Application of Artificial Neural Networks Sensitivity Analysis for the Pre-Identification of Highly Significant Factors Influencing the Yield and Digestibility of Grassland Sward in the Climatic Conditions of Central Poland

Gniewko Niedbała 1,*, Barbara Wróbel 2, Magdalena Piekutowska 3, Waldemar Zielewicz 4, Anna Paszkiewicz-Jasińska 2, Tomasz Wojciechowski 1 and Mohsen Niazian 5

1 Department of Biosystems Engineering, Faculty of Environmental and Mechanical Engineering, Poznań University of Life Sciences, Wojska Polskiego 50, 60-627 Poznań, Poland; tomasz.wojciechowski@up.poznan.pl
2 Institute of Technology and Life Sciences-National Research Institute, Falenty, 3 Hrabska Avenue, 05-090 Raszyn, Poland; b.wrobel@itp.edu.pl (B.W.); a.paszkiewicz@itp.edu.pl (A.P.-J.)
3 Department of Geocology and Geoinformation, Institute of Biology and Earth Sciences, Pomeranian University in Słupsk, Partyzantów 27, 76-200 Słupsk, Poland; magdalena.piekutowska@apsl.edu.pl
4 Department of Grassland and Natural Landscape Sciences, Poznań University of Life Sciences, Dojazd 11, 60-632 Poznań, Poland; waldemar.zielewicz@up.poznan.pl
5 Crop and Horticultural Science Research Department, Kurdistan Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEEO), Sanandaj 6616936311, Iran; mniazian@ut.ac.ir

* Correspondence: gniewko.niedbala@up.poznan.pl

Abstract: Progressive climate changes are the most important challenges for modern agriculture. Permanent grassland represents around 70% of all agricultural land. In comparison with other agroecosystems, grasslands are more sensitive to climate change. The aim of this study was to create deterministic models based on artificial neural networks to identify highly significant factors influencing the yield and digestibility of grassland sward in the climatic conditions of central Poland. The models were based on data from a grassland experiment conducted between 2014 and 2016. Phytophenological data (harvest date and botanical composition of sward) and meteorological data (average temperatures, total rainfall, and total effective temperatures) were used as independent variables, whereas qualitative and quantitative parameters of the feed made from the grassland sward (dry matter digestibility, dry matter yield, and protein yield) were used as dependent variables. Nine deterministic models were proposed Y_G, DIG_G, P_G, Y_GB, DIG_GB, P_GB, Y_GC, DIG_GC, and P_GC, which differed in the input variable and the main factor from the grassland experiment. The analysis of the sensitivity of the neural networks in the models enabled the identification of the independent variables with the greatest influence on the yield of dry matter and protein as well as the digestibility of the dry matter of the first regrowth of grassland sward, taking its diverse botanical composition into account. The results showed that the following factors were the most significant (rank 1): the average daily air temperature, total rainfall, and the percentage of legume plants. This research will be continued on a larger group of factors influencing the output variables and it will involve an attempt to optimise these factors.

Keywords: sensitivity analysis; deterministic modeling; artificial neural networks; meadow sward; botanical composition of sward; air temperature; precipitation

1. Introduction

Permanent grasslands (PG) are grass ecosystems that cover slightly more than 3 billion ha worldwide, which is about 70% of all agricultural land [1,2]. In Poland, they cover...
about 3.1 million ha, i.e., 21.4% of agricultural land, where the majority (about 73%) are permanent grasslands [3,4]. It is a much smaller area than in other European countries, e.g., on the British Isles or in the Alpine countries, where permanent grasslands predominate in agricultural areas [5]. Grasslands provide a wide range of key ecosystem services, including support (e.g., water and nutrient cycle), supply (e.g., food production), regulation (e.g., climate regulation), cultural services (e.g., recreational services), and biocontrol services (e.g., source of predatory organisms) [6–9]. Grasslands also have scenic and aesthetic functions [10,11]. However, their most important function is the production of feed.

Grasslands play a crucial role in the nutrition of ruminants such as cattle, goats, sheep, and other herbivores [12]. They provide a natural, valuable feed, which is rich in protein, soluble sugars, carotene, vitamins, microelements, and other substances that catalyse the conversion of roughage into milk and other animal products. The productivity of grassland is influenced by several factors, such as the type of soil and climate (the amount of rainfall, temperature, latitude, and height above sea level) [13,14]. In Poland, the average annual yield from permanent grassland is about 50.0 dt·ha⁻¹. The first spring regrowth, which is usually harvested for hay, has the greatest share in the annual yield of grassland [4]. The yield volume may be modified by the intensity of use (management) of grassland, i.e., mowing frequency, grazing intensity, fertilization, and irrigation) [15,16].

The productivity of animals depends on the nutritional value of feed. Therefore, it is not only the quantity of biomass obtained from permanent grasslands but also its quality that is important. Feed quality includes various characteristics, such as the chemical composition, energy concentration, digestibility, protein content, etc. Protein is the main nutrient, and it is essential for the growth and development of animals, because it builds their tissues and organs. Feed digestibility is also an important factor affecting the efficiency of nutrition.

The maturity of plants at the harvest time, especially in the first windrow, is a basic factor influencing the nutritional value of feed [17–19]. The composition and nutritional value of sward changes during the growth of plants, with a tendency to decrease during the growing season. The rate and scope of these changes depends on the dominant species in the botanical composition of grasslands.

Grasslands in Poland are multi-species grassy ecosystems, where grasses, legumes, and other species of dicotyledons grow. The species composition also affects the nutritional value of feed [20]. Legumes are particularly valuable components of the sward. They have a higher content of nitrogen, lignin, and minerals, and a lower content of neutral detergent fibre (NDF), especially hemicellulose, than grasses of similar organic matter digestibility [21].

Climate change, which increases temperatures in most European regions as well as changes in the amount and distribution of rainfall, causes changes in the species composition, productivity (for biomass accumulation) and, consequently, in the quality of grasslands [22–28]. Recent climate change, manifest by both higher average daily air temperatures and lower total rainfall, causes extreme weather phenomena such as droughts [29]. Grasslands are more susceptible to long-term water shortages than other ecosystems [30]. Some researchers have already analysed the effects of forecasted climate change on grassland productivity in specific regions or in larger areas, such as Northern Europe [31] or Mediterranean regions [32]. However, there have been few studies on the effect of climatic trends on local grassland productivity. There may be big differences in the influence of droughts on the yield of grasslands, depending on the terrain conditions, such as the type of soil, landform, and microclimatic conditions before the drought [33,34].

