Authors response to the reviews in the interactive discussion of the manuscript submitted to ESSD: "Soil moisture and matric potential – An open field comparison of sensor systems" by C. Jackisch et al.

We thank all reviewers and the editor for their careful work on our manuscript and the constructive comments. Along the lines drafted in our responses to the individual reviews, we have substantially revised our manuscript.

The most substantial change is the inclusion of a brief cross-correlation analysis of the recorded data. As the manuscript is intended as data publication, we include this to elucidate the dataset rather than a throughout evaluation and discussion of the implications of our observations. Although it is often proposed to refer to the TDR systems, we do not see any value in a more or less arbitrary attribution of some sensors to be a true reference. Consequently, we seek to highlight the deviations between records of the scientific instruments, which actually should report the same true state dynamics. Depending on the application (e.g. a combination of data from different sensors or a selection of a sensor system for a forthcoming study) each reader might select other features to base an evaluation on.

As suggested by the anonymous referee #2, we have updated the table and description of the used sensor systems and added specific details about the actual measurement technique for each sensor. Moreover, we added a brief discussion of the data to the manuscript.

Our revisions led to a modification of figure 2 and 3. We have added 5 correlation matrix figures which hold a lot of information. Instead of a repetition of our line-by-line comments from the replies to the referees, we have added a comparison version of our manuscript with highlighted changes (excluding the figures).

On behalf of all co-authors,
Conrad Jackisch

Soil moisture and matric potential – An open field comparison of sensor systems

Conrad Jackisch[1], Kai Germer[2], Thomas Graeff[3,4], Ines Andrä[2], Katrin Schulz[2], Marcus Schiedung[2], Jaqueline Haller-Jans[2], Jonas Schneider[2], Julia Jaquemotte[2], Philipp Helmer[2], Leander Lotz[2], Andreas Bauer[3], Irene Hahn[3], Martin Šanda[5], Monika Kumpan[6], Johann Dorner[6], Gerrit de Rooij[7], Stefan Wessel-Bothe[8], Lorenz Kottmann[9], Siegfried Schittenhelm[9], and Wolfgang Durner[2]

[1]Karlsruhe Institute of Technology KIT, Institute of Water Resources and River Basin Management, Chair of Hydrology, Kaiserstr. 12, 76131 Karlsruhe, Germany; current address: Technische Universität Braunschweig, Institute of Geoecology, Dept. Landscape Ecology and Environmental Systems Analysis, Langer Kamp 19c, 38106 Braunschweig, Germany.
[2]Technische Universität Braunschweig, Institute of Geoecology, Dept. Soil Science and Soil Physics, Langer Kamp 19c, 38106 Braunschweig, Germany.
[3]University of Potsdam, Institute of Earth and Environmental Science, Potsdam, Germany
[4]Umweltbundesamt (Federal Environment Agency), Department IV 2.1, Dessau-Rosslau, Germany
[5]Czech Technical University, Faculty of Civil Engineering, Department of Irrigation, Drainage and Landscape Engineering, Thakurova 7, 166 29, Praha 6, Czech Republic
[6]Federal Agency for Water Management, Institute for Land & Water Management Research, 3252 Petzenkirchen, Austria
[7]Helmholtz Centre for Environmental Research UFZ, Department of Soil System Science, Theodor-Lieser-Str. 4, 06120 Halle, Germany
[8]ecoTech Umwelt-Meßsysteme GmbH, Nikolausstr. 7, 53129 Bonn, Germany
[9]Julius Kühn-Institute, Federal Research Centre for Ciltivated Plants, Institute for Crop and Soil Science, Bundesallee 58, 38116 Braunschweig, Germany

Correspondence: Conrad Jackisch (c.jackisch@tu-braunschweig.de)

Abstract. Soil water content and matric potential are central hydrological state variables. A large variety of automated probes and sensor systems for state monitoring exists and is frequently applied. Most ~~studies~~ applications solely rely on the calibration by the manufacturers. Until now, there is no commonly agreed calibration procedure. Moreover, several opinions about the capabilities and reliabilities of specific sensing methods or sensor systems exist and compete.

5   A consortium of several institutions conducted a comparison study of currently available sensor systems for soil water content and matric potential under field conditions. All probes have been installed in 0.2 m depth below surface following best practice procedure. We present the setup and the recorded data of 58 probes of 15 different systems measuring soil moisture and further 50 probes of 14 different systems for matric potential. We briefly discuss the limited coherence of the measurements in a cross-correlation analysis.

