Field Calibration of TDR to Assess the Soil Moisture of Drained Peatland Surface Layers

Tomasz Gnatowski 1, 2*, Jan Szatyłowicz 1, *, Bogumila Pawluśkiewicz 1, Ryszard Oleszczuk 1, Maria Janicka 2, Ewa Papierowska 3, and Daniel Szejba 1

1 Department of Environmental Improvement, Faculty of Civil and Environmental Engineering, Warsaw University of Life Sciences—SGGW, Nowoursynowska 159, 02-776 Warsaw, Poland; jan_szatyłowicz@sggw.pl (J.S.); bogumila_pawluskieicz@sggw.pl (B.P.); ryszard_oleszczuk@sggw.pl (R.O.); daniel_szejba@sggw.pl (D.S.)
2 Department of Agronomy, Faculty of Agriculture and Biology, Warsaw University of Life Sciences—SGGW, Nowoursynowska 159, 02-776 Warsaw, Poland; maria_janicka@sggw.pl
3 Laboratory Water Center, Faculty of Civil and Environmental Engineering, Warsaw University of Life Sciences—SGGW, Nowoursynowska 159, 02-776 Warsaw, Poland; ewa_papierowska@sggw.pl
* Correspondence: tomasz_gnatowski@sggw.pl; Tel.: +48-22-593-53-63

Received: 8 October 2018; Accepted: 29 November 2018; Published: 13 December 2018

Abstract: The proper monitoring of soil moisture content is important to understand water-related processes in peatland ecosystems. Time domain reflectometry (TDR) is a popular method used for soil moisture content measurements, the applicability of which is still challenging in field studies due to requirements regarding the calibration curve which converts the dielectric constant into the soil moisture content. The main objective of this study was to develop a general calibration equation for the TDR method based on simultaneous field measurements of the dielectric constant and gravimetric water content in the surface layers of degraded peatlands. Data were collected during field campaigns conducted temporarily between the years 2006 and 2016 at the drained peatland Kuwasy located in the north-east area of Poland. Based on the data analysis, a two-slopes linear calibration equation was developed as a general broken-line model (GBLM). A site-specific calibration model (SSM-D) for the TDR method was obtained in the form of a two-slopes equation describing the relationship between the soil moisture content and the dielectric constant and introducing the bioindices as covariates relating to plant species biodiversity and the state of the habitats. The root mean squared error for the GBLM and SSM-D models were equal, respectively, at 0.04 and 0.035 cm$^3$ cm$^{-3}$.

Keywords: degraded peat soils; dielectric constant; TDR calibration curve; broken-line model; bioindices of habitats

1. Introduction

Accurate determination of soil moisture is important for various hydrological, agricultural and environmental applications. In hydrology, soil moisture content plays a major role in the water cycle by controlling hydrological fluxes, such as infiltration, runoff and groundwater recharge [1]. In climate studies, soil moisture governs the energy flux exchange between the land surface and the atmosphere through its impact on evapotranspiration [2]. Determination of the soil moisture content is very important for hydrological studies in peatland soils, the actual condition and preservation of which depends on water-related processes [3–5]. Peatland soils occupy about 400 million hectares, which is about 3% of the world’s land [6]. Most of these soils are located in the northern part of the globe, including large areas of North America, Europe and Russia. It is estimated that peat soils contain about one-third of global soil organic carbon stock [7]. The last GIS-based estimation of peatland areas in Europe shows that this type of ecosystems occupies approximately 593,727 km$^2$ and 46% of this area.
Water 2018, 10, 1842 2 of 22

is classified as peatland where the peat forming process was stopped [8]. Stopping the forming process was associated with artificial land drainage in response to agricultural, forestry and horticultural demands [9]. The great majority of drained peatland is used as meadow and pasture, and only a limited percent as arable land [10]. Drainage of peat soils leads to the formation of a new organic soil material in the surface layers known as muck (moorsh), which has a granular structure where the original plant remains are not recognizable and contains more mineral matter than peat [11].

Soils contain twice as much organic carbon than the amount of C in the atmosphere [12], therefore any changes in the terrestrial pool are usually associated with the impact of global warming [13,14]. The importance of peatlands in global warming relates to the fact that 14% to 20% of these areas globally are currently used for agriculture, mainly as meadows or pastures, and oxidized soils of these ecosystems are a source of carbon, which has an influence on the global increase of CO₂ emissions [15].

A literature review shows that there are three major factors used to quantify the amount of CO₂ emission from drained peat soils, i.e., soil temperature [16–18], soil moisture content status [17,19,20] and water table depth [21]. The increase of soil temperature and optimal soil moisture status create favorable conditions for the increase of CO₂ efflux. Soil moisture content is one of the main hydrological factors to be assessed in the calibration and validation of activities in global soil moisture products (spaceborne), as well as in order to increase the understanding of water, energy and carbon flux exchanges at the active surface on a local scale [22].

Over the last four decades, the most popular in situ soil moisture measuring technique was based on measurements of the dielectric properties of the medium in time domain. This method is named Time Domain Reflectometry (TDR) and its first application in soil science was reported in 1980 [23]. The theory behind this measurement technique is based on travel time analysis of the short electromagnetic step pulse which formed an electrical field along a waveguide (e.g., two parallel rods of a given length) surrounded by the soil [24]. The time required to travel the length of the waveguide (there and back) is recorded and used to calculate the dielectric constant (Kₐ) of the medium [25]. The attractiveness of the TDR method results from the simple non-destructive and quick measurement of the dielectric constant which can be transformed using the calibration curve to the volumetric soil moisture content. However, many TDR probes on the market are often not properly tested for the high dielectric permittivity values of peat soils and that the algorithms for the TDR waveform analysis might fail in that range [26]. The relationship between the measured Kₐ and moisture content can be derived either by using the dielectric mixing model [27,28] or by empirical equations [25]. In peat soils, in most cases, the calibration curves are empirical equations obtained by fitting a mathematical model of the measured dielectric constant to the gravimetrically measured soil moisture content data. The majority of calibration equations for peat soils are third-order polynomials with respect to the raw measured dielectric constant [22,23,29–32] or its transformed form as a square root [33]. The polynomial calibration equations are very flexible and more efficient than the manufacturer default relationships (determined by developers of the TDR measurement systems) but they are appropriate as a site- or soil-specific equations which enables improved accuracy of soil moisture content estimation [34]. A more general calibration equation for the organic materials can also be expressed in the form of a nonlinear relationship between the soil moisture content and the dielectric constant [22,35,36]. Another popular group of empirical equations constitutes the linear relationship between the soil moisture content and the square root of the dielectric constant. Malicki et al. [37] developed an empirical model which incorporates soil bulk density as second covariate, in addition to the Kₐ⁰.5, thereby improving the moisture content predictions for a broad range of soils, also including organic soils. Oleszczuk et al. [38], using a similar data analysis technique, obtained a general empirical model describing soil moisture content as a function of the dielectric constant, bulk density and ash content for different top organic soil layers from the Biebrza River Valley. However, this type of calibration procedure, apart from the dielectric constant measurements, requires extra knowledge regarding other soil properties in order to predict its moisture content status. Despite the large number of calibration equations for determination of the soil moisture content using TDR,
they were developed mostly under laboratory conditions and their use is limited to specific sites or dedicated soils.

Peat soils are characterized by soil volume changes due to soil moisture content changes. Peat soil shrinks during drying and its bulk density increases while wetting, causing soil swelling and a decrease in bulk density [36,39–42]. In the majority of laboratory calibrations using TDR, bulk density changes are not considered as a factor influencing the dielectric constant.