In recent years, the trend of increasing air temperatures and falling rainfall has also been noticeable in Poland, which is a visible proof of climate change [35–38]. Climate change has negative effects such as causing increasingly frequent water shortages and droughts, especially in the summer, when the demand of grasslands for rainfall is 500 mm. Grasses may be particularly sensitive to rising temperatures and rainfall deficits, because many of them are shallow-rooted and short-lived. Therefore, they quickly react to fluctuations in
climatic conditions [39]. This poses a threat to grasslands, especially in the regions where water is a factor limiting the yield of crops [40,41].

Recently there has been increasing interest in data analysis methods used for research in various life sciences, including agricultural sciences [42]. Currently, a large part of precision agriculture and elements of sustainable agricultural production are based on machine learning and artificial intelligence. Artificial neural networks (ANNs) are a tool used in the implementation of predictive [43–45], classification [46–48], and deterministic problems. Artificial neural networks are based on the assumption that they are a set of simple computing units that process data, communicate with each other, and work in parallel. ANNs are a computational tool simulating the operation of the human brain in a simplified way. The neuron is a basic and simple unit of computation and a specific object capable of signal processing and cooperation with other neurons [49,50]. The real value of neural networks and their simplicity lies in a very important process, i.e., the learning of neural networks. Neurons are able to process large amounts of information because specific numerical values, the so-called weights, are assigned to connections between them. During the operation of the neural network these weights are modified to keep the network learning error at the lowest possible level. This mechanism is referred to as the learning process [51]. The most important advantages of neural networks include: the ability to learn on the basis of preset specific examples, the ability to classify and group, and the ability to broadly interpret phenomena and dependencies on the basis of an incomplete set of experimental data [52]. It is important to remember that artificial neural networks operate on the “black box” principle—they do not provide complete information about the method of getting specific answers or detailed relations between the input and output variables. In order to fully recreate and use a previously created model, it is necessary to access a specific file in the network.

ANNs have already been used in grassland ecosystem research for more than two decades [53]. During these twenty years, ANNs have been used and are still used today to analyze and interpret data from various sources. Passive and active satellite imaging, UAV imaging, short-range and contact VIS-NIR spectrometry, laser scanning, soil and plant sensors, and classical agrotechnical data are the most common [54–59]. In the mentioned studies, ANN models address problems such as estimating pasture cover by classifying sward composition, estimating biomass, predicting meadow sward yield, optimizing agrotechnical processes, estimating vegetation dynamics, estimating emissions and forecasting emissions, and many others related to climate change impacts on grassland ecosystems. For predictive models, they are also an effective tool in analyzing the input factors of the network by sensitivity testing [60]. The utilization of information carried by the results of a neural network sensitivity analysis is an interesting and practical application of ANNs. It enables the identification and ordering of the most important factors that can explain a dependent variable’s variability [50,61,62].

The authors of this study assumed that environmental conditions, such as the rate of progression of the growing season, the increase in temperature, and the changes in moisture levels and soil conditions, affect the nutritional value of species in the sward. Neural networks could help to develop models of the growth and yield of grassland sward and be applied for the management and efficient use of these changes in grassland ecosystems.

The aim of the pilot study was to develop preliminary deterministic models of the yield and digestibility of the first regrowth of grassland sward based on publicly available weather data and allowing for the botanical composition of the sward. In this study, the following realization phases can be distinguished: (i) field research and acquisition of meteorological data, harvest date of the first regrowth, and meadow sward species composition as well as the determination of yield characteristic parameters such as dry matter yield, dry matter digestibility and protein yield, (ii) the generation of nine neural models for the attributes acquired in the first stage during the field research, (iii) the sensitivity analysis of the neural networks, which resulted in a ranking of the most important factors that lead to
an explanation of the network output variability, (iv) the generation of network response plots, which show the process of change in the value of the analysed traits as a function of the selected independent variable, (vi) and the analysis and discussion of the obtained results of the pilot study together with a summary and conclusions.

2. Materials and Methods

2.1. Research Site

Between 2014 and 2016, an experiment was conducted at the Institute of Technology and Life Sciences-National Research Institute in Falenty, Poland (52°8’27.27” N 20°55’39.426” E) on a three-cut permanent meadow located on mineral soil. In March 2012, a meadow (2.4 ha) was delineated in three 0.8 ha plots. A mixture of grasses and bird’s foot trefoil (Lotus corniculatus) was undersown in the first plot. A mixture of grasses and red clover (Trifolium pratense) was undersown in the second plot. No plants were undersown in the third plot. As a result of this undersowing, three types of meadow sward differing in the botanical composition were obtained.

Every year the experimental meadow was fertilised with mineral NPK fertilisers at the following doses: 60 kg N (ammonium nitrate), 30 kg P (granulated triple superphosphate 46% P₂O₅), and 60 kg K (potassium salt 60% K₂O) per ha.

2.2. Botanical Composition of Meadow Sward

Every year in mid-May, the botanical composition of the sward was analysed with the Klapp’s method [63] by estimating the share of species in the sward with an accuracy of 1%. Three types of meadow sward were distinguished based on the percentage share of the main plant species: grasses (G), grasses and bird’s foot trefoil (GB), and grasses and red clover (GC). The percentage of the main plants species in the meadow sward is shown in Figure 1.

![Figure 1](image-url)  
**Figure 1.** The percentage of the dominant plant species in the three types of meadow sward (G, GB, GC) in study years (2014–2016).

2.3. Herbage Sampling and Analysis

Between 2014 and 2016 ten herbage samples were collected from each part of the meadow at seven-day intervals. Samples from an area of 1 m² were hand-cut with scissors at the height of 5 cm. Every year, sampling started at the end of April and was continued until the end of June. Herbage samples were collected for chemical analyses and to determine the dry mass yield (DMY).
The DMY was assessed on the basis of the fresh weight of biomass collected from 1m². The dry matter content was measured with the oven method (100 g of sample was dried in an oven for 24 h at a temperature of 105 °C) and expressed as tonnes of dry matter per hectare. The crude protein (CP) content in harvested herbage was used to calculate the crude protein yield (CPY). The DMY and CPY were expressed as tonnes per hectare.

