

Article

An ANN Model Trained on Regional Data in the Prediction of Particular Weather Conditions

Aleksandra Bączkiewicz ^{1,2}, Jarosław Wątróbski ^{1,*}, Wojciech Sałabun ³ and Joanna Kołodziejczyk ³

¹ Institute of Management, University of Szczecin, Cukrowa 8, 71-004 Szczecin, Poland; aleksandra.baczkiwicz@phd.usz.edu.pl

² Doctoral School of University of Szczecin, Mickiewicza 16, 70-383 Szczecin, Poland

³ Research Team on Intelligent Decision Support Systems, Department of Artificial Intelligence and Applied Mathematics, Faculty of Computer Science and Information Technology, West Pomeranian University of Technology in Szczecin ul. Żołnierska 49, 71-210 Szczecin, Poland; wojciech.salabun@zut.edu.pl (W.S.); joanna.kolodziejczyk@zut.edu.pl (J.K.)

* Correspondence: jaroslaw.watrobki@usz.edu.pl

Abstract: Artificial Neural Networks (ANNs) have proven to be a powerful tool for solving a wide variety of real-life problems. The possibility of using them for forecasting phenomena occurring in nature, especially weather indicators, has been widely discussed. However, the various areas of the world differ in terms of their difficulty and ability in preparing accurate weather forecasts. Poland lies in a zone with a moderate transition climate, which is characterized by seasonality and the inflow of many types of air masses from different directions, which, combined with the compound terrain, causes climate variability and makes it difficult to accurately predict the weather. For this reason, it is necessary to adapt the model to the prediction of weather conditions and verify its effectiveness on real data. The principal aim of this study is to present the use of a regressive model based on a unidirectional multilayer neural network, also called a Multilayer Perceptron (MLP), to predict selected weather indicators for the city of Szczecin in Poland. The forecast of the model we implemented was effective in determining the daily parameters at 96% compliance with the actual measurements for the prediction of the minimum and maximum temperature for the next day and 83.27% for the prediction of atmospheric pressure.

Keywords: Artificial Neural Networks; Multilayer Perceptron; backpropagation algorithm; weather prediction

check for
updates

Citation: Bączkiewicz, A.; Wątróbski, J.; Sałabun, W.; Kołodziejczyk, J. An ANN Model Trained on Regional Data in the Prediction of Particular Weather Conditions. *Appl. Sci.* **2021**, *11*, 4757. <https://doi.org/10.3390/app11114757>

Academic Editors: Christian W. Dawson and Amerigo Capria

Received: 11 March 2021

Accepted: 19 May 2021

Published: 22 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Prediction is one of the basic goals of Data Mining [1], and weather forecasting plays a significant role in meteorology [2]. Artificial Neural Networks (ANNs) belong to non-linear and non-parametric tools for modeling actual processes and phenomena, i.e., problems that are difficult to solve using classical methods. ANNs are widely used in engineering practice with the possibility of their effective modeling in software. They are an extremely useful alternative to traditional statistical modeling techniques in many scientific disciplines [3].

This makes it possible to use neural networks widely, not only in research on brain functions but further to analyze data in areas as diverse as economics [4–6], automation [7], the energy industry [8–11], the natural sciences [12], and medicine [13,14]. ANNs are a tool used in machine learning. They have great possibilities for recording and presenting complex relationships between input and output data [15].

ANNs are parallel computational models, comprising interconnected adaptive data processing units. The adaptive nature of networks, where “learning by example replaces programming”, makes ANN techniques widely used to solve highly non-linear phenomena [16]. The advantage of neural networks is that they can represent both linear and non-linear relationships that exist between data.

1.1. The Essence, Significance, and Complexity of Weather Forecasting

Weather forecasting is the use of science and technology to predict the state of the atmosphere and the associated meteorological phenomena, concerning a specific area and time period. Weather forecasts are performed based on scientific knowledge of atmospheric processes and historical quantitative data as typical meteorological conditions. The chaotic and complex nature of the atmosphere, the enormous computing power required by computers to solve atmospheric equations, and an incomplete understanding of atmospheric processes make predictions less accurate as the range of timely predictions increases [17].

Meteorologists use different methods to forecast the weather. Lewis Fry Richardson proposed the possibility of creating the first numerical weather prediction models in 1922 [18]. The practical use of numerical models began in 1955 due to the development of programmable electronic computers [19]. Data for forecasting come from observations of atmospheric pressure, temperature, wind speed, wind direction, humidity, and precipitation. Trained observers make these observations close to the ground surface or automatic weather stations are used for this purpose [16].

At present, it is common and established to use numerical models for weather forecasting, which are uncompetitive but still not sufficiently effective at this stage of research regarding weather forecast models [20]. The basic condition for the usefulness of a weather forecast is its high verifiability. Users expect accurate local forecasts, broken down into hourly data and a very dense grid, i.e., for specific geographical points.

Weather forecasting is one of the essential tools for planning, risk management, and decision making in sectors of the economy and everyday life. The sectors exposed to weather risk are energy, agriculture, food industry, construction, entertainment and tourism, transport, and defense—together, the lion's share of the national economy [21,22].

Therefore, despite the colossal costs generated by weather forecasting, technological infrastructure for testing alternative methods and creating better forecasting models is still being developed. These expenses are justified by ensuring the safety of people and their property from natural disasters. However, precise weather forecasting is extremely difficult. Weather is a chaotic phenomenon, characterized by temporal and spatial irregularity in which successive states vary and are difficult to predict [23].

New technologies can change this situation. The world of meteorology is exploring Artificial Neural Networks for the generation of non-linear and non-parametric tools for modeling processes and genuine phenomena, i.e., problems that are difficult to solve with classical methods. Such networks have some brain properties. They learn from examples and apply this knowledge to problem solving, i.e., they can generalize. They are useful for tasks that are not very precisely defined formally.

They can function properly at a certain level of damage and despite partially incorrect input. They also have a relatively high speed of operation (information processing). Their current level of development means that they are not yet competitive with numerical models. However, their speed is predicted to compete with standard supercomputers, whose current speed reaches 2.8×10^{14} elementary operations per second [24–26].

Currently, models based on machine learning using Artificial Neural Networks and based on historical data are being tested. According to global reports, these models produce forecast results that far exceed the precision and efficiency of numerical models based on current weather data. These models improve the forecasting efficiency by 40–50% for air temperature and 129–169% for precipitation [27].

1.2. Research Gap

To find and implement an efficient solution with low computational complexity for predicting selected weather indicators while ensuring high verifiability of forecasts it is necessary to maintain a balance between the level of complexity of the model architecture and the type and number of methods used to improve the accuracy. The verifiability of modern numerical weather prediction (NWP) models is good; however, it requires considerable computing resources. This is important for making probabilistic beam forecasts, in which

the models repeatedly (several dozen times) generate the course of future meteorological conditions. Models using machine learning may prove less demanding in this respect [28].

The question was, therefore, asked whether Artificial Neural Networks can be applied to weather forecasting—a phenomenon that is extremely complicated and chaotic in nature. Achieving this goal was quite a challenge as the model was planned to be applied in difficult forecasting conditions, which are characterized by high variability, closely related to the specificity of the climate occurring in Poland. In addition, the city of Szczecin is an area with a specific climate, which causes difficulties in effective weather forecasting, which is described in the following part of the study.

1.3. Aim of the Study

The purpose of this study was to obtain the basis for the statement that a multilayer Artificial Neural Network (or, to be more precise, a Multilayer Perceptron), even one with an uncomplicated structure and implementation, can effectively predict basic weather conditions and whether the MLP model can be used as an adjunctive and corrective tool for forecasting weather conditions at the local scale. We intended to create a unidirectional multilayer Artificial Neural Network model trained on a regional dataset and used to forecast selected weather conditions for the city of Szczecin (Poland) and then to compare the effectiveness of model predictions with forecasts available in the archives of the internet weather service using statistical metrics selected based on a literature review.

We assumed that, with a reliable and sufficiently large weather dataset; properly constructed, established, and proven neural network model; and properly prepared computer environment it would be possible to effectively forecast selected weather indicators. This survey also covered the practical contribution, which is the building of an MLP model for a specific location and characterized by high climatic complexity affecting the difficulty of weather forecasting, which is significant from a pragmatic point of view.

The rest of the report is organized as follows: Section 2 contains a description of the materials and methods used in the study. Section 3 introduces the results of the study. A discussion of the results obtained and the influencing factors are given in Section 4. In Section 5, we present the conclusions and future directions of our workshop.

2. Literature Review

Our literature review contains studies of machine learning methods and their applicability to weather data along with their relevant statistical properties. Numerical Weather Prediction (NWP) models play a key role in operational weather forecasting, especially for longer prediction times. Utilizing NWP with deep learning to improve the accuracy of weather forecasting systems is a fruitful avenue to consider. An article published in 2020 presents results from the Ensemble of Spatial-Temporal Attention Network and Multi-Layer Perceptron (E-STAN-MLP) for forecasting meteorological elements to predict the surface temperature, humidity, wind speed, and direction at 24 automatic meteorological stations in Beijing.

This research can be generalized to local weather prediction in other regions [29]. Multilayer Perceptrons (MLPs) belong to the common type of feed-forward networks used for the future predictions of rainfall and temperature [30]. MLP has been continuously used for many years to solve problems, such as the prediction of weather conditions. MLP is important for real-life problems [31]. In the study described in the article, in 2020, it was used as a predictive model for wind speed forecasting (wind power) in Villonaco alongside a Long Short-Term Memory (LSTM) and Convolutional Neural Network model (CNN) [32].

The continued popularity and usefulness of the MLP is also demonstrated by its use as a prediction model for time series of meteorological tsunamis next to Evolved Radial Basis Function (ERBF) in Evolved Neural Networks (ENN) in 2020 for Accurate Meteorological Forecasting Applications in Vietnam [33]. In a paper published in 2021, F. Dupuy and others presented an application of the ANN technique in the form of an MLP model as a tool to

correct and complement a numerical model for forecasting the wind speed and direction at the local scale, including the Cadarache valley, which is localized in southeastern France).

As a measure of network accuracy, the Mean Absolute Error (MAE) was taken [34]. The satisfactory results obtained by the authors confirmed the usefulness of this tool in weather condition prediction. A Multilayer Perceptron has been successfully used as a predictive model for various natural conditions, such as soil temperature [35], compared to the Support Vector Machine (SVM) for rainfall prediction [36,37]. This recent literature showed that the use of a Multilayer Perceptron for predicting weather conditions was confirmed and, thus, indicated the possibility of its further use.

To research weather condition predictions performed in Poland and worldwide as well as the methods used, we reviewed the literature. The subject of the work by M. Hayati and Z. Mohebi from 2007 was the application of an Artificial Neural Network (Multilayer Perceptron) in forecasting the mean temperature for the next day for Kermanshah in Iran [15]. The authors of the study trained and tested the MLP model using previous meteorological data. The chosen weather data were divided into two randomly selected groups: the training group, corresponding to 67% of the patterns, and the test group, corresponding to 33% of the patterns. MAE was used as a measure of the network accuracy. The training data were selected as the mean temperature, wind speed, humidity, pressure, and sunshine. The dataset was normalized by converting its values to a range between -1 and 1 .

In a paper from 2009, Y. Radhika and M. Sashi projected the maximum temperature for the next day [38]. The MLP prediction was compared with SVM. For MLP, an algorithm of backpropagation was used. A training dataset from 5 years and test data from 1 year were used. In the case of SVM, the Radial Basis Function Kernel (RBF), which is a popular kernel function used in various kernelized learning algorithm functions, was used.

There are three layers in the ML: input, hidden, and output, as well as an algorithm for the backpropagation of errors. A sigmoid function was selected as the activation function. Before the test, the data were normalized, and the measure of the accuracy of the models was the Mean Squared Error MSE. In this study, SVM performed better than MLP. A comparison of the results of air temperature forecasting using a MLP neural network model for prediction using the SVM model is the subject of many studies [39]. This shows that models with such structures are useful in forecasting weather variables.

In the largest of the studies available in the literature, the most frequently predicted conditions were the mean, minimum, and maximum temperature. The research concerned a specific place in the world. S. S. Baboo and K. Shereef, in their work from 2010, used an Artificial Neural Network with a backpropagation algorithm to predict the mean temperature [40], as did M. Hossain and others in a study conducted in 2015 [41].

Among the studies using Artificial Neural Networks for the prediction of the mean temperature, we paid attention to the work of an author from Poland, I. Białobrzewski from 2005 [42]. Another study on temperature prediction using an Artificial Neural Network model is presented in the article [43]. By far, the largest number of studies described mean temperature predictions using ANNs, and one of the most frequently used network learning algorithms was the backpropagation algorithm [44–47].

In the literature, there were additionally studies regarding forecasting the maximum temperature using ANNs [48,49] as well as the minimum temperature [50]. This problem is discussed in an article regarding the prediction of the maximum and minimum temperature with linear regression [51].

A study comparing the effectiveness of the Support Vector Machine Regressor (SVR) and k -Nearest Neighbors in predicting the wind power used in renewable electricity production is presented in the article [52]. Wind speed is also the subject of research presented in the Polish study [53]. This study predicted wind speed as one determinant of the energy consumption in buildings. The tests were carried out on neural network models based on Multilayer Perceptron architecture (MLP), Generalized Regression Networks (GRNN), and networks with Radial Base Functions (RBFs).

Another study with the aim of determining the annual mean wind speed by using a Multilayer Perceptron and backpropagation algorithm for training the network in foreign literature [54], showed the popular use of this network model for wind speed forecasting. Many types of neural network models can be used in forecasting weather conditions.

For this purpose, networks with relatively low complexity, such as MLP, and networks with a more complex architecture, i.e., deep networks [55,56], Recurrence Neural Networks (RNN), Conditional Restricted Boltzmann Machine models (CRBM), and Convolutional Neural Networks (CNN) are suitable [57]. The application of a gated recurrent unit neural network (GRUNN) modification of RNNs to forecast wind power, which is one of the largest renewable energy sources, is described in a paper published in 2019 by M. Ding and others [58].

LSTM network models for predicting the precipitation based on meteorological data from 2008 to 2018 in Jingdezhen City were deployed in the research described by J. Kang and others in 2020. LSTM is a special kind of RNN, capable of learning long-term dependence. It was introduced by Hochreiter and Schmidhuber [59]. RNNs are special ANNs that are connected in a feedback structure between units of an individual layer. They are called recurrent because they perform the same operation on all elements in the sequence.

They make it possible to model data with time-series characteristics by complementing the limits of non-recurrent ANNs, which independently assume the relationship between inputs [60]. In this article, a study involving temperature prediction published by B. Kwon and others in 2020 is presented. The difference between LSTM and RNN is that it adds a “processor”, which is called the cell state, to the algorithm to judge whether the information is useful or not. A network where information enters the LSTM can be judged by rules. Only the information that accords with the authentication will be left behind, and the discrepant information is forgotten by the forget gate [61].

In addition to ANNs, many other approaches are frequently used in weather forecasting, such as multiple regression, SVM, decision trees, and the k-Nearest Neighbor model. The disadvantage of MLP and deep learning models over other methods is that they are time-consuming during the training process. Although SVM requires intensive training and experimentation on different kernel functions as well as other parameters, it is significantly faster compared to MLP and deep learning models [62].

The SVM algorithm was developed by Vapnik and is based on statistical learning theory [63]. SVM includes efficient algorithms for a wide range of regression problems because they not only take into account the approximation of the error to the data but also provide a generalization of the model—that is, its ability to improve the prediction of the data when a new data set is evaluated by it [39].