Interdependence of peat properties and vegetation cover is often observed in peatlands [43]. The linkage between the physical and chemical properties of soil and the state of plant cover is often described using diversity indices representing plant species [44] or using more complex environmental indicators [45]. Biodiversity is a useful metric for quantifying the composition and abundances of the plant species [46] providing information about community structure and it may also be used as an indicator of the ecosystem’s health [47]. In Europe, the Shannon–Wiener index (H’) is a commonly used bioindicator for ecological assessment of the quality of ecosystems [48]. According to its value, the ecosystem’s status can be classified from “Bad” (H’ ≤ 1) to “High” (H’ > 4). In between there are three classes of ecosystem quality: “Poor” (1 < H’ ≤ 2), “Moderate” (2 < H’ ≤ 3) and “Good” (3 < H’ ≤ 4). The species composition of grassland is considered a good indicator of habitat properties which are frequently quantified using species indicator values [49]. Ellenberg indicator values (EIV) [45] are the most commonly used ecological indicators in Central Europe because they summarize complex environmental factors including: Groundwater level, soil moisture content, precipitation, humidity, etc. in one figure and they are calculated on the basis of complete species composition of plant communities [50]. The system of indicators, as compiled by [45], considers vascular plant species with respect to: Light intensity (L), mean annual air temperature (T), continentality (K), substrate moisture (F), substrate reaction (R), substrate nutrient availability (N) and salt content of the substrate (S). Six of these indicators have a nine-point scale and 12 levels have been established for the F value. The present definition of the EIV system and the re-interpretation of the L indicator were presented in the study by [51]. Despite the ordinal character of the EIV indicators [50], their average estimators of a vegetation record provide helpful environmental site information, which can substitute costly and time-consuming measurements [52–54].

Taking into account the importance and range of occurrence of drained peatland ecosystems in the northern hemisphere, we have tried to develop a TDR calibration model which enables the monitoring of soil moisture content being the most important factor controlling environmental processes at the active surface. We have assumed that it is possible to establish an accurate relationship between soil moisture content and soil dielectric properties based on the field measurements of both features. In light of the literature review, we have also supposed that the soil moisture content prediction using Ks values can be improved at the site scale by introducing the bioindicators describing the biodiversity of plant species and the state of habitats into the calibration equation.

2. Materials and Methods

2.1. Study Area and Site Characteristics

The study area was located in the Middle Basin of the Biebrza River Valley (22° 30’–23° 60’ E and 53° 30’–53° 75’ N) within the Kuwasy drainage sub-irrigation system [55]. The first drainage system in this area was created between 1933 and 1939 using open channels. Then, the drainage sub-irrigation ditches were constructed between 1951 and 1960. Due to the high rate of peat soils subsidence they were deepened in between 1977 and 1988. The Kuwasy Channel is the main part of the land reclamation system (about 15 km long) which connects the Raigrodzkie Lake (the main reservoir of water used for the irrigation of the Kuwasy peatlands) with the Rudzki Channel. In this study, six experimental sites characterized by different soil properties and vegetation patches were selected (Figure 1). The term “site” used in this study is defined as a part of the field quarter covered by representative and relatively homogenous vegetation patches. All the studied sites were used as meadows. Two sites
were located near Biebrza Village (site I and II) and the soil profiles were developed from alder peat. Next, two profiles were situated near the Stoczek Stream (III, IV) where soil upper layers had evolved from sedge-reed peat deposits. The last two experimental sites (V, VI) were located at the border of the Kuwasy system near the Jegrznia River. The reed peat was the parent material of these soil profiles. The underlying peat soil layers (below 15 cm) in sites I and II were classified as sapric Histosols horizons whereas peats from the other soil profiles were classified as hemic Histosols [56]. The upper layers (up to 15 cm) of the studied soils were formed as a result of long-term water table lowering. Land drainage resulted in the soil mucky process which in turn led to developing a new organic soil material with specific chemical, physical and biological properties [57]. The surface layers are often referred to as muck (moorsh) material [58].

Figure 1. Location of study sites (based on http://mapy.geoportal.gov.pl).

2.2. Field Measurements and Basic Soil Properties

The set of field data was collected during 23 field campaigns conducted between 2008 and 2016 for sites I, III–VI, and 12 field campaigns during the vegetation seasons between 2006 and 2007 for site II. During the measurement dates at the selected sites, a soil pit was dug to the depth limited by the ground water table position. The measurements were done in the top layers (0–15 cm) recognized as the surface soil layers dominated by moorsh material. The dielectric constant was measured in the soil profiles at three locations spaced horizontally by 30 cm at three different depths: 5, 7.5 and 12.5 cm. A slightly different measurement methodology was applied at site II, where measurements were conducted at a depth of 10 cm. In each measurement spot, the two rods of a TDR probe (10 cm long) were installed horizontally for recording the dielectric constant. The TDR probe was a part of the measurement system (field operator meter), which also comprises the pulse generator
module producing a Gaussian needle pulse, with rises/falls at about 200 ps [59]. After the dielectric constant measurements, the soil profiles were sampled for laboratory determination of gravimetric soil water content and soil bulk density. Undisturbed soil samples were collected into horizontally-oriented plastic cores (5.4 cm diameter and 11.8 cm height, 270 cm³ volume), coaxially inserted in the place of Kₐ measurements. The collected soil cores were sealed with a rubber cover at both ends and packed into thermal plastic bags to avoid water loss through evaporation and stored in a cooler at 4 °C until laboratory tests were conducted. In laboratory conditions, each soil sample was weighed and was oven dried at 105 °C for 24 h after air drying. Soil bulk density was calculated by dividing the oven-dry sample weight by the core volume. Gravimetric moisture content was calculated for each sample and then converted to volumetric moisture content by multiplying the gravimetric moisture content by soil bulk density. In total, 278 measurements regarding the soil moisture content, soil bulk density and dielectric constants were obtained (the measurement data for each site are shown in Table 1).

Additionally, in the laboratory conditions, the average ash content (AC) for each soil depth was determined using the ignition method. For this reason, the disturbed sub-samples (the number of samples are shown in Table 1) were ignited in a muffle furnace at 550 °C for 4 h. The specific soil particle density (ρₚ) of the analyzed upper layers of six soil profiles was calculated assuming that an increase of ash content of about 1% causes an increase of the ρₚ parameter of about 0.011 g cm⁻³ with respect to the nominal value equal to 1.451 g cm⁻³ [60,61].

### Table 1. Basic soil properties of mucky layers.

<table>
<thead>
<tr>
<th>Sites</th>
<th>Actual Soil Bulk Density (g cm⁻³)</th>
<th>Ash Content (%)</th>
<th>Particle Density (g cm⁻³)</th>
<th>Porosity (cm³ cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Min</td>
<td>Max</td>
<td>SD ²</td>
</tr>
<tr>
<td>I</td>
<td>81</td>
<td>0.3025</td>
<td>0.2256</td>
<td>0.3883</td>
</tr>
<tr>
<td>II</td>
<td>23</td>
<td>0.3123</td>
<td>0.2770</td>
<td>0.3540</td>
</tr>
<tr>
<td>III</td>
<td>45</td>
<td>0.3541</td>
<td>0.2830</td>
<td>0.4271</td>
</tr>
<tr>
<td>IV</td>
<td>27</td>
<td>0.3432</td>
<td>0.3102</td>
<td>0.3780</td>
</tr>
<tr>
<td>V</td>
<td>51</td>
<td>0.2968</td>
<td>0.2211</td>
<td>0.3680</td>
</tr>
<tr>
<td>VI</td>
<td>51</td>
<td>0.3082</td>
<td>0.2361</td>
<td>0.3712</td>
</tr>
</tbody>
</table>

¹ N1, N2—measurement numbers of actual soil bulk density and ash content respectively. ² SD—standard deviation of selected basic properties.

