals. Schweizerische Zeitschrift für Forstwesen, 147 (12), 925-938

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Bastiaensen, J., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D.M., Hergarten, C., Isaeva, A., Pandey, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G. and Zhumanova, M. (2014) Citizen science in hydrology and water resources : opportunities for knowledge generation , ecosystem service management , and sustainable development. Frontiers in Earth Science, 2 (26), 1-21

Cohen, J. (1960) A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, Vol 20 (1), 37-46

Cordner, G. W. and Foreman, D. I. (2009) Nonparametric Statistics for Non-statisticians: A Step-by-Step Approach. John Wiley & Sons, Inc., Hoboken, New Jersey

Decagon Devices, Inc. (2014a) Em50 / Em50R / Em50G - Em50 Series Data Collection System, manual, Pullman WA

Decagon Devices, Inc. (2014b) GS3 - Water Content, EC and Temperature Sensors, manual, Pullman WA

Delta-T Devices Ltd (n.y.) ML3 ThetaProbe Soil Moisture Sensor, Cambridge UK

Delta-T Devices Ltd (2013a) ML3 ThetaProbe, manual, Cambridge UK

Delta-T Devices Ltd (2013b) Moisture Meter type HH2, manual, Cambridge UK

Dobriyal, P.,Qureshi, A., Badola, R. and Hussain, S.A. (2012) A review of the methods available for estimating soil moisture and its implications for water resource management

Dunne, T. and Black, R. D. (1970) Partial area contributions to storm runoff in a small New England Watershed. Water Resources Research, 6, 1296–1311

Dunne, T., Moore, T. R., and Tayloer, C. H. (1975) Recognition and prediction of runoff producing zones in humid regions. Hydrological Sciences Bulletin, 20, 305–327

Dyck, S., Peschke, G. (1995) Grundlagen der Hydrologie.Verlag für Bauwesen GmbH, Berlin

Egli, M., Brandová, D., Schreiner, C., Woodhatch, I. and Pichler, B. (2013) Geochronology laboratory methods. Department of Geography, Zürich

Famiglietti, J. S., Rudnicki, J. W. and Rodell, M. (1998) Variability of surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. Journal of Hydrology, 210, 259-281

FAO (2006) Guidelines for soil description, Fourth edition, Rome

Ferré, P.A. and Topp, G.C. (2000) Time-domain Reflectometry Techniques for Soil Water Content and Electrical Conductivity Measurements. Sensors Updae, 7, 277-300

Gee, G. W. and Bauder, J. W. (1986) Particle-size Analysis, in Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods – Agronomy Monograph no. 9 (2nd Edition), American Society of Agronomy, Inc. Soil Science Society of America, Inc., Madison

Google Inc. (2013): Google Earth. Version 6.1.7601.1. Available online at http://www.google.ch/intl/de/earth/index.html (Accessed 12 April 2015)

IMKO Micromodultechnik GmbH (2009) PICO-BT mit Software PICO-TALK, manual, Ettlingen

IMKO Micromodultechnik GmbH (2013) TRIME-PICO 64/32, manual, Ettlingen

iMoMo (Innovative Technologies for Monitoring, Modeling and Managing Water) www.imomohub.ch (Accessed 14 February 2015)

Inamdar, S. P. and Mitchell, M. J. (2007) Contributions of riparian and hillslope waters to storm runoff across multiple catchments and storm events in a glaciated forested watershed. Journal of Hydrology, 341, 116–130

Jones, S. B., Wraith, J. M. and Or, D. (2002) Time domain reflectometry measurement principles and applications. Hydrological Processes, 16, 141-153

Karamouz, M., Nazif, S. and Falahi, M. (2012) Hydrology and Hydroclimatology: Principles and Applications. CRC Press

Komakech, H. C. and Van der Zaag, P. (2011) Understanding the Emergence and Functioning of River Committees in a Catchment of the Pangani Basin, Tanzania. Water Alternatives, 4(2): 197-222

Krippendorff, K. (2011) Computing Krippendorff's Alpha-Reliabiliy. Retrieved from http://repository.upenn.edu/asc_papers/43

Kruskal, W. H. and Wallis, A. (1952) Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association, 47 (260), 583-621

Kulasova, A., Beven, K. J., Blazkova, S. D., Rezacova, D., and Cajthaml, J. (2014) Comparison of saturated areas mapping methods in the Jizera Mountains, Czech Republic. Journal of Hydrology and Hydromechanics, 62, 160–168

Lowry, C.S. and Fienen, M.N. (2013) CrowdHydrology: Crowdsourcing Hydrologic Data and Engaging Citizen Scientists. Ground Water, 51 (1), 151-156