10   The measuring campaign was conducted in the growing period of 2016. The monitoring data, results from pedophysical analyses of the soil and laboratory reference measurements for calibration are published in Jackisch et al. (2018, https://doi.org/10.1594/PANGAE

1 Introduction

15 Soil water content is defined as volumetric proportion of water in the multiphase bulk soil. Since the proposition of soil moisture determination based on relative electrical permittivity of the bulk soil in the 1970s (presumably starting with Davis et al., 1966; Geiger and Williams, 1972; Chudobiak et al., 1979) many commercially-available systems have been developed. They can be roughly grouped into time-domain reflectometry (TDR), ~~frequency-domain reflectometry (FDR) and capacitive sensors~~ mostly impedance-based determination of the capacitance, and time domain transmission (TDT) techniques, which all rely on the

20 strong contrast of the relative electrical permittivity of water (80) compared to air (1) and minerals (3-5) in the ~~soil bulk~~ bulk soil. However, the relative electrical permittivity is also influenced by temperature (Roth et al., 1990; Wraith and Or, 1999; Owen et al., 2002; Rosenbaum et al., 2011), soil texture (Ponizovsky et al., 1999) and organisation of thin water film layers (Wang and Schmugge, 1980). In addition, there is a frequency dependency of such measurements~~emphasising~~. While low measurement frequencies might be dominated by bulk electrical conductivity (Schwartz et al., 2013), also higher frequencies

25 emphasise more or less on effects of solutes, clay surfaces and organic matter on the conductor and dielectric properties (Loewer et al., 2017).

~~Besides the theoretical aspects~~In addition to the theoretical concerns, the sensing systems ~~have~~ need to solve a series of technical issues ~~associated with the~~ e.g. the sensor wiring and coupling, the facilitation of the signal propagation from the sensor into the soil, and stability of the measurements themselves to corrosion~~and shielding . Thus~~, shielding and temperature. Thus

30 one has to be aware that the theoretically more appropriate TDR technology might not per se deliver more precise readings, when technical issues obscure the actual measurement. ~~Equally,~~Another common assumption relates to a large sensing volume ~~is neither guaranteeing more precise readings integrating over soil heterogeneity nor is~~ of being more favourable. However, neither effects of a change of the sensed soil volume with changing bulk permittivity nor the influence of ~~water equally distributed within the sensed volume~~ the distribution of water within this volume can be usually specified.

35 Accompanying soil water content, matric potential is the second central hydrological state variable of soils. It ~~measures~~ is an integral over the macroscopic interfacial tension of the pore-scale menisci ~~integration over~~ of all air-water-soil interfaces. It was introduced by Buckingham (1907); Gardner and Widtsoe (1921) as capillary potential and combines the effects of soil water content, pore space characteristics and the respective configuration of the soil water in the pore space. Tensiometers are ~~employed~~ used since over a century to directly measure the capillary tension (Or, 2001). Because the measurement is limited

40 to the vaporisation point of water in the tensiometer against an atmospheric pressure at approx. $1000\,\mathrm{hPa}$, polymer-based versions (van der Ploeg et al., 2010) and alternative sensing techniques measuring matric potential indirectly through water content detection in a porous ceramic material with known retention properties have been developed.

In order to identify conceptual limits and technological issues of currently available systems for measurement of soil water content and matric potential we conducted a comparison study under field conditions. For this, a large number of sensors has

2

Figure 1. Sensor comparison field site after sensor installation.

Figure 2. Layout of the sensor comparison field site. Each grid cell covers 0.5 x 0.5 m. All sensors are placed in 0.2 m depth. Positions of soil samples marked with dots. Total covered area 14 m in east-west and 4 m in north-south extent.

45 been installed in a specifically homogenised and levelled agricultural field with loamy sandy soil. Vegetation effects have been excluded by glyphosate treatment. The test has been conducted from May to November 2016.

2   Study setup

2.1   Site description and study layout

The study site is located on an agricultural test site of the Julius Kühn Institute, Braunschweig, Germany ($52.2964°$ N, $10.4361°$
50 E, Fig. 1). The site is characterised by loess and sand depositions over marls of the last glaciations in a plain with very little relief. The soil is a very homogeneous sandy loam with a gravel content below 3%. The plot had been prepared by harrowing, plowing and compacting. In order to keep the system as simple as possible, vegetation was suppressed by glyphosate application.

The sensors were installed in a grid of $0.5\,\mathrm{m}$ distance (Fig. 2) in $0.2\,\mathrm{m}$ depth following the best practice recommendation
55 of the manufacturers. The total covered area amounts to $14\,\mathrm{m}$ in east-west and $4\,\mathrm{m}$ in north-south extent. Whenever the probe design allowed (round shape with suitable diameter) insertion from the surface with minimal disturbance using an auger tilted by $45°$ was chosen. Alternatively, probes were positioned horizontally below undisturbed surface from a shallow access pit. Probes which included their access tubes have been installed vertically. In order to avoid compaction of the surface by walking on it, plywood panes were temporally placed along the access paths during the field work. Installation took place in several
60 campaigns in April and May, 2016. The field was exposed to natural weather conditions until August 24, 2016. After that date, a tunnel green house was installed for protection against rain in order to reach lower matric potentials. Adverse effects of drainage from the tunnel near the edges of the tunnel appear to have occurred later in the year. The full dataset is included in the repository. However, we focus on the period until August 24 in this study.

