

Compare Results

Old File:

egusphere-2024-407-manuscript-version2.pdf

34 pages (1.80 MB)

01/07/2024 12:12:06

versus

New File:

22-07-final.pdf

33 pages (1.70 MB)

23/07/2024 06:43:05

Total Changes

170

Content

130 Replacements

17 Insertions

17 Deletions

Styling and Annotations

6 Styling

0 Annotations

[Go to First Change \(page 2\)](#)

Prediction of Hysteretic Matric Potential Dynamics Using Artificial Intelligence: Application of Autoencoder Neural Networks

Nedal Aqel^{1*}, Lea Reusser^{1,2,3}, Stephan Margreth², Andrea Carminati¹, and Peter Lehmann¹

¹Physics of Soils and Terrestrial Ecosystems, ETH Zurich, Switzerland

²Office for the Environment, canton of Solothurn, Switzerland

³now at the Swiss Academy of Sciences, Forum Landscape, Alps, Parks (FoLAP), Switzerland

*Corresponding author, nedal.aqel@usys.ethz.ch, Universitätstrasse 16, 8092 Zurich, Switzerland

Abstract

Information on soil water potential is essential to assess soil moisture state, to prevent soil compaction in weak soils, and to optimize crop management. In lack of direct measurements, the soil water potential values must be deduced from soil water content dynamics that can be monitored at plot scale or obtained at larger scale from remote sensing information. Because the relationship between water content and soil water potential in natural field soils is highly ambiguous, the prediction of soil water potential from water content data is a big challenge. The hysteretic relationship observed in nine soil profiles in the region of Solothurn (Switzerland) is not a simple function of texture or wetting and drainage cycles but depends on seasonal patterns that may be related to soil structural dynamics. Because the physical mechanisms governing seasonal hysteresis are unclear, we developed a deep neural network model that predicts water potential changes using rainfall, potential evapotranspiration, and water content time series as inputs. To adapt the model for multiple locations, we incorporated a Deep Autoencoder Neural Network as a classifier. The autoencoder compresses the water content time series into a site-specific feature that is highly representative of the underlying water content dynamics of each site and quantifies the similarity of dynamic patterns. By adding the Autoencoder's output as an additional input and training the neural network model with three stations located in three major classes founded by the autoencoder, we predict matric potential for other sites. This method has the potential to deduce the dynamics of matric potential from water content data (including satellite data) despite strong seasonal effects that cannot be captured by standard methods.

34 1. Introduction

35 The soil water characteristics curve SWC relates the matric potential (MP) and water content (WC) and
36 is the key physical property to quantify soil water dynamics (Tuller and Or, 2023). The SWC (also
37 denoted as soil water retention curve or pressure-saturation relationship) depends on both soil texture
38 and structure and differs with soil types and soil textural classes (Rawls et al., 2003; Shwetha and Varija,
39 2015). The SWC contains information on the pore size distribution and allows the assessment of flow
40 and transport properties for different hydration states (Menon et al., 2020; Rostami et al., 2015). To
41 provide a complete characterization of the actual soil moisture state and flow regimes, information on
42 both the matric potential and the water content must be specified. Information on volumetric water
43 content is needed to assess the free storage capacity, optimize water management, and to formulate
44 mass balance. The matric water potential is a component of the total and hydraulic soil water potential
45 and determines the water flow in direction of decreasing water potential to achieve equilibrium with its
46 surroundings (Ma et al., 2022). The matric potential is also of particular interest to assess mechanical
47 stability of a soil (Holthusen et al., 2010; Lu et al., 2010). The capillary and adsorptive forces expressed
48 with the matric potential define the unsaturated soil strength mitigating soil compaction by heavy
49 machinery in construction work, farming, and forestry (Smith et al., 2001). For example, matric
50 potential thresholds are defined in various regions of Switzerland to prevent mechanical damage and
51 regulate the maximum load linked to factors like soil type, texture, and vehicle impact (Bundesamt für
52 Energiewirtschaft, 1997). Other important potential thresholds are the wilting point and the field
53 capacity, characterizing the plant available water (Gupta et al., 2023).

54 It would be optimal to determine the soil moisture status relative to these potential thresholds based on
55 information of water content using the SWC, without direct measurement of the matric potential. In that
56 case, matric potential dynamics could be deduced from remote sensing water content data that are
57 available at various scales. However, the application of this procedure is limited by two effects. Firstly,
58 under saturated conditions, the water potential can change without modifying the volumetric water
59 content. The transition of conditions with negative water potential within the capillary fringe to positive
60 pressures below a water table is crucial for the triggering of landslides (Gallipoli et al., 2003). Secondly,

61 the SWC under field conditions is often an ambiguous relationship between potential and water content
62 due to hysteretic and dynamic effect as will be discussed next.

63 The SWC is typically measured in the lab as series of equilibrium states obtained during drainage, with
64 one water content value assigned to the applied pressure. The results of such small-scale experiments
65 are not sensitive to structural pores that can be found at the field scale (Romero-Ruiz et al., 2018) and
66 can thus be expressed as function of basic soil properties (texture, bulk density, content of organic
67 material) using pedotransfer functions (PTF; Zuo and He, 2021). Because these PTFs ignore the effects
68 of soil structures including macropores and cracks (Basile et al., 2019) and are trained with data from
69 small samples with artificially high initial saturation conditions, their applicability to model dynamic
70 processes in the field is limited. Another limitation is the underlying assumption of an unambiguous
71 relationship between water content and matric potential (and hydraulic conductivity). In all land surface
72 models, water content is linked by an unambiguous relationship between water content and matric
73 potential. In reality, this relationship is highly ambiguous under field conditions as was analyzed in
74 detail by Hannes et al. (2016) and as we will show later in this paper as well.

75 Hannes et al. (2016) analyzed long-term experiments and concluded that the high variation of matric
76 potential values for the same water content are a result of hysteresis, dynamic effects, and structural
77 changes during the season. Hysteresis is related to differences in wetting and drying cycles (Capparelli
78 and Spolverino, 2020) as controlled by different pore structures controlling air- or water invasion and
79 differences in receding or advancing wetting angles (Fomin et al., 2023). Hysteresis is often manifested
80 in coarse textured soils and occurs as well during slow processes. Another process resulting in an
81 ambiguous pressure-saturation relationship is dynamic effects with water contents that are not in
82 equilibrium with the quickly changing potential (Ross and Smettem, 2000). Finally, the size of
83 structural pores is not constant with time but changes with season, water content, and soil formation
84 processes (Fu et al., 2021). The combined effect of hysteresis, non-equilibrium, and structural changes
85 makes it extremely challenging to deduce soil matric potential from information on water content. Also,
86 the implementation of these combined effects in physically-based models of unsaturated water flow is
87 not straightforward. As an alternative approach to physically-based models, machine learning can be

88 applied to simulate the complex relationship between matric potential and water content under field
89 conditions. In this study, we will apply a deep neural network (DNN).

90 Deep neural networks (DNN) have demonstrated their effectiveness as a powerful numerical tool for
91 resolving complex patterns. Their ability to learn from data and recognize intricate relationships makes
92 them valuable in various fields, including the modeling of soil water characteristics. For example, Jain
93 et al. (2004) and Achieng (2019) used artificial neural network (ANN) models to predict the hysteretic
94 water content from observed matric potential values. However, both publications simulated lab data
95 under equilibrium conditions and cannot be applied for the more complex dynamic processes in the
96 field. In addition, the models were site-specific and needed both water content and matric potential
97 information for the training. Here we will apply a different DNN using an autoencoder approach. As
98 we will explain in the theory section, the autoencoder condenses the complexity of temporal (and
99 spatial) patterns into a single (or a few) number(s). The hypothesis of this study is that the autoencoder
100 value is a new and unique characterization of the soil moisture dynamics and can be used to predict
101 matric potential dynamics from observed water content data. The paper is organized as follows: in
102 section 2, the study sites and the basics of the deep neural network with the autoencoder approach are
103 presented. The results section compares the model performance of site-specific deep neural network
104 (DNN) and shows the possibility to build a generalized DNN using the autoencoder analysis as model
105 input. Limits and possible applications of the model approach are discussed in section 5.

106 2. Material and methods

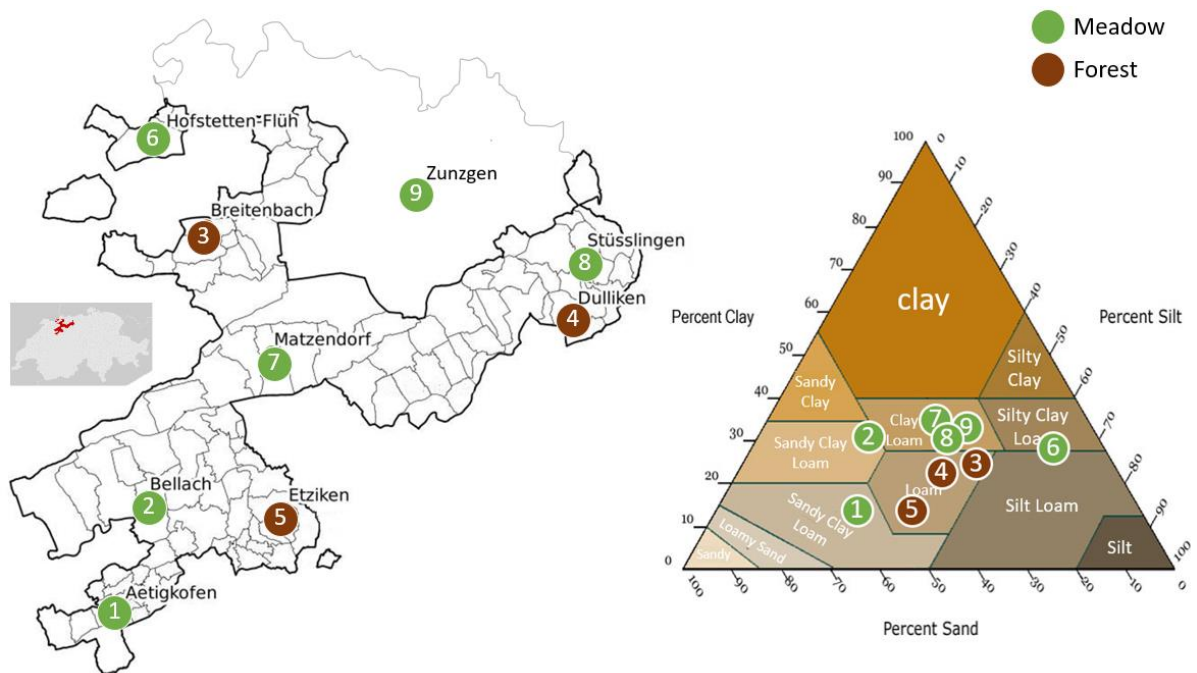
107 In a first step, matric potential time series were simulated at nine sites in the region of Solothurn
108 (Switzerland) using site specific ANN model, to proof that the ANN models can predict matric potential
109 from water content dynamics with site specific training. In the next step, the autoencoder analysis of
110 water content dynamics of all sites was conducted. Finally, the site-specific ANN model was enhanced
111 and transformed into a multisite model by combining two deep neural networks. This transformation
112 allowed for a more comprehensive and versatile predictive framework of matric potential as function
113 of water content.

114 2.1 Study area and soil moisture data

115 The study area covers mainly the canton of Solothurn in Switzerland (Fig.1), and thus an area of
116 approximately 629 km². The climate in Solothurn is classified as oceanic climate (Cfb) according to
117 Koppen and Geiger climate classification, with an average yearly temperature of 9.5 °C and annual
118 precipitation of around 1400 mm. Approximately half of the annual precipitation in the canton
119 undergoes the process of evaporation (Spreafi and Weingartner, 2005). During the year, the average
120 temperature varies by 19 °C with the highest temperature occurring in the month of July and the lowest
121 average temperature in January. Regarding precipitation patterns, the month of June has the highest
122 level of precipitation, while March stands out as the driest month. Soil moisture dynamics (see below)
123 were studied for the period from 2011 to 2022. For this period, climatic data was available on the data
124 portal of MeteoSwiss (IDAWEB, 2024) . The data was gathered from the closest meteorological stations
125 to each of the nine sites in the Solothurn region.