### 2.3. Soil Water Retention Characteristics

The undisturbed soil samples (standard volume 100 cm³) were collected in three replications from the top layers (at depth 7.5 cm) of the six studied peat-muck soil profiles in order to determine soil water retention characteristics (pF curves). The moisture content values in the range of predefined matric potential from 0.4 to 2 pF were determined on a sand table, whereas the amount of water retained at pF: 2.2, 2.5, 2.7, 3.4 and 4.2 was measured in pressure chambers [62]. The pF curves for each soil profile are represented in this study by average values of moisture content at nine predefined soil matric potentials.

### 2.4. Botanical Composition of Plant Cover

For each of the selected sites, the detailed quantitative botanical composition of plant species was measured in 2008 before the first cut (end of May or beginning of June). The full developed plants were cut from four randomly selected areas of 1 m², each located next to the place of the soil property measurements. The biomass samples were stored in paper bags and then the dry weight of each species was determined in laboratory conditions. Based on the gravimetric measurements, the mass percentage of the particular species as a total dry mass was calculated. The list of 70 vascular species in total was made for all selected sites can be seen in Table S1. The floristic type of the meadow
community was determined based on the dominant species, which accounted for at least 30% of the biomass. “Pure” types (domination of one species), “mixed” types (predomination of two species together constituting at least 30% of the biomass) were specified. To characterize the habitat conditions for all meadow communities, the weighted means of Ellenberg’s indices [50,52] were calculated. The proportion (%) of vascular plants in biomass samples was multiplied by the Ellenberg index of this species, i.e., light—L (1–9); soil moisture—F (1–12); soil acidity—R (1–9); soil nitrogen—N (1–9) according to [45], and then the results of multiplication for all species of plant community were summed. The value obtained was divided by 100 and, in this way, the values of individual features for the whole community were obtained. Additionally, the species composition of each site was characterized using the biodiversity Shannon–Weiner’s index [63,64]. The calculation results of the Ellenberg indices (L, F, R, N) and Shannon–Weiner’s indices (H’) are presented in the Supplementary Materials shown in Table S2.

2.5. Data Analysis

Segmented linear regression was applied to establish the relationship between the volumetric soil moisture content ($\theta_v$) and the dielectric constant ($K_a$) for all sites and the whole range of the analyzed dataset. This type of regression is a piecewise linear model, which describes the relationship between response and explanatory variables and is represented by at least two straight lines crossed in a breakpoint [65]. In this study, it was assumed that the established model is continuous, including the breakpoints (Ψ). When there is only one breakpoint, the simplest piecewise regression model can be written as follows [66]:

$$\theta_v = \beta_0 + \beta_1 K_a + e \text{ for } K_a \leq \psi,$$

or

$$\theta_v = \beta_0 + \beta_1 K_a + \beta_2 (K_a - \psi) + e \text{ for } K_a > \psi,$$

where: $\beta_0$ is intercept, $\beta_1$ and $\beta_1 + \beta_2$ are the slopes of the segmented lines ($\beta_2$ can be interpreted as a difference in slopes), and $e$ values are assumed to be independent errors with a zero mean and constant variance. The normality of $e$ values has been tested using the Shapiro–Wilk test and the $p$-value statistic. The iterative techniques included in the R segmented package [67], was applied to identify the breakpoint (Ψ) and to fit the regression model into the collected data. The broken-line model (segmented regression) was used to establish the general broken-line model (GBLM) and site-specific broken-line model (SSBLM).

The root mean square error (RMSE) was used as a measure of the differences between the moisture content predicted by a model(s) and the actually observed values. The RMSE statistic is frequently used to evaluate different calibration curves for the TDR method [25] or for comparisons of the empirical models dependent on different sets of explanatory variables [68]. Additionally, the bias representing the mean difference between the expected (using the TDR calibration model) and the measured values of soil moisture content was determined, as indicated in [22]. The t-test was used for comparison of the statistical difference between the data subsets.

The applicability of the bioindicators for soil moisture prediction was studied using a correlation matrix and linear regression. Moreover, we analyzed the influence of the EIV and H values on the hydro-physical properties of the mucky soil.

Based on the results, with respect to the field soil dielectric properties and bioindicators of the site, the model for soil moisture content determination for upper layers of degraded peatlands was proposed in the form of: Site-specific model discontinuous (SSM-D) and site-specific model with re-fitting of $\psi_r$ value (SSM-R). The calculations were made using the RWeka package [69]. For this purpose, we used the M5Rules algorithm [70,71] belonging to the regression tree method. The basic assumption of this type of algorithm is dividing the entire set of samples by selecting one independent input variable among all and then performing the binary split into at least two child sets, which should increase
the purity of the data compared with its single parent node [72]. The nodes which cannot be further partitioned are called leaves. After growing the tree and the pruning process (removing the leaf with insignificant influence on purity) the linear regression can be fitted for the growing leaf in order to improve the fit of the model [70,73]. The tree construction process, as described above, provides an approximation of function specific for each final node which can lead to sharp discontinuities occurring between adjacent leaves and requires smoothing of the tree after the full set of rules has been produced [74]. During the learning process, we applied a cross-validation scheme [75] to establish the SSM-D model. We used all six possible different combinations of the data from five sites for the training of the M5Rules algorithm, the remaining dataset was used for testing. Each analyzed dataset included continuous variables, i.e., soil moisture content (θv) and dielectric constant (K_a) measured for all investigated sites, as well as categorical covariates represented by bioindices (F, L, N, R, H'). The M5Rules algorithm was run six times to fit the data and the best SSM-D model was selected using the RMSE value for testing of the data. After that the SSM-D was iteratively smoothed and the threshold value ψ_r and θ_vr was determined, finally resulting in SSM-R model determination.

3. Results and Discussion

3.1. Soil Properties and Botanical Composition of Plant Cover

Table 1 summarizes the basic soil properties represented by the measured value of soil bulk density and ash content for the studied mucky horizons (0–15 cm). The average soil particle density as well as the porosity of the analyzed top soil layers was calculated using the basic physical properties of the soil.

The higher values of this feature compared to peat soils can be attributed to the intensive cultivation practice and soil compaction. Despite cultivation differences of peat soils in European countries, the analyzed average soil bulk density values (ρ_b) in the upper layers were comparable with peat soils agriculturally cultivated in England [76], Sweden [77], Norway [78] and Germany [79]. The ash content in the analyzed mucky soil was low and this results from the lack of temporary river flood conditions on the selected sites during moorsh soil development. The average porosity of the studied moorsh horizons was lower, approximately 0.05 to 0.10 cm$^3$ cm$^{-3}$ compared to the underlying original peat deposits [40,42,61].

The mucking process therefore, apart from the basic physical property changes of the soil, seriously affected the soil moisture retention characteristics (Table 2). The presented pF curves of the mucky soils are characterized by a relatively low amount of gravitational water (the difference between porosity and soil moisture content at pF = 2.0) compared to the peat soils.

Table 2. Average soil moisture content at predefined soil matric potentials pF.