~~Sensor comparison field site after sensor installation.~~
65 ~~Layout of the sensor comparison field site. All sensors are placed in 0.2 m depth.~~

Table 1. Employed sensor systems in the comparison study. All information according to the respective manufacturer.

3

2.2 Sensor systems

In total 58 probes of 15 different systems measuring soil moisture and 50 probes of 14 different systems for matric potential have been used. Each sensor of a system has two to four replicates. An overview about the ~~sensors~~ sensor systems is given in table 1. All sensors are utilised with the manufacturer's calibration.

Soil moisture is measured based on electromagnetic estimates of the bulk soil relative permittivity, which is mostly controlled by water. For this, four systems use the TDR (time domain reflectometry) method. They differ in the sampling and evaluation technique of the travel time of a pulse along the waveguides. While the Trase system evaluates the intersection of tangent lines of the incoming and reflected pulse voltages (Soilmoisture Equipment inc., 1996), the Trime systems perform time measurements at distinct voltage levels (Stacheder et al., 1997). Eight systems measure the capacitance of the bulk soil by impedance of oscillating pulses. They differ mostly in the applied frequency and geometry of the setup. Two systems use a TDT (time domain transmission) technique, which estimates the soil bulk permittivity by counting of received pulses emitted at high frequency (Wild et al., 2019). One system extends the TDT technique by determining the oscillation frequency of a TDT system (Bogena et al., 2017).

Matric potential is measured either directly through a pressure transducer in tensiometers or through soil moisture measurements in an EPM (equivalent porous medium) with known soil water retention properties. For the EPM, two systems measure the electric resistance of a gypsum EPM which is related to its moisture. Three systems use the impedance measurement described above. Three systems estimate the moisture by measuring the dissipation of a heat pulse in the EPM. One system uses a hydrophile polymer instead of water to extend the measurement range of the tensiometers (Bakker et al., 2007).

2.3 Pedophysical analyses

Eight undisturbed ring samples ($250\,\mathrm{mL}$) taken at $0.2\,\mathrm{m}$ depth near the probes (dots in Figure 2) have been analysed for soil water retention properties (HYPROP and WP4C, Meter Group). ~~In three samples also~~ The first batch of samples was taken during the sensor installation on April 21, 2016. The three samples were taken at the end of the experiment on December 6, 2016. These were also analysed for saturated hydraulic conductivity (KSAT, Meter Group), texture (sedimentation method after DIN ISO 11277), organic matter (ignition loss after DIN EN 13039), and pH (in a suspension in $0.01\,\mathrm{mol\,L^{-1}}\ CaCl_2$ using a WTW pH electrode after DIN ISO 10390)~~was measured~~.

2.4 Laboratory reference

As reference measurement intended for a posteriori calibration, an undisturbed, cylindrical soil monolith of $15.7\,\mathrm{L}$ ($0.3\,\mathrm{m}$ height, $0.26\,\mathrm{m}$ diameter) has been sampled in $0.05\text{--}0.35\,\mathrm{m}$ depth and equipped with six soil moisture sensors of three of the employed systems (Pico32, 10HS and 5TM). The probes were ~~installed~~ selected based on the plausibility of their records in the field, their size for the laboratory installation and availability. We installed the probes vertically in the same depth, referenced to the centre of the probes. The monolith was initially saturated and exposed to free evaporation for three weeks set up on a weighing scale in the lab. Referenced against the dry weight of the whole setup, this delivers time series of gravimetric soil water content

Figure 3. Soil water retention data from ~~seven~~ eight 250 mL ring samples analysed in HYPROP and WP4C apparatus. Fitted van Genuchten model with free m parameter and diffusive hydraulic conductivity k_{sat}^{*} estimate

plus the readings from six sensors. In order to avoid overly strong internal soil moisture gradients due to evaporation at the surface, the sample was periodically covered.

3 Data description

The data and some exemplary analysis are hosted in the PANGAEA repository Jackisch et al. (2018, https://doi.org/10.1594/PANGAEA.8923

3.1 Pedophysical data

The pedophysical data is given in *pedophysical_data.xlsx* as table of analyses of eight undisturbed ring samples. Samples no. 1–5 have been taken on April 21, 2016 during the first sensor installation campaign. Samples no. 6–8 were taken on December 6, 2016 during the first sensor removals. Bulk density (BD) was determined by referring the dry weight of the bulk soil after oven drying at 105 °C for 3 days to the sample volume of 250 mL. Porosity was estimated based on the soil water content at full saturation at the onset of the retention curve measurements using the free evaporation method of the HYPROP apparatus referring the total weight under saturated conditions to the dry weight. Over the course of the measurement of the HYPROP, tensions in the sample are referred to the total weight resulting in the retention curve from pF 0 to pF 2.5. ~~Three samples were also~~ To measure pairs of total weight and matric potential, the samples were processed in a WP4C chilled mirror potentiometer ~~measuring pairs of total weight and matric potential at higher suction heads~~. An overview is given in Figure 3. The resulting measurements were processed in the HYPROP-FIT software (ver. 3.5.1, Meter Group, original files given as *hyprop.zip*, exported derivatives are stored in *vG_JKI_params.xlsx, ku_obs.xlsx, retention_obs.xlsx and hyprop.xlsx*) for fitting of the original ~~(Van Genuchten, 1980)~~ Van Genuchten (1980) pedotransfer model with free m-parameter. The resulting parameters for saturated soil water content (θ_{sat}, ~~m³ m⁻³~~ $\mathrm{m^3/m^3}$), residual soil water content (θ_{res}, ~~m³ m⁻³~~ $\mathrm{m^3/m^3}$), α (~~m⁻¹~~ $\mathrm{1/m}$), n, m and diffusive flow estimated saturated hydraulic conductivity (k_{sat}^{*}, ~~m s⁻¹~~ $\mathrm{m/s}$) are reported for each sample.