126 Soil moisture data were downloaded from the ‘soil monitoring network’ (BODENMESSNETZ, 2024)
127 collecting data from 65 stations distributed over eleven cantons of Switzerland. The network’s primary
128 objective is to provide real-time soil moisture information for mitigating soil compaction.
129 BODENMESSNETZ also plays a role in raising awareness among farmers and foresters about soil
130 compaction, providing a tool to assess the current situation and adjust the use of heavy machinery based
131 on weather conditions. As the network has been running since 2011, it now serves as a valuable resource
132 by offering long-term diverse information, including land use, precipitation amounts, and matric
133 potential measured at various depths (20 and 35 cm depth in most of the stations, using T8 and T32
134 tensiometers from METER group). Only at nine sites that are located in the region of Solothurn, the
135 water content was measured at 20 cm depth (Stevens Hydra Probe). For these nine sites, daily values in
136 volumetric water content (20 cm), matric potential (20 cm) and precipitation values were used. The
137 matric potential in the downloaded data was given in kPa and was transferred to matric potential head
138 with units of cm (1 cm is 0.1 kPa), considering a water density of 1000 kg m⁻³ and gravity acceleration
139 of 10 m s⁻².

140 As the soil moisture decreases, water is drawn from the tensiometer, creating a negative pressure or
 141 tension. During dry periods, cavitation may occur, causing water vaporization and air bubble formation
 142 (Mendes and Buzzi, 2013), or tensiometers had to be refilled (Sadeghi et al., 2018). To address these
 143 challenges and ensure accurate data collection, various data preprocessing and filtering techniques were
 144 implemented. These techniques involved identifying and removing outliers, systematically excluding
 145 data points with water potential values within the problematic dry ranges and filtering out data points
 146 with extremely low or high water content values. The study also flagged abrupt changes in volumetric
 147 water content (VWC) and matric potential (MP) for further investigation, as these could indicate
 148 measurement anomalies. Additionally, a thorough analysis of weekly trends in the data was conducted
 149 to identify systematic variations over time (see Appendix A).

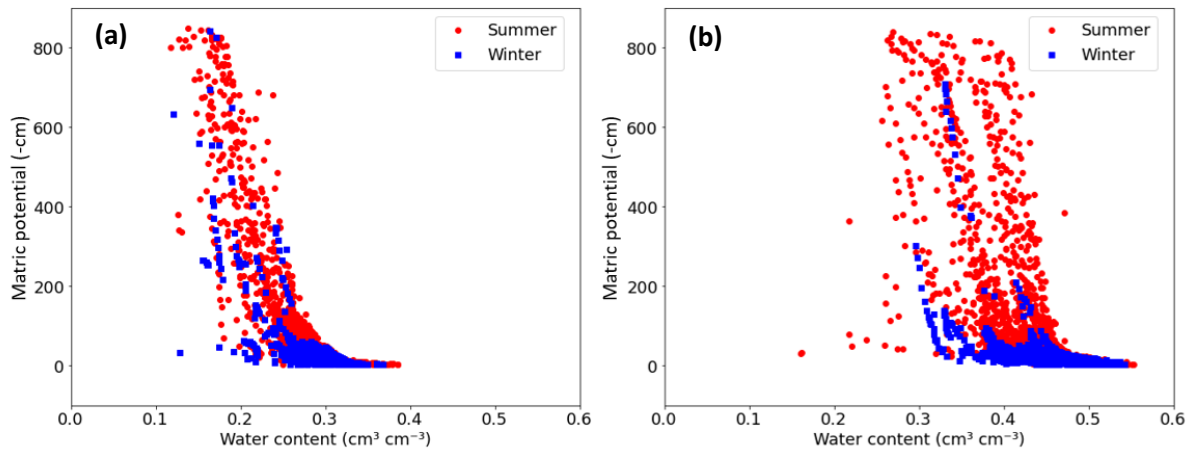


150
 151 **Figure 1** Overview of the study area with site locations, soil texture, and land cover. The primary focus
 152 is on the canton of Solothurn, outlined by the black border on the map, with an additional site from the
 153 canton of Basel (site 9, Zunzgen). Within this region, three sites are categorized as forests, while the
 154 remaining six sites are designated as meadows. The analyzed soil horizons (20 cm depth) of the study
 155 area encompasses five soil textural classes as shown in the soil texture triangle.

156 The analyzed soil horizons of the selected locations can be assigned to five different soil textural classes
 157 (figure 1) and two different land covers (meadow and forest). The location denoted as Matzendorf (site
 158 #7) contains the highest clay content, whereas locations such as Aetigkofen (site #1) are predominantly
 159 sandy. Across these nine locations, different relationships between matric potential and water content

160 were deduced from field data as shown in Figure 2 for two sites with low and high variations in water
 161 content for similar potential values. To show the relevance of seasonal patterns, we differentiate
 162 between summer (April to September) and winter period (remaining months).

163



164

165 **Figure 2** Soil-Water characteristics curve (SWC) measured in the field at two sites classified into
 166 summer (April to September) and winter period (remaining months) from 2012 to 2023. **(a)** The Etziken
 167 site (site #5) shows small changes in the SWC dynamics over the years, for both the warm and cold
 168 period. **(b)** A contrasting scenario was found for the site in Bellach (site #2) that was characterized by
 169 a wide range of water content for similar potential values. The unit of matric potential, represented as -
 170 cm, is equivalent to -0.1 kPa.

171 2.2 Deep neural network (DNN)

172 A basic artificial neural network (ANN) comprises one or two hidden interconnected layers, with each
 173 layer tasked with the conversion of an input vector (\mathbf{x}) into a hidden state vector (\mathbf{h}), as described by
 174 (Bertels and Willems, 2023). This conversion is accomplished with eq. (1):

$$175 \mathbf{h} = f(\mathbf{x}) = \text{act}(\mathbf{W} \cdot \mathbf{x} + \mathbf{b}) \quad (1)$$

176 Where $f(\mathbf{x})$ represents the transformation function applied to the input vector(\mathbf{x}), with a weight matrix
 177 (\mathbf{W}) and a bias vector (\mathbf{b}), integrated with an activation function (denoted as "act").

178 To construct a deep neural network (DNN), multiple layers (more than two hidden layers) are
 179 interconnected to form a 'multilayer perceptron.' The training process involves finding optimal values
 180 for the weights and biases in the network using suitable optimization techniques (Bertels and Willems,
 181 2023). In this study, DNN was built to predict the daily MP for the nine sites. The process involved
 182 several key steps. First, in the design of the neural network, activation functions were carefully selected

183 and integrated to introduce non-linearity into the model's transformations (Montesinos López et al.,
184 2022). The Rectified Linear Unit (ReLU) activation function was employed to mitigate vanishing
185 gradient problem and enhance the model's ability to handle noisy input. The inclusion of ReLU was
186 motivated by considerations of computational efficiency, with some attention given to the potential
187 issue of "dying ReLU" (Lu, 2020; Montesinos López et al., 2022).

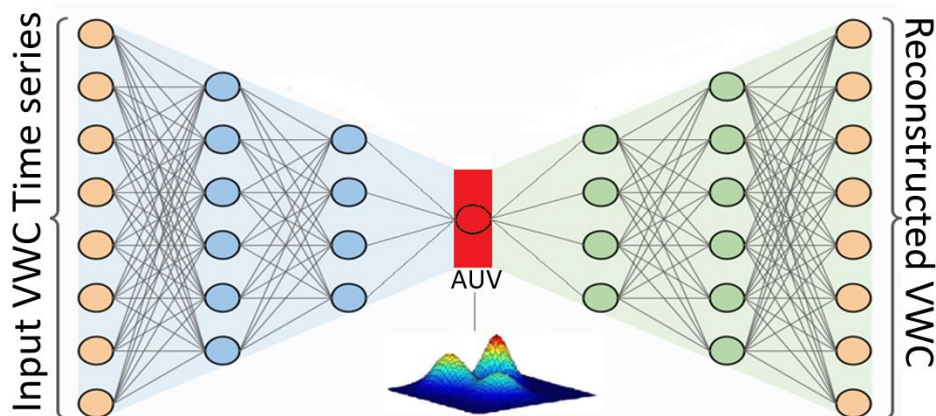
188 Next, the neural network was structured with a total of six layers, including four hidden layers as
189 suggested by Achieng (2019). All layers were densely connected, fostering strong information flow
190 between neurons. Crucially, batch normalization was incorporated after the second hidden layer. Batch
191 normalization is a technique that normalizes the activations within a layer during training, which can
192 help mitigate issues like internal covariate shift and accelerate convergence (Ioffe and Szegedy, 2015).
193 The choice of the optimization method was the Adam optimizer, a powerful tool for training neural
194 networks. It adaptively adjusted learning rates, thereby optimizing the learning process, and enabling
195 rapid convergence while employing Mean Squared Error (MSE) as the loss function (Kingma and Ba,
196 2014). To prevent overfitting by the Adam optimizer, an early stopping mechanism was implemented.
197 This mechanism continuously monitored the loss function for the hold out data during training, ceasing
198 the process if no improvement or a sudden increase was detected over a predetermined number of
199 consecutive epochs.

200 The initial deep neural networks (DNN) were configured with 4 input parameters and the daily
201 logarithmic scaled matric potential (MP) value as output. The input parameters consisted of
202 precipitation, potential evapotranspiration, measured VWC, and the weekly percentage change in
203 VWC. As the prediction process progressed, two major issues were identified. Firstly, the influence of
204 the VWC measurements on the training process was found to be predominant. Consequently, a decision
205 was made to increase the weight of precipitation and potential evapotranspiration in the calculation
206 process by incorporating three new input parameters: the weekly total precipitation and
207 evapotranspiration (the sum of the current day and the preceding six days), along with the difference
208 between these two new components. Secondly, the use of logarithmic scaled MP values was found to
209 be highly sensitive to data availability. Therefore, a decision was made to retrain the model using

210 absolute linear MP values (see Appendix B). In total, the final model was equipped with 7 input
211 parameters to predict the absolute linear MP values for a given location. For each site, a site-specific
212 DNN was built. The extent of the training data is predominantly influenced by site-specific
213 characteristics. For instance, sites characterized by sandy soils necessitated a shorter training duration
214 in contrast to sites with a higher clay content. Typically, the training dataset spanned a duration of 4 to
215 7 years. During this period, 70% of the data were randomly selected for training, while the remaining
216 30% were set aside as holdout data (Gholamy et al., 2018). The extra years of data beyond the initial
217 training period were reserved for validation purposes.

218 2.3 Autoencoder neural network

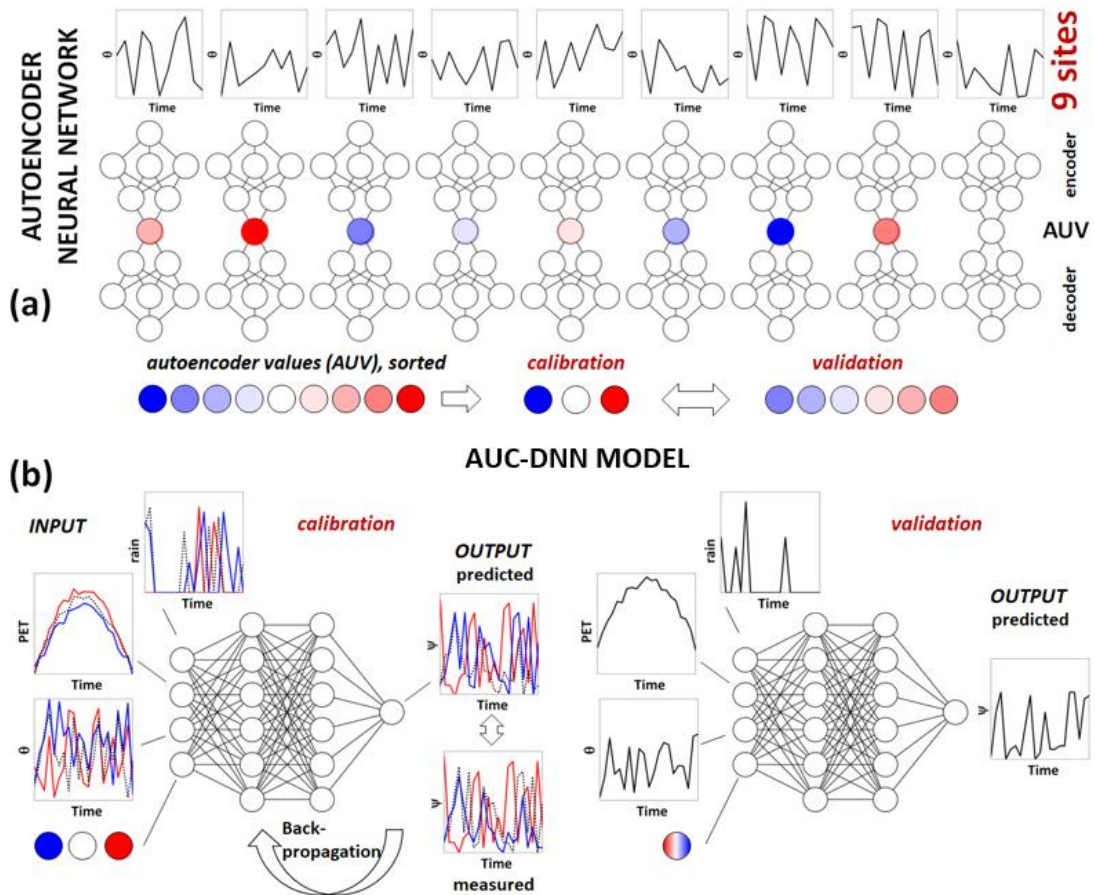
219 The autoencoder, consisting of an encoder and a decoder, is an unsupervised deep neural network that
220 learns how to efficiently compress input data into a meaningful representation and subsequently
221 reconstruct the original data from this compressed form (Chen and Guo, 2023). By connecting the
222 encoder and decoder, the autoencoder effectively captures important patterns and variations present in
223 the data, enabling comprehensive analysis and interpretation (Chen and Guo, 2023). In this study, an
224 autoencoder neural network (figure 3) was built to analyze the measured VWC time series at 20 cm
225 depth for the 9 sites.