ANNs, including Multilayer Perceptron, deep ANNs, and other machine learning models, are constantly being improved and widely used for the forecasting of air temperature [64], rainfall [36,37,65], cloudiness [66], and wind speed [67], which proves that these are forward-looking models that are worthy of constant research and improvement for forecasting purposes. To summarize, the most frequently forecasted condition in the above works was temperature; several studies have similarly found the use of models for wind speed prediction, and the most frequently used machine learning model for this purpose was a multilayer Artificial Neural Network. Other models chosen by researchers include

- Support Vector Machines and
- Linear Regression.

The most commonly used learning algorithm for neural networks is the backpropagation algorithm [2]. Datasets are usually large and contain data from many previous years—optimally 10 years. Before starting the training and prediction, the data are normalized. The authors typically verify the accuracy of the model predictions using measures, such as MAE and MSE.

The research results presented in the analyzed papers motivated us to implement a Multilayer Perceptron as the prognostic model and to investigate its effectiveness using, among others, MAE and the Mean Squared Error (MSE) as measures of the model effective-

ness. As the network learning algorithm, the backpropagation algorithm was chosen. We decided that the dataset would contain data from 10 years. The maximum and minimum temperatures, atmospheric pressure, wind speed, and daily precipitation for the next day were planned as the forecast conditions.

2.1. *The Complexity of Weather Forecasting in Poland*

In Poland, the Institute of Meteorology and Water Management is statutorily responsible for preparing weather forecasts, using numerical models. Their verifiability is on average: for short-term forecasts—90–95%, and for medium-term forecasts—70–75%. On the other hand, due to the different methods of preparing and presenting long-term forecasts, it is difficult to discuss their verifiability, as these forecasts rather show the trend of thermal changes or the probability of precipitation. The analysis of long-term weather forecasts in terms of their verifiability is not currently being carried out (<https://forum.mazury.info.pl/viewtopic.php?t=15409> accessed on 23 November 2020).

The nature of the weather in Poland was described by the atmosphere physicist Prof. Teodor Kopcewicz: “Poland lies in a zone of moderate climate, with unmoderated weather changes”. Poland is one of the more difficult places in the world to forecast the weather. This is because of Poland’s climate, which is characterized by high weather variability and significant fluctuations in the course of the seasons in subsequent years. The physical and geographical location of the country means that various masses of air crash over its area, which influences the weather and the climate of Poland.

Frequently moving atmospheric fronts, attrition, and the exchange of various air masses (hot and cold) cause the weather to change frequently and creating great problems with weather forecasting [68]. We can find confirmation of the difficulties in inaccurate weather forecasting in Poland in a recent paper that draws attention to the seasonality of the weather in Poland throughout the year [69].

The atmospheric circulation in this part of Europe is characterized by relatively high annual variability, which causes significant temperature and precipitation fluctuations during the year [70]. There is a paper presenting a workshop on the occurrence of tornadoes in Poland, which confirmed the wide spectrum of weather phenomena observed in Poland, which makes it difficult to accurately predict them [71].

2.2. *The Specific Climate of Szczecin and Its Influence on Weather*

The weather of Szczecin is reflected in the specific climate of Szczecin’s Climatic Land (VI and X), which is influenced by its location near the sea, many lakes, a large river basin, landforms, large forest areas, parks and meadows, street greenery, and relief (valleys and hills as illustrated on the Map in Figure 1). The climate of Szczecin and its surroundings is shaped primarily by the advection of polar and sea air masses. The proximity of large water reservoirs, i.e., the Baltic Sea and the Szczecin Lagoon, results in the formation of local breeze circulation affecting the course of the weather.

The Baltic Sea and the Szczecinski Lagoon have a warming effect in the winter, and cause cooling in the summer. Important climatic factors include the latitude, terrain, and elevation above sea level. From the southwest to the northeast, through the center of the province extends the front moraine shaft, which clearly differentiates the spatial distribution of sunshine, temperature, precipitation, and wind speed on its northwestern and southeastern side. The main part of the baric systems moves from western directions. The shifting lowers with atmospheric fronts, which causes weather changes and strong and stormy winds. Spring and summer blades, although many, are less active and strong storm and glare winds.

The Skagerrak low-pressure formations because of a wave disorder on a front approaching Norway—they typically move in a southeastern direction. Deepening rapidly, they cause storm weather. These occur most often in the winter and spring. Atmospheric circulation is formed under the influence of the Icelandic low (especially in winter) and the Azores high (mainly in summer). The climatic conditions in winter are significantly influenced

by the strong seasonal Siberian High. Due to the northwestern extension, time differences between the Baltic coast and the southern ends of Poland are clearly visible. In summer, the day is more than an hour longer than in the uplands of southern Poland. In winter, the day lasts equally as long. The shift of about 40 min likewise occurs between the eastern and western parts of the country [72,73].

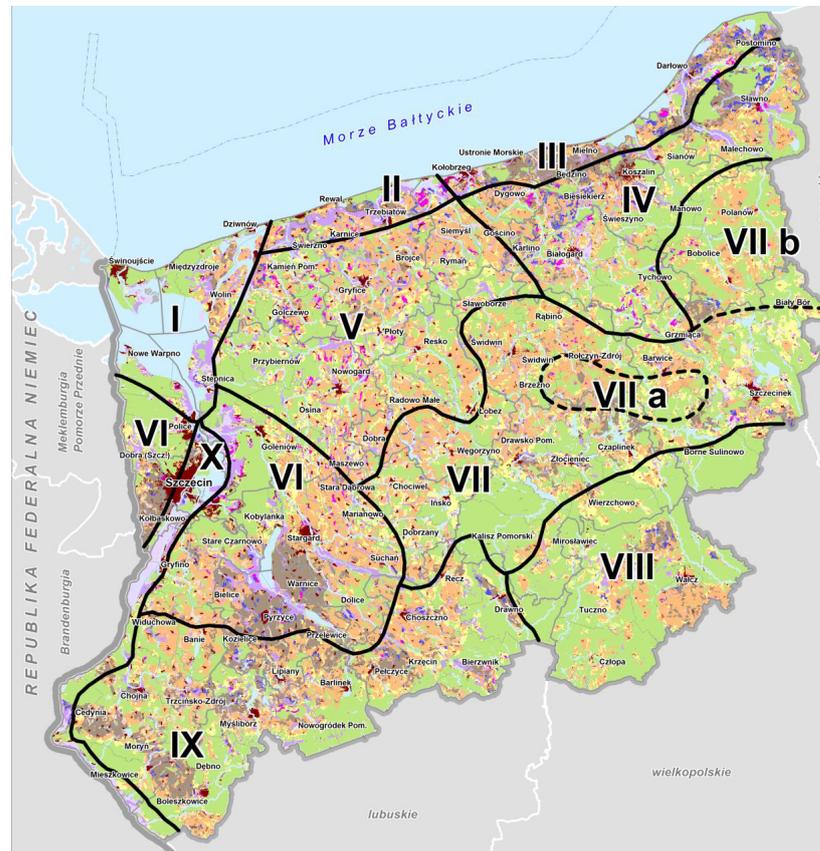


Figure 1. Climatic land map of Szczecin (VI and X) on the map of climatic lands (I–X) and sub-lands (VII a, VII b) of the West Pomeranian Province, Source: Spatial Development Plan for West Pomeranian Province Volume I. The conditions for shaping the spatial policy of the province of 24 June 2020.

Due to the described specificity of the climate in the area of the city of Szczecin, which causes difficulties in exact weather forecasting, the authors of the work decided that it is worth checking how the task of prediction of selected weather conditions for this area will perform with a MLP model, whose use for the forecasting of selected weather conditions was presented in many research articles cited in the literature review.

The location where the study was performed and the number of different predicted weather conditions included in the study are novel compared to the previous studies described in the literature that used machine learning (ML) methods. It is also a new and interesting approach of the authors to study the performance of the implemented MLP model for multiple different weather conditions, whereas the reviewed publications on this issue usually presented studies on a single weather parameter.

3. Materials and Methods

The Multilayer Perceptron (MLP) is a popular and commonly used model among neural networks. The literature study presented in the previous section enabled us to identify many papers from recent years that confirmed the actuality, usefulness, and popularity of using this model in forecasting various weather conditions for different locations around the world. This type of neural network is referred to as a supervised network [15].

Supervised learning comprises training the model using a training dataset, i.e., a set of samples containing known expected output signals. One type of supervised learning used in forecasting results with continuous values is regressive analysis. The model described contains explanatory (training) and explained (predicted) variables. Both types of variables are continuous values. The purpose of this type of network is to create a model that correctly maps the input data to the output data using historical data so that the model can then predict the output data when the desired output data is not known [15]. It enables detecting relationships between variables and predicting future results [74,75].

A multilayer neural network can approximate any function with continuous values between input and output vectors of data by picking the appropriate set of weights. The described properties of neural networks allow solving the problems of forecasting phenomena occurring in the natural world based on collected historical data, such as weather phenomena. The fundamentals of Artificial Neural Networks and the Multilayer Perceptrons with descriptions of their structures and explanations of the principles of operations, including mathematical formulae, are contained in Appendix A.

3.1. Overview of the Available Datasets and Selection of a Training Dataset

The dataset from the meteorological station with code 205 located at Szczecin was taken from the archives of Institute of Meteorology and Water Management (IMGW) (https://dane.imgw.pl/data/dane_pomiarowo_obserwacyjne/dane_meteorologiczne/dobowe/synop/ accessed on 27 April 2021). This source of meteorological data was considered reliable due to the status of IMGW as a state research unit. The real-time forecasting of meteorological phenomena is one of the basic tasks of modern meteorological services. In Poland, official weather forecasts are prepared and developed by the Institute of Meteorology and Water Management-National Research Institute (IMiGW-PIB)—a state unit supervised by the minister in charge of water management. The mission of IMiGW-PIB is to inform society and organizations about the weather—meteorological and hydrological, climate change, and all factors influencing the current weather in Poland.

The major task of the IMGW-PIB is to provide meteorological cover for Poland. For this purpose, the IMGW-PIB Monitor was created—a service for all national operational services and administrative bodies. As part of its statutory activities, the IMGW-PIB prepares and delivers meteorological forecasts, warnings against dangerous phenomena occurring in the atmosphere, and dedicated announcements and bulletins. IMGW is responsible for collecting, storing, processing, and making available domestic and foreign measurements and observation materials.

The Institute develops and distributes weather forecasts and warnings. IMGW is a member of many international organizations. It represents Poland on the forum of WMO (World Meteorological Organisation) or Eumetsat (European Organisation for the Exploitation of Meteorological Satellites). IMGW stores, and makes available in its database, measurement and observation data collected since 1960 for 2112 measuring stations in Poland (https://danepubliczne.imgw.pl/data/dane_pomiarowo_obserwacyjne/dane_meteorologiczne/ accessed on 27 April 2021).

A review of the literature on the use of multilayer neural networks in the prediction of selected weather conditions demonstrated that, for many studies, ten years of previous meteorological data were used [74,75]. We, therefore, decided to use meteorological data from 10 years (2011–2020). In this study, daily data from 7 years (2011–2017) were used as a training dataset with 10 training features for the 3 days before the day for which the forecast was performed.

The test dataset included data from 3 years (2018–2020). One sample contained selected weather conditions for each day. Due to the many samples in the dataset, we chose to apply a simple validation (train/test split) to check the effectiveness of the model. The dimensions of the datasets used for model training and testing are presented in Table 1.

Table 1. Dimensions of the datasets used for training and testing the Multilayer Perceptron (MLP) model.

Dataset	Number of Samples	% of Total Dataset
Training	2554	69.97%
Test	1096	30.03%
Total	3650	100%

The ten features selected to create a training dataset, together with their respective units, are listed in the Table 2.

Table 2. Features of the training dataset (inputs for the MLP model).

	Training Feature	Unit
1.	Maximum daily temperature	degrees Celsius [$^{\circ}\text{C}$]
2.	Minimum daily temperature	degrees Celsius [$^{\circ}\text{C}$]
3.	Mean daily temperature	degrees Celsius [$^{\circ}\text{C}$]
4.	Minimum daily temperature at ground level	degrees Celsius [$^{\circ}\text{C}$]
5.	Daily sum of precipitation	millimeters [mm]
6.	Daily dew duration	hours [h]
7.	Mean daily general cloud cover	octants [octants]
8.	Mean daily wind speed	meters per second [m/s]
9.	Mean daily relative humidity	percentages [%]
10.	Mean daily atmospheric pressure at the station level	hectopascals [hPa]

These variables from 1, 2, and 3 days before the prediction date were the inputs for the MLP model. The time horizon of 3–5 days before the day for the forecast was used, among others, by Rasp and others in a study published in 2020 [76]. Table 3 contains the conditions planned for forecasting using the MLP model in this study. They are the outputs of the MLP model.

We selected five weather conditions that are interesting for people from the point of view of watching the weather forecast and are most often provided in weather forecasts, i.e., the maximum and minimum temperature, pressure, wind speed, and precipitation. We further chose these because, in the weather's archive service, which was our point of reference, these conditions were given. The service here presented draws from the OpenWeather service; therefore, it was useful to treat it as a reference point locally for the city of Szczecin.

Table 3. The weather conditions selected for the prediction in the MLP model.

	Predicted Condition	Unit
1.	Maximum daily temperature	degrees Celsius [$^{\circ}\text{C}$]
2.	Minimum daily temperature	degrees Celsius [$^{\circ}\text{C}$]
3.	Mean daily atmospheric pressure at station level	hectopascals [hPa]
4.	Mean daily wind speed	meters per second [m/s]
5.	Daily sum of precipitation	millimeters [mm]

3.2. Data Preprocessing

The input data were normalized, and the five selected daily weather conditions were forecasted for the full 2018, 2019, and 2020 years. The study includes daily short-term forecasts for the next day. MLP was used as a prediction model. To evaluate the prediction accuracy, the results were compared with the values of measurements made available by the Institute of Meteorology and Water Management (IMGW). One stage of data preparation, preceding the use in machine learning, was the preprocessing.

With data concerning the daily precipitation sum, the lack of precipitation was recorded in files downloaded from IMGW as the lack of a value, which is interpreted in Python as a NaN value (Not a Number—value not being a number). It was, therefore, necessary to take measures to obtain the correct results of the implemented models with the data used.

Missing values, with precipitation data, were completed with 0. No missing data were found for the other parameters available in the dataset. However, if some weather data were missing, a suitable solution to fill them in would be to use the k -Nearest Neighbors algorithm for the specificity of successive weather variables in time. In the next step, the dataset was divided into the training and test dataset.

Normalization of the Input Data

Neural networks generally provide better performance when working on normalized data. Using the original data as the input to a neural network may cause algorithm convergence problems. The absence of a normal distribution is a common occurrence for environmental data [77]. Python provides several classes in which data scaling procedures are implemented. One of these procedures is a method that allows for parametric, monotonous transformations, the aim of which is to map data from any distribution to as close as possible to the Gaussian distribution to stabilize the variance and minimize the skew. This method belongs to a new power transformation family that is well defined on the whole actual line and is appropriate for reducing skewness to approximate normality. It has properties similar to those of the Box–Cox transformation for positive variables.

The large-sample properties of the transformation were investigated in a single random sample [78]. We estimated the transformation parameter by minimizing the weighted squared distance between the empirical characteristic function of transformed data and the characteristic function of the normal distribution [79].