Figure 3 presents the van Genuchten parameters fitted to all retention measurements. Not surprisingly, the greatest differences relate to a spread in porosity. However given the few exceptions with tensions below pF 2 (100 hPa) in the recorded time series (Fig. 4C), the very strong coherence of the retention curves at higher tensions corroborates the high homogeneity at the site.

3.2 Monitoring data

The monitoring data of all sensors is compiled in the files *Theta20.xslx*, *Psi20.xlsx* and *T20.xlsx* holding volumetric soil water content (~~m³ m⁻³~~ $\mathrm{m^3/m^3}$), matric potential (hPa) and soil temperature (°C) respectively. The data of all individual sensors was merged into the common tables aggregated to 30 min averages. The data was filtered for obvious measurement errors outside the physically possible ranges. Initial inconsistencies of the time stamps from different loggers have been removed by time-shift correction based on an analysis of the phase-coherence of the diurnal temperature signal.

Figure 4. A: Meteorological forcing (~~top panel,~~ DWD station 662), B: volumetric soil water content dynamics~~(mid panel)~~, and C: matric potential~~(bottom panel)~~. 30 min median of all sensors of a system as solid line, variance as shade. Sensor systems with non-plausible oscillations and value ranges are given as "other".

Figure 5. Soil water content in undisturbed monolith. Sensor values vs. gravimetric reference.

Meteorological reference is reported from the German Weather Service (DWD) Station 662, Braunschweig through their Climate Data Centre ftp://ftp-cdc.dwd.de/pub/CDC//observations_germany/climate/hourly referring to the station number. In addition, records of a weather station $100\,\mathrm{m}$ East from the plot is reported in *meteo_jki.xlsx*. It holds half-hourly records of solar radiation (~~W m^{-2}~~ W/m^2), wind direction (°), wind speed (~~m s^{-1}~~ m/s), precipitation (~~mm m^{-2}~~ mm/0.5 h), air temperature (°C) and relative air humidity (%). In addition calculated values for dew point (°C) and cumulative precipitation (~~mm m^{-2}~~ mm) are given.

The measurements of volumetric soil water content and matric potential exhibit plausible dynamics in general. The sensors react to events and recover their ranks later on (Figure 4). However, the different sensor classes deviate substantially with regard to their absolute values, the intensity of the reaction to events and to the existence and amplitude of diurnal cycles~~(Figure 4)~~.

3.3 Laboratory reference

Between February 23 and March 21, 2017, an undisturbed soil monolith was transferred to the laboratory, initially saturated and later left for drying. The monitored gravimetric soil water content and the readings from the installed sensors (all ~~m^3 m^{-3}~~ m^3/m^3, 10 min means) are given in file *lab_mono.xlsx*. The gravimetric and sensed soil water content are not all linearly related. Interestingly, the capacitative sensors (5TM and 10HS) provide a better linear fit although they deviated more strongly from the absolute values, while the TDR sensors present a non-linear relation (Figure 5). In accordance with our records in the field experiment, the sensors of one system are highly congruent compared to larger differences between the systems. Also the ranking with the 10HS > Pico32 > 5TM is the same. Moreover, the capacitive sensors show reoccurring shifts of the measurements over time, which coincide with the repeated coverage of the sample with an aluminium lid. These shifts resulted from changed configurations of the "capacitor" consisting of the sample in a stainless steel ring and a metal lid. From a point of view of measurement principles this application flaw should be considered in future calibration approaches, omitting effects of electrical currents through the sampling material.