226
227 **Figure 3** Autoencoder deep neural network for volumetric water content dynamic analysis. In this
228 illustration, a densely connected autoencoder is utilized to compress the dynamic information of
229 Volumetric Water Content (VWC) into a singular value, AUV, highlighted in red. The process begins
230 with the encoder, depicted in blue, extracting the AUV from the measured volumetric water content
231 time series (left orange layer). Subsequently, the densely connected decoder, represented in green,
232 utilizes the AUV to reconstruct the VWC (orange layer at the right). Both the encoder and decoder,
233 characterized by dense connections, optimized the AUV value by minimizing the error between the
234 measured VWC and the reconstructed VWC.

235 The process was as follows. Firstly, an encoder neural network was created for each site. Its objective
236 was to take the VWC time series as input and gradually reduce its dimensionality through hidden layers
237 (Chen and Guo, 2023). The encoders' output was a single site-specific latent representation, called
238 Autoencoder Value (AUV), and captures essential features of the VWC dynamics (Chen and Guo,
239 2023). Subsequently, a decoder neural network was developed to utilize the AUV value as reference to
240 reconstruct the original VWC time series data. The success of this reconstruction depends on the
241 training process, which aimed to optimize the AUV value by minimizing the error between the original
242 VWC time series and its reconstructed counterpart by minimizing the mean squared error (MSE) value
243 to less than 0.1.

244 After the optimization process, for each site one autoencoder value (AUV) was obtained. These AUV
245 were scaled and then used to build a combined model (Figure 4) as follows. The AUV were sorted into
246 three categories. Subsequently, one site from each category was selected. Finally, the data from the
247 three chosen sites, each representing one category, were used to train the combined AUC-DNN model.
248 The final combined model was thus equipped with 8 input parameters to predict the dynamic MP for a
249 specific location. These parameters consisted of the same 7 inputs employed in the DNN model (section
250 2.2), complemented by the AUV. The neural network structure, as detailed in section 2.2, remained
251 unchanged, employing the same optimization techniques.



252

253 **Figure 4** Application of two different types of deep neural network for the prediction of matric potential
 254 ψ . In this conceptual example, the water moisture dynamics of nine sites is considered. (a) The
 255 autoencoder neural network captures the characteristic features of the soil water content (θ) dynamics,
 256 assigning an autoencoder value (AUV) to each site. These values are sorted to AUV classes (one site
 257 from each class was used for calibration, remaining sites for validation). (b) The combined AUC-DNN
 258 model is built using the calibration sites with rainfall, potential evapotranspiration (PET), water content,
 259 and AUV as part of the 8-input parameters. The predicted matric potential (ψ) is compared to measured
 260 values for backpropagation. The calibrated DNN is then used to predict ψ for the remaining sites.

261 Initially, 70% of the data from each of the training sites were randomly selected for the training dataset.

262 Subsequently, the remaining 30% of the data were set aside as the holdout dataset, serving as a

263 benchmark for assessing model performance. The developed AUC-DNN was then applied for the other

264 six sites (with the same input variables including AUV) to predict the entire datasets of those unseen

265 sites. The combined model has thus the strengths of both components—the DNN' ability to understand

266 dynamic MP patterns and the feature extraction capabilities of the autoencoder. This shift in the model's

267 strength extends it from being site-specific to encompassing multiple sites, enabling it to gain a broader

268 understanding of how the dynamic MP and AUV values relate.

269 2.4 Statistical evaluation

270 The evaluation of model performance is carried out by comparing the model predictions to the measured
271 data. While there is no universal consensus on a standardized evaluation procedure, it is widely
272 recognized that a multi-objective approach should be adopted e.g., (Boyle et al., 2000; Willems, 2009).
273 In this study, a combination of four evaluations tools was adopted. First, a scatter plot of observations
274 against simulated values was utilized to visualize the degree of alignment with the identity line (often
275 referred to as the 1:1 line). This graphical approach allowed for a qualitative assessment of model
276 performance. A closer concentration of data points near the 1:1 line indicated higher agreement between
277 calculated and observed values. Moreover, this graphical method includes the 95 % confidence interval
278 area which help in scrutinizing the model's consistency across different prediction ranges and detecting
279 potential biases within the model's performance (Ritter and Muñoz-Carpena, 2013). The second criterion
280 evaluates the distribution of (signed) prediction errors (eq(2)). Ideally, the error distribution should be
281 centered around zero, following a normal distribution pattern around this point with low standard
282 deviation. Such a distribution indicates an unbiased model with errors that tend to balance out.
283 Deviations from this pattern may suggest model bias or other unexpected characteristics in the
284 prediction errors PE (den Ouden et al., 2012).

$$285 \quad PE = O_i - P_i \quad (2)$$

286 with observed O_i and predicted matrix potential value P_i . The third evaluation metric was the root means
287 squared error (RMSE; eq (3a)). RMSE with a value of zero indicates perfect fit, while higher RMSE
288 value means worse model performance (Ritter and Muñoz-Carpena, 2013). The final criterion for model
289 evaluation involved the use of the dimensionless goodness-of-fit indicator (eq (3b)), known as the
290 coefficient of efficiency (NSE; Nash and Sutcliffe, 1970). NSE, which ranges from negative infinity to
291 1, serves as an indicator of model performance, with a value of 1 indicating a perfect fit, while a negative
292 NSE suggests that using the means of the observed values is a better representative for the data than the
293 evaluated model itself (Gupta and Kling, 2011; Ritter and Muñoz-Carpena, 2013). A NSE value > 0.75
294 indicates a very good model, while an NSE value < 0.5 signifies unsatisfactory results (Moriasi et al.,

295 2007). In Gupta et al. (1999) a threshold NSE-value of 0.80 was used for good model performance and
296 is applied here as well. The RMSE and NSE are defined by:

$$297 \quad RMSE = \sqrt{\frac{\sum(O_i - P_i)^2}{N}} \quad (3a)$$

$$298 \quad NSE = 1 - \frac{\sum(O_i - P_i)^2}{\sum(O_i - \bar{o})^2} \quad (3b)$$

299 where O_i represents the measured value, P_i the simulation output, and \bar{o} the mean of the observed values,
300 all within the context of a sample size N .

301 3. Results

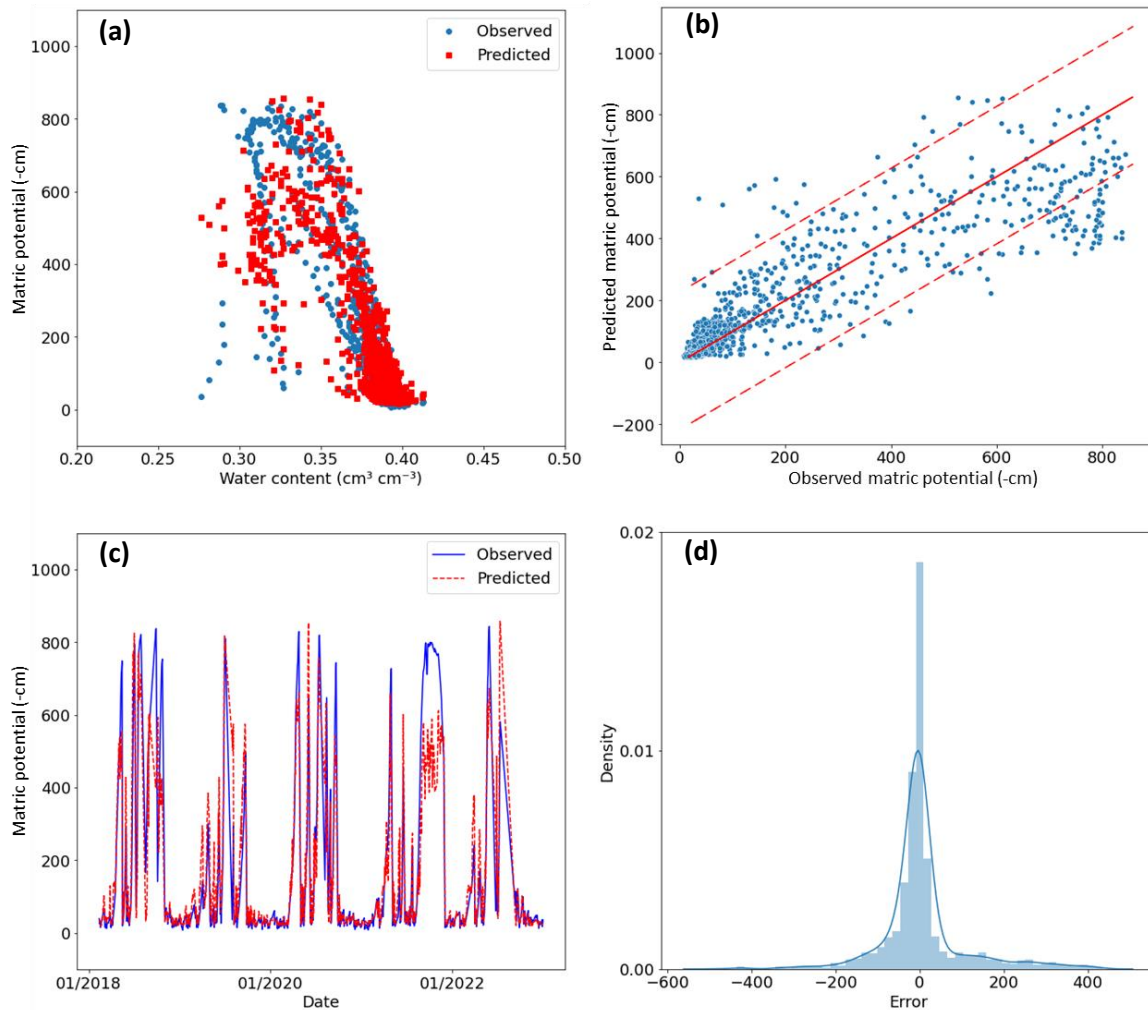
302 Following the model discussion in section 2.2 and 2.3, we present first the results of the site-specific
303 tests of predicting matric potential dynamics with a deep neural network (water content, rainfall and
304 evapotranspiration as input data), before the role of autoencoder value is considered.

305 3.1 Deep neural network modeling without autoencoder

306 The site-specific DNN model was used to simulate the time series for all nine sites. In Figure 5, the
307 results are shown for the Stüsslingen site (size #8, clay loam, meadow). The model was trained on data
308 that had 1825 days of observations from January 2012 to January 2020. The data was split randomly
309 into two parts: 1) a calibration dataset that had 1277 days and 2) a holdout dataset that had 548 days.
310 The model was then validated on data from February 2018 to January 2023 (1379 days). A strong
311 agreement between the model and the observed data was discovered in both the training and validation
312 datasets (figure 5c) as reflected by the low RMSE value and the high NSE value (table 1). Furthermore,
313 it was noticed that the error distribution exhibited a predominantly normal pattern with minimal bias
314 towards higher observed values compared to the predicted values (figure 5d). These findings suggest
315 that the site-specific DNN-model was not only able to be generalized well to unseen data but also
316 demonstrated a reliable ability to predict MP.