McSweeney, C., New, M. and Lizcano, G. (2010) UNDP Climate Change Country Profiles: Tanzania. Available: http://country-profiles.geog.ox.ac.uk/ (Accessed 6 March 2015)

MeteoSchweiz, Bundesamt für Meteorologie und Klimatologie, IDAweb https://gate.meteoswiss.ch/idaweb/ (Accessed 31 March 2015)

Natural Resources Conservation Service. (1998) Estimating Soil Moisture by Feel and Appearance. Program Aid 1619, 1-7

Patel, N. R., Anapashsha, R., Kumar, S., Saha, S. K. and Dadhwal, V. K. (2009) Assessing potential of MODIS derived temperature/vegetation condition index (TVDI) to infer soil moisture status. International Journal of Remote Sensing, 30 (1), 23-39

Penna, D., Tromp-van Meerveld, H., Gobbi, A., Borga, M. and Dalla Fontana, G. (2011) The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. Hydrology and Earth System Sciences, 15, 689-702

Pepin, S., Plamondon, A.P. and Stein, J. (1992) Peat water content measurement using time domain reflectometry. Canadian Journal of Forest Research, 22, 534-540

Petropoulos, G. P., Griffiths, H. M., Dorigo, W., Xaver, A. and Gruber, A. (2014) Surface Soil Moisture Estimation: Significance, Controls, and Conventional Measurement Techniques. In Petropoulos, etc etc.

Rajkai, K. and Rydén, B.E. (1992) Measuring areal soil moisture distribution with the TDR method. Geoderma, 52, 73-85

Reynolds, S. G. (1970) The gravimetric method of soil moisture determination, I: A study of equipment, and methodological problems. Journal of Hydrology, 11, 258-273

Rinderer, M., Kollegger, A., Fischer, B. M. C., Stähli, M. and Seibert, J. (2012) Sensing with boots and trousers – qualitative field observations of shallow soil moisture patterns.

Rinderer, M., Komakech, H., Müller, D. and Seibert, J. (2015) Qualitative soil moisture assessment n semi-arid Africa: the role of experience and training on inter-rater reliability. Hydrology and Earth System Sciences Discussions, 12, 3029-3058

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S. and Namkhai, A. (2000) The Global Soil Moisture Data Bank. Bulletin of the American Meteorological Society, 81 (6), 1281-1299

Rothe, A., Weis, W., Kreutzer, K., Matthies, D., Hess, U. and Ansorge, B. (1997) Changes in soil structure caused by the installation of time domain reflectometry probes and their influence on the measurement of soil moisture. Water Resources Research, 33 (7), 1585-1593

Schlichting, E., Blume, H.P. and Stahr, K. (1995) Bodenkundliches Praktikum : eine Einführung in pedologisches Arbeiten für Oekologen, insbesondere Land- und Forstwirte, und für Geowissenschlaftler. Blackwell Wissenschafts-Verlag, Berlin

Simons, D. B. and Sentürk, F. (1992) Sediment Transport Technology: Water and Sediment Dynamics. Water Resources Publications, Littleton

SNIFFER (2009) A Functional Wetland Typology for Scotland – Field Survey Manual, 1st edn., Scotland and Northern Ireland Forum for Environmental Research (SNIFFER), Edinburgh

Stadt Zürich, Tiefbau- und Entsorgungsdepartement: Irchelpark https://www.stadt-zuerich.ch/ted/de/index/gsz/natur-_und_erlebnisraeume/park-_und_gruenanlagen/irchelpark.html (Accessed 10 March 2015)

Stahr, K., Herrmann, L. and Jahn, R. (2000) Praktikumsanleitung zur Beschreibung, Deutung und ökologischen Bewertung von Böden. Institut für Bodenkunde und Standortslehre, Universität Hohenheim, Institut für Bodenkunde und Pflanzenernährung, Universität Halle.

Strahler, A.H. & Strahler, A.N. (1999) Physische Geographie. Ulmer, Stuttgart (Hohenheim)

Swisstopo (Bundesamt für Landestopografie) (2007) map.swisstopo.admin.ch (Accessed 15 April 2015)