This method provides two types of transformations: Yeo–Johnson transformation and the Box–Cox transformation. The Box–Cox conversion can only be used for strictly positive data; therefore, the first variant should predict temperatures that may be negative or zero. The Yeo–Johnson transformation is defined by the Formula (1):

$$x_i^{(\lambda)} \begin{cases} [(x_i + 1)^\lambda - 1]/\lambda & \text{if } \lambda \neq 0, x_i \geq 0, \\ \ln(x_i) + 1 & \text{if } \lambda = 0, x_i \geq 0 \\ -[(-x_i + 1)^{2-\lambda} - 1]/(2 - \lambda) & \text{if } \lambda \neq 2, x_i < 0, \\ -\ln(-x_i + 1) & \text{if } \lambda = 2, x_i < 0, \end{cases} \quad (1)$$

while the Box–Cox transformation is described by Formula (2):

$$x_i^{(\lambda)} \begin{cases} \frac{x_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln(x_i) & \text{if } \lambda = 0. \end{cases} \quad (2)$$

Standardization of a dataset is a common requirement for many machine learning estimators. Typically, this is done by removing the mean and scaling to the unit variance. However, outliers can often influence the sample mean/variance negatively. In such cases, the median and the interquartile range often provide better results. The interquartile range (IQR) is robust to the outliers that can occur with weather data. In this method, the median is removed and then the data are scaled according to the range of quantiles between the first quartile (25th quantile) and the third quartile (75th quantile) [80]. Due to the nature of the dataset studied, we decided that it was necessary to compare the performance of the prediction model for the two methods described.

3.3. Design and Implementation of the Multilayer Perceptron Model

All stages of the study (data preparation, implementation of the Multilayer Perceptron model, and model testing) were performed in Python programming language.

3.3.1. The MLP Parameters and Their Values

The structure of the neural network (the number of layers, number of neurons in particular layers, type of activation function, and topology of networks) as a tool for modeling proper objects was determined in relation to the problem to be resolved. As a starting point, which is as good as any other solution, a network with one hidden layer can be adopted. The best results are obtained by selecting the number of layers and neurons in the layers empirically. This choice is properly arbitrary and depends on the model creator. A grid search including cross-validation was used to determine the values of the ML model hyperparameters [81]. The range of parameter values for the cross-validation procedure was selected based on [75].

- λ —the L2 weight adjustment parameter to reduce model overtraining, making the model more simple and less susceptible to over-matching. The regularization consisted of applying penalties to the network parameters.
- Number of epochs—the number of iterations of the algorithm across a set of the training dataset.
- η —the network learning rate was used for updating the weights.
- α —a parameter for momentary learning that defines the part of the value of the previous gradient added to the updated weights to speed up the learning of the network.
- decrease_const—a constant of reduction d as part of an adaptive learning rate that decreases in subsequent epochs for higher convergence.

3.3.2. Methods Implemented in MLP Model

The most important implemented procedures that were run sequentially when training and testing the MLP network are presented in Figure 2 [75]. The methods implemented in the Multilayer Perceptron model are listed below:

- Initialization of weights.
- Sigmoid activation function.
- The derivative of the sigmoid activation function.
- Adding a bias as a vector of ones to the first column or first row.
- Forward propagation.
- Backpropagation including regularization procedure.
- Prediction.
- Fit—network training. In this method, the following procedures were initiated iteratively in subsequent epochs:
 - Forward propagation.
 - Backpropagation.
 - Calculation of the error made by the network during learning.
 - Updating of weights.

The weights were initialized randomly, with the assumption of a uniform distribution in the numerical range $[-0.001, 0.001]$.

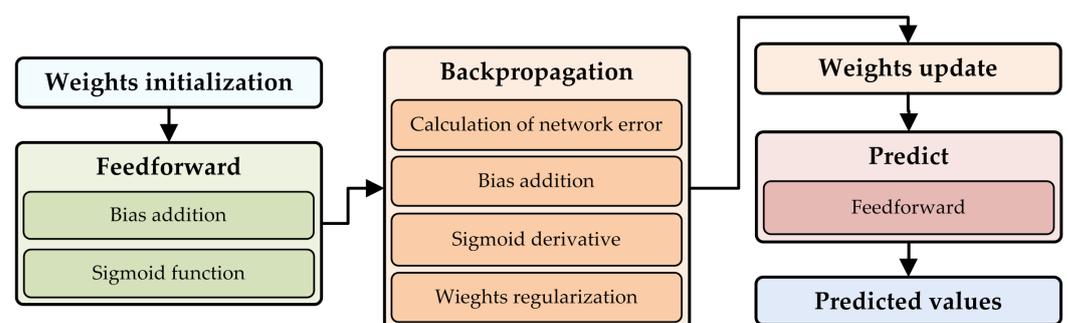


Figure 2. The procedures implemented in the MLP model.

3.3.3. Regularization—Preventing over-Adjustment of the Model

Excessive fitting or over-training of the model is one of the most common problems that appears in machine learning. It occurs when the model works well for the training data but does not generalize the learned rule sufficiently for the unknown test data. One technique used to prevent model over-training is to adjust the complexity of the model by regularization. This is based on introducing additional data and penalizing large scale values. The most common type of regularization is L2 regularization, which is also called the decomposition of weights. This regularization is defined by the Formula (3):

$$\frac{\lambda}{2} \|\mathbf{w}\|^2 = \frac{\lambda}{2} \sum_{j=1}^m w_j^2, \quad (3)$$

where λ is the adjustment parameter. Adjustment is the reason why scaling the features (e.g., normalization) is important. To carry out the adjustment correctly, all features must be adjusted to a uniform scale [74,75]. After the prediction is made, the data is denormalized to present the results of the model with the appropriate values and units for the predicted condition.

3.3.4. Architecture of the Multilayer Perceptron Model

A network architecture was designed with inputs in the input layer, whose role was performed by the most suitable variables selected from 10 weather indicators (features) for the training dataset from 1, 2, and 3 days prior to the day for which a given weather condition is predicted. The decision to use the values of the training features from the 3 days preceding the forecast date was taken based on a review of studies on the forecast of the selected weather parameters available in the literature and online sources [82].

We considered this variant of the training dataset containing the values of each of the training parameters from 1, 2, and 3 days before the day on which the prediction will be made to be sufficient and reasonable. The selected weather parameters for the training dataset are presented in Table 2.

The number of neurons of the hidden layer was established experimentally after implementing the neural network model, by applying the cross-validation method, individually for each predicted parameter. In the case of a neural network, a smaller number of hidden units was beneficial if it did not adversely affect the accuracy of the results generated by the network, as this reduces the learning time, making the network more efficient. The output layer contains a vector with expected values for the input data.

3.4. Sensitivity Analysis

The procedure of sensitivity analysis was used to investigate the effect of each parameter on the outputs. Sensitivity analysis using the change of MSE ranked the input variables in a given dataset according to the change of MSE when each input was deleted from the dataset in the training phase. Therefore, the variables that made the largest change in the MSE were considered as the most important [83].

3.5. Methods to Evaluate the Effectiveness of Regressive ML Models and to Measure the Correlations between Variables

To objectively assess the performance of the implemented predictive model, we used two measures of the error committed by the network, such as the MAE, MSE, and R^2 , applied in the evaluation of the prediction accuracy of regressive models. To answer the question of why some weather variables were predicted better and others worse, we examined the relationships between the variables included in the model, since ML models are data-driven, and consequently the prediction accuracy naturally depends on the strength of these correlations [18].

We used Pearson's correlation coefficient to examine the correlation between weather conditions. The foundations of the methods used to assess the effectiveness of the regressive models and Pearson's correlation coefficient are presented in Appendix B.

4. Results

4.1. Weather Conditions Prediction Accuracy Using Different Data Preprocessing Methods

Table 4 presents a comparison of the weather condition prediction accuracy metrics from the MLP model for the two normalization methods Yeo–Johnson transformation and IQR. Our analysis of the results allowed us to conclude that the accuracy of the MLP prediction for the compared methods differed depending on the forecasted weather condition and, for temperature, the atmospheric pressure and wind speed were higher for the Yeo–Johnson transformation, while, for daily precipitation, better results were obtained using normalization taking into account the IQR method.

Table 4. Comparison of the weather condition prediction accuracy using the Yeo–Johnson transformation and IQR in data preprocessing.

Predicted Condition	Preprocessing Method	MAE	MSE	R ²
Maximum temperature	IQR	2.0184	6.7045	0.9562
	Yeo–Johnson trans.	2.0136	6.5105	0.9575
Minimum temperature	IQR	1.8037	4.9932	0.9397
	Yeo–Johnson trans.	1.7524	4.7432	0.9427
Atmospheric pressure	IQR	4.0796	29.7486	0.8292
	Yeo–Johnson trans.	4.0458	29.3731	0.8313
Wind	IQR	0.9159	1.4239	0.5797
	Yeo–Johnson trans.	0.8951	1.3627	0.5977
Precip.	IQR	1.8417	10.5228	0.5196
	Yeo–Johnson trans.	1.2230	11.4712	0.4763

4.2. Results of Data Preprocessing

Figures A1–A4 in Appendix C display histograms of the input data before and after the data preprocessing. The shape of the histograms reflects the nature of the data resulting from the specific climate of the city of Szczecin. The presentation of histograms before and after data normalization next to each other allows for a clear visualization of the effect of data normalization on the the histogram shape change. The histogram of the maximum daily temperature in Figure A1a demonstrated an interesting characteristic. This plot shows that the data was multimodal, which may be due to two different sets of environmental circumstances.

The maximum temperature swings in the area of the city of Szczecin can be quite extreme, particularly between seasons. The histogram for the minimum daily temperature shown in Figure A1c after normalization obtained a shape more similar to a histogram with a normal distribution as shown in Figure A1d. The histogram representing the mean daily temperature seen in Figure A2a indicates that this data was somewhat multimodal. Normalization eliminated this effect as can be seen in Figure A2b.

In the case of the minimum daily temperature at ground level, a slight shift toward the high values observed in the histogram in Figure A2c was canceled out after applying normalization, as shown in Figure A2d. For the daily sum of the precipitation, the outliers visible in the histogram in Figure A2e are easy to explain because dry days (days without precipitation) were more frequent. Outliers are also visible for the data in the histogram for the daily dew duration in Figure A3a.

Szczecin's climate is characterized by high cloudiness, humidity, and fog, which is reflected in the histograms presenting the input data representing the weather parameters [72]. The histogram of the mean daily general cloud cover in Figure A3c shows that

there were more samples with high values as a result of the large number of days per year with high cloud cover due to the specific climate of the city of Szczecin. In the case of the histogram for the data representing the mean daily humidity as seen in Figure A4a, normalization reduced the shift toward high values as can be seen in Figure A4b.

The histogram for the data representing the mean daily wind speed as shown in Figure A3e presents that more samples had low values due to fewer days with strong winds. After normalization, this effect was significantly reduced, as shown in Figure A3f. The histogram representing the data for the mean daily atmospheric pressure in Figure A4c shows that the data exhibited a normal distribution because the histogram has a characteristic bell shape. The mean value is located in the central part of the histogram.

4.3. Values of the MLP Hyperparameters

Table 5 presents the MLP hyperparameters as determined using a grid search, including the k-fold cross-validation procedure—the results of which are shown in Table A7 in Appendix E.

Table 5. The MLP hyperparameters as determined using a grid search.

Parameter	Max. Temp.	Min. Temp.	Pressure	Wind	Precip
Hidden neurons	80	80	90	40	50
Learning rate	0.001	0.001	0.001	0.001	0.001
Number of epochs	500	750	250	500	250
L2	0.001	0.001	0.01	0.001	0.001
Decrease const	0.00001	0.0001	0.00001	0.0001	0.00001
Alpha	0.001	0.0001	0.001	0.01	0.0001

4.4. Training Variables Selection

The selection of the most suitable subset of training variables was performed by predicting different targets by using the same variables coming from the 3 previous days for the forecasted day [84]. The results of this procedure as performed for prediction of the maximum daily temperature are contained in Table 6. The values of the MSE are ranked in ascending order. The largest error value was obtained for the MLP training and testing with the daily sum of the precipitation, mean daily wind speed, and mean daily general cloud cover. These variables were, therefore, excluded from the predictive model for the maximum daily temperature.

The results of the variable selection for other predicted weather conditions are included in Tables A8–A11 in Appendix E, and the excluded variables are presented in Table 7. Variables that were not excluded became inputs of the models.

Table 6. The results of the selection of the most suitable subset of the training variables for the prediction of the daily maximum temperature.

Variable	Training MSE	Testing MSE
Maximum daily temperature	7.6154	8.0929
Mean daily temperature	8.7251	9.4972
Minimum daily temperature	20.34	23.781
Min. daily temp. at ground level	25.0103	28.7659
Daily dew duration	44.6071	48.8083
Mean daily relative humidity	55.5796	54.0218
Mean daily atm. pressure	58.2674	60.8849
Mean daily general cloud cover	60.6825	56.0763
Mean daily wind speed	69.7158	68.5501
Daily sum of precipitation	73.4619	77.2748

Table 7. The training variables excluded from the model based on the variable selection.

Predicted Condition	Excluded Variables
Maximum daily temperature	Daily sum of precipitation, mean daily wind speed, mean daily general cloud cover
Minimum daily temperature	Daily sum of precipitation, mean daily wind speed, mean daily general cloud cover
Mean daily atmospheric pressure	Mean daily relative humidity, maximum daily temperature, daily dew duration
Mean daily wind speed	Mean daily relative humidity, maximum daily temperature, mean daily general cloud cover
Daily sum of precipitation	Mean daily relative humidity, mean daily general cloud cover, mean daily wind speed

4.5. Results of Sensitivity Analysis

The results of the sensitivity analysis for the maximum temperature prediction with the values of the MSE ranked in descending order for the training and testing datasets are presented in Table 8. The results for other predicted weather conditions are included in Tables A12–A15 in Appendix E. In the case of the prediction of the maximum temperature, training and testing MLP model without the mean daily temperature, daily dew duration, and mean daily atmospheric pressure resulted in a remarkable increase in the error values.

In the event of the prediction of the minimum temperature, training and testing the MLP without the mean daily temperature, minimum daily temperature, and mean daily relative humidity caused a rise in the error values. When the mean atmospheric pressure was predicted, training and testing the MLP without the mean daily atmospheric pressure, mean daily temperature, and mean daily general cloud coverage resulted in growth of the error values.

For the wind speed prediction, training and testing the MLP without the mean daily wind speed, the daily sum of precipitation, and the mean daily temperature resulted in a remarkable increase in the error values. In the case of the prediction of the daily sum of the precipitation, training the MLP without the daily sum of precipitation, mean daily atmospheric pressure, and daily dew duration caused a rise in the error values.

Table 8. The results of the sensitivity analysis for the prediction of the maximum daily temperature.

Inputs Out	Training MSE	Testing MSE
Mean daily temperature	7.0542	7.6108
Daily dew duration	6.6200	6.9386
Mean daily atm. pressure	6.5968	7.0749
Min. daily temp. at ground level	6.4327	6.8041
Mean daily relative humidity	6.4202	6.7626
Maximum daily temperature	6.3103	6.8945
Minimum daily temperature	6.2282	6.6756

The results of the sensitivity analysis conducted for the model containing all input variables before selecting the most suitable set are included in Appendix D.

4.6. Comparison of the Results Obtained by MLP, Two Other ML Models, and the Weather Service

To check the effectiveness of the forecast obtained with the use of our MLP model compared to the accuracy of the weather service forecast, the obtained results of the forecast of the selected weather conditions by the MLP model for the next day were compared with the results of the weather service forecast for 2018 for the city of Szczecin.

The archival data of weather service forecasts especially for the city of Szczecin provided by OpenWeather (<https://openweathermap.org/> accessed on 29 April 2021) are available free of charge on the website under the link (https://www.ekologia.pl/pogoda/polska/zachodniopomorskie/szczecin/archiwum,zakres,01-01-2019_31-01-2019,calosc accessed on 29 April 2021). We used archived data from 1 January 2018 to 31 December 2020. OpenWeather is an open-source weather data service that requires no license to download.