317 The statistical evaluation (Table 1) reveals a consistent performance across both the training and
318 validation periods for the Stüsslingen site, offering compelling evidence that the model avoids
319 overfitting. Additionally, when it comes to predicting MP values, the 95 % confidence interval indicates

320 that the model can capture well the overall dynamics (Figure 5b). However, the model performance
 321 exhibits higher deviations for values exceeding 400 cm and consistently underestimates values higher
 322 than 600 cm (figure 5b), which could explain the mild positive skewness observed in the distribution
 323 of prediction errors in figure 5d.



324
 325 **Figure 5** Graphical evaluation of the performance of the site-specific deep neural network (DNN) for
 326 validation for the Stüsslingen site (site #8) for the validation period 2018 to 2022. **(a)** Comparison
 327 between the simulated and measured soil water characteristics curve. **(b)** Scatter plot comparing
 328 simulated and measured matric potential values, providing a visual representation of the level of
 329 conformity to the identity line. The two dashed lines represent the 95% confidence interval around the
 330 identity line, providing a visual assessment of the level of agreement. **(c)** Model validation presenting
 331 time series with the observed and predicted matric potential. **(d)** Analysis of the distribution of
 332 prediction errors (observed minus predicted values) with positively mild skewed distribution

333 Comparing the performance for the ‘holdout’ period (randomly chosen days between 2012 and 2019)
 334 of the nine site-specific DNN models, the NSE index is larger than 0.55 (‘good’) for all and larger than
 335 0.80 (‘optimal’) for six sites. For all sites it was thus possible to build a DNN model with good model

336 performance for the randomly chosen test days. However, for the validation period, only four showed
 337 optimal performance (NSE > 0.80). For two forest sites with an optimal performance for the holdout
 338 period (Dulliken, site #4, and Etziken, site #5), the NSE dropped from a range between 0.82 and 0.88
 339 to a range between 0.73 and 0.75 (table 1). Obviously, the model captured the overall short term
 340 dynamics during training (randomly chosen days) but faced problems in the precise prediction of the
 341 long validation period. An extended training period may be necessary to enhance the model's accuracy
 342 for these specific sites. Three grassland sites (Bellach, site #2, Matzendorf, #6, and Hofstetten-Flüh, #5)
 343 showed good but not optimal performance already during the holdout period. As discussed in the next
 344 section, this may be related to large variations of the pressure values for similar water contents and the
 345 corresponding large AUV. Notably, the lower performance observed in the holdout period for
 346 Hofstetten-Flüh could be also linked to data limitations, as only 1200 days were used to train the model
 347 for this specific site (compared to 1825 sites for the other sites).

348 **Table 1** Statistical assessment of calibration (1825 days, until year 2019/2020) and validation results
 349 (years 2018/2019/2020 until years 2020/2021/2022) for nine sites. The holdout dataset was part of the
 350 training period and includes 548 days (30 % of calibration).

| Location | AUV (-) | Training (holdout) | | Validation | |
|--------------------------------|---------|--------------------|------------|------------|------------|
| | | NSE (-) | RMSE (-cm) | NSE (-) | RMSE (-cm) |
| 1 Aetigkofen | 1.95 | 0.92 | 48 | 0.89 | 60 |
| 2 Bellach | 7.00 | 0.70 | 98 | 0.62 | 125 |
| 3 Breitenbach ^{a, b} | 3.56 | 0.86 | 82 | 0.83 | 96 |
| 4 Dulliken ^a | 2.19 | 0.82 | 55 | 0.73 | 103 |
| 5 Etziken ^a | 1.90 | 0.88 | 56 | 0.75 | 70 |
| 6 Hofstetten-Flüh ^b | 5.59 | 0.76 | 90 | 0.63 | 123 |
| 7 Matzendorf | 6.39 | 0.76 | 83 | 0.59 | 133 |
| 8 Stüsslingen | 4.49 | 0.80 | 71 | 0.80 | 98 |
| 9 Zunzgen | 6.44 | 0.87 | 62 | 0.83 | 73 |

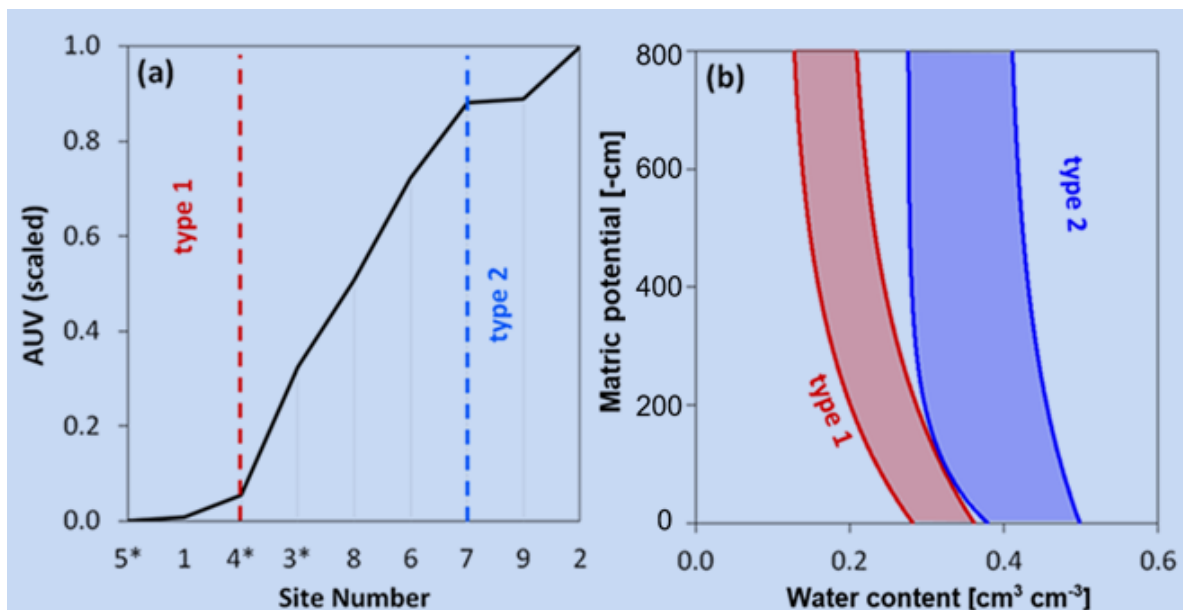
351 ^a forest sites.

352 ^b Sites with limited available data. For those sites, only 1200 days were used for training; Within this training period, a subset
 353 of 360 randomly selected days was designated as a holdout dataset; the validation period for those specific sites was from
 354 2018/2019 to 2022.

355 3.2 Autoencoder DNN

356 The Autoencoder values (AUV) deduced from the time series analysis of the volumetric water content
 357 for the period 2012-2022 can be classified in three main groups (figure 6). Soil water characteristics
 358 curves (SWC) with low water content at saturated conditions and a small variation of water content for

359 similar potential values are assigned to ‘type 1’, contrasting ‘type 2’ with large water content values
 360 and variations. These types of SWC are related to small (‘type 1’) and high (‘type 2’) autoencoder
 361 values (AUV). Sites with AUV between these two classes, are denoted in the following as ‘transitional’
 362 type. As shown in Table 1, the AUV of forest soils are small (mainly ‘type 1’) with large NSE values.
 363 In contrast to the forest soils, there are grassland sites with high AUV (‘type 2’) but small NSE.
 364 Probably, the high variations of the SWC curve for ‘type 2’ require longer training periods to capture
 365 the high variations in the pressure-saturation relationship.



366

367 **Figure 6** Autoencoder value (AUV) and its relation to the soil water characteristics curve (SWC). (a)
 368 The AUV of the nine sites with three sites of small (type 1) and three sites of high (type 2) AUV. (b)
 369 The type 1 of the SWC has small water contents close to saturation and a narrower range of water
 370 contents for similar water contents compared to type 2 with high water content values and variations.
 371 Type 1 shows the data range of Aetigkofen (site #1) and Type 2 for Bellach site (#2). The site numbers
 372 are chosen in alphabetic order and as shown in Figure 1 (Aetigkofen (1), Bellach (2), Breitenbach (3),
 373 Dulliken (4), Etziken (5), Hofstetten-Flüh (6), Matzendorf (7), Stüsslingen (8), Zunzgen (9); sites with
 374 forest are marked with *).

375 3.3 Deep neural network using the autoencoder value (AUC-DNN)

376 As mentioned in the previous section, the nine sites could be grouped into three main types according
 377 to the scaled autoencoder value (AUV). Consequently, it was assumed that the creation of a DNN
 378 model, which incorporates AUV in conjunction with the previously built site-specific neural network,
 379 could enable predictions for unseen sites. Ideally, the model should be trained with a balanced dataset,
 380 including one site from the ‘type 1’ category, one site from the ‘type 2’ category, and a few sites from
 381 the ‘transitional’ category to capture the full transition between the ‘type 1’ and ‘type 2’. However, due

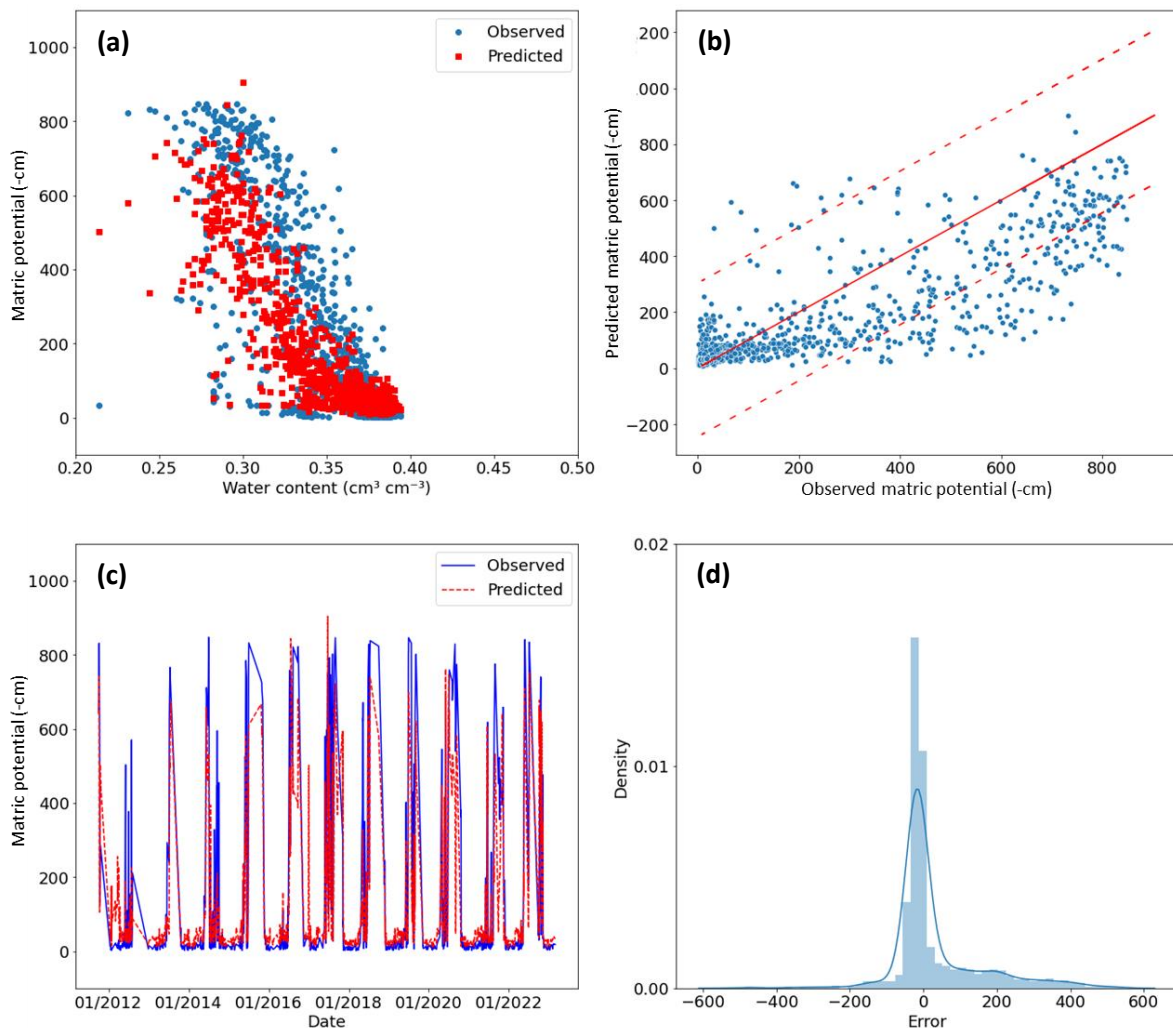
382 to the data limitation, the model was trained for only three sites representing the three types (Etziken,
383 site #5, for ‘type 1’; Bellach, #2, for ‘type 2’; Stüsslingen, #8, for the ‘transitional type’) and was then
384 used to predict the six unseen sites. The impact of the small training set (only one site for transitional
385 type) was clear in the model results, which exhibited some instability, changing from one run to another
386 as the model was not able to assume the same transitional function between sites consistently. Therefore,
387 the model was run 20 times, then the average result for these runs was taken as a representative
388 outcome. The application of the new DNN model with AUV to predict the dynamic of matrix potential
389 is shown in Figure 7 for Breitenbach (site #3, loam, forest) as unseen site. The model was found to fall
390 slightly behind the previously designed DNN model, but still can predict the dynamic in a good way.
391 Notably, the NSE value for this model for Breitenbach site was 0.71 over the entire period from 2012
392 to 2022 (Table 2).