4 Data evaluation and discussion

4.1 ~~First data assessment and evaluation of experimental hypotheses~~ Time series cross-correlation

~~The repository also holds a script (*Sensor_Comparision_EEMD.ipynb*) which provides direct access to general data visualisation and processing using Python. With respect to the experimental homogeneity assumption~~ Most systems recorded plausible

data which however differ strongly with respect to absolute value, event reaction and seasonal trend (see Figure 4B&C). In order to evaluate the presented data, we present a brief cross-correlation analysis in this section. The following pairwise correlations measures are calculated: A linear regression model and its Pearson correlation coefficient evaluating the linear distribution of the residuals (r), the Spearman rank correlation (rho) as a measure of the overall time series coherence, and the Kling-Gupta-Efficiency (KGE) as a measure for the time series dynamics and its coherence of absolute values. The parameters for the linear regression model give a first impression about the scaling (a) and deviation of the mean value (i, intercept) of the respective pairs. The figures 6 and 7 show the distribution plots of the means of each sensor system in the diagonal axis. The lower panels show the scatter points (blue), the linear regression (orange line), the 0.95 predictive uncertainty bands (grey dashed lines) and the 1:1 line reference (red). In the upper panels the correlation measures are given es values and in a simple bar plot in the background. Here, a is plotted as deviation from unity. High correlation is signalled when the first two bars are small and the latter three bars are tall.

It should be noted that high or low correlation does not necessarily reflect on the performance of a sensor system. Given the general assumption that TDR systems are superior, one might be interested in a comparison between the Pico and Trase systems. While the deviation from the 1:1 line is relatively small and moderate to high correlation exists, this is still far from being called a perfect match. The predictive uncertainty bands range around the often reported 3 vol.% accuracy. Other applications might use different sensors of the Meter/Decagon family (10HS, ~~a close-up on reactions of the tensiometers to rain events shows emerging redistribution structures at the surface which imprint on the~~ 5TM, EC5), which are all based on the same measurement principle but show substantial deviations in offset and scaling, too. Likewise, the high correlation but strong scaling of the TDR systems Pico32 and T3P (Figure A1), which are based on the same technique but the first is a rod probe and the latter a tube probe, hints to a systematic issue (which has lead to a revision of the tube sensors already).

As seen in the time series plot (Figure 4C), the tensiometers correlate very well. This is recovered in the correlation analysis (Figure 7) with very high coefficients, although the offset might be an issue to look at. The indirect sensors systems using an equivalent porous medium present larger deviations but appear to be quite capable in general. Surprisingly, the developments in the MPS-family, which are basically differing in the number of calibration points of the retention characteristics of the medium, do not show a clear superiority of the latest MPS6. One should note the high offset for the MPS6 and TM compared to the tensiometers.

The sensing systems which did not deliver plausible data are shown in the appendix Figures A1 and B1. It has to be highlighted that we cannot exclude issues with probe storage, installation failure or technical interferences leading to the registered performance. These systems can well be capable to outperform other sensors, when such issues are resolved. In addition, the systems which include a temperature sensor are reported in the appendix Figure C1. There, high correlation of the records of most systems can be seen – except for systems which are not fully buried in the soil.

4.2 Evaluation of experimental hypotheses

We selected and prepared the site to be as homogenous as possible for an agricultural soil system. In order to evaluate the homogeneity assumption in our experiment, we focus on the reactions to some rain events over the monitoring period. We

Figure 6. Correlation matrix of all plausible sensor systems measuring soil moisture (given in $\mathrm{m}^3/\mathrm{m}^3$). Diagonal panels give histogram and kernel density distribution of $0.5\,\mathrm{h}$ means of all sensors of one system. Lower half give scatter plot (blue dots), linear regression (orange line), 0.95 predictive uncertainty bands (grey dashed lines), and the 1:1 line (red). The upper half reports the respective correlation measures with a and i as scaling factor and intercept of the linear regression model, r as Pearson correlation coefficient, rho as Spearman rank correlation coefficient, and KGE as Kling-Gupta-Efficiency. These values are plotted as bars in the same order in the background, where a is plotted as deviation from unity.

Figure 7. Correlation matrix of all plausible sensor systems for matric potential (given in hPa). Diagonal panels give histogram and kernel density distribution of $0.5\,\mathrm{h}$ means of all sensors of one system. Lower half give scatter plot (blue dots), linear regression (orange line), 0.95 predictive uncertainty bands (grey dashed lines), and the 1:1 line (red). The upper half reports the respective correlation measures with a and i as scaling factor and intercept of the linear regression model, r as Pearson correlation coefficient, rho as Spearman rank correlation coefficient, and KGE as Kling-Gupta-Efficiency. These values are plotted as bars in the same order in the background, where a is plotted as deviation from unity and i is scaled with 0.01.

185   selected the tensiometers as most direct measurement techniques as references for the soil water states~~in 0.2 m depth~~. When comparing the individual sensor reaction to four rain events over four months, a strong deviation from the initially high congruency over time is apparent (Figure 8). While the tensiometers recorded highly consistent values in the early phase of the experiment, the sensor readings divert irrespectively to their sensor system over the course of the experiment. Since we observed emerging redistribution structures at the surface, we attribute them to imprint on the soil water states in $0.2\,\mathrm{m}$ depth.

190   The effect of emerging structures on the overall system properties can also be seen in the in-situ retention curves of some systems (which is left to further analyses).

Moreover, one has to be aware of the bare-soil field conditions which resulted in relatively large diurnal temperature amplitudes in the soil, including related soil water processes and a potential exaggeration of local heterogeneity.