<table>
<thead>
<tr>
<th>Sites</th>
<th>pF = 0.4</th>
<th>pF = 1.0</th>
<th>pF = 1.5</th>
<th>pF = 2.0</th>
<th>pF = 2.2</th>
<th>pF = 2.5</th>
<th>pF = 2.7</th>
<th>pF = 3.4</th>
<th>pF = 4.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.803</td>
<td>0.783</td>
<td>0.728</td>
<td>0.645</td>
<td>0.606</td>
<td>0.564</td>
<td>0.547</td>
<td>0.385</td>
<td>0.309</td>
</tr>
<tr>
<td>II</td>
<td>0.808</td>
<td>0.799</td>
<td>0.786</td>
<td>0.694</td>
<td>0.660</td>
<td>0.608</td>
<td>0.571</td>
<td>0.441</td>
<td>0.360</td>
</tr>
<tr>
<td>III</td>
<td>0.775</td>
<td>0.758</td>
<td>0.730</td>
<td>0.649</td>
<td>0.599</td>
<td>0.553</td>
<td>0.530</td>
<td>0.379</td>
<td>0.330</td>
</tr>
<tr>
<td>IV</td>
<td>0.790</td>
<td>0.774</td>
<td>0.727</td>
<td>0.677</td>
<td>0.576</td>
<td>0.502</td>
<td>0.476</td>
<td>0.356</td>
<td>0.303</td>
</tr>
<tr>
<td>V</td>
<td>0.819</td>
<td>0.812</td>
<td>0.724</td>
<td>0.616</td>
<td>0.539</td>
<td>0.521</td>
<td>0.504</td>
<td>0.335</td>
<td>0.248</td>
</tr>
<tr>
<td>VI</td>
<td>0.802</td>
<td>0.750</td>
<td>0.723</td>
<td>0.658</td>
<td>0.612</td>
<td>0.583</td>
<td>0.562</td>
<td>0.393</td>
<td>0.343</td>
</tr>
</tbody>
</table>

Values of Basic Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean 1</th>
<th>SD 2</th>
<th>CV 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.7995</td>
<td>0.0152</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td>0.7793</td>
<td>0.0237</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td>0.7363</td>
<td>0.0245</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>0.6565</td>
<td>0.0270</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>0.6020</td>
<td>0.0347</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>0.5552</td>
<td>0.0391</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>0.5317</td>
<td>0.0362</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>0.3815</td>
<td>0.0361</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>0.3155</td>
<td>0.0392</td>
<td>4.11</td>
</tr>
</tbody>
</table>

1 mean—arithmetic average (cm$^3$ cm$^{-3}$). 2 SD—standard deviation (cm$^3$ cm$^{-3}$). 3 CV—coefficient of variation (%).
In the present study, these soil water retention indicators were in the range between 0.11 and 0.20 cm$^3$ cm$^{-3}$, whereas in the study of [42], in the peat soils, the amount of gravitational water was higher and equal to approximately 0.34 cm$^3$ cm$^{-3}$. However, the soil moisture content at the permanent wilting point (pF = 4.2) of the studied mucky layers was relatively high and remains in the same range as in the peat soils [42]. The obtained results of hydro-physical properties of the mucky layers show an increment of mineral compounds in the soil matrix compared to the peat soils, which indicates that the structure of this organic material dramatically changed due to drainage [80] and cultivation practice, and evolved from fibrous into a grainy or humic structure [61].

The plant communities have been classified in the five floristic types, two “pure” (II, V–VI sites) and three “mixed” (Table 3). I, V and VI sites were characterized by the predominant occurrence of low grasses mainly *Festuca rubra*, and at site VI also sedge (*Carex panicea*).

<table>
<thead>
<tr>
<th>Sites</th>
<th>Ellenberg’s Indices</th>
<th>H$'$</th>
<th>Floristic Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L$^1$</td>
<td>F$^2$</td>
<td>R$^3$</td>
</tr>
<tr>
<td>I</td>
<td>4.96</td>
<td>5.62</td>
<td>2.39</td>
</tr>
<tr>
<td>II</td>
<td>6.80</td>
<td>6.22</td>
<td>4.43</td>
</tr>
<tr>
<td>III</td>
<td>5.50</td>
<td>5.03</td>
<td>1.03</td>
</tr>
<tr>
<td>IV</td>
<td>6.37</td>
<td>5.12</td>
<td>2.83</td>
</tr>
<tr>
<td>V</td>
<td>3.43</td>
<td>5.03</td>
<td>3.35</td>
</tr>
<tr>
<td>VI</td>
<td>3.14</td>
<td>5.50</td>
<td>3.21</td>
</tr>
</tbody>
</table>

Table 3. The habitat conditions and the Shannon–Weiner’s index of the meadow communities in the examined sites (mean values).

The Shannon–Wiener index in the analyzed plant communities was moderate [48] and oscillates (apart from site VI) in the range from 2.09 to 2.43. VI site was characterized by low diversity indices mainly because of the lowest number of species (20 species) and the predominance of two main species (about 73% of biomass). The highest values of the EIV measures were noted for the II site. The floristic composition of plant communities of the studied sites indicates the diversity of ecological conditions of the habitats (Table 3) thereby confirming the assumptions formulated in the materials and methods section focusing on the biological diversity of the plant cover.

3.2. Analysis of Existing Relationships between Dielectric Constant and Soil Moisture Content in Organic Soils

The relationship between the measured soil moisture content and the dielectric constant in respect to the field site is presented in Figure 2. Data are illustrated together with the TDR calibration curves for peat soils available in the literature listed in Table 4. A simple graphical inspection of the presented features indicates that any of the applied calibration equations do not fit the whole range of our dataset. The best prediction of the TDR soil moisture content for moorsh soil from the Biebrza River Valley can be approximated using equations P4 and PS2 [22].

The RMSE for these models was nearly 0.06 cm$^3$ cm$^{-3}$. Despite the reasonably good results of the RMSE statistic, the mentioned equations gave a satisfactory prediction of soil moisture content at different ranges of dielectric constant changes. The most appropriate literature model P4 enabled a reasonably good approximation of the soil moisture content in the range of dielectric constant between 10 and 20. The highest overestimation of soil moisture content (of about 0.15 cm$^3$ cm$^{-3}$) occurred when the dielectric constant was equal to 30. Figure 2 also shows the measured data of the soil moisture content and the dielectric constant and their fitting using the logarithmic model (NLOwn) The following form of equation was obtained:

$$\theta_v = 0.2417 \cdot \ln(K_a) - 0.2173.$$  \hspace{1cm} (4)
The calculated RMSE value was essentially lower (0.0452 cm\(^{-3}\)) in comparison to the P4 and PS2 models. However, Equation (4) overestimated the moisture content by as much as 0.15 cm\(^{-3}\) in the middle range of the dielectric constant changes.

![Figure 2](image-url)

**Figure 2.** Measured data of dielectric constant and soil moisture content for different sites together with own empirical relationship (NL\(_{own}\)) and relationships listed in Table 4.

**Table 4.** Selected literature calibration models for time domain reflectometry (TDR) methods and root mean square error (RMSE) statistics with respect to all collected dataset, as well as for data of the “dry” and “wet” segment of the general broken-line model (GBLM).