393 **Table 2** AUC-DNN Model performance for the period 2012-2022. Three training sites were used to
394 build the AUC-DNN model that was then applied for the other six sites. The sites are listed according
395 to the corresponding autoencoder value (AUV). The asterisks mark the sites with forest; The AUV
396 was scaled from 1.9 to 7.0 to simplify input. Alternatively, scaled values ranging from 0 to 1 could
397 also be utilized.

| Location | AUV | AUV (type) | used as | NSE (-) | RMSE |
|-------------------|------|--------------|-----------------|---------|------|
| 5 Etziken* | 1.90 | Type 1 | Training site | 0.82 | 70 |
| 1 Aetigkofen | 1.95 | Type 1 | Validating site | 0.76 | 88 |
| 4 Dulliken* | 2.19 | Type 1 | Validating site | 0.65 | 100 |
| 3 Breitenbach* | 3.56 | Transitional | Validating site | 0.71 | 73 |
| 8 Stüsslingen | 4.49 | Transitional | Training site | 0.85 | 116 |
| 6 Hofstetten-Flüh | 5.59 | Transitional | Validating site | 0.60 | 113 |
| 7 Matzendorf | 6.39 | Type 2 | Validating site | 0.58 | 123 |
| 9 Zunzgen | 6.44 | Type 2 | Validating site | 0.69 | 104 |
| 2 Bellach | 7.00 | Type 2 | Training site | 0.71 | 104 |

398
399 It was noticed that the error distribution exhibited a predominantly normal pattern with a bias towards
400 higher observed values compared to the predicted values (figure 7d). The analysis indicates the model's
401 proficiency in forecasting dynamic trends rather than precise values (figure 7c). The results align with
402 the anticipated scenario as the AUV for Breitenbach (3.56) was relatively close the Stüsslingen AUV
403 value (4.49). Therefore, the underestimation detected in Stüsslingen for the site-specific DNN (figure
404 5b) is expected to exist in Breitenbach as well. The average model performance for all sites is presented

405 in Table 2. The NSE values was > 0.55 for the 6 unseen sites (validating sites) and provided strong
 406 evidence that the model can be relied upon for the dynamic MP predictions.

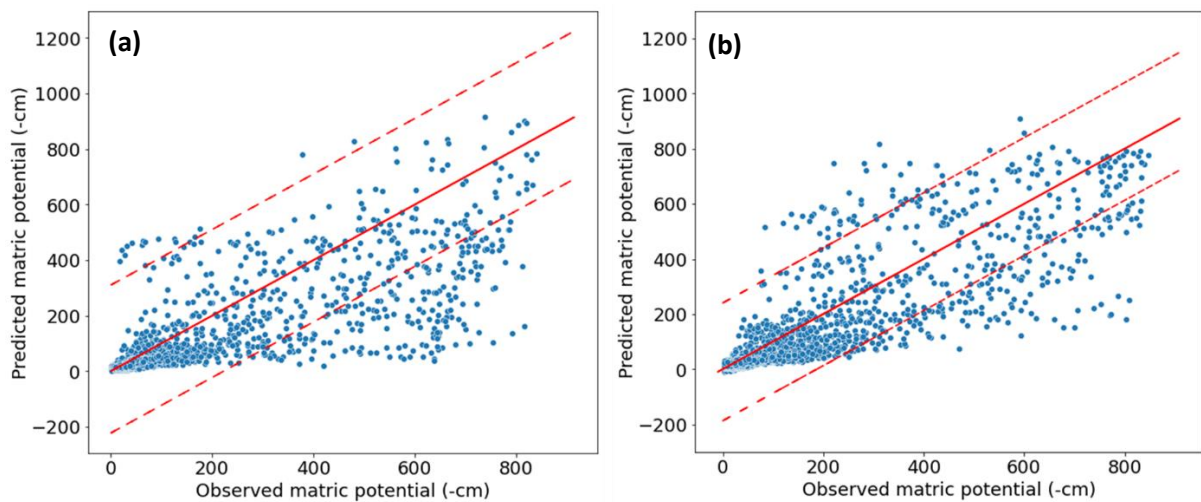


407

408 **Figure 7** Evaluation of the Deep Neural Network with Autoencoder (AUC-DNN) model performance
 409 at the Breitenbach site for the period 2012-2022. **(a)** Comparison between the expected Soil Water
 410 characteristics curve (SWC) and the observed SWC. **(b)** Scatter plot that compares observed data points
 411 with their corresponding simulated values, providing a visual representation of the level of conformity
 412 to the identity line. The two dashed lines represent the 95% confidence interval around the identity line,
 413 providing a visual assessment of the level of agreement. **(c)** Time series comparison showing the
 414 observed and predicted matric potential for the entire period. **(d)** Analysis of the distribution of
 415 prediction errors (observed minus modelled value) using positively mild skewed distribution.

416 The NSE values for the unseen sites (validating sites) varied from 0.58 to 0.76, indicating a spectrum
 417 of model performance, ranging from acceptable to good. The low NSE values observed for Matzendorf
 418 (site #7) suggest that the model's utility is more suited for capturing overall trends and dynamics rather
 419 than precise values. This evaluation was further supported by examining a scatter plot (Figure 8) that
 420 compares the observed data points with their corresponding simulated values for the sites scored the

421 lowest and the highest NSE, Matzendorf (site #7) and Aetigkofen (site #1). The plot revealed a wider
422 95% confidence interval for Matzendorf (figure 8a) in comparison to Aetigkofen (figure 8b), indicating
423 that the lower the NSE value is, the more challenging it became for the model to predict the exact MP
424 values. However, the model performance indicated the ability of the AUC-DNN model to predict
425 dynamic MP without the necessity of site-specific training data, marking a transition from the DNN
426 site-specific nature to a more versatile multi-site model.



427
428 **Figure 8** Comparison between observed data points and their corresponding simulated values for two
429 sites with lowest and highest efficiency coefficient NSE. **(a)** Matzendorf (site #7) with NSE of 0.58. **(b)**
430 Aetigkofen (site #1) with NSE of 0.76. The solid lines mark the 1:1 correspondence, the dashed lines
431 the 95% confidence interval.

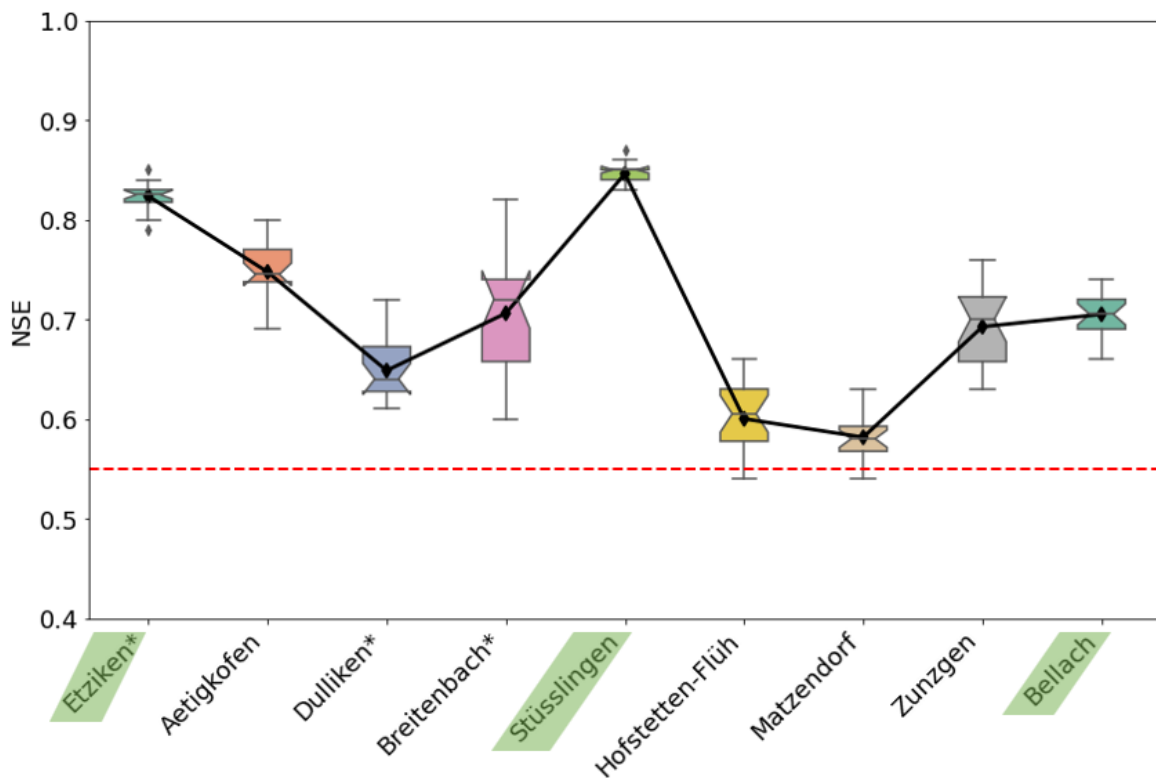
432 4. Discussion

433 Based on the analysis of the simulation results presented in section three, it can be asserted that the
434 model was successfully built. However, as discussed in the next subsection, the model is expected to
435 have certain drawbacks due to the limited number of available sites. In the other subsections, the
436 relationship between the autoencoder value and soil properties and its application for satellite data will
437 be discussed.

438 4.1 Limits of the deep neural network with autoencoder value (AUC-CNN)

439 First, the model's statistical evaluation revealed that the matric potential (MP) at a depth of 20 cm could
440 be simulated with acceptable precision. However, a high variability in the evaluation is indicated by the
441 NSE values for the unseen sites. This variance is attributed to the model's limited generalization

442 capacity, as it was trained on just three sites. Furthermore, the model was not able to catch the whole
 443 dynamic for the training sites due to the limited length of available data. For example, Bellach (site #2),
 444 a training site that has a high AUV, had NSE value of 0.71 for the training period (table 2), which
 445 indicates that the model was able to catch the general trend for this site, but still can't predict the exact
 446 value of the MP. The effect of this result was obvious on the sites that are closed to AUV 'type two'
 447 category (e.g., Hofstetten-Flüh and Matzendorf, sites #6 and #7, with NSE of 0.60 and 0.58,
 448 respectively).
 449 The stability of the AUC-DNN model was insufficient, as the model showed different prediction quality
 450 upon running the model repeatedly for the same training sites (figure 9). This variability in the outcomes
 451 indicates that the model can find different MP dynamics scenarios inside the training data. Therefore,
 452 it is recommended to train the model for more than one site in the same AUV type.



453
 454 **Figure 9** Variation of prediction results for 20 Runs for the AUC-DNN model quantified with the
 455 efficiency coefficient NSE. The highest variation was with the unseen sites in the transitional and type
 456 2 categories. Each box represents the interquartile range, with the line inside denoting the median. The
 457 black diamond markers connect the mean values for each station, providing insight into the central
 458 tendency of the data. Notches on the boxplots offer a visual indication of the uncertainty around the
 459 median. The red dashed line represents the defined threshold for the NSE, set at 0.55 ; sites with forest
 460 are marked with *; training sites are highlighted in green.