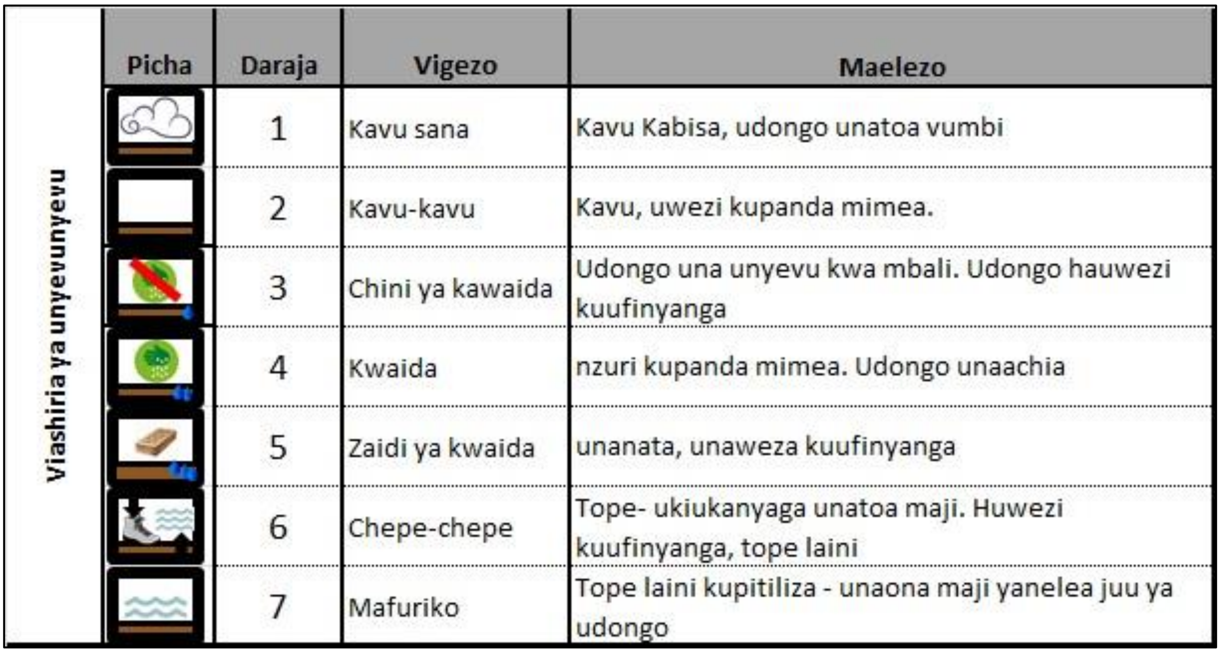
Syvitski, P. M. (ed.) (1991) Principles, Methods and Application of Particle Size Analysis. Cambridge University Press, Cambridge

Turner, D.S. and Richter, H.E. (2011) Wet/Dry Mapping: Using Citizen Scientists to Monitor the Extent of Perennial Surface Flow in Dryland Regions. Environmental Management, 47, 497-505

Western, A. W., Zhou, S., Grayson, R. B., McMahon, T. A., Blöschl, G. and Wilson, D. J. (2004) Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. Journal of Hydrology, 286, 113-134

Wikipedia: Seite “Tansania”. In Wikipedia, Die freie Enzyklopädie. Bearbeitungsstand: 3. April 2015, 21:13 UTC. URL: http://de.wikipedia.org/w/index.php?title=Tansania&oldid=140544396 (Accessed 12 April 2015)

9. Appendix



Viashiria ya unyevunyevu	Picha	Daraja	Vigezo	Maelezo
		1	Kavu sana	Kavu Kabisa, udongo unatoa vumbi
		2	Kavu-kavu	Kavu, uwezi kupanda mimea.
		3	Chini ya kawaida	Udongo una unyevu kwa mbali. Udongo hauwezi kuufinyanga
		4	Kwaida	nzuri kupanda mimea. Udongo unaachia
		5	Zaidi ya kwaida	unanata, unaweza kuufinyanga
		6	Chepe-chepe	Tope- ukiukanyaga unatoa maji. Huwezi kuufinyanga, tope laini
		7	Mafuriko	Tope laini kupitiliza - unaona maji yanelea juu ya udongo

Figure 39: Soil wetness classification scheme for dry environmental conditions in Swahili.

Formular zur Bodenfeuchte Einschätzung

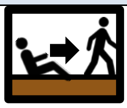
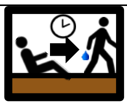
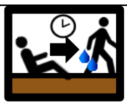
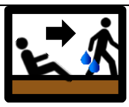
Name:

Alter:

Beruf / Ausbildung:

Datum & Uhrzeit:*

(Die mit einem * markierten Angaben müssen ausgefüllt sein. Die Namen der Testpersonen werden nicht veröffentlicht.)