The remainder of the sample was left for chemical analyses. After drying and grinding the samples, the crude protein content (CP) and dry matter digestibility (DMD) were estimated with the NIRS method on a NIRFlex N-500 near-infrared spectrometer (Büchi Labortechnik AG, Flawil, Switzerland) with ready-made INGOT® meadow hay calibrations.

2.4. Weather Conditions

Meteorological data were gathered from the Warsaw-Okęcie weather station (EPWA 12375) located approximately 4 km away from the research site. The Selyaninov hydrothermal coefficient (HTC) [64] was used for detailed assessment of rainfall and temperature in growing seasons (from March to June). The following formula was used for calculations:

\[ \text{HTC} = \frac{P \times 10}{\Sigma t} \]  

where:
- P—total monthly rainfall (mm),
- \( \Sigma t \)—sum of monthly average daily air temperatures > 0 °C.

The available meteorological data were used to calculate the average daily temperature (T_AV), total rainfall (PREC_S), and cumulative degree-day values (T_S) for each date of herbage sampling. The date of the beginning of the growing season in each year was calculated from the average daily temperatures and the total rainfall according to the mathematical formula developed by Gumiński [65].

During the period under study, the weather conditions were diverse, especially the amount and distribution of rainfall (Figure 2). In the first year (2014), the amount and distribution of rainfall was favourable for plant growth. In 2015 April it was fairly humid, May was dry, and June was very dry. April and June were fairly dry in the third year, and May was very dry (Table 1).

Figure 2. Average decade air temperatures and precipitation in period March–June in study years (2014–2016).
The following variables could be found at the network output in the models presented in the protein yield model for the sward consisting mostly of grasses, \( Y_{-GB} \), a yield model for the sward consisting of grasses, \( Y_G \), and a protein yield model for the sward consisting of grasses and bird’s foot trefoil, \( Y_{-GC} \).

Table 1. Values of the Selyaninov hydrothermal coefficient (HTC).

<table>
<thead>
<tr>
<th>Year</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>1.82 (quite humid)</td>
<td>1.45 (optimum)</td>
<td>2.35 (humid)</td>
<td>1.23 (quite dry)</td>
</tr>
<tr>
<td>2015</td>
<td>1.00 (dry)</td>
<td>1.62 (quite humid)</td>
<td>0.94 (dry)</td>
<td>0.36 (extreme dry)</td>
</tr>
<tr>
<td>2016</td>
<td>2.37 (very humid)</td>
<td>1.27 (quite dry)</td>
<td>0.65 (very dry)</td>
<td>0.95 (dry)</td>
</tr>
</tbody>
</table>

2.5. Organisation and Division of Experimental Data

The first stage of work on the concept of the development of deterministic models involved the creation of a spreadsheet database, which contained all the data from the field experiment conducted between 2014 and 2016. In total, 840 experimental cases were collected. It is important to note that an experiment replicate was treated as a separate case. One of the research assumptions was not to include the cases with incomplete experimental information or with data raising objections in the final analyses. Figure 3 shows the concept of development of deterministic models based on the field experiment.

![Diagram of experimental data organization](image)

**Figure 3.** The deterministic models developed in the study, including the independent variables and three experimental variants.

As shown in Figure 3, the authors of this study proposed the creation of nine deterministic models, differing in the output variable and the main factor from the field experiment. The following variables could be found at the network output in the models presented in the study: dry matter yield (t·ha\(^{-1}\)), dry matter digestibility (%), and protein yield (t·ha\(^{-1}\)). In addition, the experiment was diversified in terms of the species composition in the sward. The sward consisted of grasses (G), grasses and bird’s foot trefoil (GB), and grasses and red clover (GC). The adoption of the concept enabled the creation of the following deterministic models: \( Y_G \)—a yield model for the sward consisting mostly of grasses, \( DIG_G \)—a model of dry matter digestibility percentage for the sward consisting mostly of grasses, \( P_G \)—a protein yield model for the sward consisting mostly of grasses, \( Y_{-GB} \)—a yield model for the sward consisting of grasses and bird’s foot trefoil, \( DIG_{-GB} \)—a model of dry matter digestibility...
digestibility percentage for the sward consisting of grasses and bird’s foot trefoil, P_GB—a protein yield model for the sward consisting of grasses and bird’s foot trefoil, Y_GC—a yield model for the sward consisting of grasses and red clover, DIG_GC—a model of dry matter digestibility percentage for the sward consisting of grasses and red clover, and P_GC—a protein yield model for the sward consisting of grasses and red clover.

2.6. Deterministic Models—Construction Methodology

The deterministic models for the grassy sward (Y_G, DIG_G, and P_G), the sward with grasses and bird’s foot trefoil (Y_GB, DIG_GB, and P_GB), and the sward with grasses and clover (Y_GC, DIG_GC, and P_GC) were based on the same input variables (Table 2). The output variable was the differentiating factor.

Table 2. The structure and ranges of data used for the construction of deterministic models, including the symbol of each variable.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Variable Name</th>
<th>Unit of Measure</th>
<th>Data Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Independent variable—network input</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_D</td>
<td>Harvest date</td>
<td>Number of days since the beginning of the year</td>
<td>114–176</td>
</tr>
<tr>
<td>T_AV</td>
<td>The average air temperature</td>
<td>°C</td>
<td>7.94–13.5</td>
</tr>
<tr>
<td>PREC_S</td>
<td>The sum of precipitation</td>
<td>mm</td>
<td>46.23–223.26</td>
</tr>
<tr>
<td>T_S</td>
<td>Sum of effective temperatures</td>
<td>°C</td>
<td>331.80–1344.90</td>
</tr>
<tr>
<td>%_G</td>
<td>Grasses %</td>
<td>%</td>
<td>49–79</td>
</tr>
<tr>
<td>%_L</td>
<td>Fabaceae %</td>
<td>%</td>
<td>5–33</td>
</tr>
<tr>
<td>%_HW</td>
<td>Herbs and Weeds %</td>
<td>%</td>
<td>6–43</td>
</tr>
<tr>
<td></td>
<td>Dependent variable—network output</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%_DMD</td>
<td>Dry matter digestibility</td>
<td>%</td>
<td>17.05–73.73</td>
</tr>
<tr>
<td>Y_DM</td>
<td>Dry matter yield</td>
<td>t·ha⁻¹</td>
<td>0.51–17.14</td>
</tr>
<tr>
<td>P_Y</td>
<td>Protein yield</td>
<td>t·ha⁻¹</td>
<td>0.02–3.41</td>
</tr>
</tbody>
</table>