OpenWeather's API provides a user-friendly way to download data for a selected city without having to enter the geographic coordinates, provides current and archived data

for 40 years back, and includes Polish diacritical marks, which is important for the location for which our study was conducted. The API of this service is also recommended by the Polish government service (<https://www.gov.pl/web/popcwsparcie/standard-api-dla-u-dostepniana-danych> accessed on 1 March 2021).

Table A16 in Appendix E presents sample values of the parameters forecasted from 1 January 2018 to 31 December 2020. The predicted parameters are listed below:

- Maximum daily temperature in °C.
- Minimum daily temperature in °C.
- Mean daily atmospheric pressure at station level in hPa.
- Mean daily wind speed in m/s.
- Daily sum of precipitation in mm.

We used the MAE, MSE, and R^2 score as a measure of the accuracy of the MLP model prediction and weather service, calculated for the test data as these are frequently used and proven measures found in the literature for the prediction of weather parameters [32,35–37]. The smaller the values of MAE and MSE were, the more accurate the model was in predicting the tested parameters. High values of the R^2 score, close to 1, were evidence of the high efficiency of the tested model.

The tables with values of the weather conditions measured according to the IMGW, predicted by the MLP, and predicted by the service, include sample data from 1 January 2018 to 5 January 2018. This demonstrated the values of the parameters that we are investigating; however, the comparison was made for data from 1 January 2018 to 31 December 2020, and thus our study covers the whole 2018, 2019, and 2020 years. To compare the effectiveness of the MLP model and the weather service, quantitative model effectiveness models, such as MAE, MSE, and the determination factor R^2 score, were used.

Charts of the parameters measured and projected by the MLP model and weather service were compared in terms of the visual differences in the graphs. The tables also contain the prediction results of the studied weather conditions obtained using two other ML predictive models. One of the benchmark models was the LSTM neural network with a complexity of architecture and parameters comparable to the MLP model used in the study. The second model used for comparison was SVR.

These models were chosen because they are often used in weather prediction studies, as mentioned in our literature review. The ANN models were trained and tested 10 times; therefore, in the tables, the mean and standard deviation values of the results are provided. The results of the evaluation models on the training and testing dataset are presented in Tables 9–13.

Based on the small differences between the accuracies for the training and test data, there was no overtraining of the models or over-fitting to the training data. In the next step, the prediction results of the tested models and the weather service were compared for the test data.

Table 9. Comparison of the effectiveness of the MLP model, two other models, and the weather service in forecasting the maximum temperature.

	MLP				LSTM				SVR		Service
	Train		Test		Train		Test		Train	Test	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std			
MAE	1.9842	0.0095	2.0201	0.0212	2.0871	0.0339	2.0466	0.0180	1.6747	2.0792	2.3453
MSE	6.3276	0.0646	6.5448	0.0716	6.9771	0.2054	6.7315	0.1205	4.9096	7.0436	9.2178
R^2	0.9570	0.0004	0.9573	0.0005	0.9525	0.0014	0.9561	0.0008	0.9666	0.9540	0.9398

Table 10. Comparison of the effectiveness of the MLP model, two other models, and the weather service in forecasting the minimum temperature.

	MLP				LSTM				SVR		Service
	Train		Test		Train		Test		Train	Test	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std			
MAE	1.8022	0.0172	1.7588	0.0156	1.9652	0.0192	1.9054	0.0269	1.5738	1.9338	2.8084
MSE	5.1024	0.0925	4.7804	0.0685	6.0593	0.1361	5.7054	0.1724	4.2701	5.8238	14.1628
R ²	0.9440	0.0010	0.9423	0.0008	0.9335	0.0015	0.9311	0.0021	0.9531	0.9297	0.8290

Table 11. Comparison of the effectiveness of the MLP model, two other models, and the weather service in forecasting the mean daily atmospheric pressure.

	MLP				LSTM				SVR		Service
	Train		Test		Train		Test		Train	Test	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std			
MAE	3.7830	0.0087	4.0537	0.0094	4.2301	0.0564	4.0602	0.0138	3.1029	4.2030	4.4267
MSE	24.1300	0.0553	29.1347	0.1087	29.4683	0.7534	29.3869	0.1467	18.6167	32.2005	37.4616
R ²	0.8427	0.0004	0.8327	0.0006	0.8079	0.0049	0.8312	0.0008	0.8787	0.8151	0.7849

Table 12. Comparison of the effectiveness of the MLP model, two other models, and the weather service in forecasting the mean daily wind speed.

	MLP				LSTM				SVR		Service
	Train		Test		Train		Test		Train	Test	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std			
MAE	0.9097	0.0181	0.8795	0.0119	0.9354	0.0048	0.8551	0.0035	0.8042	0.8765	1.2526
MSE	1.3807	0.0559	1.3453	0.0276	1.4480	0.0135	1.2764	0.0041	1.1987	1.3334	2.6179
R ²	0.6719	0.0133	0.6029	0.0081	0.6559	0.0032	0.6232	0.0012	0.7152	0.6064	0.2272

Table 13. Comparison of the effectiveness of the MLP model, two other models, and the weather service in forecasting the daily precipitation sum.

	MLP				LSTM				SVR		Service
	Train		Test		Train		Test		Train	Test	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std			
MAE	2.0415	0.0063	1.8236	0.0143	2.0196	0.0374	1.7475	0.0633	1.5551	1.2230	1.0505
MSE	15.6428	0.0512	10.5480	0.0374	16.2313	0.1525	10.5223	0.0419	17.8979	11.5570	10.0980
R ²	0.5371	0.0015	0.5184	0.0017	0.5197	0.0045	0.5196	0.0019	0.4703	0.4724	0.5390

4.6.1. Maximum Temperature Prediction

Table A17 in Appendix E includes sample maximum temperature values as measured, predicted by the MLP, and predicted by the weather service. Table 9 presents the quantitative values of the MLP and two other predictive model effectiveness measures as well as the weather service in forecasting the maximum temperature. Figure 3 illustrates a comparison of the effectiveness of the MLP model and the weather service in forecasting the maximum temperature.

The MAE was equal to 2.0201 for the MLP and this is lower by 0.3252 compared with the MAE calculated for the weather service, which was equal to 2.3453. This shows that the MLP was more accurate compared with the weather service in terms of this measure of prediction accuracy. The MSE calculated for the MLP was lower than MSE calculated

for the weather service by 2.673, which shows the higher accuracy of the MLP over the weather service in terms of the MSE values.

The value of the R^2 score for MLP was higher by 0.0175 compared with for the weather service; therefore, in terms of the R^2 score, the accuracy of MLP in forecasting the maximum temperature was higher. The accuracy of the MLP model for the maximum temperature forecasting was higher when compared with the weather service. Comparison with the LSTM and SVR model based on the accuracy metrics used showed a slightly higher accuracy of the MLP model over the compared models.

All of the models tested were able to achieve greater accuracy in predicting the maximum temperature compared with the weather service. A slightly higher effectiveness of the MLP model compared with for the weather service is noticeable on the graph. The differences between the measured and predicted values in the graph are small, both for the MLP model and for the weather service. This shows the high accuracy of the MLP model for the maximum temperature forecasting for the next day.

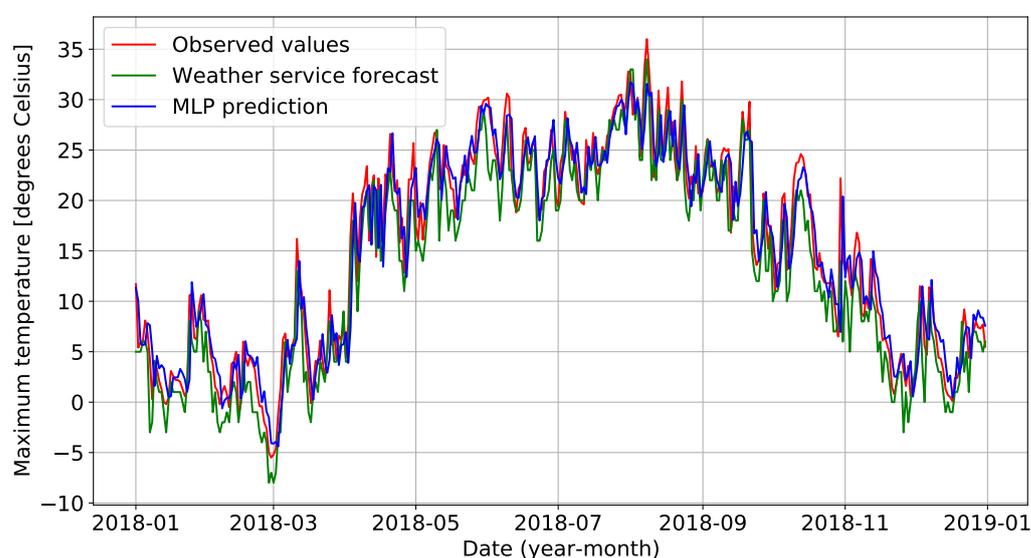


Figure 3. Comparison between the observed values and those predicted by the MLP and weather service maximum temperature.

4.6.2. Minimum Temperature Prediction

Table A18 in Appendix E contains examples of the minimum temperature observed values, the values predicted by the MLP, and those predicted by the weather service. Table 10 includes the quantitative values of the effectiveness measures of the MLP, two other predictive models, and the weather service in forecasting the minimum temperatures. Figure 4 displays a comparison of the effectiveness of the MLP and weather service in forecasting the minimum temperatures.

Here, the effectiveness of the MLP was higher than that of the weather service. The MAE was 1.0496 lower for the MLP, and the MSE had a value of 9.3824 less for the MLP than for the weather service. The determination factor R^2 score used to interpret the model effectiveness was 0.1133 higher for the MLP compared with for the weather service. These results show the much higher effectiveness of the MLP model over the weather service in forecasting the minimum temperature.

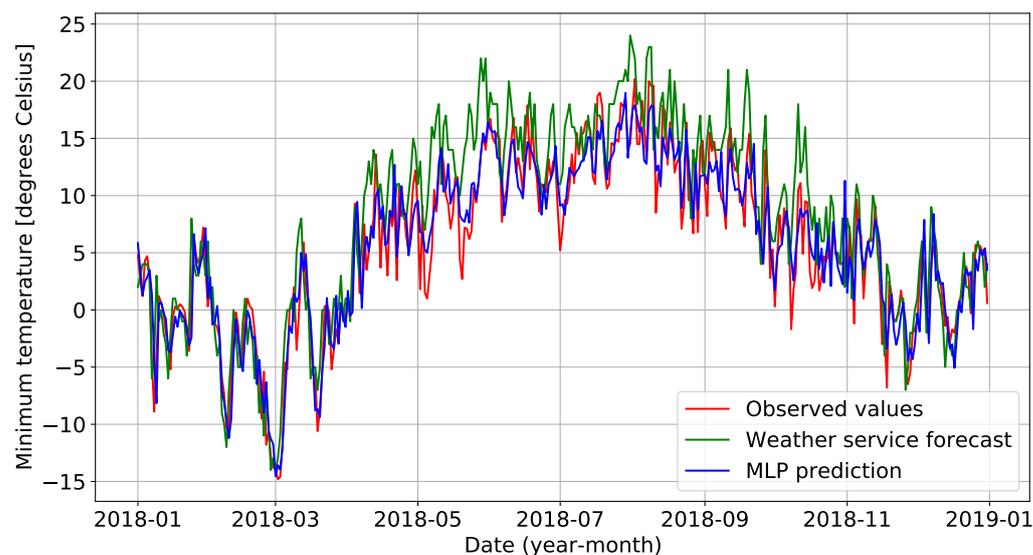


Figure 4. Comparison between the observed values and those predicted by the MLP and weather service minimum temperature.

For the compared models, the prediction accuracy of the daily minimum temperature was slightly higher for the MLP, while the LSTM model performed better than the SVR. The higher effectiveness in forecasting the minimum temperature for the MLP model compared with for the weather service is evidenced because the chart of values predicted by the MLP is closer to the chart of observed values, especially for the time range from April to October. In the spring and summer months, the weather service forecasts show greater divergence from the observed values than they do in the autumn and winter months, i.e., from January to March and from November to December. The MLP model showed high accuracy throughout the year.

4.6.3. Forecast of Mean Daily Atmospheric Pressure

Table A19 in Appendix E contains example values of the mean daily atmospheric pressure as measured, predicted by the MLP, and forecasted by the weather service. Figure 5 presents a comparison of the effectiveness of the MLP model and the weather service in forecasting the mean daily atmospheric pressure. Table 11 includes quantitative values of the effective measures of the MLP, two other predictive models, and the weather service in forecasting the mean daily atmospheric pressure.

The effectiveness of the MLP model was higher than that of the weather service when forecasting this parameter. The MAE value was 0.373 less for the MLP model, and the MSE value was 8.3269 less for the MLP model than for the weather service. The value of the determination factor R^2 score was 0.0478 higher for the MLP model than for the weather service, which shows the higher effectiveness of the MLP in predicting the atmospheric pressure.

However, the calculated value of R^2 score is not as high as the value of this factor calculated for the MLP when forecasting the maximum and minimum temperatures. This observation indicates that the accuracy of the forecasting of atmospheric pressure by the MLP was lower than that of the forecasting of temperatures. For the prediction of the daily atmospheric pressure values, the MLP model enabled slightly higher accuracy compared with the LSTM and SVR. A higher effectiveness in forecasting the atmospheric pressure was observed for the MLP model compared with for the weather service overall in the months in the surveyed year.

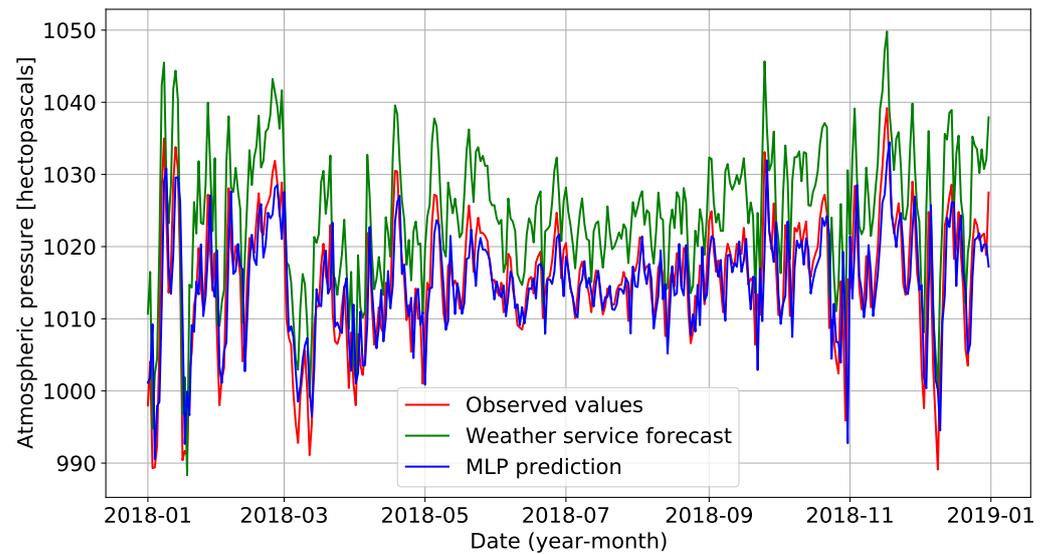


Figure 5. Comparison between the observed values and those predicted by the MLP and the weather service mean atmospheric pressure.