Klasse		Qualitative Indikator Kriterien
1		Die Hosen einer auf dem Boden sitzenden Person bleiben trocken
2		Die Hosen einer auf dem Boden sitzenden Person würden nach einigen Minuten feucht werden
3		Die Hosen einer auf dem Boden sitzenden Person würden nach einigen Minuten nass werden
4		Die Hosen einer auf dem Boden sitzenden Person würden sofort nass
5		Ein schmatzendes Geräusch beim auf den Boden treten, Wasser tritt jedoch keines aus
6		Wasser quillt aus dem Boden beim Auftreten mit einem Schuh
7		Wasser steht auf der Bodenoberfläche

- Gehen Sie nun mit dem Formular los. Die markierten Stellen (gelb-orange Holzpfosten mit Nummern von 1-10) befinden sich entweder direkt neben dem Wanderweg oder in Sichtweite davon. Die Reihenfolge ist dabei unwichtig, nur die Standorte auf dem Formular müssen am richtigen Ort eingetragen sein.

- Für die Einschätzung der Bodenfeuchte soll der Boden gleich beim Holzpfosten betrachtet werden (es sollte wenn möglich keine Vegetation oder Blätterauflage vorhanden sein).
- Sie brauchen NICHT auf den Boden zu sitzen. Testen Sie ruhig mit der Hand oder Ihrem Schuh. Achten Sie darauf, dass dabei kein Bodenmaterial entfernt wird!
- Das ausgefüllte Formular kann bei einer der beiden Informationstafeln in die Kiste gelegt werden.

Standort	Bodenfeuchteklasse (1-7)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Vielen Dank für Ihre Hilfe!

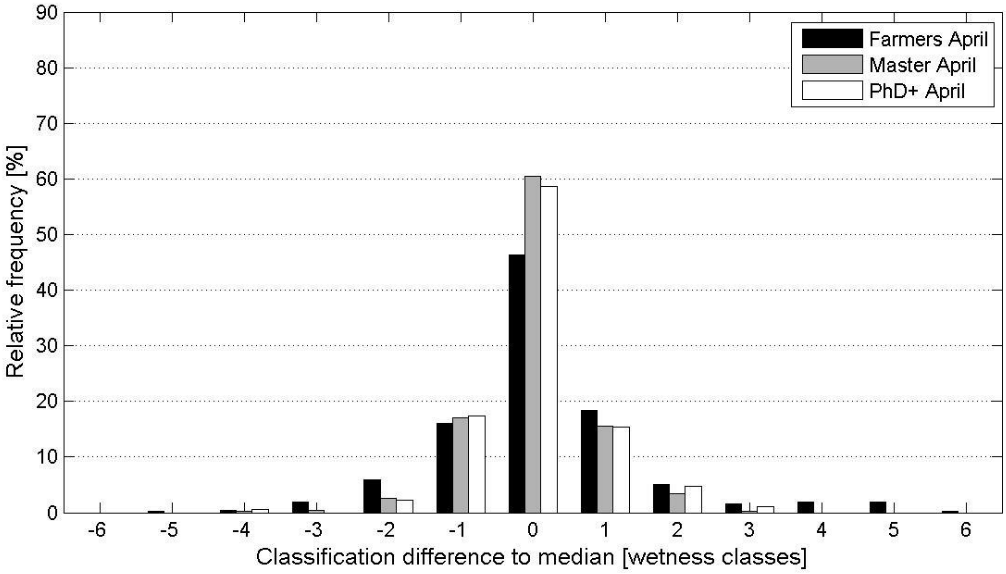


Figure 40: Deviation of wetness classifications relative to the median of each sampling point for farmers, master students and PhD+ for the test in April plotted as relative frequency distribution.