Three types of factors represented the independent variables: phytopenological factors, meteorological factors, and factors related to the percentage share of various plant groups in the sward composition. The phytopenological variable, i.e., the harvest date (H_D), was expressed as the number of days since the beginning of the calendar year. The meteorological variables, i.e., the average daily temperature (T_AV), total rainfall (PREC_S), and total effective temperatures (T_S) in each year of the study were calculated for the period ranging from the spring onset of vegetation to the harvest date. The variables related to the percentage share of various groups of plants in the sward were calculated based on the botanical composition of the sward, which was determined in each year in mid-May. The output variables, i.e., the dry matter digestibility (%) (%_DMD), dry matter yield (t·ha⁻¹) (Y_DM), and protein yield (t·ha⁻¹) (P_Y) were determined in each year after the harvest.

It is difficult to choose the right architecture of neural networks due to their different specificity, the data selected for analysis and the selected learning method. Therefore, the generalisation and approximation abilities of networks based on previously established measures of their quality were taken into account when selecting the best topology and learning method.

Various network architectures were tested with the Automatic Network Designer implemented in the Statistica v7.1 software [66]. Ten thousand neural networks were tested for each of the issues under consideration. The analyses enabled the selection of the best neural network architecture. A multi-layer perceptron network with two hidden layers was selected for all models tested: Y_G, DIG_G, P_G, Y_GB, DIG_GB, P_GB, Y_GC, DIG_GC, and P_GC. This type of network is often used to implement prognostic and deterministic issues in agricultural production [43,67]. When selecting the networks that will be used at further stages of the analyses, the best values of parameters relating to their quality were
taken into account, i.e., standard error deviation, mean value of error modules, standard deviation quotient, and correlation coefficient. If the results were ambiguous or difficult to assess, networks with a high correlation coefficient and a low value of the mean absolute error were searched. Table 3 lists the structures of selected neural networks.

Table 3. The structures of neural networks in the deterministic models and the learning methods.

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of Neurons</th>
<th>Network Learning Method *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y_G</td>
<td>7</td>
<td>BP100, CG57b</td>
</tr>
<tr>
<td>DIG_G</td>
<td>7</td>
<td>BP100, CG87b</td>
</tr>
<tr>
<td>P_G</td>
<td>7</td>
<td>BP100, CG58b</td>
</tr>
<tr>
<td>Y_GB</td>
<td>7</td>
<td>BP100, CG90b</td>
</tr>
<tr>
<td>DIG_GB</td>
<td>7</td>
<td>BP100, CG119b</td>
</tr>
<tr>
<td>P_GB</td>
<td>7</td>
<td>BP100, CG70b</td>
</tr>
<tr>
<td>Y_GC</td>
<td>7</td>
<td>BP100, CG62b</td>
</tr>
<tr>
<td>DIG_GC</td>
<td>7</td>
<td>BP100, CG50b</td>
</tr>
<tr>
<td>P_GC</td>
<td>7</td>
<td>BP100, CG113b</td>
</tr>
</tbody>
</table>

* Abbreviations for neural network learning methods: BP—error backpropagation method, CG—conjugate gradient method, _b—the epoch with the best network learning result.

To train and verify the selected multilayer perceptron networks, each data set containing experimental cases, were randomly divided into three sets: training set (70% of cases), validation set (15% of cases), and test set (15% of cases). The data collected in the training set enabled the calculation of the gradient, weight, and value of all loads for the network. The isolation of the validation set made it possible to monitor the network learning error during the learning stage continuously. If the error of the validation set tended to increase for several consecutive epochs, the learning process was stopped. In general, this set is attributed a very important role—it prevents the overlearning of the neural network. The test set was intended to carry out a one-time verification and control after the completion of learning to check whether, despite periodic validation, the neural network had not lost the generalisation ability.

Two neural network learning methods were used: backward propagation of errors (BP) and conjugate gradient method (CG). The results listed in Table 3 show that there were different conditions at the end of the network learning process. All the networks were trained with the backpropagation method for 100 epochs. The learning with the conjugate gradient method was diversified. The learning of the MLP 7:7-11-5-1:1 network in the Y_G model was the shortest at the second stage—57 epochs. The learning of the MLP 7:7-9-5-1:1 network in the P_GC model was the most time-consuming—it took 113 learning cycles with the conjugate gradient method.

2.7. Neural Networks—Sensitivity Analysis

After learning the neural network, its sensitivity was analysed to show which input data were the most important in explaining the variability of the dependent variable. This information was obtained by analysing the increase in error while eliminating individual variables from the input data. This task is accomplished by using the error quotient represented by the rank used. The error quotient expresses the ratio of the error to the total error of all independent variables. As its value increases, so does the significance of a particular variable. If the quotient value is below 1 for any of the independent variables, it is best to remove it from the model to improve its quality. The rank shows the position of the variable on the “ranking list”—it indicates the characteristics according to the decreasing error. The rank value of 1 means that this variable contributed the most to the explanation of variability of the dependent variable [44,68,69].
3. Results

3.1. Neural Network Sensitivity Analysis—Summary

Table 4 shows the results of the analysis of sensitivity of neural networks in the Y_G, DIG_G, P_G, Y_GB, DIG_GB, P_GB, Y_GC, DIG_GC, and P_GC models for three variables ranked from 1 to 3. The corresponding values of the error quotients were considered.