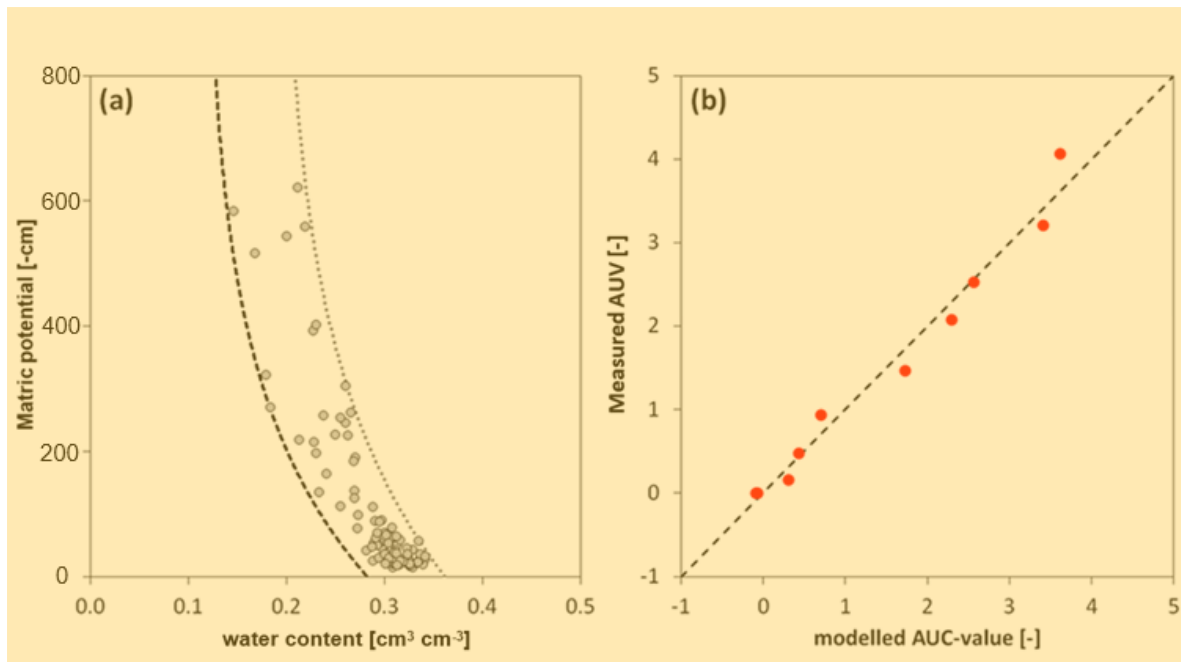
461 Especially for the ‘transitional type’, choosing a site in the beginning, in between, and in the end of the
462 category would stabilize the modeling results. However, in this study, there was no possibility to
463 provide the model with extra data to solve the prediction instability. Therefore, a solution was
464 implemented by 1) closely monitoring the model manually to ensure it captures the dynamic from all
465 three sites. This involved training the model with nearly identical time periods for each site and visually
466 confirming comprehensive coverage of the cloud of points for the retention curve of each site, avoiding
467 concentration on specific patterns during training. The process also includes 2) running the model for
468 20 times, then averaging the results. Additionally, the statistical evaluation plots as shown in Figure 8,
469 were used to detect instances with very low or very high MP prediction values.

470 For the set of sites analysed in this study, the model showed good generalization capacity and stability.
471 However, the nine sites were similar with respect to climate and geology and the range of soil textural
472 classes (see Figure 1) was relatively narrow. In a future study, the AUC approach will be applied for
473 sites differing in climate and soil textural classes. We expect that the model can predict the dynamic
474 matric potential for a new site as long as the autoencoder value falls within the range of AUV of the
475 training sites. To predict the soil moisture dynamics for soils with autoencoder values outside of the
476 range of training data, the model must be re-built using additional training data.

477

478 4.2 Interpretation of AUV and its relationship to physical soil properties

479 As discussed in section 3.2, the autoencoder value (AUV) is low for soil water characteristics curves
480 (SWC) with low saturated water content and low variations of water content for a certain matric
481 potential value (type 1) and high AUV for large values and variations of water content (type 2). To
482 provide a more quantitative relationship between SWC and AUV, the SWC data were characterized as
483 follows: the time average of volumetric water content (VWC) and SWP were calculated for 15 days for
484 the period 2015 to 2022. The envelope of these data was then calculated by fitting a minimum and
485 maximum pressure saturation relationship including the averaged data (see Figure 10a).



486

487 **Figure 10** Relationship between autoencoder value (AUV) and soil water characteristics curve (SWC).
 488 (a) 15-days average of SWC data for Aetigkofen (symbols; site #1). The two lines are exponential
 489 functions building the envelope of the SWC curve. (b) Linear model for the nine sites linking the
 490 parameters of the exponential model with the ‘measured’ AUV (deduced from measured water content
 491 data).

492 The two boundary lines of the SWC were then characterized by a ‘saturated’ and ‘residual’ water
 493 content and a shape parameter defining an exponential decrease of water content with increasing
 494 absolute matric potential values. The SWC of each site can thus be described by six parameters (three
 495 parameters per boundary line). As shown in Figure 10b, a linear model expressing the AUV as function
 496 of these six parameters can be built. Simpler models with less parameters could not reproduce the AUV
 497 of all sites. Despite the positive correlation between AUV and average water content, the average water
 498 content alone is not sufficient to explain the range of AUV for all sites. Also combining average water
 499 content with soil texture information could not reproduce the AUVs of all sites, indicating that the soil
 500 moisture dynamics represented by AUV is not only dependent on static soil textural attributes but
 501 seasonal structural features as well.

502 Accordingly, there is no simple interpretation of AUV based on texture and average water content, but
 503 the dynamic variation of water content must be considered as well. Due to the relevance of the variation
 504 in water content for similar matric potential value, the use of a variational autoencoder (VAE) instead
 505 of the typical autoencoder could be considered. In contrast to the typical autoencoder that maps the

506 input information into a single point (or a few points), the VAE produces a probability distribution
507 capturing the variability (second moment) of the data. This could be specifically of interest for clay
508 soils with high water contents (much larger than the residual water content) for the entire range of matric
509 potential values. By including a probabilistic approach in the compressing and decompressing step, the
510 variability of the data could be captured more efficiently using VAR.

511

512 4.3 Application for satellite data

513 The AUC-DNN model was used to analyze satellite-based volumetric water content (VWC) satellite
514 data, including SMAP L4 and L3, SMOS products, and Sentinel data. Subsequently, a comparison was
515 carried out for the AUV for both site-specific measurements and earth observation (EO) measurements
516 for the same region. The initial findings highlighted a disparity between the dynamics captured by EO
517 products and the actual dynamics. Therefore, if the objective is to establish a robust system capable of
518 detecting changes in water retention dynamics on a regional scale, it is considered necessary to enhance
519 the calibration of EO in Europe. Only with EO-data that can reproduce the essential of the soil moisture
520 dynamics as manifested in the AUV, the matric potential dynamics can be deduced from EO-data. For
521 future EO-data with improved capacity to capture regional soil moisture dynamics, the concept
522 presented in this study (AUC-DNN) could be used to predict matric potential dynamics at global scale
523 (see Appendix C).

524 5. Summary and conclusions

525 The soil water potential (SWP) determines water flow direction, water ability for plants, and mechanical
526 stability. Because it cannot be measured directly by remote sensing techniques at larger scales, it is
527 often deduced from water content information, assuming an unambiguous relationship between water
528 content and SWP. However, this relationship under dynamic field conditions is highly ambiguous due
529 to hysteresis, dynamic effects, and soil structural changes that cannot be modeled with a physically-
530 based model. To enable prediction of SWP from soil water content, we apply a deep neural network
531 (DNN) with an autoencoder to define unique features of the soil moisture dynamics. By inserting the

532 autoencoder value (AUV) together with climatic data and water content measured at nine sites in the
533 region of Solothurn (Switzerland) in a deep neural network (AUC-DNN), the soil water potential could
534 be predicted. The main findings of the study can be summarized as follows:

- 535 • The SWC of the nine sites can be classified in three types based on the width of pressure-
536 saturation relationship and the water content close to saturation
- 537 • These SWC-types are manifested in different autoencoder values (AUV)
- 538 • The AUV is not a simple function of average water content or soil texture but includes structural
539 effects as well
- 540 • The AUC-DNN model could predict successfully the SWP dynamics of sites without site-
541 specific training

542 The autoencoder value (AUV) is thus a new descriptor of the complex soil moisture dynamics that
543 cannot be captured with physically based models. Future satellite generation may be sensitive enough
544 to measure the AUV from remote sensing water content data. The approach presented in this paper will
545 then enable the prediction of the soil matric potential at the global scale using remote sensing water
546 content data.

547 Appendix A: Data Quality Assurance and Trend Analysis



548 As a precaution for data quality, the Absolute Matric Potential (AMP) and volumetric water content
549 (VWC) data were scrutinized to identify potential errors the data. The process includes different steps
550 that were necessary to discover anomalies, checking the integrity of the data, and detecting systematic
551 changes with time.

552 **1- Flagging Abrupt Changes in VWC and MP:**

553

554 **VWC Flagging and removing:**

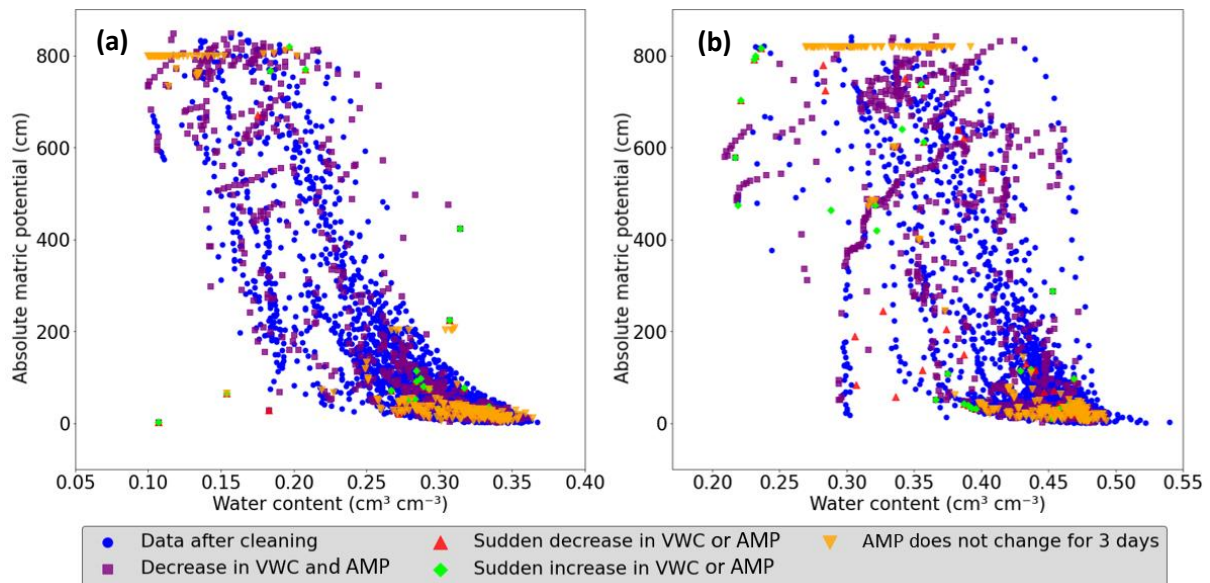
- 555 • Differences between consecutive (daily) time steps in the water content time series were
556 calculated.