<table>
<thead>
<tr>
<th>NO</th>
<th>Calibration Models</th>
<th>Soil</th>
<th>(\theta_v) Range</th>
<th>(K_a) Range</th>
<th>Independent Variable</th>
<th>Function</th>
<th>Root Mean Square Error RMSE (cm(^{-3}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overall Range of Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>“dry”</td>
</tr>
<tr>
<td>1</td>
<td>P1 [23]</td>
<td>O</td>
<td>0.03–0.55</td>
<td>–3–35</td>
<td>(K_a)</td>
<td>p</td>
<td>0.4262</td>
</tr>
<tr>
<td>2</td>
<td>P2 [29]</td>
<td>SM</td>
<td>0.20–0.90</td>
<td>–5–60</td>
<td>(K_a)</td>
<td>p</td>
<td>0.0704</td>
</tr>
<tr>
<td>3</td>
<td>P3 [30]</td>
<td>O</td>
<td>0.00–0.75</td>
<td>–1–60</td>
<td>(K_a)</td>
<td>p</td>
<td>0.0748</td>
</tr>
<tr>
<td>4</td>
<td>P4 [31]</td>
<td>SC</td>
<td>0.30–0.80</td>
<td>–10–65</td>
<td>(K_a)</td>
<td>p</td>
<td>0.0571</td>
</tr>
<tr>
<td>5</td>
<td>P5 [32]</td>
<td>B</td>
<td>0.60–0.90</td>
<td>–29–75</td>
<td>(K_a)</td>
<td>p</td>
<td>0.1130</td>
</tr>
<tr>
<td>6</td>
<td>PS1 [33]</td>
<td>SM</td>
<td>0.00–0.98</td>
<td>–3–75</td>
<td>sqrt ((K_a))</td>
<td>p</td>
<td>0.1090</td>
</tr>
<tr>
<td>7</td>
<td>PS2 [33]</td>
<td>SM</td>
<td>0.00–0.98</td>
<td>–3–75</td>
<td>sqrt ((K_a))</td>
<td>p</td>
<td>0.0586</td>
</tr>
<tr>
<td>8</td>
<td>LS1 [81]</td>
<td>O</td>
<td>0.10–0.80</td>
<td>–2–40</td>
<td>sqrt ((K_a))</td>
<td>l</td>
<td>0.0958</td>
</tr>
<tr>
<td>9</td>
<td>LS2 [82]</td>
<td>Fo</td>
<td>0.00–0.70</td>
<td>–1–40</td>
<td>sqrt ((K_a))</td>
<td>l</td>
<td>0.0908</td>
</tr>
<tr>
<td>10</td>
<td>LS3 [83]</td>
<td>SM</td>
<td>0.00–0.95</td>
<td>–2–73</td>
<td>sqrt ((K_a))</td>
<td>l</td>
<td>0.0653</td>
</tr>
<tr>
<td>11</td>
<td>NL(_{own}) *</td>
<td>Mu</td>
<td>0.22–0.81</td>
<td>6–61</td>
<td>ln((K_a))</td>
<td>nl</td>
<td>0.0452</td>
</tr>
<tr>
<td>12</td>
<td>GBLM *</td>
<td>Mu</td>
<td>0.22–0.81</td>
<td>6–61</td>
<td>(K_a)</td>
<td>ls</td>
<td>0.0405</td>
</tr>
</tbody>
</table>

\(\theta_v\)—volumetric soil moisture content (cm\(^{-3}\)); \(K_a\)—dielectric constant (-); p—polynomial, l—linear; nl—nonlinear; ls—segmented linear; *—this study; O—organic soils, SC—Sphagnum, Carex peat SM—Sphagnum moss, B—blanket peat, Fo—forest soils, Mu—muck soils.

### 3.3. General and Site-Specific Relationships between Dielectric Constant and Soil Moisture Content

Based on the graphical inspection and simple statistical analysis, we decided to select the continuous broken-line model as an approximation to describe the non-linear relationship between soil moisture content and dielectric constant of the mucky soil (for all measured data of \(\theta_v\)
and $K_a$). For this purpose, the segmented linear regression was applied and the model parameters (Equations (1) and (3)) were estimated iteratively [67]. The results of the statistical computation indicated that the established continuous general broken-line model (GBLM) included one threshold parameter ($\Psi$), and can be written as:

$$\theta_v = 0.2335 + 0.01178 \cdot K_a \text{ for } K_a \leq \psi,$$

$$\theta_v = 0.2335 + 0.01178 \cdot K_a - 0.0046 \cdot (K_a - \psi) \text{ for } K_a > \psi.$$  

After rearrangement the Equation (6) can be written in its final form as follows:

$$\theta_v = 0.2335 + 0.01178 \cdot \psi + 0.007179 \cdot (K_a - \psi) \text{ for } K_a > \psi.$$  

The break-point ($\Psi$) for the analyzed dataset was determined for $K_a$ equal to 31.49 (Figure 3). The 95% confidence interval for this parameter was wide and ranged from 27.22 to 35.76 (Figure 3). Residuals from the model (Equations (5) and (6), or (5) and (7)) were normally distributed around the average which was confirmed by the Shapiro–Wilk testing statistic ($W = 0.9919$) and the $p$-value (0.1302). The RMSE value for the GBLM equations equals to 0.0405 cm$^3$ cm$^{-3}$ and is essentially lower than the estimates determined for the equations indicated in Table 4. This type of broken-line relationship is often used to assess the threshold value where the effect of the covariate changes. In the fields of hydrology and soil hydrology, the piecewise linear regressions (PLR) and the determined threshold values are used to explain the influence of the vegetation cover fraction on the reduction of runoff [84] and the erodibility coefficient [85]. In the water study, the two-segmented broken-line equation is applied for predicting water electrical conductivity using a portable X-ray fluorescence meter [86]. The meaning of a break-point in this study of the TDR calibration curve can be explained as the point depending on the soil water retention characteristic (horizontal lines on Figure 4). The threshold value ($\Psi$) with the confidence interval in general was included between horizontal lines which indicated the moisture contents at matric potential between pF values from 2.0 to 2.7. The average moisture content at break point ($\theta_v = 0.6044$ cm$^3$ cm$^{-3}$) was approximately equal to soil moisture content at pF = 2.2 ($\theta_v = 0.6020$ cm$^3$ cm$^{-3}$). The validity of the threshold value can be explained by bulk density changes (Figure 4).

![Figure 3](image_url)

**Figure 3.** Field measurements and fitted GBLM for the relationship between soil moisture content and dielectric constant. The horizontal dotted black lines indicate the soil moisture content at predefined pressure head. The red vertical line indicates the average value of the breaking point ($\Psi = 31.49$) whereas the red dotted vertical line is the 95% confidence interval.
The soil bulk density of the two segments was statistically different. For the “wet” range (segment wet above the threshold \( \Psi \)), the average bulk density was equal to 0.31 g cm\(^{-3}\), whereas at the “dry” range (segment dry) its estimates are equal to 0.33 g cm\(^{-3}\). It can be concluded that below the breakpoint of the \( \theta_v(K_a) \) characteristic, the significant increase of bulk density indicates the essential soil volume changes. The field study conducted for similar soils confirmed that, apart from the soil moisture contents changes [87], the soil bulk density also changes in space and time [40].

Despite improvements in the GBLM model in comparison to the literature equations, the bias of the calibration equation depends on site-specific data measurements of \( \theta_v(K_a) \). The segmented model gave a weak prediction of the soil moisture content for site II (bias equal to approximately 0.044 cm\(^{-3}\)) as well as in the dry range of the soil moisture content in sites VI, V and IV.

![Figure 4. Boxplot of soil bulk density for “dry” and “wet” segments of the measured dataset (dotted line indicates overall average soil bulk density). The scattered point indicates the soil bulk density value with respect to sites. The dotted horizontal line indicates the overall soil bulk density average. The diamond symbol indicates the average soil bulk density for each segment of the dataset. The red dots indicate the outliers.](image)

The analysis of the GBLM model limitations forced us to develop the site-specific broken-line model (SSBLM). For this purpose, the linear regression lines were fitted for the analyzed sites in respect to data below and above the threshold value (Figure 3). The obtained slopes (\( \beta_1, \beta_1 + \beta_2 \)) and intercepts (\( \beta_0 \)) enabled the calculation of the breakpoint value separately for the soil from the specific sites. The parameters of the SSBLM models are summarized in Table 5.