<table>
<thead>
<tr>
<th>Model</th>
<th>Rank 1 Variable</th>
<th>Quotient</th>
<th>Rank 2 Variable</th>
<th>Quotient</th>
<th>Rank 3 Variable</th>
<th>Quotient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y_G</td>
<td>T_AV</td>
<td>3.322</td>
<td>T_S</td>
<td>3.053</td>
<td>PREC_S</td>
<td>2.293</td>
</tr>
<tr>
<td>DIG_G</td>
<td>T_AV</td>
<td>1.946</td>
<td>T_S</td>
<td>1.828</td>
<td>%_L</td>
<td>1.518</td>
</tr>
<tr>
<td>P_G</td>
<td>T_AV</td>
<td>5.924</td>
<td>T_S</td>
<td>3.758</td>
<td>%_L</td>
<td>2.708</td>
</tr>
<tr>
<td>Y_GB</td>
<td>T_AV</td>
<td>2.155</td>
<td>H_D</td>
<td>1.403</td>
<td>PREC_S</td>
<td>1.351</td>
</tr>
<tr>
<td>DIG_GB</td>
<td>T_AV</td>
<td>2.830</td>
<td>T_S</td>
<td>2.172</td>
<td>PREC_S</td>
<td>1.359</td>
</tr>
<tr>
<td>P_GB</td>
<td>PREC_S</td>
<td>2.445</td>
<td>T_S</td>
<td>1.991</td>
<td>T_AV</td>
<td>1.885</td>
</tr>
<tr>
<td>Y_GC</td>
<td>T_AV</td>
<td>2.254</td>
<td>T_S</td>
<td>2.098</td>
<td>PREC_S</td>
<td>2.090</td>
</tr>
<tr>
<td>DIG_GC</td>
<td>%_L</td>
<td>1.707</td>
<td>T_S</td>
<td>1.610</td>
<td>%_G</td>
<td>1.388</td>
</tr>
<tr>
<td>P_GC</td>
<td>T_AV</td>
<td>4.816</td>
<td>PREC_S</td>
<td>3.899</td>
<td>T_S</td>
<td>2.435</td>
</tr>
</tbody>
</table>

Data in Table 4 shows that the average daily air temperature (T_AV), calculated for the period ranging from the spring onset of vegetation to the harvest date, was the independent variable with the greatest influence on the values of most of the tested dependent variables. Other important rank one variables were: the total rainfall (P_GB model) and the percentage of legumes in the sward (%_L). Rank two in the analysis of the sensitivity of neural networks was most often assumed by the independent variable T_S, i.e., the total temperatures in the period ranging from the spring onset of vegetation to the harvest date. The other rank two variables were the harvest date (H_D) in the Y_GB model and the total rainfall (PREC_S) in the P_GC model. Rank three in all the models was diversified. It was assigned four times to the total rainfall variable (PREC_S) in nine models. The variable referring to the percentage share of legumes in the sward (%_L) was identified twice. It is noteworthy that in each of the analyses the significance of all tested variables was confirmed—the value of the quotient was above one for each variable.

3.2. Response Charts for the Most Significant Variables

Response charts show how the predicted value changes as a function of the selected independent variable. The comparison of the response charts with the ranks provided by the sensitivity analysis shows which parameters, and to what extent, influence the value of the dependent variable. Figures 4–12 show the relations between the rank one variables from the neural network sensitivity analyses and the values of the dependent variables for the Y_G, DIG_G, P_G, Y_GB, DIG_GB, P_GB, Y_GC, DIG_GC, and P_GC models.

In the deterministic model for the grass sward (Y_G) shown in Figure 4, the dry matter yield increased along with the average daily air temperature.

The data presented on Figure 5 show that an increase in the average daily air temperature caused a decrease in the digestibility of the dry matter of grass sward (%).

As the average daily air temperature increased, so did the yield of protein from the grass sward (Figure 6).

Figure 7 shows the relation between the yield of dry matter from the sward with the predominance of grasses and bird’s foot trefoil (Y_GB model) and the average daily air temperature (T_AV). The average daily air temperature in the period ranging from the spring onset of vegetation to the harvest date was the factor with the greatest influence on the yield of dry matter of green forage. As the value of this parameter increased, so did the value of the dependent variable.
Figure 4. The Y_G model. The response chart for variables T_AV (rank 1) and Y_DM.

Figure 5. The DIG_G model. The response chart for variables T_AV (rank 1) and %_DMD.
Figure 6. The P_G model. The response chart for variables T_AV (rank 1) and P_Y.

Figure 7. The Y_GB model. The response chart for variables T_AV (rank 1) and Y_DM.

The dependence between the percentage of digestibility of the dry matter from the sward with the predominance of grasses and bird’s foot trefoil and the two variables with the greatest influence on explaining the variability of this trait (DIG_GB model) is presented on Figure 8. An increase in the average daily air temperature caused a decrease in the dry matter digestibility.

Figure 9 shows the relation between the yield of protein from the sward with the predominance of grasses and bird’s foot trefoil and the total rainfall in the test period (PREC_S). As seen from the data, the higher the rainfall in the period ranging from the spring onset of vegetation to the harvest date was, the higher the yield of protein from the harvested biomass was.
The relation between the yield of dry matter from the sward with the predominance of grasses and red clover (Y_GB model) and the average daily air temperature (T_AV) is shown on Figure 10. The average daily air temperature in the period ranging from the spring onset of vegetation to the harvest date stimulated the yield of dry matter from the sward with the predominance of grasses and red clover.

![Graph](image1)

**Figure 8.** The DIG_GB model. The response chart for variables T_AV (rank 1) and %_DMD.

![Graph](image2)

**Figure 9.** The P_GB model. The response chart for variables PREC_S (rank 1) and P_Y.
Figure 10. The Y_GC model. The response chart for variables T_AV (rank 1) and Y_DM.

Figure 11. The DIG_GC model. The response chart for variables %_L (rank 1) and %_DMD.

Figure 11 shows the data referring to the dependence between the percentage of digestibility of the dry matter from the sward with the predominance of grasses and red clover (DIG_GC model). An increase in the percentage share of legumes (%_L) caused an increase in the digestibility of the green forage dry matter.