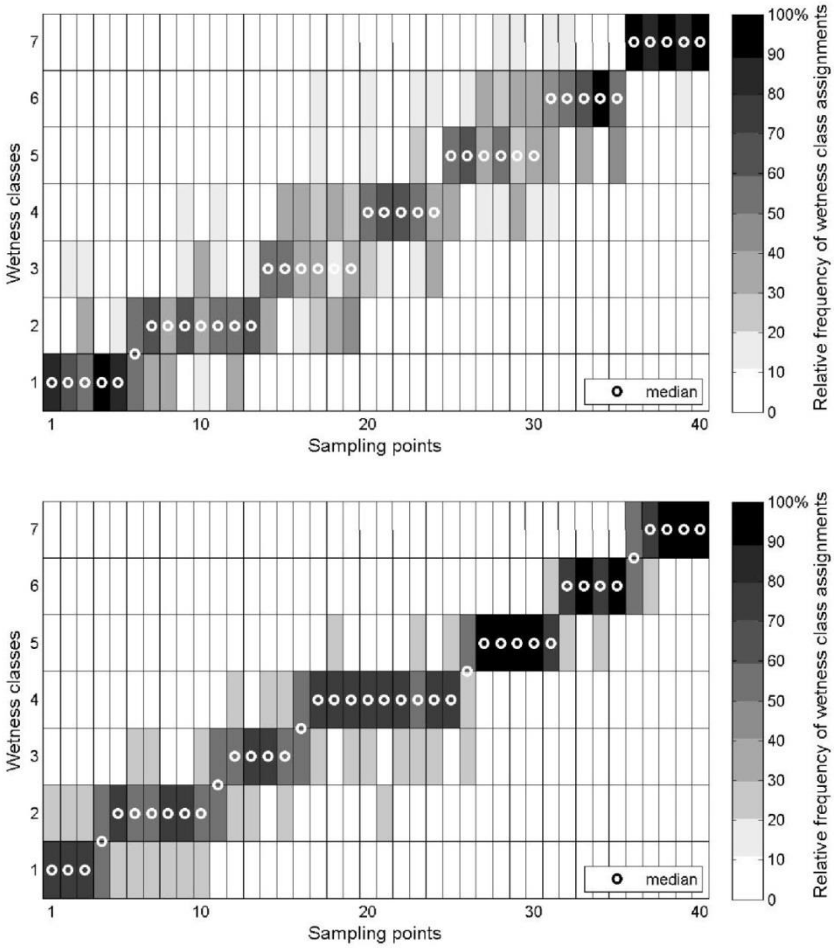


Figure 41: Wetness class assignments for each of the 40 sampling points numbered and sorted by the median of class assignments for each sampling point for not informed (top) and informed (bottom) master students for the test in April. White circles show the median as reference and grey shades indicate relative frequency of wetness class assignments.

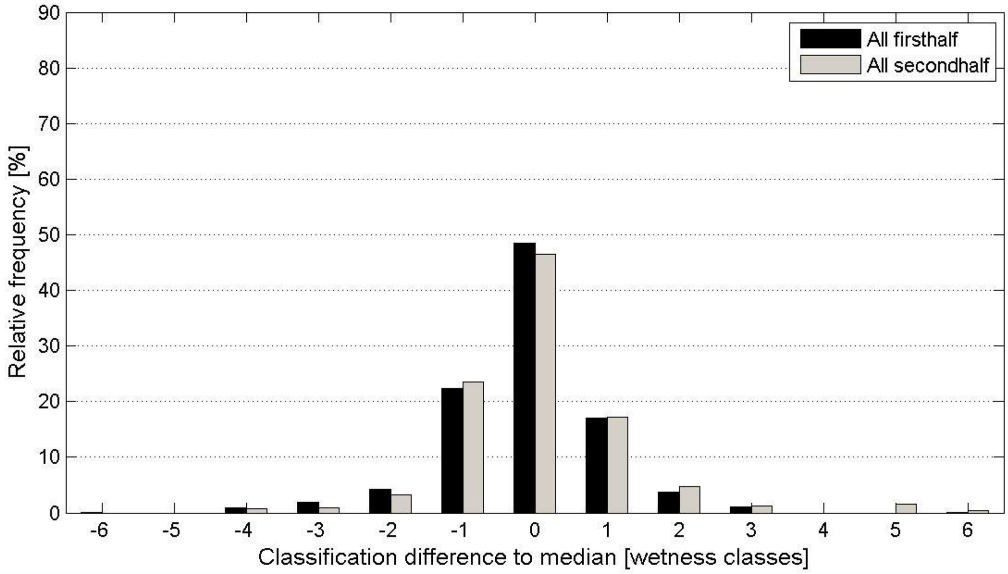


Figure 42: Deviation of wetness classifications relative to the median of each sampling point for Fi*all* and Se*all* for the test in April plotted as relative frequency distribution.

Table 6: Deviation of wetness classifications relative to the median of each sampling point for all raters of the group. For Alpthal all the deviation of wetness classification is relative to the reference of each sampling point.