Fitted lines below the threshold (“dry” range) were characterized by similar slope values (average value equal to 0.0136 and CV = 6.31%). In the “wet” range, slopes (\( \beta_1 + \beta_2 \)) were generally flatter than in the “dry” range (\( \beta_1 \)) and relatively more variable (CV = 24%) mainly because of the lowest slope of the \( \theta_v(K_a) \) relationship for site II. Except for this site, the slope parameters in the “wet” range were also not essentially different. The \( \Psi \) parameter represents a wide range of changes in the dielectric constant (Table 5) which directly influences the really wide scatter of the soil moisture content (\( \theta_v \)) estimated at break point (\( \Psi \)). Despite the low CV value, the moisture content at the break points (\( \theta_v(\Psi) \)) ranged from 0.50 to 0.70 cm\(^{-3}\) which represents a relatively wide scatter for this type of soil. The comparison of \( \theta_v(\Psi) \) values with the range of moisture content of pF curves indicated that the threshold values were included in the range of the soil matric potential approximately between pF = 2.0 (site II) to pF = 2.5 (site IV).
Table 5. Site-specific broken-line model parameters (SSBLM) for the TDR method.

<table>
<thead>
<tr>
<th>Sites</th>
<th>( \beta_0 ) (cm³ cm⁻³)</th>
<th>( \beta_1 ) (cm³ cm⁻³)</th>
<th>( \beta_1 + \beta_2 ) (cm³ cm⁻³)</th>
<th>( \Psi )</th>
<th>( \theta_{\Psi} ) (cm³ cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.2157</td>
<td>0.0125</td>
<td>0.00800</td>
<td>27.42</td>
<td>0.5585</td>
</tr>
<tr>
<td>II</td>
<td>0.2456</td>
<td>0.0129</td>
<td>0.00380</td>
<td>35.25</td>
<td>0.7003</td>
</tr>
<tr>
<td>III</td>
<td>0.1941</td>
<td>0.0144</td>
<td>0.00660</td>
<td>27.26</td>
<td>0.5866</td>
</tr>
<tr>
<td>IV</td>
<td>0.2587</td>
<td>0.0133</td>
<td>0.00640</td>
<td>18.29</td>
<td>0.5020</td>
</tr>
<tr>
<td>V</td>
<td>0.1330</td>
<td>0.0139</td>
<td>0.00810</td>
<td>34.98</td>
<td>0.6192</td>
</tr>
<tr>
<td>VI</td>
<td>0.0740</td>
<td>0.0147</td>
<td>0.00613</td>
<td>42.14</td>
<td>0.6935</td>
</tr>
</tbody>
</table>

Values of basic statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean 1</th>
<th>SD 2</th>
<th>CV 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean 1</td>
<td>0.1869</td>
<td>0.01362</td>
<td>30.89</td>
</tr>
<tr>
<td>SD 2</td>
<td>0.0709</td>
<td>0.00086</td>
<td>8.32</td>
</tr>
<tr>
<td>CV 3</td>
<td>37.93</td>
<td>6.31</td>
<td>24.12</td>
</tr>
</tbody>
</table>

\(^1 \text{mean—arithmetic average.} \ ^2 \text{SD—standard deviation.} \ ^3 \text{CV—coefficient of variation (%).}

Figure 5 shows the measured data of the soil moisture content and the dielectric constant for the specific sites and their fitting using the segmented model (SSBLM) and the logarithmic equations (NL type).

The SSBLM model developed separately for the soils from different sites indicates the reasonably good results of fitting in all ranges of the \( \theta_{\Psi}(K_a) \) relationship changes. The RMSE values were varied depending on the site data. For the calibration curves from sites III, V and VI, the values varied from 0.03 to 0.04 cm³ cm⁻³. The logarithmic models underestimated the soil moisture content by as much as 0.07 cm³ cm⁻³ in the wet range of calibration curve. From this reason for further analysis, we applied broken-line model parameters as an approximation of the non-linear relationship between \( K_a \) and \( \theta_{\Psi} \).
did not exceed 0.03 cm$^3$ cm$^{-3}$. The RMSE values for the rest of the developed SSBLM equations were slightly higher and varied from 0.036 cm$^3$ cm$^{-3}$ (site I) to 0.041 cm$^3$ cm$^{-3}$ (site IV). It should be stressed that logarithmic type models developed for specific site were less accurate and the RMSE value was generally higher (of about 0.005 to 0.012 cm$^3$ cm$^{-3}$) in comparison to the SSBLM models. The logarithmic models underestimated the soil moisture content by as much as 0.07 cm$^3$ cm$^{-3}$ in the wet range of calibration curve. From this reason for further analysis, we applied broken-line model parameters as an approximation of the non-linear relationship between $K_a$ and $\theta_v$.

3.4. Applicability Analysis of the Bio-Indices in Soil Moisture Prediction and Hydro-Physical Properties of Mucky Soils

Feasibility analysis of the calibration model development for the moorsh moisture content prediction as a function of dielectric constant with the participation of the covariates represented by bioindices is illustrated using the correlation coefficients (Figure 6).

The broken-line model parameters (Table 5) were correlated with the bioindices (Table 3). The highest correlation coefficient has been noted between intercept ($\beta_0$) and indicator L of the Ellenberg scale. However, the $\beta_0$ also correlates with “productivity” of the ecosystem (N) as well as with the Shannon–Weiner index ($H'$). The slope in the “dry” range of the calibration curves ($\beta_1$) was strongly but negatively correlated with “productivity” of the ecosystem (N) determined on the basis of the plant cover as well as positively correlated with the ash content (AC) of the analyzed soils. Nevertheless, both of the explained covariates (AC and N) were moderately and negatively correlated, which can suggest that the increase of the AC amount in the soil causes reduction of the “productivity” of the analyzed ecosystems. The slope in the “wet” range of the calibration curves ($\beta_1 + \beta_2$) can be approximated using the soil moisture substrate indicator (F), which in the case of our studied sites describes fresh soil with moderate wetness [51]. Due to the low range of the F indicator changes and moderate variation of the $\beta_1 + \beta_2$ parameter, the prediction of the slope in the “wet” range of the calibration curve can be more complex than simple application of the linear relationship. The estimation of this parameter is especially important for the organic soils where soil moisture changes in the wet range strongly influence the soil degradation processes and physical soil properties. On the other hand, the $\beta_1 + \beta_2$ parameters were highly correlated with soil pF curve characteristics. From the data presented in Figure 6, it can be observed that these parameters were also negatively correlated with soil moisture contents measured at the predefined soil matric potentials ranging from pF = 1.5 to pF = 4.2. Moreover, it should be stressed that the moisture of the soil substrate (indicator F) was also highly correlated with the moisture content of the pF curves (Figure 6) which, in general, is in agreement with the statements made in [50]. Based on the correlation analysis, it can be noted that the slope in the “wet” range of the field calibration curves can be fairly well determined using the moisture indicator of the Ellenberg scale, or more precisely using the soil moisture content at pF = 2 as a soil type (site) specific predictor.