- 557  Instances with daily differences exceeding $0.1 \text{ cm}^3/\text{cm}^3$ were flagged and denoted as sudden
558 decreases or increases in VWC.
- 559  Instances with VWC below $0.1 \text{ cm}^3/\text{cm}^3$ or exceeding $0.7 \text{ cm}^3/\text{cm}^3$ were identified and removed
560 from the dataset. These extreme values were considered as measurement anomalies or outliers
561 affecting the overall dataset's reliability.
- 562 • Instances with $\text{AMP} < 1 \text{ cm}$ was removed from the data to overcome limitations in the used
563 method. The water potential can change without modifying the volumetric water content after
564 this limit, which could make the results of the model not accurate enough.
 - 565 • The differences between consecutive time steps in AMP -time series was calculated; instances
566 with daily differences exceeding 500 cm were flagged and called sudden decreases or increases
567 in AMP (figure A1).
 - 568 • The threshold AMP-value of 850 cm was employed in a specific step, where instances with
569 AMP exceeding 850 cm were removed from the dataset, addressing the physical properties of
570 water as it starts to boil in the tensiometers under pressure after this limit.
 - 571 • Periods of concurrent decrease in AMP (indicator for wetting) and decrease in VWC (drying)
572 were flagged (figure A1).
 - 573 • Periods with matric potential values remaining constant over a three-day rolling window were
574 flagged (figure A1).

575 **2- Utilizing Index Windows for Data Manipulation and Data Removal**

576 To address flagged instances mentioned before, a systematic approach is employed. For each
577 flagged instance, three additional indices are generated around it to construct an index window,
578 spanning one day before (index_1), the flagged instance itself (index_0), and two days after
579 (index_2 and index_3). This four-day index window was eliminated from the dataset (figure
580 A1). The decision to eliminate this window was informed by a visual assessment of
581 measurements as it was noticed that when a measurement error occurs, the accuracy of the
582 preceding day is affected. Furthermore, it was assumed that the device requires two subsequent

583 days to restore normal measurement precision. This process contributes to a refined dataset,
584 providing a more accurate representation of the underlying trends in AMP and VWC.



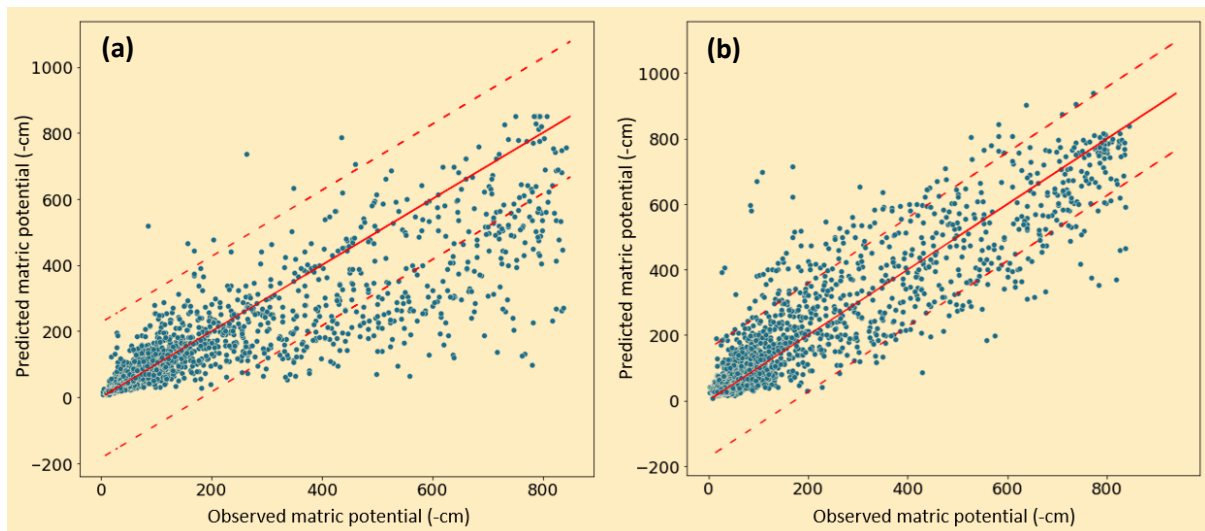
585
586 **Figure A1** Comparison of data before and after cleaning procedure: the blue circles depict the remaining
587 data after applying the cleaning criteria. Each distinct marker represents eliminated points, each
588 corresponding to a specific criterion (e.g., the square purple marker for simultaneous decrease in
589 volumetric water content (VWC) and the absolute matric potential (AMP), the red upward-pointing
590 triangle is the marker for sudden decreases, the lime diamond for sudden increases, and the orange
591 downward-pointing triangle marks periods of unchanged AMP). This provides insights into the reasons
592 for data removal and illustrates the profound impact of the data cleaning process in retaining high-
593 quality data points. In (a) the cleaning process for sandy clay loam site in Aetigkofen (site #1) is shown,
594 in (b) the cleaning process for the Matzendorf site (site #9, clay loam soil).

595 Appendix B: Running the model with Logarithmic MP value.

596 The AUC-DNN showed a good performance in predicting the dynamic MP for the different 6 unseen
597 sites. However, it was clear that the model prioritizes tends to focus on capturing significant changes in
598 values rather than accurately representing the values themselves. This tendency is attributed to the
599 substantial difference between the highest and lowest absolute values (approximately 850 cm), leading
600 the model to emphasize major fluctuations while neglecting minor ones. To address this issue and
601 enhance the model's precision in capturing the exact AMP, a suggestion has been made to train the
602 model for the same three sites but with the logarithmic value for the AMP. This modification aims to
603 strike a better balance, ensuring that both major and minor changes are effectively captured while
604 maintaining accuracy in representing the specific values of MP.

605 To qualitatively assess the model training performance under the logarithmic scale, a scatter plot (Figure
606 B1) was generated, comparing observations against simulated values for the second training site
607 (Stüsslingen). The reason for choosing a training site was to understand how the model captures the
608 dynamics when trained with logarithmic matric potential. The results suggest that using logarithmic
609 scale, the model prioritized the prediction of the exact absolute value of matric potential (AMP), which
610 makes the model to optimize predictions for the absolute values between 0 to 200 cm. This approach is
611 giving the same importance to small and large changes in the AMP, which causes that the model
612 assigned a higher weight to small changes according to their higher frequency, while neglecting less
613 frequently occurring major dynamic shifts. Consequently, the model's accuracy went down beyond 200
614 cm (figure B1a) when compared to the model trained on non-logarithmic AMP-values (figure B1b). To
615 maintain a balanced consideration of changes, logarithmic MP was avoided in the main part of the
616 paper.

617



618

619 **Figure B1** Visual comparison of model performance, comparing the observed and simulated values for
620 the Stüsslingen training site. (a) the model trained with logarithmically scaled AMP-values, while in
621 (b) The model trained with absolute linear matric potential (AMP) values. The solid line denotes the
622 1:1 correspondence, and dashed lines represent the 95% confidence interval.

623 Appendix C: SMAP data and Autoencoder for global scale analysis

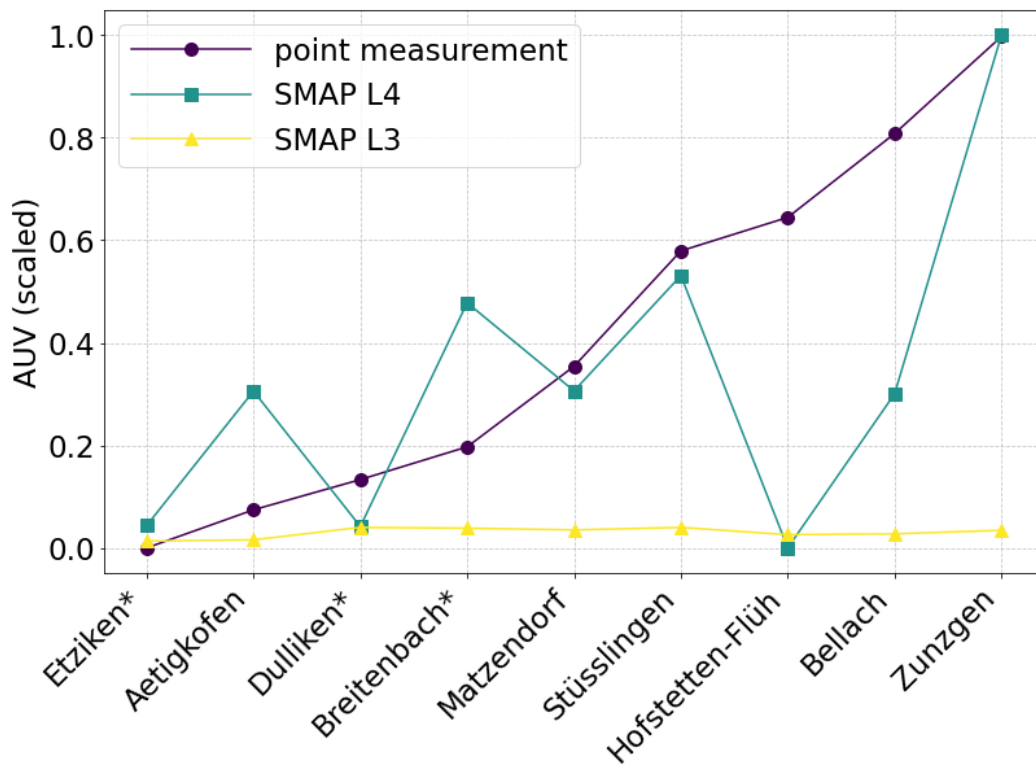
624 SMAP (Soil Moisture Active Passive) is a NASA satellite mission that was established to help in
625 improving weather forecasts and global drought monitoring. SMAP data products are available at
626 different levels of processing, from Level 1 (L1; instrument measurements) to Level 4 (L4; model-

627 derived value-added products). For this study, SMAP L3 and SMAP L4 products for measuring
628 moisture content were used. The main difference between the two products is that SMAP L3 depends
629 on the passive radiometer measurements, while SMAP L4 products are derived from a data assimilation
630 system that combines the L-band brightness temperature observations from SMAP with a land surface
631 model and meteorological forcing data (Reichle et al., 2019). SMAP L3 products for moisture content
632 are primarily affected by vegetation and surface roughness, allowing them to capture surface soil
633 moisture variations. In contrast, the incorporation of land surface models in SMAP L4 products reduces
634 its sensitivity to vegetation covers and surface roughness, making the products more representative of
635 the profile soil moisture conditions (Reichle et al., 2019).

636 The autoencoder's encoded representations offer a unique opportunity to compare the spatial patterns
637 inherent in "point measurement" with remote sensing data such as SMAP L3 and SMAP L4 data. The
638 autoencoder method could illuminate how these diverse data streams align or diverge, providing crucial
639 insights into the compatibility and complementarity of ground and satellite measurements. The process
640 was applied for the data between the years 2015 to 2022. All the data (SMAP L4, SMAP L3, and on-
641 site measurements) were given to the autoencoder neural network together. Subsequently, the resulting
642 autoencoder values were scaled. Finally, a comparison was made to show if the satellite measurements
643 and the on-site measurements have the same measured dynamics.

644 The autoencoder analysis of SMAP L3 (figure C1) indicates that satellite measurements struggle to
645 capture the dynamic change of the water content, as all locations yield approximately the same
646 Autoencoder Value (AUV). In contrast, the SMAP L4 product (figure C1) exhibits fluctuations in AUV
647 results. For instance, Stüsslingen and Matzendorf align closely with on-site measurements in terms of
648 AUVs. However, for Hofstetten-Flüh, the SMAP L4 product indicates a very small AUV, suggesting
649 an expected dynamic in line with a type 1 soil water retention curve (figure 6b). In contrast, on-site
650 measurements indicate a higher AUV for Hofstetten-Flüh, suggesting a closer association with a type
651 2 soil water retention curve. These findings underscore the imperative for developing a new
652 methodology to calibrate satellite data in the Switzerland area. The prevalent uniformity in SMAP L3
653 results and the notable disparities between on-site measurements and satellite data across various

654 products highlight the need for a more refined approach to ensure accurate and reliable dynamic soil
655 moisture assessments.



656
657 **Figure C1** Comparative analysis of Autoencoder Neural Network results for SMAP L3 and SMAP L4
658 satellite data, alongside with profile measurements. The fluctuating AUV values indicate varying
659 degrees of alignment with on-site measurements across different locations. Sites with forest are marked
660 with *.

661 Code and data availability

662 The related input data for the AUC-DNN model and Python code are openly accessible under
663 <https://doi.org/10.5281/zenodo.10600669> and <https://doi.org/10.5281/zenodo.10602397> respectively.

664 The input for the autoencoder and its python codes are openly accessible under
665 <https://www.doi.org/10.5281/zenodo.10605108>

666 Author contributions

667 NA, AC, and PL designed the research. NA and PL performed the research. NA and MR analyzed the
668 soil moisture time series. SM was responsible for the soil moisture network. NA wrote the codes and
669 built the model. NA and PL wrote the manuscript with substantial input from all co-authors.