	Classification difference to median [wetness classes]												
Group	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
Tanzania April all	0.07	0	0.77	1.46	3.69	22.91	47.49	17.3	4.25	1.18	0	0.77	0.28
Tanzania April Inf*all*	0	0	0	2.34	5.56	18.13	50	18.13	1.17	0.88	0.58	3.22	0
Tanzania April Not*all*	0	0	0.7	0.53	2.8	24.69	49.56	14.01	5.08	1.75	0	0.35	0.35
Tanzania April F*all*	0	0.22	0.44	1.97	5.92	16.01	46.27	18.42	5.04	1.54	1.97	1.97	0.22
Tanzania April M*all*	0	0	0.22	0.43	2.59	17.06	60.48	15.55	3.46	0.22	0	0	0
Tanzania April PhD+	0	0	0.5	0	2.27	17.38	58.69	15.37	4.79	1.01	0	0	0
Tanzania April Fi*all*	0.14	0	0.84	1.96	4.21	22.3	48.53	16.97	3.79	1.12	0	0	0.14
Tanzania April Se*all*	0	0	0.69	0.97	3.18	23.51	46.47	17.29	4.7	1.24	0	1.52	0.41
Tanzania June all	0	0	0	0.76	2.67	12.48	62.57	19.33	1.62	0.1	0.1	0.29	0.1
Tanzania June Inf*all*	0	0	0	0	2.12	12.17	69.05	16.14	0.26	0	0	0	0.26
Tanzania June Not*all*	0	0	0	2.12	5.29	12.43	65.87	11.38	1.85	0	0	1.06	0
Tanzania June F*all*	0	0	0	1.08	3.25	11.25	65.99	16.53	1.22	0	0.14	0.41	0.14
Tanzania June E*all*	0	0	0	0	2.72	16.67	58.84	20.07	1.7	0	0	0	0
Tanzania June Fi*all*	0	0	0	0.95	1.71	12.76	61.33	20.95	1.71	0	0.19	0.19	0.19
Tanzania June Se*all*	0	0	0	0.57	3.62	12.19	63.81	17.71	1.52	0.19	0	0.38	0
Zürich all	0	0	0	0	0.95	13.1	72.5	13.1	0.36	0	0	0	0
Zürich Inf*all*	0	0	0	0	0	12.28	73.68	13.6	0.44	0	0	0	0
Zürich Not*all*	0	0	0	0	0.56	12.22	75	12.2	0	0	0	0	0
Zürich B*all*	0	0	0	0	0.56	13.33	71.39	14.17	0.56	0	0	0	0
Zürich M*all*	0	0	0	0	1.39	11.39	74.44	12.5	0.28	0	0	0	0
Zürich Fi*all*	0	0	0	0	1.43	12.86	70.24	14.76	0.71	0	0	0	0
Zürich Se*all*	0	0	0	0	0.48	13.33	74.76	11.43	0	0	0	0	0
Alpthal all	0	0	0	2.9	24.1	37.5	30.6	4.6	0.2	0	0	0	0

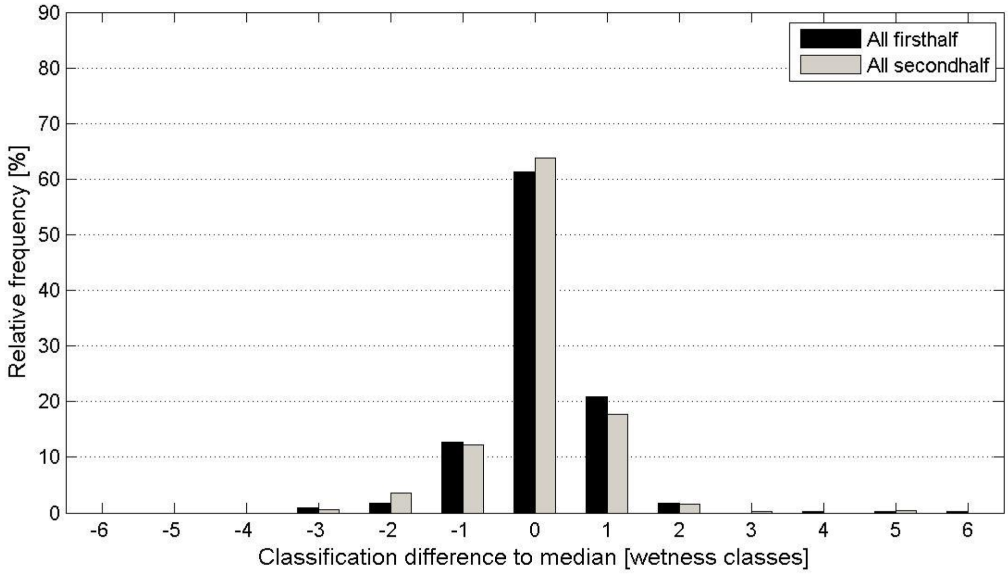


Figure 43: Deviation of wetness classifications relative to the median of each sampling point for Fiall and Seall for the test in June plotted as relative frequency distribution.