The broken-line model parameter $\Psi$ was correlated with bioindicators of soil cover, physical soil properties and soil water retention characteristic. The highest r value was determined between $\Psi$ and the Shannon–Weiner index. This could suggest that the variety of the plant species, which are typical for the site location, influences the slope separation point of the TDR calibration curves. This was especially visible in the case of two sites: VI and IV. In the former, the $\Psi$ value was the highest and soil cover is mainly occupied by two plant species (*Festuca rubra* and *Carex panicea*) which affected the lowest value of the $H'$ index. On the other hand, at site IV the highest $H'$ value was observed (2.4) which indicates a condition of greater biodiversity (with a predominance of dicotyledonous plants), however, the $\Psi$ value was the lowest (18.29). Although the $\Psi$ value is was highly correlated with the Shannon–Weiner index, its value can also be carefully approximated by physical soil properties. Our data indicate that the increase of $\Psi$ value was related to the increase of average soil porosity and soil moisture status at pF = 2.5 and 2.7. The $\Psi$ value was also highly correlated with the ratio between
3.4. Applicability Analysis of the Bio-Indices in Soil Moisture Prediction and Hydro-Physical Properties of Water

parameters (Equations (1) and (3)) using bioindices (L, F, N, H') determined on the basis of the general calibration curve for predicting moorsh moisture content using the TDR method for a specific approach indicates that, for general purposes, both slopes (linear equations plotted in Figure 7a–d. On the other hand, the simplification of the segmented line studied sites describes fresh soil with moderate wetness [51]. Due to the low range of the F indicator changes and moderate variation of the prediction as a function of dielectric constant with the participation of the covariates represented by soil substrate (indicator F) was also highly correlated with the moisture content of the pF curves potentials ranging from pF = 1.5 to pF = 4.2. Moreover, it should be stressed that the moisture of the plant cover as well as positively correlated with the ash content (AC) of the analyzed soils. Nevertheless, both of the explained covariates (AC and N) were moderately and negatively correlated, which can suggest that the increase of the AC amount in the soil causes reduction of the productivity vs. the slope  \( \beta \) of SSBLM; (b) the Ellenberg index F – moisture vs. the slope \( \beta_1 + \beta_2 \) of SSBLM; (c) the Ellenberg index N – productivity vs. the slope \( \beta_1 \) of SSBLM; (d) the biodiversity index – H' vs. the breakpoint (Y) of SSBLM.

Figure 6. Correlation graph of the analyzed variables. Arrows in green and red indicate statistically significant positive and negative correlations. Dark yellow arrows indicate not statistically significant correlations. The labelled cells indicate statistically significant correlation between different types of variables regarding the state of the habitats, SSBLM and soil hydro-physical properties.

Figure 7a–d shows simple regression lines allowing prediction of the broken-line model parameters (Equations (1) and (3)) using bioindices (L, F, N, H') determined on the basis of the measured plant species’ composition for the selected sites.

Figure 7. Linear regression lines between parameters of the broken-line model and explanatory variables: (a) the Ellenberg index L - light vs. the intercept \( \beta_0 \) of SSBLM; (b) the Ellenberg index N – productivity vs. the slope \( \beta_1 \) of SSBLM; (c) the Ellenberg index F – moisture vs. the slope \( \beta_1 + \beta_2 \) of SSBLM; (d) the biodiversity index – H’ vs. the breakpoint (Y) of SSBLM.
3.5. Calibration Model Development of Soil Moisture Content as a Function of Dielectric Constant and Botanical Indices of Soil Cover Using the Regression Tree Method

Based on the discussion presented in Section 3.4, we can state that the development of the general calibration curve for predicting moorsh moisture content using the TDR method for a specific site is very complex. We can predict each of the broken-line parameters one by one using simple linear equations plotted in Figure 7a–d. On the other hand, the simplification of the segmented line approach indicates that, for general purposes, both slopes ($\beta_1$, $\beta_1 + \beta_2$) can be averaged (Equations (5) and (7)), however, this requires finding out the threshold $\Psi$ values. The quality of both attempts depends mainly on the prediction of the SSBLM parameters based on the bioindicators of the soil cover as the simplest explanatory variables. However, error in the estimation of the parameters affects the accuracy of the soil moisture prediction. Therefore, taking into account the temporal stability of the investigated ecosystems (supported by water supply and farmer practice), we developed a calibration equation which in one approach includes the relationship between soil moisture content and dielectric constant and the bioindicators’ influence. For this reason, we used the MSRules algorithm belonging to regression tree methods (see description in Section 2.5).

Table 6 summarizes the cross-validation scheme. In total we trained six models SSM-D1, . . . , SSM-D6. The results of the model selection approach indicate that, depending on training dataset, the SSM-D models included different explanatory variables for soil moisture prediction. The performance of the cross-validation split models is described by RMSE$_{\text{train}}$. The main independent variable in each training model, which binary split our dataset into two child sets, was dielectric constant ($K_a$). In the second stage of the cross-validation scheme we reduced the number of variables to the commonly appearing covariates in the training model stage. For this reason, different sets of explanatory variables were tested. Finally, we found the appropriate combination of the variables i.e., $K_a$, F, L and H’ (SM—set of model in the model for which at each split of the data the RMSEcv values were close to the RMSE$_{\text{train}}$ (Table 6). Additionally, in Table 6 the RMSEcv values determined in cross-validation procedure performed for the GBLM model are also included. A last SSM-D model fitting was conducted using the mentioned combination of variables (SM) as single fit to all data. Fortunately, the resulted SSM-D model includes two nodes for which the linear regression models have been established (Figure 8).

Table 6. Summary of the cross-validation approach for establishing the site-specific model discontinuous (SSM-D) calibration model for the TDR method in mucky soils.

<table>
<thead>
<tr>
<th>Model</th>
<th>Training Dataset</th>
<th>$n_{\text{train}}$</th>
<th>$n_{\text{test}}$</th>
<th>Model Variables</th>
<th>RMSE$_{\text{train}}$ (cm$^3$ cm$^{-3}$)</th>
<th>RMSE$_{\text{cv}}$ (cm$^3$ cm$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSM-D1</td>
<td>I, II, III, IV, V</td>
<td>227</td>
<td>51</td>
<td>$K_a$, F, L, N, H’</td>
<td>0.0351</td>
<td>0.0351</td>
</tr>
<tr>
<td>SSM-D2</td>
<td>II, III, IV, V, VI</td>
<td>196</td>
<td>82</td>
<td>$K_a$, F, N, R, H’</td>
<td>0.0321</td>
<td>0.0326</td>
</tr>
<tr>
<td>SSM-D3</td>
<td>I, III, IV, V, VI</td>
<td>256</td>
<td>22</td>
<td>$K_a$, F, N, L, R, H’</td>
<td>0.0318</td>
<td>0.0426</td>
</tr>
<tr>
<td>SSM-D4</td>
<td>I, II, IV, V, VI</td>
<td>233</td>
<td>45</td>
<td>$K_a$, F, L, N</td>
<td>0.0362</td>
<td>0.0363</td>
</tr>
<tr>
<td>SSM-D5</td>
<td>I, II, III, V, VI</td>
<td>251</td>
<td>27</td>
<td>$K_a$, F, L, N, H’</td>
<td>0.0301</td>
<td>0.0304</td>
</tr>
<tr>
<td>SSM-D6</td>
<td>I, III, IV, VI</td>
<td>227</td>
<td>57</td>
<td>$K_a$, F, L, N, R, H’</td>
<td>0.0363</td>
<td>0.0368</td>
</tr>
</tbody>
</table>

I, II, III, IV, V, VI—indicate the site data selected for model training; RMSE$_{\text{train,cv}}$ (cm$^3$ cm$^{-3}$) for training model and for cross-validation scheme respectively; SM—set of model parameters including $K_a$, F, H’ and L variables.

This final fitting resulted in a two-slopes calibration model (SSM-D) split by the parent node ($K_a = 30.845$) similarly to the segmented regression model (Equations (5) and (7)). It should be stressed that the value of 30.845 is included in the confidence interval of the GBLM model. The slope of the line below the splitting point (dry range) was equal to 0.0127 (regression parameter at $K_a$ variable) and was also comparable with its value determined using segmented regression (GBLM). The position of this line along the soil moisture content axis (the value of intercept) was improved by the H’ and L values (specific measures) of the site. Above the $K_a = 30.845$, the slope was equal to 0.0075 (nearly the same as in GBLM), however, the position of the line can be determined using covariates represented
by F, L, bioindicators. However, it should be stressed that the influence of the L value was marginal for soil moisture prediction (wet range) which results from the regression coefficient at this variable.