4.6.4. Mean Daily Wind Speed Prediction

Table A20 in Appendix E contains example values of the mean daily wind speed as measured, predicted by the MLP, and predicted by the weather service. Figure 6 illustrates a comparison of the effectiveness of the MLP model and weather service in forecasting the mean daily wind speed.

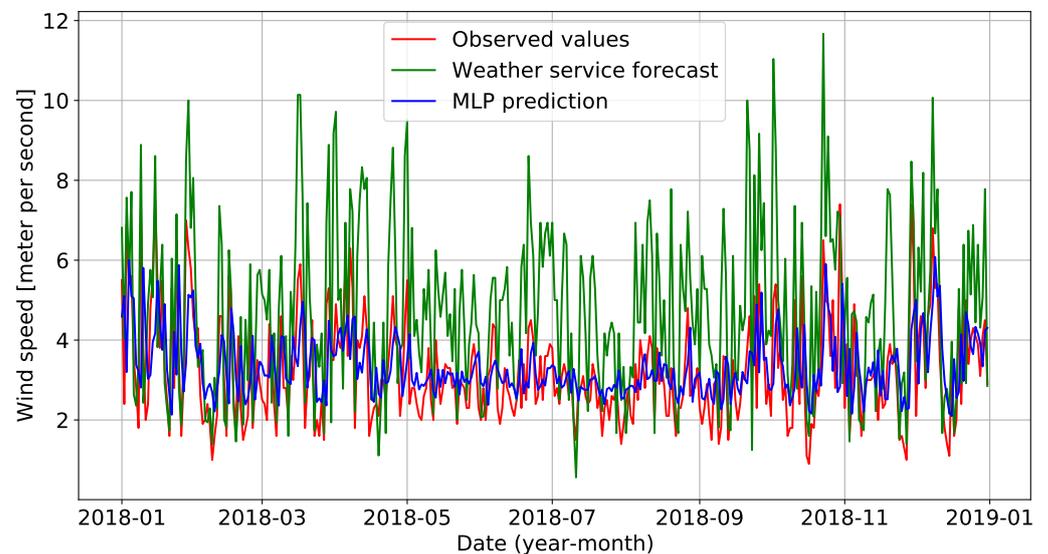


Figure 6. Comparison between the observed values and those predicted by the MLP and the weather service mean wind speed.

The effectiveness in the wind speed forecasting was higher for the MLP model than for the weather service as, in the graph, the values predicted by the MLP show much smaller deviations from the observed values compared with the values predicted by the weather service. Table 12 presents the quantitative values of the MLP, two other predictive model effectiveness measures, and the weather service in forecasting the mean daily wind speed. The MAE value was lower for MLP by 0.3731, and the MSE value was lower for MLP by 1.2726 compared with for the weather service; therefore, there was a higher accuracy of the MLP model compared to the weather service results.

The value of the determination factor R^2 score also showed a higher efficiency of the MLP model compared with the weather service for forecasting the mean daily wind speed as this value was higher by 0.3757 for MLP compared with for the weather service. However, this is lower compared to the value of the R^2 score calculated for the MLP model for temperature and atmospheric pressure prediction, which indicates a lower effectiveness of the MLP for wind speed prediction. For the prediction of the daily wind speed, the two other prediction models obtained higher prediction accuracy. The best results were obtained using the LSTM model, and the SVR model also showed a small advantage over the MLP model.

4.6.5. Daily Precipitation Prediction

Table A21 in Appendix E includes examples of the daily precipitation values as measured, predicted by the MLP, and forecast by the weather service. Figure 7 illustrates a comparison of the effectiveness of the MLP model and the weather service forecasting the daily precipitation sum.

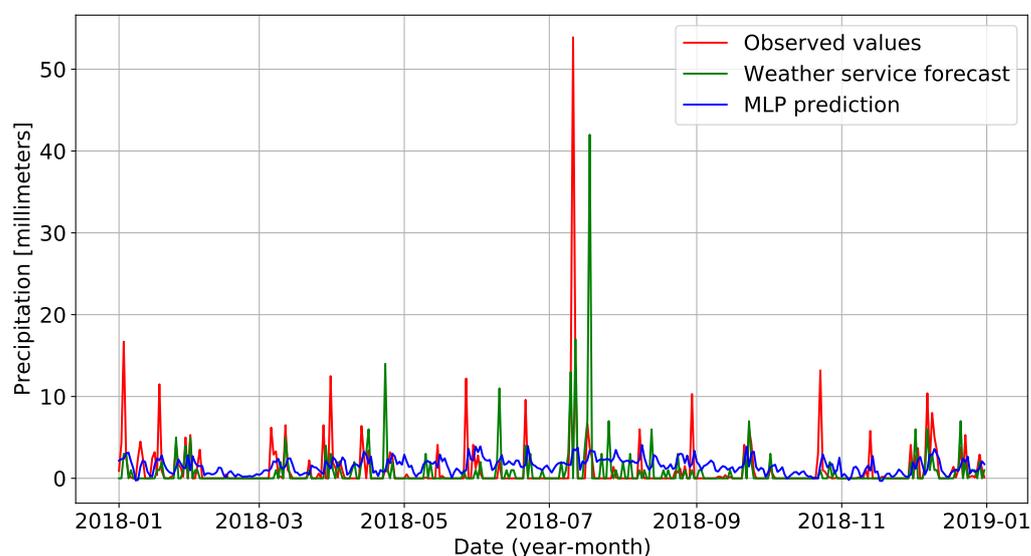


Figure 7. Comparison between the observed values and those predicted by the MLP and the weather service mean precipitation.

The effectiveness of the MLP model and the weather service forecasting the daily sum of the precipitation was lower than with the other weather parameters. Table 13 presents quantitative measures of the effectiveness of the MLP, two other predictive models, and the weather service in forecasting the daily precipitation sum. For the prediction of the daily precipitation, the values provided by the weather service were more accurate than the values predicted by the MLP model and the other models.

The value of MAE was 0.7731 higher for the MLP, and the value of the MSE was 0.45 higher for the MLP compared with for the weather service. The value of the R^2 score was lower by 0.0206 for the MLP when compared with the weather service. The prediction accuracy of the daily precipitation was comparable for the MLP and LSTM models, while it was lower for the SVR model. The effectiveness of all the predictive models was lower for the sum of the daily precipitation than it was for the other weather parameters analyzed in this survey.

Figure 8 shows a bar graph comparing the values of the prediction effectiveness measure R^2 score for all tested weather parameters for the MLP and weather service.

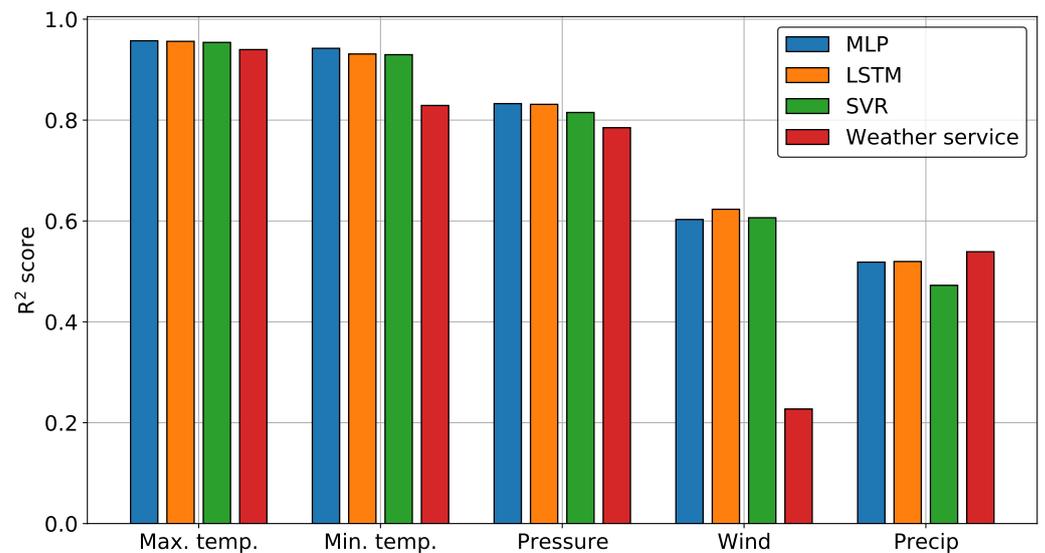


Figure 8. Comparison of the values of the R^2 score for all predicted parameters obtained by the MLP and weather service.

For the MLP, the highest predictive performance was observed at the maximum (0.9573) and minimum (0.9423) temperatures. This was followed by atmospheric pressure, for which the R^2 score was 0.8327, followed by wind speed (0.6029), and lastly by precipitation (0.5184). For the prediction of the daily maximum temperature, the predictions of the compared models and the weather service had high accuracy and were comparable; however, all models achieved higher accuracy compared with the weather service. For the daily minimum temperature, the prediction accuracy of the benchmarked models was high and comparable. They were further superior to the weather service than was observed with the maximum temperature.

In the case of prediction studies of the daily mean atmospheric pressure, the accuracy of the tested models was lower than for the temperatures. The MLP and LSTM models showed the highest prediction accuracy, and the prediction accuracy by SVR was slightly lower. However, all the models tested provided a more accurate prediction than the weather service did. The greatest advantage of the predictive models over the weather service was observed for the daily average wind speed. The LSTM model obtained the highest accuracy in this case, followed by SVR, and the accuracy of the MLP was minimally lower in comparison.

The daily precipitation is the weather condition that was the most difficult to accurately predict for the models used in this study. It is the only weather condition in this study for which the weather service forecast was more accurate than the other predictive models. MLP and LSTM were superior to the SVR model in terms of the daily precipitation prediction accuracy.

The output data histograms of the weather conditions predicted with the MLP model are presented in Appendix C in Figure A5. The observed shape of the histograms for the output data reflect the shapes and nature of the input data before normalization. The maximum daily temperature shows a multimodal distribution that was also observed for the input data representing this feature.

In the case of the mean daily wind speed, we observed that more samples were shifted toward low values on the histogram, which was similar to the input data. The mean daily atmospheric pressure demonstrated a normal distribution, which also reflects the nature of the input data for this parameter. For the daily sum of the precipitation, more samples had low values, similar to what was observed for the input data representing this parameter.

To investigate the reason that the predictions for target variables, such as the atmospheric pressure, wind speed, and daily precipitation sum, were less effective than the

predictions for temperature, we checked whether their values in the time series depend on the features based on which they were predicted or whether they are rather random. To check the dependence of the target weather variables on the explanatory variables (i.e., the traits), the strength of the correlation between the forecast weather parameters and the traits in the training dataset was examined using the r-Pearson correlation coefficient.

We examined the number of characteristics for which the absolute value of the Pearson's coefficient of correlation with the predicted parameter was equal to or greater than 0.6, as this value is considered to be a strong correlation threshold. Tables 14–16 includes characteristics that reached an r-Pearson correlation coefficient greater than or equal to 0.6 for the eight weather parameters that were training features. In the column next to the names of these features, the values of the r-Pearson correlation coefficient are presented. For predicted parameters, such as the wind and precip, there were no training characteristics where the correlation was greater than 0.6.

Table 14. The parameters with the strongest correlation to the predicted mean daily atmospheric pressure.

Parameter	Mean Daily Atmospheric Pressure
Mean daily atmospheric pressure 1 day ago	0.8014

Table 15. The parameters with the strongest correlation to the predicted maximum temperature.

Parameter	Maximum Daily Temperature
The minimum daily temperature on the ground 3 days ago	0.7492
The minimum daily temperature on the ground 2 days ago	0.7692
Minimum daily temperature 3 days ago	0.7894
The minimum daily temperature on the ground 1 day ago	0.7937
Minimum daily temperature 2 days ago	0.8109
Minimum daily temperature 1 day ago	0.8396
Mean daily temperature 3 days ago	0.8691
Maximum daily temperature 3 days ago	0.8770
Mean daily temperature 2 days ago	0.8969
Maximum daily temperature 2 days ago	0.9029
Mean daily temperature 1 day ago	0.9417
Maximum daily temperature 1 day ago	0.9493

Table 16. The parameters with the strongest correlation to the predicted minimum temperature.

Parameter	Minimum Daily Temperature
The minimum daily temperature on the ground 3 days ago	0.7665
Minimum daily temperature 3 days ago	0.7954
The minimum daily temperature on the ground 2 days ago	0.8028
Maximum daily temperature 3 days ago	0.8292
Minimum daily temperature 2 days ago	0.8307
Mean daily temperature 3 days ago	0.8461
Maximum daily temperature 2 days ago	0.8551
The minimum daily temperature on the ground 1 day ago	0.8674
Mean daily temperature 2 days ago	0.8791
Maximum daily temperature 1 day ago	0.8867
Minimum daily temperature 1 day ago	0.8873
Mean daily temperature 1 day ago	0.9279

The above analysis demonstrated that parameters, such as the maximum temperature and minimum temperature, had a strong correlation with many other parameters that were features of the training dataset. For these variables, the best prediction results were

obtained using the MLP model. The results of the above analysis were different for the mean daily atmospheric pressure.

In the case of this variable, there was a strong correlation with only one parameter that was a feature from the training dataset. This was the atmospheric pressure from one day ago. The effect of the existence of only one parameter from a set of features with a strong correlation for this variable is visible as the reduced accuracy of the prediction of this parameter. For variables, such as the mean daily wind speed and daily precipitation sum, there was no parameter from the feature set that achieved a strong correlation. For these three variables, the prediction was even less accurate.

5. Discussion

The study described in this report aimed to show the possibility of using a unidirectional multilayer neural network to forecast selected weather indicators and to compare the results achieved with other forecasting models. We assumed that, with a reliable and sufficiently large set of weather data, a properly constructed neural network model, and appropriate software, it would be possible to effectively forecast selected weather indicators.

The results of an application using ANN as presented in this report confirm that Artificial Neural Networks can be useful as a tool for forecasting weather indicators. Although a simple construction of the MLP model—comprising three layers of neurons—was used in the study, the results obtained were satisfactory.

The surprising effectiveness of our implementation of the MLP model with comparable or, in most cases, higher effectiveness compared with the other forecasting models proves that this forecasting model was properly designed and implemented and is a model suitable for the assumed purpose, i.e., forecasting selected weather parameters, and that the data used to train the model were properly prepared.

However, there were some difficulties in applying the proposed weather forecasting model. As reported in the literature [18], neural networks can be susceptible to learning false relationships between data. A pure data-based weather forecasting model may fail to respect basic physical principles and, thus, generate false forecasts because it does not take into consideration that every atmospheric process is affected by physical laws.

There are specific properties of weather data for which classical ML concepts (which work for typical problems solved by ML, such as computer vision and speech recognition) are not effective enough in a complete weather prediction system. The reason for this is the necessity for the model to handle the complexity of the meteorological data and feedback processes to provide accurate prediction results. Another difficulty encountered when using MLPs for weather forecasting is because ANNs are good interpolators but poor extrapolators. The dataset used to train the ANN must, therefore, contain numerous and heterogeneous examples to cover the widest range of cases that the ANN is expected to predict [34].

A separate training and testing dataset containing regional data would be required for each location where the method would be applied. The training dataset needs to be updated regularly due to occurring climate changes, and the network model needs to be re-trained due to the changing climate patterns in the world [85]. Another disadvantage of MLPs and deep learning models over other methods is that their training process takes a long time [62].

Forecasting time series, which include a prediction of the weather parameters changing over time, is an important area of machine learning. The time component provides useful information that is used in the construction of the machine learning model, but it also brings with it problems that make it difficult to accurately predict certain variables. If a time series' data are correlated over time, it is much easier to obtain an accurate prediction because the model uses historical values in the machine learning process and then generates a forecast for the future from these.