4.3   Data quality of soil water sensing

195   Overall, the data raise substantial questions about the data quality of state of the art measurement systems of soil water content based on relative electrical permittivity of the bulk soil without specific, in-situ calibration. Despite delivering plausible signals, neither the absolute values nor the relative reactions to events appear to be very accurate. Given the non-linear relation of gravimetric and TDR-sensed soil water content in the laboratory, and given the highly different monitoring records of the different TDR systems, the general believe of their superiority might deserve more detailed examination.

Figure 8. ~~Close-up on reactions~~ Reactions of all tensiometers (T4, T5, T8) to four events. Emerging redistribution structures at the surface lead to growing deviation of the soil states across relatively short distances.

The a posteriori calibration approach did not succeed. To not disturb the system, we did not take any soil samples at different states in the field, which could have been a better means for sensor calibration.

Although several studies evaluated soil moisture sensing systems (e.g. Walker et al., 2004; Mittelbach et al., 2011; Chow et al., 2009) and sensor calibration is known to be an issue (e.g. Rowlandson et al., 2013; Bogena et al., 2017; Rosenbaum et al., 2011), the scientific application still lacks a common procedure for evaluation and calibration of such data. When data from different sites and sensor systems shall be combined (e.g. Dorigo et al., 2011), our findings should raise awareness that deviations are not always a matter of soil heterogeneity.

Only few sensors allow for a recording of their raw signal. No sensor records the relative electrical permittivity of the bulk soil. However, the internal conversion functions and parameters are rarely accessible to the users. As the systems apply a higher order polynomial function related to the Topp et al. (1980) equation, reverse calculation of the raw values or relative electrical permittivity is very difficult. For scientific application the sensor systems should provide such raw values as prerequisite for a common calibration and evaluation procedure.

5 Conclusions

The data reported ~~from~~ in this study is intended to compare currently available systems for measurement of soil water content and matric potential under field conditions~~in order to identify conceptual limits and technological issues~~. While most systems did deliver plausible data, the records do neither agree on a specific absolute value range nor are the relative values in accordance or rank-stable during events. Thus, mere plausibility checks of such data appear to be insufficient and cannot replace thorough calibration efforts and maintenance. Without calibration, the precious sensor data is risked to render futile for all types of sensing systems. The capability of laboratory calibration on exemplary sensors appears to be insufficient.

6 Code and data availability

The data of the sensor comparison study is hosted in the PANGAEA repository Jackisch et al. (2018, https://doi.org/10.1594/PANGAEA.8923 It is given under creative commons license (CC BY-NC-SA 3.0) without any liability. The repository also holds a script (*Sensor_Comparision_EEMD.ipynb*) which provides direct access to general data visualisation and processing using Python given under a general public license (GNU GPL 3).

Appendix A: Additional data visualisation

The following figures A1, B1 and C1 are given to complement the general findings above.

Figure A1. Correlation matrix of two references and the less plausible soil moisture sensor systems (values given in $\mathrm{m^3/m^3}$). Diagonal panels give histogram and kernel density distribution of $0.5\,\mathrm{h}$ means of all sensors of one system. Lower half give scatter plot (blue dots), linear regression (orange line), 0.95 predictive uncertainty bands (grey dashed lines), and the 1:1 line (red). The upper half reports the respective correlation measures with a and i as scaling factor and intercept of the linear regression model, r as Pearson correlation coefficient, rho as Spearman rank correlation coefficient, and KGE as Kling-Gupta-Efficiency. These values are plotted as bars in the same order in the background, where a is plotted as deviation from unity. The ES system did not record data in the desired density (only 27 values compared which is 0.5% of the data series).

Figure B1. Correlation matrix of two references and the less plausible sensor systems for matric potential (given in hPa). Diagonal panels give histogram and kernel density distribution of $0.5\,\mathrm{h}$ means of all sensors of one system. Lower half give scatter plot (blue dots), linear regression (orange line), 0.95 predictive uncertainty bands (grey dashed lines), and the 1:1 line (red). The upper half reports the respective correlation measures with a and i as scaling factor and intercept of the linear regression model, r as Pearson correlation coefficient, rho as Spearman rank correlation coefficient, and KGE as Kling-Gupta-Efficiency. These values are plotted as bars in the same order in the background, where a is plotted as deviation from unity and i is scaled with 0.01.

Author contributions. All authors contributed in the field work and data collection. WD, IA and KG designed the experimental setup and primarily coordinated the field study. CJ and KG did the laboratory analyses. CJ compiled the data and performed the pre-analysis. CJ, KG, TG and WD wrote the manuscript. All authors contributed valuable feedback during this process.

Competing interests. The authors declare no competing interests.

230 *Disclaimer.* All data is reported as measured in the field and laboratory. We explicitly warn to use the data for nomination of any "best" system as local heterogeneity, emerging surface structures, inappropriate storage of probes before installation, malfunction of single probes and meanwhile revisions of hard- and software of the systems cannot be excluded. Some of the sensors have been upgraded by the manufacturers based on preliminary feedback from the comparison study already.