670 Competing interests

671 The contact author has declared that none of the authors has any competing interests.

672 Financial support

673 “This research is part of the project Artificial Intelligence for Soil Health. Funded by the European
674 Union. Views and opinions expressed are however those of the author(s) only and do not necessarily
675 reflect those of the European Union or of the Research Executive Agency (REA). Neither the European
676 Union nor the granting authority can be held responsible for them.”

677 Acknowledgements

678 NA acknowledges the utilization of ChatGPT to enhance coherence within certain sections of the
679 manuscript.




680

References

681

- 682 Achieng, K. O.: Modelling of soil moisture retention curve using machine learning techniques:
683 Artificial and deep neural networks vs support vector regression models, *Comput Geosci*, 133,
684 104320, <https://doi.org/10.1016/j.cageo.2019.104320>, 2019.
- 685 Basile, A., Bonfante, A., Coppola, A., De Mascellis, R., Falanga Bolognesi, S., Terribile, F., and Manna,
686 P.: How does PTF Interpret Soil Heterogeneity? A Stochastic Approach Applied to a Case Study on
687 Maize in Northern Italy, *Water (Basel)*, 11, 275, <https://doi.org/10.3390/w11020275>, 2019.
- 688 Bertels, D. and Willems, P.: Physics-informed machine learning method for modelling transport of a
689 conservative pollutant in surface water systems, *J Hydrol (Amst)*, 619, 129354,
690 <https://doi.org/10.1016/j.jhydrol.2023.129354>, 2023.
- 691 BODENMESSNETZ: <https://www.bodenmessnetz.ch/messwerte/datenabfrage>, last access: 22 July
692 2024.
- 693 Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models:
694 Combining the strengths of manual and automatic methods, *Water Resour Res*, 36, 3663–3674,
695 <https://doi.org/10.1029/2000WR900207>, 2000.
- 696 Bundesamt für Energiewirtschaft: Richtlinien zum Schutze des Bodens beim Bau unterirdisch
697 verlegter Rohrleitungen, Schweiz, 1997.
- 698 Capparelli, G. and Spolverino, G.: An Empirical Approach for Modeling Hysteresis Behavior of
699 Pyroclastic Soils, *Hydrology*, 7, 14, <https://doi.org/10.3390/hydrology7010014>, 2020.
- 700 Chen, S. and Guo, W.: Auto-Encoders in Deep Learning—A Review with New Perspectives,
701 *Mathematics*, 11, 1777, <https://doi.org/10.3390/math11081777>, 2023.
- 702 D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith: Model
703 Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Trans*
704 *ASABE*, 50, 885–900, <https://doi.org/10.13031/2013.23153>, 2007.
- 705 den Ouden, H. E. M., Kok, P., and de Lange, F. P.: How Prediction Errors Shape Perception, Attention,
706 and Motivation, *Front Psychol*, 3, <https://doi.org/10.3389/fpsyg.2012.00548>, 2012.
- 707 Fomin, D. S., Yudina, A. V., Romanenko, K. A., Abrosimov, K. N., Karsanina, M. V., and Gerke, K. M.:
708 Soil pore structure dynamics under steady-state wetting-drying cycle, *Geoderma*, 432, 116401,
709 <https://doi.org/10.1016/j.geoderma.2023.116401>, 2023.
- 710 Fu, Y. P., Liao, H. J., Chai, X. Q., Li, Y., and Lv, L. L.: A Hysteretic Model Considering Contact Angle
711 Hysteresis for Fitting Soil-Water Characteristic Curves, *Water Resour Res*, 57,
712 <https://doi.org/10.1029/2019WR026889>, 2021.
- 713 Gallipoli, D., Gens, A., Sharma, R., and Vaunat, J.: An elasto-plastic model for unsaturated soil
714 incorporating the effects of suction and degree of saturation on mechanical behaviour,
715 *Géotechnique*, 53, 123–135, <https://doi.org/10.1680/geot.2003.53.1.123>, 2003.
- 716 Gholamy, A., Kreinovich, V., and Kosheleva, O.: Why 70/30 or 80/20 Relation Between Training and
717 Testing Sets: A Pedagogical Explanation, University of Texas at El Paso., 1–6 pp., 2018.

- 718 Gupta, H. V. and Kling, H.: On typical range, sensitivity, and normalization of Mean Squared Error and
719 Nash-Sutcliffe Efficiency type metrics, *Water Resour Res*, 47,
720 <https://doi.org/10.1029/2011WR010962>, 2011.
- 721 Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Status of Automatic Calibration for Hydrologic Models:
722 Comparison with Multilevel Expert Calibration, *J Hydrol Eng*, 4, 135–143,
723 [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)), 1999.
- 724 Gupta, S., Lehmann, P., Bickel, S., Bonetti, S., and Or, D.: Global Mapping of Potential and Climatic
725 Plant-Available Soil Water, *J Adv Model Earth Syst*, 15, <https://doi.org/10.1029/2022MS003277>,
726 2023.
- 727 Hannes, M., Wollschläger, U., Wöhling, T., and Vogel, H. -J.: Revisiting hydraulic hysteresis based on
728 long-term monitoring of hydraulic states in lysimeters, *Water Resour Res*, 52, 3847–3865,
729 <https://doi.org/10.1002/2015WR018319>, 2016.
- 730 Holthusen, D., Peth, S., and Horn, R.: Impact of potassium concentration and matric potential on soil
731 stability derived from rheological parameters, *Soil Tillage Res*, 111, 75–85,
732 <https://doi.org/10.1016/j.still.2010.08.002>, 2010.
- 733 IDAWEB: <https://gate.meteoswiss.ch/idaweb/login.do>, last access: 22 July 2024.
- 734 Ioffe, S., and Szegedy, C.: Batch normalization: Accelerating deep network training by reducing
735 internal covariate shift, *arXiv [preprint]*, arXiv:1502.03167, 11 February 2015.
- 736 Jain, S. K., Singh, V. P., and van Genuchten, M. Th.: Analysis of Soil Water Retention Data Using
737 Artificial Neural Networks, *J Hydrol Eng*, 9, 415–420, [https://doi.org/10.1061/\(ASCE\)1084-
738 0699\(2004\)9:5\(415\)](https://doi.org/10.1061/(ASCE)1084-0699(2004)9:5(415)), 2004.
- 739 Kingma, D. P., and Ba, J.: Adam: A method for stochastic optimization, *CoRR*, arXiv:1412.6980, 22
740 December 2014.
- 741 Lu, L.: Dying ReLU and Initialization: Theory and Numerical Examples, *Commun Comput Phys*, 28,
742 1671–1706, <https://doi.org/10.4208/cicp.OA-2020-0165>, 2020.
- 743 Lu, N., Godt, J. W., and Wu, D. T.: A closed-form equation for effective stress in unsaturated soil,
744 *Water Resour Res*, 46, <https://doi.org/10.1029/2009WR008646>, 2010.
- 745 Ma, Y., Liu, H., Yu, Y., Guo, L., Zhao, W., and Yetemen, O.: Revisiting Soil Water Potential: Towards a
746 Better Understanding of Soil and Plant Interactions, *Water (Basel)*, 14, 3721,
747 <https://doi.org/10.3390/w14223721>, 2022.
- 748 Mendes, J. and Buzzi, O.: New insight into cavitation mechanisms in high-capacity tensiometers
749 based on high-speed photography, *Canadian Geotechnical Journal*, 50, 550–556,
750 <https://doi.org/10.1139/cgj-2012-0393>, 2013.
- 751 Menon, M., Mawodza, T., Rabbani, A., Blaud, A., Lair, G. J., Babaei, M., Kercheva, M., Rousseva, S.,
752 and Banwart, S.: Pore system characteristics of soil aggregates and their relevance to aggregate
753 stability, *Geoderma*, 366, 114259, <https://doi.org/10.1016/j.geoderma.2020.114259>, 2020.
- 754 Montesinos López, O. A., Montesinos López, A., and Crossa, J.: Fundamentals of Artificial Neural
755 Networks and Deep Learning, in: *Multivariate Statistical Machine Learning Methods for Genomic
756 Prediction*, Springer International Publishing, Cham, 379–425, [https://doi.org/10.1007/978-3-030-
757 89010-0_10](https://doi.org/10.1007/978-3-030-89010-0_10), 2022.

- 758  Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion
759 of principles, *J Hydrol (Amst)*, 10, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6), 1970.
- 760 Rawls, W. J., Pachepsky, Y. A., Ritchie, J. C., Sobecki, T. M., and Bloodworth, H.: Effect of soil organic
761 carbon on soil water retention, *Geoderma*, 116, 61–76, [https://doi.org/10.1016/S0016-](https://doi.org/10.1016/S0016-7061(03)00094-6)
762 [7061\(03\)00094-6](https://doi.org/10.1016/S0016-7061(03)00094-6), 2003.
- 763 Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., Ardizzone, J. V.,
764 Bosch, D., Colliander, A., Cosh, M., Kolassa, J., Mahanama, S. P., Prueger, J., Starks, P., and Walker, J.
765 P.: Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product, *J Adv Model Earth Syst*,
766 11, 3106–3130, <https://doi.org/10.1029/2019MS001729>, 2019.
- 767 Ritter, A. and Muñoz-Carpena, R.: Performance evaluation of hydrological models: Statistical
768 significance for reducing subjectivity in goodness-of-fit assessments, *J Hydrol (Amst)*, 480, 33–45,
769 <https://doi.org/10.1016/j.jhydrol.2012.12.004>, 2013.
- 770 Romero-Ruiz, A., Linde, N., Keller, T., and Or, D.: A Review of Geophysical Methods for Soil Structure
771 Characterization, *Reviews of Geophysics*, 56, 672–697, <https://doi.org/10.1029/2018RG000611>,
772 2018.
- 773  Ross, P. J. and Smettem, K. R. J.: A Simple Treatment of Physical Nonequilibrium Water Flow in Soils,
774 *Soil Science Society of America Journal*, 64, 1926–1930,
775 <https://doi.org/10.2136/sssaj2000.6461926x>, 2000.
- 776 Rostami, A., Habibagahi, G., Ajdari, M., and Nikooee, E.: Pore Network Investigation on Hysteresis
777 Phenomena and Influence of Stress State on the SWRC, *International Journal of Geomechanics*, 15,
778 [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000315](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000315), 2015.
- 779 Sadeghi, H., Chiu, A. C. F., Ng, C. W. W., and Jafarzadeh, F.: A vacuum-refilled tensiometer for deep
780 monitoring of in-situ pore water pressure, *Scientia Iranica*, 27, 596–606,
781 <https://doi.org/10.24200/sci.2018.5052.1063>, 2018.
- 782  Shwetha, P. and Varija, K.: Soil Water Retention Curve from Saturated Hydraulic Conductivity for
783 Sandy Loam and Loamy Sand Textured Soils, *Aquat Procedia*, 4, 1142–1149,
784 <https://doi.org/10.1016/j.aqpro.2015.02.145>, 2015.
- 785 Smith, C. W., Johnston, M. A., and Lorentz, S. A.: The effect of soil compaction on the water
786 retention characteristics of soils in forest plantations, *South African Journal of Plant and Soil*, 18, 87–
787 97, <https://doi.org/10.1080/02571862.2001.10634410>, 2001.
- 788 Spreafi, M. and Weingartner, R.: The Hydrology of Switzerland Selected aspects and results, *Reports*
789 *of the FOWG, Water Series*, Bern, 2005.
- 790 Tuller, M. and Or, D.: Soil water retention and characteristic curve, in: *Encyclopedia of Soils in the*
791 *Environment*, Elsevier, 187–202, <https://doi.org/10.1016/B978-0-12-822974-3.00105-1>, 2023.
- 792 Willems, P.: A time series tool to support the multi-criteria performance evaluation of rainfall-runoff
793 models, *Environmental Modelling & Software*, 24, 311–321,
794 <https://doi.org/10.1016/j.envsoft.2008.09.005>, 2009.
- 795 Zuo, Y. and He, K.: Evaluation and Development of Pedo-Transfer Functions for Predicting Soil
796 Saturated Hydraulic Conductivity in the Alpine Frigid Hilly Region of Qinghai Province, *Agronomy*, 11,
797 1581, <https://doi.org/10.3390/agronomy11081581>, 2021.