When data values change randomly over time, the model cannot predict future changes based on historical events with great accuracy [82]. The implications of this are highly accurate prediction results for weather conditions that show a high correlation with the other variables included in the prediction model and poor results for weather conditions for which no such correlations exist. In most of the analyzed cases, the MLP model achieved higher or comparable results to the forecasts from the internet weather service.

The highest prediction efficiency was observed for the maximum and minimum temperature. These are conditions for which there is a strong correlation in time with many other weather indicators. Thus, there is less risk of them adopting random values. Therefore, these are the parameters that are particularly suitable for prediction by artificial neural networks.

Our conclusion is that it is worthwhile to build models of machine learning for the prediction of time series with values that are strongly correlated with each other because then there is a high probability of obtaining an accurate prediction. These models will be useful, for example, as a correcting tool to forecast the weather conditions at the local scale when numerical modeling is insufficient due to specific local conditions—for instance, as described earlier in the paper, where numerical modeling is performed at a resolution too high to account for the influence of the local topography of the Cadarache valley (localized in southeastern France) [34].

In contrast, the overestimation of the predictive capacity of the model for certain parameters results in its low competitiveness compared to the possibilities offered by alternative methods [86]. Our investigation confirmed the facts available in the literature review, in which most of the research concerned the use of machine learning methods to forecast temperatures for which the predictions are accurate because of their strong correlation with other weather indicators.

6. Conclusions and Future Research Directions

This study presents the use of an Artificial Neural Network with a Multilayer Perceptron (MLP) architecture regressive model to forecast selected weather parameters. The MLP model was built, and its effectiveness was compared with the forecasts available in the website archive. The developed model is a lightweight solution and can be an element of any application that would require forecasting of the above weather parameters. The most important aims achieved by us are listed below:

- This study presents a successful attempt to use an application based on a model of a unidirectional multilayer Artificial Neural Network to forecast selected weather conditions for a selected location, i.e., the city of Szczecin.
- The application used in this survey was successful for a local scale study.
- We obtained satisfactory results, i.e., accurate forecasts, with the simple design of the MLP model, which comprised three layers of neurons.
- We confirmed that this forecasting model was properly designed and implemented, that this is a model suitable for forecasting selected weather parameter, and that the data used to train the model were properly selected and prepared.
- We analyzed and explained the reasons for the different forecasting accuracy results for the different weather parameters.

Our approach in this study was limited to using data and performing forecasts of the selected weather conditions for the local area of Szczecin city in Poland. The weather data set limited to the territory of the city of Szczecin was a small set. As our study was focused on a regional territory and dataset, the usefulness for a larger area would be limited. The achieved results encourage future investigations to generalize whether the satisfactory scores obtained by this application working on local data will be repeatable and useful in the case of other regions of Poland.

The next direction of further work is to compare the performance of this application for a larger dataset including Poland and the European area. It would also be desirable for the direction of future work to include comparing the effectiveness of our method with applications working on other, newer, and more advanced machine learning models to see if the results obtained with our proposed method working in a weather forecasting application can be improved.

The MLP provided slightly more accurate predictions of temperature and atmospheric pressure compared with the LSTM and SVR. This indicates that an appropriate choice of data set, learning features, and parameters for the MLP training can provide results comparable to more complex ML models. For the wind speed prediction, LSTM and SVR showed a slight advantage, and for the precipitation sum prediction, LSTM. Thus, we concluded that, for weather conditions that are more difficult to predict, more complex ML models, especially LSTM, represent an opportunity to obtain more accurate predictions, and it is worthwhile to focus on research aimed at adjusting their structure to further improve the prediction performance.

As the MLP model did not achieve equally effective results for all the predicted weather conditions, further work is needed to improve the accuracy and expand the model. To improve the accuracy of the prediction of the weather parameters with the MLP model, more hidden layers in the network model can be used, i.e., a deep network model can be used because the quality of the multilayer network is higher compared to traditional simple MLP networks with only a few hidden layers. For more accurate forecasts, the precision of the information provided by the training data can be increased by using more densely spaced samples in the training dataset, e.g., every hour instead of every 24 h.

To obtain more accurate forecasts, more training samples can be used for the day that the weather parameters are forecast, e.g., as a training sample of data from ten days ago. Further work includes testing the effectiveness in the problem of weather parameter prediction for more advanced models, such as deep networks, CNN models, LSTM models, RNN models, and SVR models, as well as comparing the accuracy of the results obtained by these with the results of the implemented MLP. Such a study, in the opinion of the authors, could provide interesting results and enable empirical evaluation of the effectiveness of the MLP model from this work.

Author Contributions: Conceptualization, A.B., J.W., J.K. and W.S.; methodology, A.B., J.W., J.K. and W.S.; software, A.B.; validation, A.B., J.K. and W.S.; formal analysis, A.B. and J.W.; investigation, A.B., J.W., J.K. and W.S.; resources, A.B. and J.K.; data curation, A.B., J.W. and J.K.; writing—original draft preparation, A.B.; writing—review and editing, A.B.; visualization, A.B.; supervision, J.W., J.K. and W.S.; project administration, J.W.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the project financed within the framework of the program of the Minister of Science and Higher Education under the name “Regional Excellence Initiative” in the years 2019–2022, Project Number 001/RID/2018/19; the amount of financing: PLN 10.684.000,00 (A.B., J.W.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the editor and the anonymous reviewers, whose insightful comments and constructive suggestions helped us to significantly improve the quality of this paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Networks
CN	Convolutional Networks
CNN	Convolutional Neural Networks
CRBM	Conditional Restricted Boltzmann Machine
ENN	Envolvred Neural Networks
Eumetsat	European Organisation for the Exploitation of Meteorological Satellites
GRNN	Generalized Regression Networks
IMGW	Institute of Meteorology and Water Management
IMGW-PIB	Institute of Meteorology and Water Management–National Research Institute
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MSE	Mean Squared Error
MLP	Multilayer Percptron
NWP	Numerical Weather Prediction
RBF	Radial Basis Function Kernel
RNN	Recurrence Neural Networks
SVM	Support Vector Machine
SVR	Support Vector Regression
WMO	World Meteorological Organisation

Appendix A. Fundamentals of Artificial Neural Networks and the Multilayer Perceptron

The concept of a neural network is defined by mathematical structures that perform calculations or signal processing by elements called neurons. Neurons are arranged in layers. Signals from each neuron of the preceding layer come to each neuron of the following layer. For predictive purposes, multilayer networks are used, containing a minimum of three layers: the input layer, the hidden layer (this number may be greater), and the output layer. The number of neurons present in the input layer corresponds to the number of input features.

Input layer elements do not perform data processing. Their role is to receive signals, i.e., input variables for the network, and then to separate them and pass them on without modifying the neurons of subsequent layers. Hidden layers, i.e., between the input and output layers, may have a different number of neurons. Their number in practice is empirically selected and is an important element influencing the quality of the results. The task of the hidden layer is essentially information processing. The task of the output layer is to process data and generate final output signals, i.e., the network response (predicted values).

A Multilayer Perceptron is a unidirectional network that contains one or more hidden layers. In this type of network, the data stream flows in one direction—from the input layer, through the subsequent hidden layers, to the output layer. In the hidden layers and the output layer, a neural model with an aggregation of signals using the scalar product and usually a non-linear activation function f (sigmoid or tangensoid functions are typically used) as using only a non-linear neural model (with a non-linear activation function) guarantees the possibility to use a Multilayer Perceptron in modeling nonlinear phenomena.

Learning takes place in a supervised mode, and the most common is backpropagation or another algorithm. Each input is associated with a parameter called the weight, which is subject to modification during the training phase of the network, while, during its operation, it is constant. If, as x is denoted by the n -elementary vector containing the input signals from the training sequence given to the network input, and, as w is an n -elementary vector with corresponding weighting values, this processing of input signals in a mathematical neuron model is described by a general rule defined by the Formula (A1):

$$y = f(g(x, w)), \quad (\text{A1})$$

where f means *activate*, while g means *aggregate*. The output signal, which takes the value of the activation function, is separated and passed on to subsequent neurons or is the final network response. An aggregation function g in neuron models is most often used for the scalar product of x and w vectors according to the Formula (A2) [87]:

$$x = x \circ w = \sum_{i=1}^n x_i w_i, \text{ i.e. } s_k = v_{k0} + \sum_{k=1}^n v_{ki} \cdot x_i. \quad (\text{A2})$$

This means that the sum of the values of the products of the input signals and the corresponding weights is calculated. Then, the resulting value is subjected to the activation function, and the result is the output value of the neuron y according to the Formula (A3) [88]:

$$y = w_0 + \sum_{k=1}^K w_k \cdot \phi(s_k). \quad (\text{A3})$$

When aggregating inputs using a scalar product, the following functions are most commonly used as activation functions for f : sigmoid (logistic), which is also called an s-shape function because of the characteristic shape of the graph represented by the Formula (A4) and tangensoid (hyperbolic tangent) as defined by the Formula (A5) [74,75,87]:

$$\phi(s_k) = \frac{1}{1 + e^{-s_k}}, \quad (\text{A4})$$

$$\phi(s_k) = \frac{e^{s_k} - e^{-s_k}}{e^{s_k} + e^{-s_k}}, \quad (\text{A5})$$

where s_k is the total stimulation, i.e., a linear combination of the weights and features of a function that can be calculated using the Formula (A6):

$$z = w^T x = w_0 + w_1 x_1 + \dots + w_m x_m, \quad (\text{A6})$$

where, by k , we mean any neuron with an index in the range of $1 \dots K$, and K means the number of hidden neurons. In this way, the network's response to the signal (pattern) given to its input is known. For each output, we can specify the function F that generates the network output signals. The set of parameters of the F function is a set of all network scales (W). For a specific output of j , there is a dependency described by the Formula (A7):

$$Y_j = F_j(X, W), \quad (\text{A7})$$

where X is a vector for the network input signals. The set of W encodes the knowledge about the modeled phenomenon that the network obtains during the learning process [87]. The expected output value is known because it is in the training sequence. It is possible to change the weight in the network in such a way that the value obtained at the output is close to the standard value. Then the error at the neuron output is calculated, which is the difference between the value at its output and the exact value. In this way, the error for the last layer is defined. For the hidden layers, to define the error, we use an algorithm named the backpropagation algorithm [88]. The backpropagation algorithm includes four stages:

1. Initialization of weights with low random values.
2. Forward propagation (feedforward)—at this stage, each neuron receives and sends a signal to the hidden neurons. Each latent neuron calculates the value of the activation function and sends a signal to the output unit, which calculates the output signal.
3. Backpropagation—after comparing the output values calculated by the network with the exact value, the error is calculated and sent back to all units.
4. Weight adjustment (weight correction) [38].

Modern second-order algorithms, such as the conjugated gradient method and Levenberg-Marquardt method, are faster, which could lead to more frequent use. However,

the classic backpropagation method has many important advantages, first of all, it is the simplest algorithm to understand for most users of neural networks, which makes it the most popular with neural network users [74,75]. The error value depends on the weights of the network W , and thus appropriate modification of the weights during the learning process leads to its reduction by striving to minimize the cumulative error for all elements of the training sequence. This error is determined by the Formula (A8):

$$E(W) = \sum_{i=1}^n E_i, \quad (\text{A8})$$

where n is the number of patterns in a training sequence. The minimization process is carried out using iterative methods. This begins with a randomly selected point in the W_0 weighting space and consists of determining the next points W_1, W_2, \dots so that $E(W_0) > E(W_1) > E(W_2) > \dots$ is seeking to find a global minimum of the $E(W)$ error function spanning over the multidimensional space of the W weights. Before running the backpropagation algorithm, the weights are selected randomly, assuming an even distribution within a certain numerical range.

Correction of the network weights with the backpropagation algorithm determined by the Formulas (A9) and (A10):

$$v_{ki} = v_{ki} - \eta \cdot (y - y^*) \cdot w_k \cdot \phi(s_k) \cdot (1 - \phi(s_k)) \cdot x_i, \quad (\text{A9})$$

$$w_k = w_k - \eta \cdot (y - y^*) \cdot \phi(s_k), \quad (\text{A10})$$

where y means a neural network response and y^* is the expected value for the input. Performing subsequent iterations that change the network scales will minimize the global error. The single processing of all patterns in a training sequence is referred to as epoch. To effectively train the network, it is necessary to carry out many epochs (often in the number of several thousand) [87].

In the backpropagation algorithm, the output error is propagated (backpropagated) from the back (from the output to the input layer) according to the connections of the neurons between layers and considering their activation functions. The modification of the weights that is carried out each time a vector with a sequence of training values is given to the input is called an incremental update of weights. The learning rate η is an element that significantly affects the convergence of the error backpropagation algorithm. There is no general method of determining its value; it is chosen empirically and depends on the type of problem to be solved [88].

Appendix B. Methods to Evaluate the Effectiveness of Regressive ML Models and to Measure the Correlations between Variables

Appendix B.1. Mean Absolute Error (MAE)

This is a commonly used measure of the error made by the network. The MAE is calculated by the Formula (A11):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (\text{A11})$$

where y_i is the observed value, \hat{y}_i is the network response (predicted value), and N is the total number of samples [15].

Appendix B.2. Mean Squared Error (MSE)

This is a usefully quantitative measurement of the effectiveness of the model. It is an average value of the cost function SSE (Sum of Squared Errors), which is reduced during training of the model. The formula used to calculate it is (A12):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2. \quad (A12)$$

If the value of the MSE calculated for the test samples is much higher than for the training samples, it means that the model has been overtrained.

Appendix B.3. R² Score

This can be defined as a standardized version of the MSE method. This allows a better interpretation of the model's performance. It is often treated as a measure of the quality of the regression model. Its value shows to what extent the predictors, i.e., the explanatory variables that have been introduced into the model, allow for the prediction of the values of variables in the test dataset. The R² score is calculated according to the Formula (A13) [74,75]:

$$R^2 = 1 - \frac{SSE}{SST} \quad (A13)$$

where SSE means the sum of the squares of the errors calculated from the Formula (A14):

$$SSE = \frac{1}{2} \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2. \quad (A14)$$

SST means the total sum of squares calculated using the Formula (A15):

$$SST = \sum_{i=1}^n (y^{(i)} - \mu_y)^2, \quad (A15)$$

that is a variance of the answer.

For a set of training data, the R² score takes values between 0 and 1, but with test samples, it can be negative. A value of R² equal to 1 shows a model ideally fitted to the data with a simultaneous MSE value of 0 [74,75].

Appendix B.4. Pearson Correlation Coefficient

The r-Pearson correlation coefficient is a measure of the strength of the linear correlation between the variables, which takes values in the range from −1 to 1. Correlation values in the range from 0 to 1 show an increasingly strong positive correlation. This means that the data are positively correlated as their values increase simultaneously. Correlation values between 0 and −1 show a negative correlation or the opposite [89]. The values of the r-Pearson correlation coefficient, which oscillate around zero, suggest a weak linear relationship [90]. The r-Pearson correlation coefficient is calculated using the Formula (A16) [91].

$$r(x, y) = \frac{cov(x, y)}{\sigma_x \cdot \sigma_y}, \quad (A16)$$

where

$$cov(x, y) = E(x \cdot y) - (E(x) \cdot E(y)), \quad (A17)$$

where

$r(x, y)$ —the r-Pearson correlation coefficient between the variables x and y ,
 $cov(x, y)$ —the covariance between the variables x and y ,

σ —the standard deviation, and
 E —the expected value.

Table A1 contains ranges for the absolute values of the r-Pearson correlation coefficient and their interpretations.

Table A1. The absolute values of the Pearson’s correlation coefficient and their significance.