Figure C1. Upper panel: Time series of most recorded temperature values.
Lower panels: Correlation matrix of the sensor systems with temperature recording (given in °C). Diagonal panels give histogram and kernel density distribution of $0.5\,\mathrm{h}$ means of all sensors of one system. Lower half give scatter plot (blue dots), linear regression (orange line), 0.95 predictive uncertainty bands (grey dashed lines), and the 1:1 line (red). The upper half reports the respective correlation measures with a and i as scaling factor and intercept of the linear regression model, r as Pearson correlation coefficient, rho as Spearman rank correlation coefficient, and KGE as Kling-Gupta-Efficiency. These values are plotted as bars in the same order in the background, where a is plotted as deviation from unity.

References

Bakker, G., van der Ploeg, M. J., de Rooij, G. H., Hoogendam, C. W., Gooren, H. P. A., Huiskes, C., Koopal, L. K., and Kruidhof, H.: New Polymer Tensiometers: Measuring Matric Pressures Down to the Wilting Point, *Vadose Zone Journal*, 6, 196–202, https://doi.org/10.2136/vzj2006.0110, 2007.

240 Bogena, H. R., Huisman, J. A., Schilling, B., Weuthen, A., and Vereecken, H.: Effective Calibration of Low-Cost Soil Water Content Sensors, *Sensors*, 17, 208, https://doi.org/10.3390/s17010208, 2017.

Buckingham, E.: Studies on the movement of soil moisture, Bulletin (United States. Bureau of Soils) No. 38, Washington, http://www.worldcat.org/title/studies-on-the-movement-of-soil-moisture/oclc/29749917, 1907.

Chow, L., Xing, Z., Rees, H. W., Meng, F., Monteith, J., and Lionel, S.: Field Performance of Nine Soil Water Content Sensors on a Sandy
245 Loam Soil in New Brunswick, Maritime Region, Canada, *Sensors*, 9, 9398–9413, https://doi.org/10.3390/s91109398, 2009.

Chudobiak, W. J., Syrett, B. A., and Hafez, H. M.: Recent Advances in Broad-Band VHF and UHF Transmission Line Methods for Moisture Content and Dielectric Constant Measurement, ~~IEEE Transactions on Instrumentation and Measurement~~*IEEE Transactions on Instrumentation and Measurement*, 28, 284–289, https://doi.org/10.1109/TIM.1979.4314833, 1979.

Davis, B. R., Lundien, J. R., and Williamson, A. N.: Feasibility study of the use of radar to detect sur-
250 face and ground water, Tech. rep., US Army Engineer Waterways Experiment Station, Vicksburg, Mississippi, http://cdm16021.contentdm.oclc.org/cdm/ref/collection/p266001coll1/id/2860, 1966.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Oevelen, P. v., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrology and Earth System Sciences*, 15, 1675–1698, https://doi.org/10.5194/hess-15-1675-2011, 2011.

255 Gardner, W. and Widtsoe, J. A.: ~~THE MOVEMENT OF SOIL MOISTURE~~The Movement of Soil Moisture, ~~Soil Science~~*Soil Science*, 11, 215–232, https://journals.lww.com/soilsci/Fulltext/1921/03000/THE_MOVEMENT_OF_SOIL_MOISTURE.3.aspx, 1921.

Geiger, F. E. and Williams, D.: Dielectric constants of soils at microwave frequencies, http://ntrs.nasa.gov/search.jsp?R=19720021782, 1972.

Jackisch, C., Andrä, I., Germer, K., Schulz, K., Schiedung, M., Haller-Jans, J., Schneider, J., Jaquemotte, J., Helmer, ~~Ph~~P., Lotz, L., Graeff, ~~Th~~T., Bauer, A., Hahn, I., Sanda, M., Kumpan, M., Dorner, J., de Rooij, G., Wessel-Bothe, ~~St~~S., Kottmann, L., Schit-
260 tenhelm, S., and Durner, W.: Soil moisture and matric potential - An open field comparison of sensor systems, PANGAEA~~: ,~~, https://doi.org/10.1594/PANGAEA.892319, 2018.

Loewer, M., Günther, T., Igel, J., Kruschwitz, S., Martin, T., and Wagner, N.: Ultra-broad-band electrical spectroscopy of soils and sedi-ments—a combined permittivity and conductivity model, ~~Geophysical Journal International~~*Geophysical Journal International*, 210~~(3):~~, 1360–1373, https://doi.org/10.1093/gji/ggx242, 2017.

265 Mittelbach, H., Casini, F., Lehner, I., Teuling, A. J., and Seneviratne, S. I.: Soil moisture monitoring for climate research: Evaluation of a low-cost sensor in the framework of the Swiss Soil Moisture Experiment (SwissSMEX) campaign, *Journal of Geophysical Research: Solid Earth*, 116, L18 405, https://doi.org/10.1029/2010JD014907, 2011.