The relation between the average daily air temperature (T_AV)—rank one and the protein yield is presented on Figure 12. An increase in the average daily air temperature to about 9.5 °C stimulated the accumulation of protein in the sward with the predominance of grasses and red clover. When this value was exceeded, the protein yield remained at a steady level of about 0.3 t·ha⁻¹.
3.3. Comparison of Qualitative Parameters of Generated Models

In each of the presented cases, all the deterministic models were characterised by the best values of the measures referring to the quality of the generated neural networks (Table 5). When it was difficult to identify the best network at the summary stage, the values of two qualitative parameters were taken into account: the correlation coefficient (r) and the mean absolute error. The authors followed the principle that the value of the correlation coefficient was to be the highest, whereas the value of the mean absolute error was to be low. In the nine cases under analysis the values of the correlation coefficient were high or very high, as they ranged from 0.739 (P_G model) to 0.946 (DIG_GB model). The value of the mean absolute error was the lowest for the P_G model—it amounted to 0.094 t·ha⁻¹. The highest value of the mean absolute error was noted for the DIG_GC model—3.584%. The error quotient was another important parameter of assessment of the quality of the neural networks. It was defined as the quotient of the standard deviation of the prediction errors and the standard deviation of the output variable. The value of this parameter should not exceed 0.7 for the model to be useful for prognostic purposes. Each of the nine cases under analysis met this assumption. The lowest value of the deviation quotient was noted for the DIG_GB model—0.321.

Table 5. The values of the main qualitative parameters for the neural networks in the tested models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Average Error</th>
<th>Deviation Error</th>
<th>Mean Absolute Error</th>
<th>Quotient Deviations</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y_G</td>
<td>3.777</td>
<td>2.065</td>
<td>−0.002</td>
<td>1.164</td>
<td>0.859</td>
<td>0.564</td>
<td>0.827</td>
</tr>
<tr>
<td>DIG_G</td>
<td>53.504</td>
<td>11.384</td>
<td>0.031</td>
<td>3.693</td>
<td>2.865</td>
<td>0.324</td>
<td>0.946</td>
</tr>
<tr>
<td>P_G</td>
<td>0.350</td>
<td>0.259</td>
<td>−0.004</td>
<td>0.146</td>
<td>0.094</td>
<td>0.563</td>
<td>0.827</td>
</tr>
<tr>
<td>Y_GB</td>
<td>4.887</td>
<td>2.380</td>
<td>−0.026</td>
<td>1.464</td>
<td>1.078</td>
<td>0.615</td>
<td>0.790</td>
</tr>
<tr>
<td>DIG_GB</td>
<td>52.025</td>
<td>12.468</td>
<td>−0.004</td>
<td>4.001</td>
<td>2.981</td>
<td>0.321</td>
<td>0.947</td>
</tr>
<tr>
<td>P_GB</td>
<td>0.491</td>
<td>0.236</td>
<td>0.008</td>
<td>0.160</td>
<td>0.117</td>
<td>0.679</td>
<td>0.739</td>
</tr>
<tr>
<td>Y_GC</td>
<td>4.520</td>
<td>2.531</td>
<td>−0.039</td>
<td>1.619</td>
<td>1.141</td>
<td>0.640</td>
<td>0.769</td>
</tr>
<tr>
<td>DIG_GC</td>
<td>52.396</td>
<td>12.939</td>
<td>−0.112</td>
<td>4.689</td>
<td>3.584</td>
<td>0.362</td>
<td>0.932</td>
</tr>
<tr>
<td>P_GC</td>
<td>0.464</td>
<td>0.371</td>
<td>−0.022</td>
<td>0.226</td>
<td>0.146</td>
<td>0.609</td>
<td>0.794</td>
</tr>
</tbody>
</table>
4. Discussion

Over the past decades, many models describing grass growth have been developed. They have ranged from simple empirical models [70,71] to more complex mechanistic models [72–74]. They are often based on growth, senescence, litter, and standing biomass. They use data on incoming radiation, temperature, soil moisture, day length, and altitude. The development of grass growth models in response to temperature may help to forecast the growth depending on the season and temperature and show how climate change affects grassland productivity [31]. Climate changes, which are also noticeable in Poland, lead to the systematic deterioration of humidity conditions. Particularly grasslands are sensitive to fluctuations in climatic conditions. Rising temperatures and rainfall deficits will cause changes in the species composition of grasslands, their productivity [37–39] and, consequently, in the quality of obtained forages. User-friendly grass growth prediction models provide real-time information based on forecasted meteorological data and farm management data to predict grass growth. Hence, herbage supply may allow farmers to manage a key resource on their farms.

These models facilitate a better understanding of the grass growth processes and their interactions with farm management. They also help to study the influence of the environment on grassland. Grass growth models are often sub-models or elements of larger farm system models, e.g., in APSIM [75], PaSim [76,77], EcoMod, DairyMod, and the SGS pasture model [74]. Integrated models can be used to assess the effects of climate change on grasslands production and livestock production systems. For example, the PaSim model was used to assess the effects of climate change on the productivity pastures and ruminants. It is important to assess models to make sure that the selected model can be used in a specific region or under specific conditions [78].

In our study, the average daily air temperature calculated for the period ranging from the spring onset of vegetation to the harvest date was the independent variable with the greatest influence on the values of most of the tested dependent variables. In all the deterministic models, the dry matter yield tended to increase along with the average daily air temperature. However, this increase was curvilinear and reached the maximum values when temperatures exceeded 10°C. This is consistent with other scientific reports, according to which the grass growth rate is controlled by three factors: temperature, soil moisture content, and the intensity of sunlight. Of these three factors, it is temperature that has the greatest influence on the grass growth rate, especially in the spring, because it determines the rate of plant development and the structure of the root mass and aerial parts of plants [14,79]. In Poland, the grass growing season is assumed to begin when the average daily air temperature exceeds 5°C for five consecutive days [80]. Depending on the region and the year, this moment may occur as early as the beginning of March. Taking the temperature above 5°C as the starting point, the higher the temperature, the faster the sward growth rate is until the temperature is about 10°C [81,82]. Most grass species achieve the maximum growth rate at an average daily temperature of about 10°C. A further increase in temperature has little influence on the growth rate of grassland plants. Therefore, the temperature range of 10–15°C is considered to be optimal for the good development and yield of grassland sward.