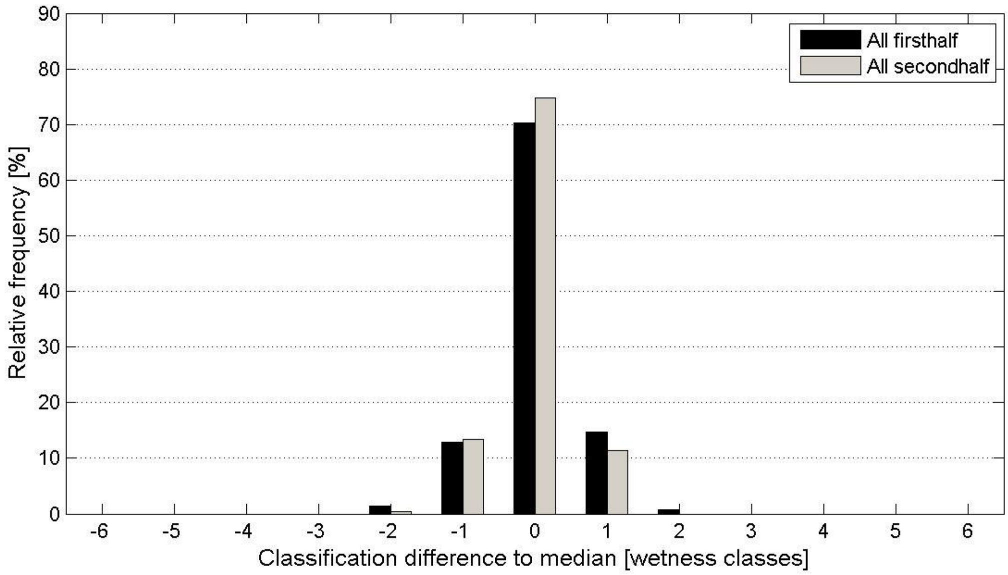


Figure 44: Deviation of wetness classifications relative to the median of each sampling point for Fiall and Seall for the test in Zürich plotted as relative frequency distribution.

Table 7: Volumetric water contents for 16 sampling points in Mungushi measured with TDR and gravimetric method.

			Volumetric Water Content [%]	
Sampling Point	Soil Sample moist [g]	Soil Sample dry [g]	Gravimetric Measurement	TDR-Measurement
1	114.9	95.3	19.6	32.2
2	115.4	98.3	17.1	9.7
3	150.8	110.8	40.0	31.4
4	149.9	112.4	37.5	32.9
6	102.6	87.6	15.0	8.4
8	161.8	131.0	30.8	33.7
9	182.2	132.6	49.6	32.3
10	115.7	98.0	17.7	14.2
11	159.9	131.6	28.3	26.9
12	151.3	124.1	27.2	28.2
13	152.0	118.7	33.3	35.1
14	145.1	111.8	33.3	39.2
15	129.9	104.5	25.5	28.2
16	119.7	96.9	22.8	23.8
17	139.1	100.8	38.3	42.4
18	142.4	97.3	45.1	51.9

Table 8: Volumetric water contents for all 42 sampling points in Kichangani measured with FDR and gravimetric method.

			Volumetric Water Content [%]	
Sampling Point	Soil Sample moist [g]	Soil Sample dry [g]	Gravimetric Measurement	FDR-Measurement
1	120.7	90.9	29.8	39.8
2	147.5	118.4	29.1	32.9
3	118.5	105.4	13.0	13.2
4	131.9	108.7	23.2	34.6
5	144.4	113.0	31.4	42.7
6	124.9	108.5	16.3	22.5
7	111.9	91.1	20.8	27.7
8		102.2	27.9	33.1

(Continuation Table 8)			Volumetric Water Content [%]	
Sampling Point	Soil Sample moist [g]	Soil Sample dry [g]	Gravimetric Measurement	FDR-Measurement
9	122.5	93.1	29.3	41.5
10	128.0	98.3	29.7	44.4
11	94.1	83.1	11.0	14.0
12	152.1	113.5	38.7	45.0
13	103.0	90.7	12.3	12.7
14	121.7	94.5	27.2	38.0
15	126.6	96.0	30.6	48.9
16	127.1	112.8	14.3	13.0
17	121.4	98.3	23.1	30.5
18	141.5	109.7	31.8	44.8
19	114.5	106.4	8.1	11.8
20	155.2	124.5	30.7	34.9
21	171.2	138.7	32.5	38.4
22	153.1	119.5	33.6	37.9
23	163.9	133.8	30.1	34.3
24	141.8	111.7	30.2	35.9
25	123.5	109.9	13.6	15.5
26	161.6	129.8	31.8	34.9
27	127.8	113.8	14.1	14.5
28	146.4	130.0	16.4	18.5
29	137.1	112.4	24.8	34.0
30	151.2	115.8	35.5	40.7
31	105.8	73.3	32.5	53.0
32	144.3	109.4	34.9	41.8
33	143.6	116.1	27.5	27.8
34	147.8	116.2	31.7	36.9
35	153.3	131.6	21.7	16.2
36	123.8	95.5	28.3	44.8
37	157.3	126.8	30.4	23.3
38	119.1	99.8	19.3	20.7