Or, D.: Who Invented the Tensiometer?, ~~Soil Science Society of America Journal~~*Soil Science Society of America Journal*, 65, 1, https://doi.org/10.2136/sssaj2001.6511, 2001.

270 Owen, B. B., Miller, R. C., Milner, C. E., and Cogan, H. L.: The Dielectric Constant of Water as a Function of Temperature and Pressure, ~~The Journal of Physical Chemistry~~*The Journal of Physical Chemistry*, 65, 2065–2070, https://doi.org/10.1021/j100828a035, 2002.

Ponizovsky, A. A., Chudinova, S. M., and Pachepsky, Y. A.: Performance of TDR calibration models as affected by soil texture, ~~Journal of Hydrology~~*Journal of Hydrology*, 218, 35–43, https://doi.org/10.1016/S0022-1694(99)00017-7, 1999.

Rosenbaum, U., Huisman, J. A., Vrba, J., Vereecken, H., and Bogena, H. R.: Correction of Temperature and Electrical Conductivity Effects on Dielectric Permittivity Measurements with ~~ECH O~~ ECH2O Sensors, ~~VADOSE ZONE JOURNAL~~*Vadose Zone Journal*, 10, 582–593, https://doi.org/10.2136/vzj2010.0083, 2011.

Roth, K., Schulin, R., Flühler, H., and Attinger, W.: Calibration of time domain reflectometry for water content measurement using a composite dielectric approach, ~~Water Resources Research~~*Water Resources Research*, 26, 2267–2273, https://doi.org/10.1029/WR026i010p02267, 1990.

Rowlandson, T. L., Berg, A. A., Bullock, P. R., Ojo, E. R., McNairn, H., Wiseman, G., and Cosh, M. H.: Evaluation of several calibration procedures for a portable soil moisture sensor, *Journal of Hydrology*, 498, 335–344, https://doi.org/10.1016/j.jhydrol.2013.05.021, 2013.

Schwartz, R. C., Casanova, J. J., Pelletier, M. G., Evett, S. R., and Baumhardt, R. L.: Soil Permittivity Response to Bulk Electrical Conductivity for Selected Soil Water Sensors, *Vadose Zone Journal*, 12, 0, https://doi.org/10.2136/vzj2012.0133, 2013.

Soilmoisture Equipment inc.: Trase Operating Instructions, 1996.

Stacheder, M., Fundinger, R., and Koehler, K.: On-Site Measurement of Soil Water Content by a New Time Domain Reflectometry (TDR) Technique, in: Field Screening Europe, pp. 161–164, Springer, Dordrecht, Dordrecht, https://doi.org/10.1007/978-94-009-1473-5_37, 1997.

Topp, G., Davis, J., and Annan, A.: Electromagnetic determination of soil water content: Measurements in coaxial transmission lines, *Water Resources Research*, 16, 574–582, https://doi.org/10.1029/WR016i003p00574, 1980.

van der Ploeg, M. J., Gooren, H. P. A., Bakker, G., Hoogendam, C. W., Huiskes, C., Koopal, L. K., Kruidhof, H., and de Rooij, G. H.: Polymer tensiometers with ceramic cones: direct observations of matric pressures in drying soils, ~~Hydrology and Earth System Sciences~~*Hydrology and Earth System Sciences*, 14, 1787–1799, https://doi.org/10.5194/hess-14-1787-2010, 2010.

Van Genuchten, M. T.: ~~CLOSED-FORM EQUATION FOR PREDICTING THE HYDRAULIC CONDUCTIVITY OF UNSATURATED SOILS~~Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils., ~~Soil Science Society of America Journal~~*Soil Science Society of America Journal*, 44, 892–898, 1980.

Walker, J. P., Willgoose, G. R., and Kalma, J. D.: In situ measurement of soil moisture: a comparison of techniques, *Journal of Hydrology*, 293, 85–99, https://doi.org/10.1016/j.jhydrol.2004.01.008, 2004.

Wang, J. R. and Schmugge, T. J.: An Empirical Model for the Complex Dielectric Permittivity of Soils as a Function of Water Content, Geoscience and Remote Sensing, ~~IEEE Transactions on~~*IEEE Transactions*, GE-18, 288–295, https://doi.org/10.1109/TGRS.1980.350304, 1980.

Wild, J., Kopecký, M., Macek, M., Sanda, M., Jankovec, J., and Haase, T.: Climate at ecologically relevant scales: A new temperature and soil moisture logger for long-term microclimate measurement, *Agricultural and Forest Meteorology*, 268, 40–47, https://doi.org/10.1016/j.agrformet.2018.12.018, 2019.

Wraith, J. M. and Or, D.: Temperature effects on soil bulk dielectric permittivity measured by time domain reflectometry: Experimental evidence and hypothesis development, ~~Water Resources Research~~*Water Resources Research*, 35, 361–369, https://doi.org/10.1029/1998WR900006, 1999.