Air temperature affects the intensity of intracellular metabolic processes during plant growth. It causes changes in the chemical composition of plants, mainly in the formation of structural carbohydrates, such as hemicellulose, cellulose, and lignin, which influence the digestibility of feed [71,83]. In our study, the average daily temperature also influenced the digestibility of the dry matter of grasses (DIG_G model) and grasses with bird’s foot trefoil (DIG_GB). There was a rectilinear decrease in the digestibility of dry matter observed along with an increase in the average daily temperature calculated from the vegetation onset to the harvest day. A similar phenomenon was observed by Żurek [84], who analysed the relationship between the digestibility of the dry matter of rescuegrass (Bromus willdenovii) and tall fescue (Festuca arundinacea) and the average daily air temperature during the growing season in Poland. Like in our study, the digestibility of the dry matter of both grass
species tended to decrease as the average daily air temperature increased. The average daily temperature also influenced the yield of protein from the grass sward (P_G model) and the grass and red clover sward (P_GC). The protein yield, being the resultant of the dry matter yield and protein content in the harvested crops, increased along with the average daily temperature in both cases. However, there were differences in this increase influenced by the sward species composition. The fastest increase in the protein yield in the grass sward was observed at 10–13 °C, which is the optimal temperature range for grasses. In the grass and clover sward, the most intense increase in the protein yield was observed within a very narrow range of temperatures, i.e., 8.5–9.5 °C. There was no increase in the protein yield at temperatures higher than 9.5 °C.

It is a well-known fact that grassland productivity and rainfall are often positively correlated with each other [85,86]. In our study the total rainfall assumed rank one only in the model of the yield of protein from the sward with grass and bird’s foot trefoil (P_GB). As the amount of rainfall increased, so did the protein yield. The total rainfall was the least significant (rank three) variable for the yield models of all three types of sward (Y_G, Y_GB, and Y_GC) and the model of digestibility of grasses with bird’s foot trefoil (DIG_GB). The insignificant or poorly significant influence of rainfall on the yield of all types of sward may have been caused by the fact that plants use water stored in the soil in autumn and winter to produce the first regrowth in spring, so the lack of rainfall in this period does not limit the growth of the sward. On the other hand, a low soil moisture content in summer disrupts the photosynthesis and transpiration processes, thus limiting the development of plants and their yield [41,87]. Grasses, which predominate in the composition of sward with a mixture of legumes and grasses in the spring regrowth, have a special ability to make good use of the water resources accumulated during the winter. In summer with periodic rainfall deficiencies, legumes predominate in the regrowth, because, unlike grasses, they better withstand periodic and short-term summer droughts [88]. During longer periods of drought, which sometimes occur in the middle of summer, plant growth is completely inhibited. For example, in summer, when the amount of rainfall was insufficient, the yield of red clover decreased by 3–34%, as compared with the yield obtained under the optimal soil moisture content with a rainfall of 350–460 mm [89].

Another important rank one variable was the percentage of legumes in the sward (%_L). The digestibility of dry matter in the sward tended to increase along with the percentage share of red clover (Trifolium pratense) in the DIG_GC deterministic model. Earlier studies showed that the presence of red clover in the sward positively affected the digestibility of feed and its protein content [90,91]. The share of legumes in mixtures with grasses increases the feed energy and protein content, contributes to more efficient absorption of these nutrients in the digestive system, and thus influences the growth of animals receiving these feeds [92,93]. Moreover, the sward composed of red clover and bird’s foot trefoil with grasses enriches the feed with vitamins, fatty acids, and carotene [94,95]. According to an earlier study [96], the best chemical composition and the highest nutritional value of grass and clover mixtures is obtained when the share of red clover in the yield is 50%. According to Staniak [97], mixtures with 60% and 80% clover content were the best balanced in terms of the protein and fibre content. In the study by Wróbel and Zielewicz [90], the optimal content of protein was guaranteed by a 30–40% share of bird’s foot trefoil and a 60–80% share of red clover. Our study showed that a higher share of legumes resulted in a higher digestibility of the sward. A 5% share of legumes in the sward resulted in the dry matter digestibility of 37–38%. However, when the share of legumes in the sward increased to 33%, the digestibility of the feed increased to 62%.

The development of plants (the sequence of consecutive development phases) depends almost exclusively on temperature and the length of the day. Therefore, one of the most popular units used in agriculture for measurements of plant growth, development, and maturity is the growing degree-day (GDD), which ensures a more accurate physiological estimate than calendar days alone [98]. The GDD is a measure of heat accumulation above a certain base temperature, which is different for each forage species. When days are warmer
than those marked as normal for a particular species, plants grow rapidly. However, when temperature drops below the base level, the GDD value does not increase. The GDD is an important measure determining the dynamics of plant development because each development stage requires a specific amount of accumulated sum of GDD temperatures.

The sum of temperatures determines the length of plants’ growing seasons, stimulates individual stages of their growth and development, and the actual production of dry matter, including the occurrence of critical periods of temperature stress [99]. In our study, total temperatures (T_S) was a rank two parameter in the analysis of the sensitivity of neural networks. The increase in total temperatures resulted in a decrease in the protein content in the sward.

GDD is a value that enables an approximate prediction of the dates when individual plant development phases will begin. However, due to climate change the total GDD may increase, whereas the number of days necessary for crops to mature may decrease and this may accelerate the sward harvest date. This may mean that due to an earlier occurrence of the necessary sum of temperatures, individual plant species will grow faster. In consequence, the sward will have to be mown and forage will have to be harvested much earlier than in other years [100,101].

The harvest date (H_D), expressed as the number of days from the beginning of the calendar year, was a significant parameter only in the model of the yield from the sward with grasses and bird’s foot trefoil (Y_GB model), where it assumed rank two (Table 3). The lack of proven significance of this variable means that this parameter should not be taken into account in the construction of grassland sward yield models. The lack of significance of this variable in most models results from the considerable variability of the course of meteorological conditions in individual growing seasons, mainly rainfall and total temperatures, which have the greatest influence on the quantitative and qualitative changes in the biomass of various plants grown in fields and grasslands [102,103].

Artificial neural networks are a suitable tool for making deterministic models, which can be used in plant production, including grassland production. The right rendering of the results generated by a model makes it possible to explain the influence of a significant factor on the dependent variable.