(Continuation Table 8)			Volumetric Water Content [%]	
Sampling Point	Soil Sample moist [g]	Soil Sample dry [g]	Gravimetric Measurement	FDR-Measurement
39	152.6	124.4	28.2	25.2
40	137.9	106.3	31.6	42.9
41	145.0	117.0	28.0	34.3
42	163.7	121.6	42.1	45.9

Table 9: Volumetric water contents for 39 sampling points in Zurich (without point 2, as the sampling point was excluded due to an incorrect gravimetric value) measured with FDR and gravimetric method.

			Volumetric Water Content [%]	
Sampling Point	Soil Sample moist [g]	Soil Sample dry [g]	Gravimetric Measurement	FDR-Measurement
1	112.1	88.8	23.3	36.5
3	176.0	111.8	64.2	54.2
4	110.7	94.0	16.7	26.8
5	154.0	119.8	34.2	44.3
6	90.7	70.4	20.3	20.8
7	113.2	106.0	7.2	21.0
8	118.7	83.4	35.3	50.2
9	124.6	99.4	25.2	37.2
10	115.0	82.3	32.8	33.1
11	122.8	71.7	51.1	66.3
12	145.4	81.9	63.6	73.3
13	117.9	104.2	13.7	30.6
14	188.3	128.3	60.0	69.9
15	187.6	124.0	63.6	64.3
16	127.7	96.5	31.2	34.5
17	196.7	146.7	50.0	61.9
18	144.7	121.8	22.9	21.7
19	121.9	113.8	8.1	15.0
20	106.7	94.8	12.0	14.1

(Continuation Table 9)			Volumetric Water Content [%]	
Sampling Point	Soil Sample moist [g]	Soil Sample dry [g]	Gravimetric Measurement	FDR-Measurement
21	88.4	70.0	18.4	13.3
22	109.5	83.9	25.6	23.3
23	137.0	128.9	8.0	21.1
24	86.0	50.5	35.5	43.7
25	94.8	73.6	21.3	27.1
26	130.5	64.6	65.9	57.5
27	90.2	53.7	36.5	42.5
28	136.1	82.5	53.7	52.4
29	148.5	79.0	69.5	52.9
30	131.8	53.7	78.0	67.7
31	125.2	59.5	65.7	53.0
32	111.6	55.0	56.7	58.5
33	90.0	58.5	31.5	44.3
34	108.0	92.6	15.4	23.3
35	106.0	77.4	28.6	36.9
36	144.6	56.6	87.9	72.6
37	126.4	54.4	72.0	61.6
38	126.0	51.5	74.5	60.2
39	125.1	52.9	72.2	71.1
40	130.7	99.5	31.2	55.2

Table 10: Selected information about the two soil profiles in Mungushi, Tanzania.

Horizon	Abundance of rock fragments [Vol.%]	Estimated bulk density [kg dm-^{3}]	Munsell-Colour	pH-value (lab)
Profile 1 15.04.2014				
1	1% - 2%	1.2 – 1.4 (Heavy loam and clay soils)	5YR 3/3	6.2
2	1% - 2%	1.0 – 1.2 (Heavy loam and clay soils)	7.5YR 4/4	6.3
3	5% - 10%	0.9 – 1.2 (Sand-, silt and light/moderate loam soils)	7.5YR 4/4	6.8
4	1% - 2%	1.4 – 1.6 (Heavy loam and clay soils)	5YR 3/4	6.6
Profile 2 15.04.2014				
Horizon	Abundance of rock fragments [Vol.%]	Estimated bulk density [kg dm-^{3}]	Munsell-Colour	pH-value (lab)
1	1% - 5%	1.4 – 1.6 (Heavy loam and Clay soils)	10YR 3/3	6.7
2	1% - 5%	1.2 – 1.6 (Sand-, silt and light/moderate Loam soils)	2.5Y 4/2	6.4
3	1% - 5%	1.2 – 1.6 (Sand-, silt and light/moderate loam soils)	2.5Y 5/2	6.9
4	1% - 5%	1.6 – 1.7 (Heavy loam and clay soils)	10YR 3/1	7

Personal declaration:

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Zürich, 30.06.15 Daniela Müller