Following the statements in [74], the regression tree indicated discontinuities occurring between adjacent leaves. In the post-processing procedure, we iteratively found out the threshold value $\psi_r$ which enables establishing a continuous soil moisture content prediction model in the whole range of the dielectric constant changes (site-specific model with re-fitting of $\psi_r$ value, SSM-R). This is especially important for monitoring of this feature. The uncertainty of this kind of the procedure was tested using the RMSE value as a criterion of losing quality in the model compared to the SSM-D model. The $\psi_r$ values are in fact the re-established criterion of the splitting of the dataset which enables, again, application of the linear equations developed for both leaves (Figure 8).

The main conclusion from the modelling attempt is that the SSM-D model is structurally similar to the GBLM model (the splitting value $K_a = 30.845$ is similar to the $\psi_r$ value of the GBLM model) and this results in discontinuities occurring between adjacent leaves. However, we propose to use the SSM-D model in the case when the TDR user does not want to perform a refitting (SSM-R). In the SSM-R model approach, additionally in the post-processing procedure, the threshold value in the form of $\psi_r$ parameter is obtained. In this sense, the SSM-R model is the smoothed form of the SSM-D model and can be treated as a site-specific model similar to SSBLM. The inconveniences of the modelling approach as well as the importance of threshold value in mucky soils should be further investigated.

Figure 8. Diagram representing graphical explanation of the SSM-D and site-specific model with re-fitting of $\psi_r$ value (SSM-R) models.

Table 7 presents the RMSE value estimation according to the different calibration models for the TDR method. The results of the data presented in the table showed that the RMSE values for the proposed model (SSM-R) produce nearly the same error estimators as the SSM-D algorithm.

Table 7. Values of RMSE for the different types of TDR calibration models established in this study.

<table>
<thead>
<tr>
<th>Sites</th>
<th>RMSE $^1$ for Different Calibration Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GBLM</td>
</tr>
<tr>
<td>overall</td>
<td>0.0405</td>
</tr>
<tr>
<td>I</td>
<td>0.0361</td>
</tr>
<tr>
<td>II</td>
<td>0.0623</td>
</tr>
<tr>
<td>III</td>
<td>0.0264</td>
</tr>
<tr>
<td>IV</td>
<td>0.0559</td>
</tr>
<tr>
<td>V</td>
<td>0.0405</td>
</tr>
<tr>
<td>VI</td>
<td>0.0349</td>
</tr>
</tbody>
</table>

$^1$ RMSE—root mean square error (cm$^3$ cm$^{-3}$).
Based on this observation, we can state that the model, which is graphically presented in Figure 8, can be used for continuous prediction of moorsh soil moisture content based on dielectric constant measurements and complex information about bioindices. The form of the proposed model is generally in agreement with the main results of this study which showed the possibility to establish a two-segmented broken-line calibration model (GBLM, Equations (5) and (7)) for the TDR method.

4. Conclusions

Monitoring of soil moisture content in peat soil is a crucial issue which enables study processes related to peat soil degradation, particularly in the vadose zone. The presented study indicated that simple regression models which relate the soil dielectric constant with moisture content did not properly address determination of soil moisture content in degraded peatlands in field conditions. This was especially apparent in the intermediate range of dielectric constant changes (from approximately 20 to 40), where the best existing empirical models presented in the literature overestimated the moisture content by as much as 0.15 cm$^3$ cm$^{-3}$. The main expectation from TDR users is to be able to obtain relatively accurate moisture content determination in field conditions. The results of our investigations showed that the proposed calibration curve (GBLM) composed of two segments meets those requirements with the acceptable error of soil moisture determination in field conditions (RMSE = 0.04 cm$^3$ cm$^{-3}$). Despite the relatively good approximation of moisture content by the broken-line models (GBLM, SSBLM) developed during this study, these are only empirical equations, which should be subject to an interpretation derived from further investigation of the segmented line parameters. It is important to stress that in the case of the GBLM model, the existence of the two independent slopes separated by the threshold value $\psi$ has been proved empirically by the statistically significant difference in soil bulk densities in both ranges of dielectric constant changes. We propose that further investigation should be focused on the assessment of the influence of soil air content on the slope of calibration equation in the dry range and assessment of the physical meaning of the threshold value $\psi$.

The conducted research also proved that, in site-specific calibration equations (SSBLM), the broken-line model parameters, in general, are correlated with the following bioindices: L, F, R, N values and biodiversity index (H$'\prime$). This means that EIV and H$'\prime$ indicators can be useful in the assessment of soil moisture content in degraded peatlands using the TDR technique. It should be stressed that an increase of slope value ($\beta_1 + \beta_2$) in the “wet” range of calibration equation was also correlated with a decrease of moisture content at pressure head corresponding to pF = 2. This tendency was also observed in the correlation between the F (substrate moisture) indicator and $\beta_1 + \beta_2$ slope. Moreover, it can be concluded that the F classifier of the habitats determined only on the basis of plant species’ composition is also highly correlated with the soil water retention characteristic at the investigated sites. Therefore, the performed research confirmed that the plant cover of the habitat depends on the water-related properties of the soil peatlands.

The analysis of the linear equations included in the SSM-D and SSM-R models indicates that the soil moisture content, aside from dielectric constant, was also dependent on the biodiversity index (H$'\prime$) and light intensity (L) in the “dry” range of calibration curves. In the “wet” range of the calibration curve, the proposed model contains, apart from the variable ($K_a$), covariates represented by the following parameters: Substrate moisture (F), and light intensity (L).

The results of our study showed that bioindices, as well as their basic application, can also help in the identification of the specific soil condition which can improve our empirical knowledge of soil moisture regimes. Therefore, we hope the quantification of the EIV and H$'\prime$ values of degraded peatlands as simply measured categorical variables will enable a TDR user to make decisions on what type of calibration equation can be selected for soil moisture monitoring in this type of ecosystem.
Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4441/10/12/1842/s1, Table S1: List of the vascular plants and their share (%) in the analyzed sites; Table S2: Calculation results of the Ellenberg indices (L, F, R, N) and Shannon-Weiner indice (H’).

Author Contributions: The main conception of this article should be attributed to T.G. and B.P.; conducted field data collection T.G., R.O., B.P.; laboratory analyses of soil hydro-physical properties J.S., E.P., D.S.; botanical analyses of plant cover and calculations of bioindicators M.J., B.P.; statistical data analysis T.G., J.S., R.O., B.P.; writing of paper T.G. and J.S.; all authors read, corrected and approved the final manuscript.

Funding: This research was founded by the Polish Ministry of Science and Higher Education for statutory activities of the Department of Environmental Improvement, Warsaw University of Life Sciences—SGGW.

Acknowledgments: We are grateful to the technical staff of the Experimental Department Biebrza of the Institute of Technology and Life Science for their assistance in the field work.

Conflicts of Interest: The authors declare no conflict of interest.

References


47. Paerl, H.W.; Valdes, L.M.; Pinckney, J.L.; Piehler, M.F.; Dyble, J.; Moisander, P.H. Phytoplankton photopigments as indicators of estuarine and coastal eutrophication. AIBS Bull. 2003, 53, 953–964. [CrossRef]


50. Schaffers, A.P.; Sykora, K.V. Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction: A comparison with field measurements. J. Veg. Sci. 2000, 11, 225–244. [CrossRef]

51. Simmel, J.; Bassler, C.; Poschlod, P. Ellenberg indicator values for macrofungi—a methodological approach and first applications. Fungal Ecol. 2017, 27, 202–212. [CrossRef]


53. Diewroew, M. Species indicator values as an important tool in applied plant ecology—A review. Basic Appl. Ecol. 2003, 4, 493–506. [CrossRef]