Absolute Value of the Correlation	Correlation Strength
0.8–1.0	very strong
0.6–0.8	strong
0.4–0.6	moderate
0.2–0.4	weak
0.0–0.2	very weak

Appendix C. Figures

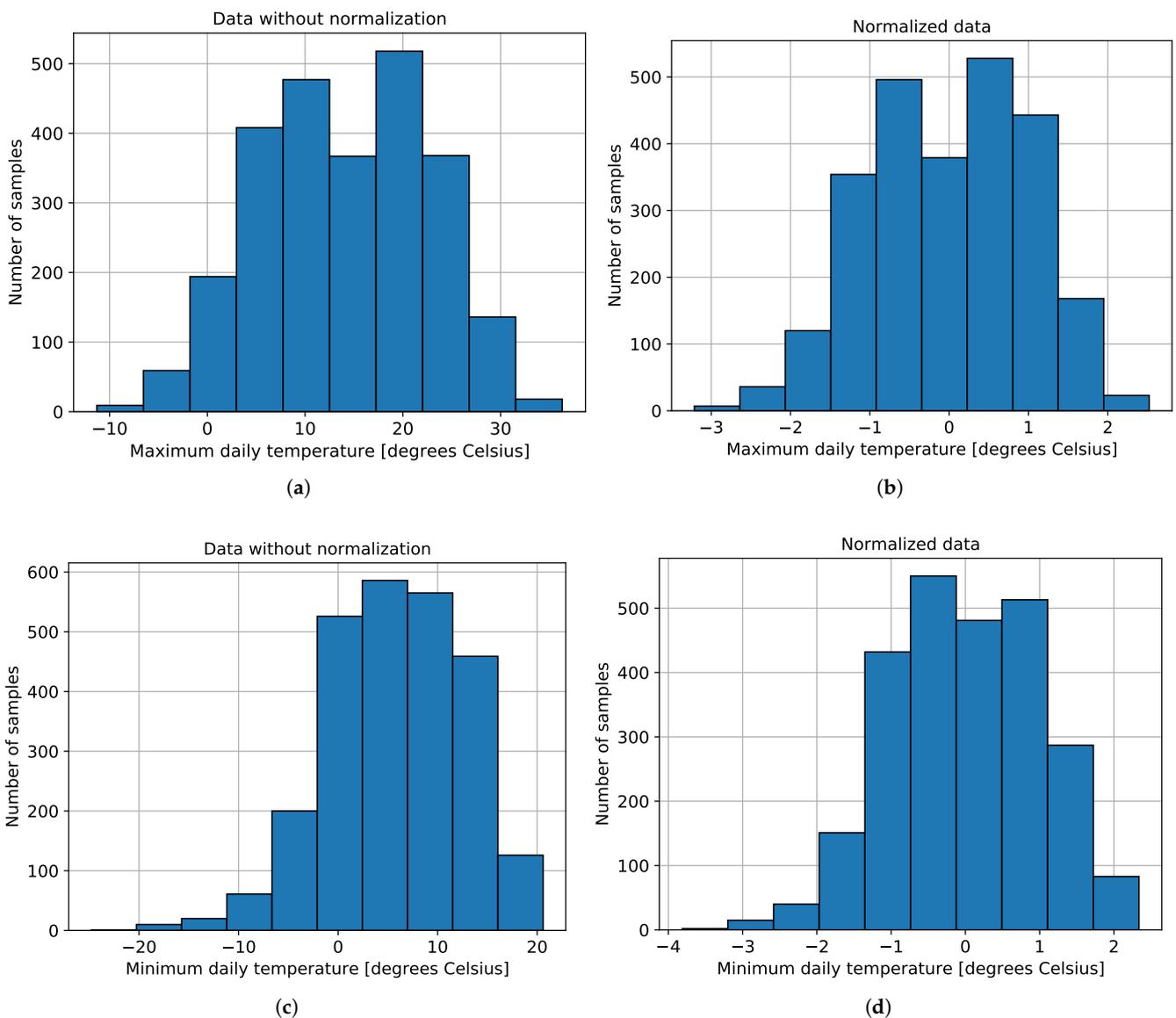


Figure A1. Input data histograms (the maximum daily temperature before (a) and after normalization (b), and minimum daily temperature) before (c) and after normalization (d).

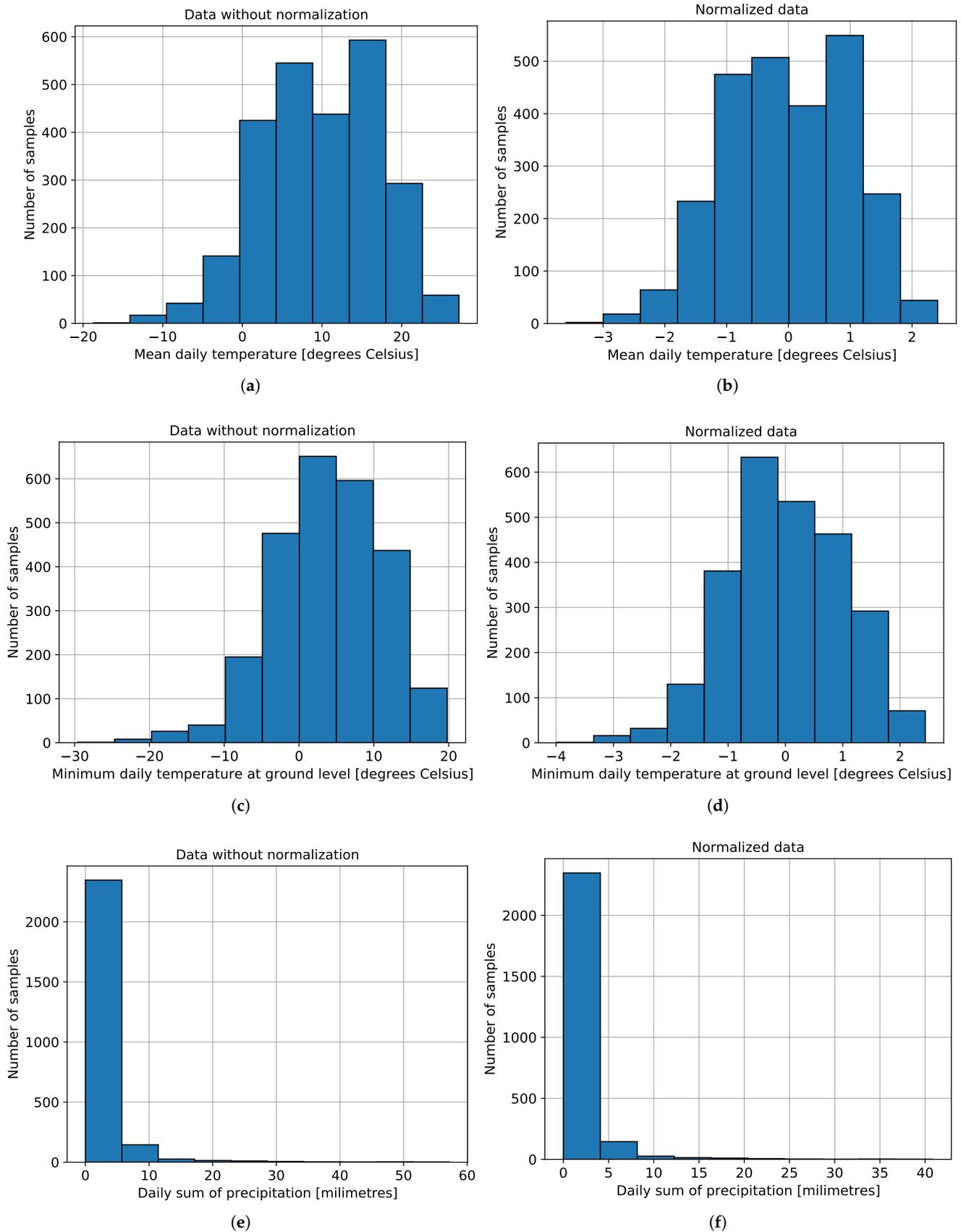


Figure A2. Input data histograms (the mean daily temperature before (a) and after normalization (b), minimum daily temperature at ground level before (c) and after normalization (d), and daily sum of precipitation before (e) and after normalization (f)).

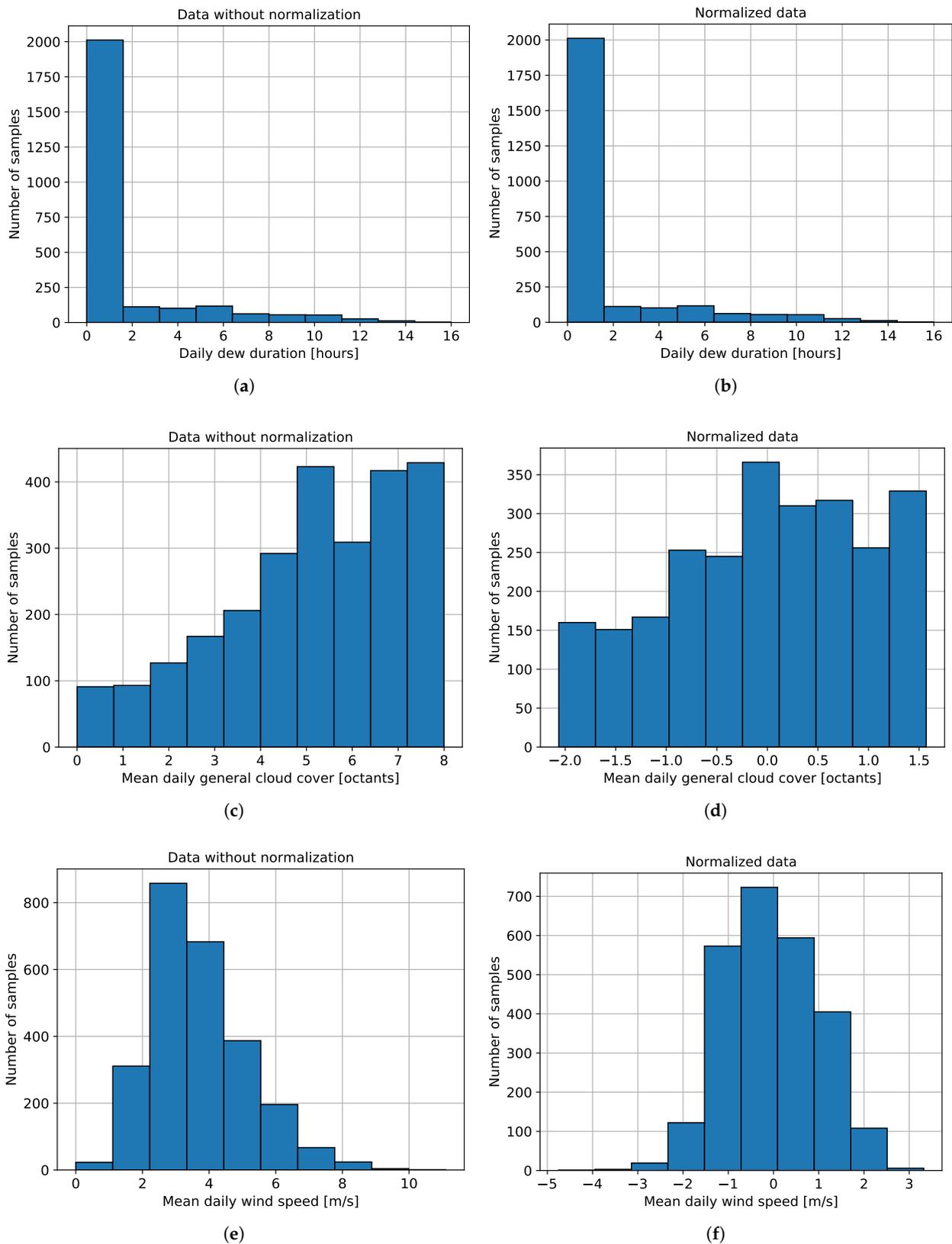


Figure A3. Input data histograms (daily dew duration before (a) and after normalization (b), mean daily general cloud cover before (c) and after normalization (d), and mean daily wind speed before (e) and after normalization (f)).

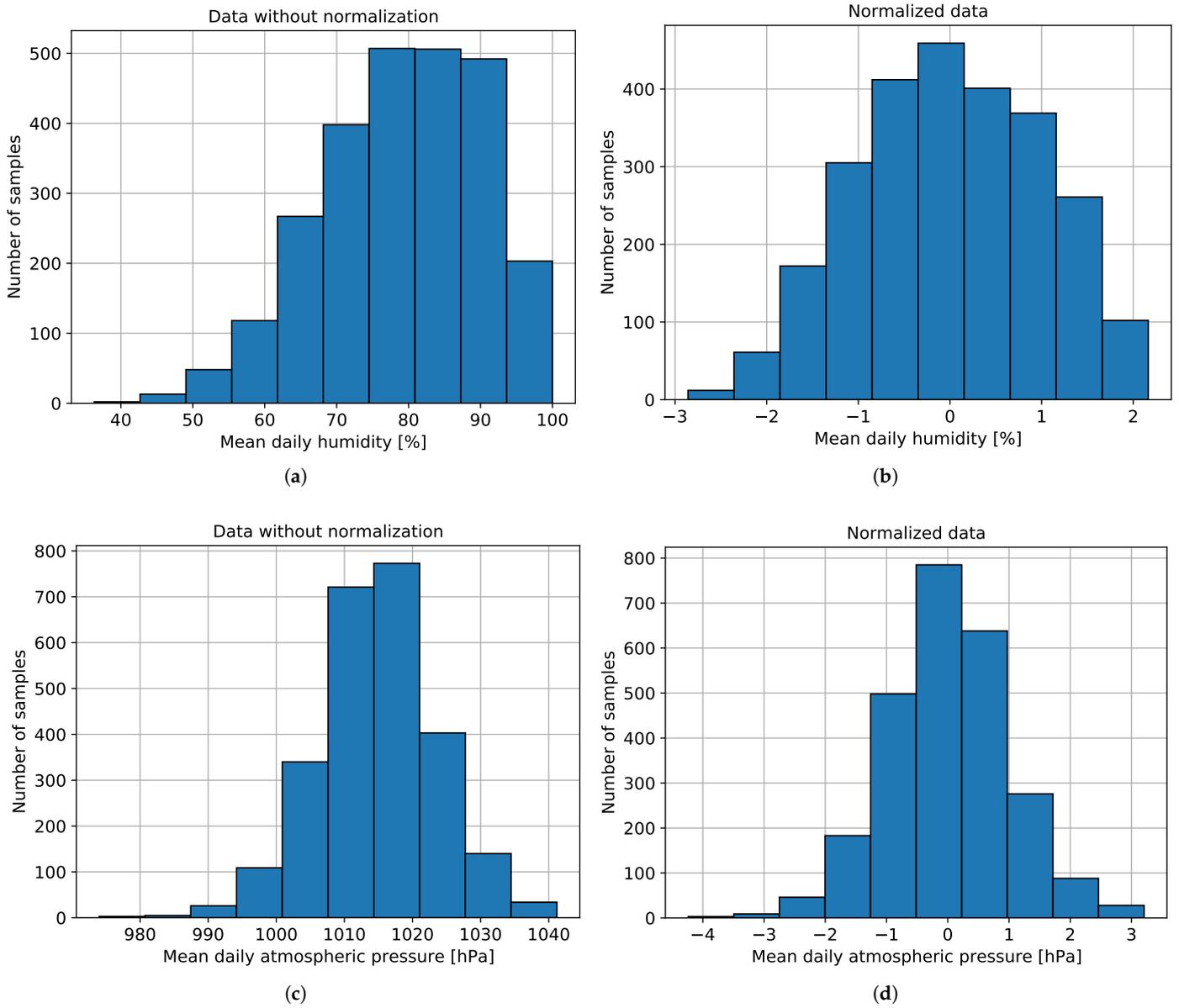


Figure A4. Input data histograms (the mean daily relative humidity before (a) and after normalization (b), and mean daily atmospheric pressure before (c) and after normalization (d)).