The choice of the right network topology and learning method is one of the key steps in working with neural networks. These parameters are closely related to the target use of this model. There are usually differences between networks selected for typical classification problems and those for prognostic and deterministic problems. In this paper the authors of neural models usually use available scientific publications discussing similar research topics, but it is one’s own experience in working with neural models and the knowledge of the research object that should be the main guidelines for finding the optimal approach to presenting and solving the problem. In our study, an MLP network with two hidden layers was finally selected.

Neural network learning makes it possible to combine certain behaviours of the model based on multiple learning cases. The model user enforces specific responses to the given input signals from the network. The network remembers the questions and responses based on learnt patterns of behaviour to provide the closest response to the original when the next question is asked. All the neural networks presented in our study were trained in two stages. The error backpropagation method was used at the first stage of the process. Each time this process was continued for 100 epochs. The learning was continued thanks to the inclusion of the conjugate gradient method. The best result of network learning was obtained after a different number of learning cycles. Thanks to this dual approach to learning it is possible to create neural networks with very favourable qualitative parameters [44,104].

The selected neural networks used as the bases for the creation of nine models: Y_G, DIG_G, P_G, Y_GB, DIG_GB, P_GB, Y_GC, DIG_GC, and P_GC, had very favourable values of qualitative parameters. The MLP 7:7-11-8-1:1 network fared particularly well—the value of the correlation coefficient was close to one, whereas the value of the deviation quotient was 0.324. These results show that the network was very well suited to the set of
cases used for its construction. This proves the high reliability of analyses. The collective analysis of the values of the qualitative parameters assumed by the tested models leads to the conclusion that each of these models has a high application potential [105].

The selection of independent variables for the construction of a deterministic model is a compromise between the experimental dataset, the author’s practical knowledge, and the requirements of the market and potential users of the model. It is important to remember that the choice of variables must enable real influence on the dependent variable and that variables must be freely accessible. Although the significance of selected variables can be checked by means of additional analyses and calculations [106], practical experience is the best way to properly adjust the independent variables to explain the complexity and variability of a particular phenomenon. Additional software gathering experimental data used in the model significantly reduces its application capability. Our study used meteorological data (T_AV, PREC_S, and T_S), phytophenological data (H_D), and data on the share of individual groups of plants in the sward (%_G, %_L, and %_HW). Each of the nine models was based on the same independent variables, defined separately for each of the three types of sward: with the predominance of grasses, with grasses and bird’s foot trefoil, and with grasses and red clover.

The analysis of the sensitivity of neural networks enabled such a detailed approach to the explanation of the influence of the weather conditions (temperature and rainfall) in the spring on the dry matter yield of the harvested plant biomass and its quality. This method enables the distinction between significant variables in the model and those that have little influence on the result of the network operation [107]. This method is widely used to select the most important variables for problems related to the phenology and yield of crop species [108–110]. The presentation of the relations (Figures 4–12) between the variables with the greatest influence on the explanation of variability of the dependent variables and the dependent variables themselves gives a chance for efficient and conscious control of plant production in the current crop growing season. It is an excellent basis for the implementation of a wide range of practical recommendations concerning agricultural production.

5. Conclusions

The pilot study presented in this paper included the successive steps of field and climate data acquisition from a 3-year study, the construction of neural networks and their sensitivity analysis, and the generation of network response plots. An important element of inference was the generation of nine neural models for which the inputs were seven independent traits, i.e., harvest date, mean daily temperature, total precipitation, total effective temperature, and the proportion of grasses, broad beans, herbs, and weeds in the sward. The output of the network included the following attributes: dry matter yield, dry matter digestibility, and protein yield. The obtained results of the analyses lead to the following conclusions:

1. The deterministic models developed and verified in our experiment enabled the implementation of the aim of the study set in this paper. This means that artificial neural networks can be used to pre-identify highly significant factors affecting the yield and digestibility of grassland sward in the climatic conditions of central Poland.
2. The analysis of the sensitivity of a neural network enables the preliminary selection of factors with the greatest influence on the dependent (output) variable, while maintaining the assumed level of significance.
3. The average daily air temperature (T_AV) is usually the factor with the greatest influence on the tested output variables: dry matter digestibility, dry matter yield, and protein yield, regardless of the species composition of the sward. The total rainfall (PREC_S) and the percentage of legumes were also significant factors.
4. Future research on the improvement of deterministic models in grassland production should concentrate on several aspects. First, it is necessary to analyse in detail the influence of other equally important independent variables, which are usually
expressed in a linguistic form. Second, other analytical methods enabling the optimization of significant (controllable) production factors should also be used, because they considerably affect the quality and quantity of the yield of grassland sward.


**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

**References**


17. Pelletier, S.; Tremblay, G.F.; Bélanger, G.; Bertrand, A.; Castonguay, Y.; Pageau, D.; Drapeau, R. Forage nonstructural carbohydrates and nutritive value as affected by time of cutting and species. *Agron. J.* 2010, 102, 1388–1398. [CrossRef]
30. Raich, J.W.; Tufekcioglu, A. Vegetation and soil respiration: Correlations and controls. *Biogeochemistry* 2000, 48, 71–90. [CrossRef]
34. Li, Q.; Hou, J.; Yan, P.; Xu, L.; Chen, Z.; Yang, H.; He, N. Regional response of grassland productivity to changing environment conditions influenced by limiting factors. *PLoS ONE* 2020, 15, e0240238. [CrossRef]
35. Łabędzki, L. Actions and measures for mitigation drought and water scarcity in agriculture. *J. Water L. Dev.* 2016, 29, 3–10. [CrossRef]
43. Piekutowska, M.; Niedbala, G. Selection of independent variables for crop yield prediction using artificial neural network models with remote sensing data. *Land* 2021, 10, 609. [CrossRef]
45. Niedbala, G. Simple model based on artificial neural network for early prediction and simulation winter rapeseed yield. *J. Integr. Agric.* 2019, 18, 54–61. [CrossRef]


48. Li, Y.; Chao, X. ANN-based continual classification in agriculture. Agriculture 2020, 10, 178. [CrossRef]


64. Selyaninov, G.T. Methods of Agricultural Climatology. Agric. Meteorol. 1930, 22L.


68. Niedbala, G. Application of Artificial Neural Networks for Multi-Criteria Yield Prediction of Winter Rapeseed. Sustainability 2019, 11, 533. [CrossRef]