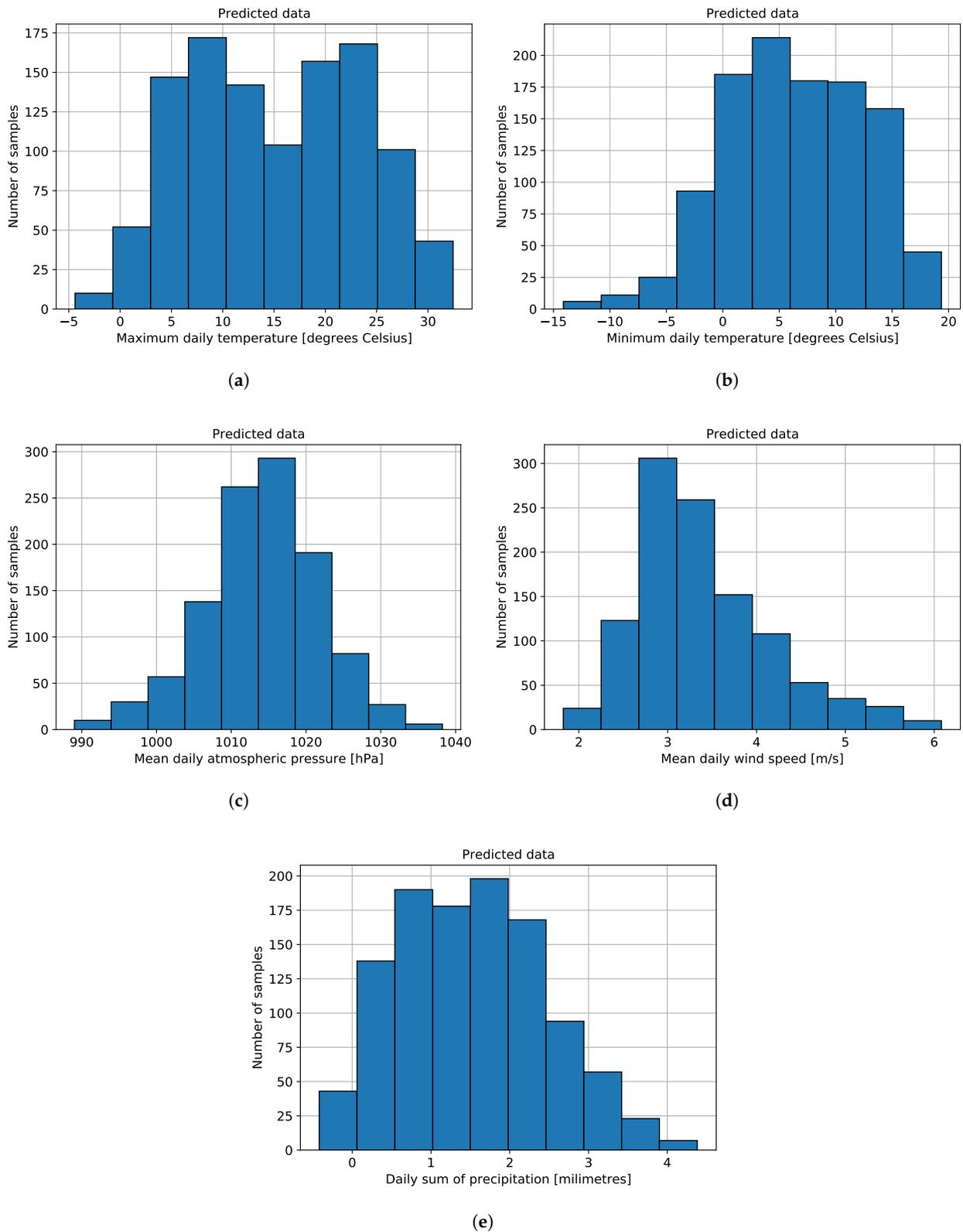


Figure A5. Output data histograms (the maximum (a) and minimum (b) daily temperature, mean daily atmospheric pressure (c), mean daily wind speed (d), and daily sum of precipitation (e)).

Appendix D. Sensitivity Analysis for Each Input Parameter Of Model

The results of the sensitivity analysis for the maximum temperature prediction with the values of MSE ranked in descending order for the training and testing datasets are

listed in Table A2. The results for the other predicted weather conditions are included in Tables A3–A6. In the case of the prediction of the maximum temperature, the training and testing MLP model without mean temperature 1 day ago, training without the sum of precipitation 1 day ago, general cloud cover 1 day ago, mean atmospheric pressure 1 day ago, and testing without mean atmospheric pressure 2 days ago demonstrated a remarkable increase in the error values.

In the event of the prediction of the minimum temperature, the training and testing MLP without minimum temperature 1 day ago, the sum of precipitation 1 day ago, general cloud cover 1 day ago, relative humidity 1 day ago, and testing without mean temperature 1 day ago caused a rise in the error values.

When the mean atmospheric pressure was predicted, the training and testing MLP without mean atmospheric pressure 2 days ago, training without dew duration 3 days ago, mean temperature 1 day ago, minimum temperature 2 days ago, testing without mean atmospheric pressure 1 day ago, minimum temperature 1 day ago, and mean temperature 3 days ago resulted in the growth of the error values.

For the wind speed prediction, the training and testing MLP without wind speed 1 day ago, training without the sum of precipitation 2 days ago, general cloud cover 2 days ago, dew duration 1 day ago, testing without wind speed 3 days ago, and mean temperature 2 days ago demonstrated a remarkable increase in the error values.

In the case of the prediction of the daily sum of precipitation, the training MLP without mean atmospheric pressure 1 day ago, dew duration 2 days ago and 1 day ago, the sum of precipitation 1 day ago, testing without mean temperature 1 day ago, and maximum temperature 1 day ago caused a rise in the error values.

Table A2. The results of the sensitivity analysis for the prediction of the maximum daily temperature.

Input Out	Training MSE	Testing MSE
Mean temp. 1 day ago	6.1293	7.1936
Sum of precip 1 day ago	6.1096	6.6900
Cloud cover 1 day ago	5.6495	6.4668
Mean pressure 1 day ago	5.6224	6.6076
Mean pressure 3 days ago	5.5493	6.6871
Min. temp. 3 days ago	5.4739	6.4487
Min. temp. 1 day ago	5.4714	6.5110
Wind speed 2 days ago	5.4314	6.4600
Max. temp. 2 days ago	5.3495	6.4757
Dew duration 2 days ago	5.3297	6.4460
Cloud cover 3 days ago	5.3293	6.4904
Wind speed 1 day ago	5.3034	6.5390
Relative humidity 1 day ago	5.2536	6.5292
Max. temp. 3 days ago	5.2047	6.4087
Mean temp. 3 days ago	5.1212	6.4370
Mean temp. 2 days ago	5.1123	6.6664
Wind speed 3 days ago	5.1070	6.4721
Min. temp. 2 days ago	5.1020	6.7268
Cloud cover 2 days ago	5.0996	6.6013
Max. temp. 1 day ago	5.0857	6.5947
Dew duration 1 day ago	5.0502	6.5090
Relative humidity 2 days ago	5.0488	6.5872
Min. temp. at ground level 3 days ago	5.0239	6.5584
Min. temp. at ground level 1 day ago	4.9977	6.7814
Min. temp. at ground level 2 days ago	4.8545	6.4719

Table A2. *Cont.*

Input Out	Training MSE	Testing MSE
Dew duration 3 days ago	4.8161	6.6515
Sum of precip 3 days ago	4.7716	6.7705
Relative humidity 3 days ago	4.6714	6.8509
Sum of precip 2 days ago	4.6357	6.9436
Mean pressure 2 days ago	4.2168	7.0869

Table A3. The results of the sensitivity analysis for the prediction of the minimum daily temperature.

Input Out	Training MSE	Testing MSE
Min. temp. 1 day ago	4.5443	5.2170
Sum of precip 1 day ago	4.4590	4.8864
Cloud cover 1 day ago	4.4176	4.9784
Relative humidity 1 day ago	4.3869	4.8971
Max. temp. 2 days ago	4.3381	4.8620
Min. temp. at ground level 1 day ago	4.3112	4.7980
Min. temp. at ground level 2 days ago	4.2294	4.7069
Mean pressure 2 days ago	4.2225	4.6790
Sum of precip 2 days ago	4.1868	4.7069
Dew duration 2 days ago	4.1481	4.6961
Dew duration 1 day ago	4.1397	4.8220
Mean temp. 3 days ago	4.1362	4.7693
Sum of precip 3 days ago	4.1100	4.8355
Min. temp. 3 days ago	4.1053	4.7684
Mean pressure 3 days ago	4.0678	4.6875
Relative humidity 3 days ago	4.0575	4.7483
Min. temp. at ground level 3 days ago	4.0523	4.8223
Cloud cover 3 days ago	4.0382	4.8339
Mean pressure 1 day ago	4.0355	4.8321
Wind speed 2 days ago	4.0264	4.7658
Relative humidity 2 days ago	3.9853	4.6970
Wind speed 3 days ago	3.9212	4.7005
Cloud cover 2 days ago	3.8736	4.7066
Wind speed 1 day ago	3.8730	4.8630
Max. temp. 1 day ago	3.8717	4.9451
Dew duration 3 days ago	3.8021	4.8410
Min. temp. 2 days ago	3.7911	4.8276
Max. temp. 3 days ago	3.7708	4.9188
Mean temp. 2 days ago	3.7592	4.8722
Mean temp. 1 day ago	3.4077	7.1279

Table A4. The results of the sensitivity analysis for the prediction of the mean daily atmospheric pressure.

Input Out	Training MSE	Testing MSE
Mean pressure 2 days ago	25.1855	31.1740
Dew duration 3 days ago	24.2178	29.4694
Mean temp. 1 day ago	24.0707	29.2241
Min. temp. 2 days ago	23.9608	29.3834
Sum of precip 1 day ago	23.9463	29.3296
Sum of precip 3 days ago	23.9220	28.8652
Mean pressure 1 day ago	23.9204	62.6607

Table A4. Cont.

Input Out	Training MSE	Testing MSE
Wind speed 1 day ago	23.8595	29.1812
Cloud cover 2 days ago	23.8532	29.0658
Cloud cover 3 days ago	23.8343	29.1343
Mean temp. 2 days ago	23.8293	29.2587
Max. temp. 3 days ago	23.8255	29.2112
Cloud cover 1 day ago	23.8225	29.3703
Wind speed 3 days ago	23.8204	28.9979
Relative humidity 1 day ago	23.8117	28.8358
Sum of precip 2 days ago	23.7623	29.2192
Min. temp. at ground level 2 days ago	23.7609	29.1715
Min. temp. at ground level 1 day ago	23.7330	29.2121
Relative humidity 3 days ago	23.7292	28.8488
Dew duration 2 days ago	23.7173	29.0915
Wind speed 2 days ago	23.7018	29.2215
Mean pressure 3 days ago	23.6990	29.2511
Min. temp. 3 days ago	23.6700	29.0244
Max. temp. 1 day ago	23.6338	28.9743
Max. temp. 2 days ago	23.5350	29.1366
Relative humidity 2 days ago	23.3966	29.0603
Min. temp. at ground level 3 days ago	22.8807	29.0599
Dew duration 1 day ago	21.0242	29.6448
Min. temp. 1 day ago	14.0007	36.1986
Mean temp. 3 days ago	13.6569	34.8080

Table A5. The results of the sensitivity analysis for the prediction of the mean daily wind speed.

Input Out	Training MSE	Testing MSE
Wind speed 1 day ago	1.0802	1.9793
Sum of precip 2 days ago	0.9902	1.5116
Cloud cover 2 days ago	0.9874	1.5600
Dew duration 1 day ago	0.9822	1.5489
Sum of precip 1 day ago	0.9791	1.5722
Wind speed 2 days ago	0.9625	1.5947
Sum of precip 3 days ago	0.9616	1.5777
Relative humidity 1 day ago	0.9574	1.6152
Mean pressure 1 day ago	0.9533	1.5846
Cloud cover 3 days ago	0.9492	1.6550
Dew duration 3 days ago	0.9476	1.5351
Mean temp. 2 days ago	0.9460	1.6596
Cloud cover 1 day ago	0.9439	1.6620
Relative humidity 3 days ago	0.9388	1.6069
Min. temp. 2 days ago	0.9374	1.6696
Relative humidity 2 days ago	0.9373	1.6385
Wind speed 3 days ago	0.9326	1.6940
Mean pressure 2 days ago	0.9315	1.5701
Max. temp. 1 day ago	0.9245	1.6162
Min. temp. 1 day ago	0.9219	1.6535
Max. temp. 3 days ago	0.9207	1.5513
Dew duration 2 days ago	0.9201	1.6023

Table A5. Cont.

Input Out	Training MSE	Testing MSE
Mean pressure 3 days ago	0.9169	1.6439
Min. temp. at ground level 3 days ago	0.9169	1.6503
Min. temp. at ground level 1 day ago	0.9160	1.6323
Mean temp. 3 days ago	0.9155	1.6772
Mean temp. 1 day ago	0.9113	1.6654
Min. temp. 3 days ago	0.9094	1.6361
Min. temp. at ground level 2 days ago	0.9080	1.5682
Max. temp. 2 days ago	0.8978	1.6278

Table A6. The results of the sensitivity analysis for the prediction of the daily sum of precipitation.

Input Out	Training MSE	Testing MSE
Mean pressure 1 day ago	10.5840	13.7433
Dew duration 2 days ago	10.5257	13.6365
Dew duration 1 day ago	10.4687	14.3820
Sum of precip 1 day ago	10.4163	14.2739
Dew duration 3 days ago	10.3711	14.3944
Wind speed 1 day ago	10.2544	13.9901
Mean pressure 3 days ago	10.2285	13.8151
Relative humidity 3 days ago	10.1896	13.7449
Cloud cover 1 day ago	10.1345	13.8556
Relative humidity 1 day ago	10.1337	14.6339
Wind speed 3 days ago	10.0379	14.0473
Sum of precip 3 days ago	9.9644	14.5009
Min. temp. at ground level 1 day ago	9.9450	13.9429
Mean temp. 2 days ago	9.9069	14.4427
Sum of precip 2 days ago	9.8581	14.4874
Min. temp. 1 day ago	9.8550	14.9830
Mean temp. 1 day ago	9.8490	15.2247
Min. temp. at ground level 3 days ago	9.7950	14.5681
Max. temp. 2 days ago	9.7279	13.9829
Cloud cover 3 days ago	9.7204	14.1295
Mean temp. 3 days ago	9.7064	14.2510
Mean pressure 2 days ago	9.6979	14.2494
Relative humidity 2 days ago	9.6714	14.5204
Min. temp. 3 days ago	9.6577	14.2914
Cloud cover 2 days ago	9.6425	14.1778
Min. temp. at ground level 2 days ago	9.6336	14.3124
Min. temp. 2 days ago	9.5946	14.4167
Wind speed 2 days ago	9.5171	14.7333
Max. temp. 3 days ago	9.2872	14.6339
Max. temp. 1 day ago	9.2770	15.1178

Appendix E. Tables

Table A7. The mean and standard deviation values of the MLP evaluation obtained with cross-validation for the predicted weather conditions.

	Max. Temp.		Min. Temp.		Pressure		Wind		Precip	
Hidden Neurons	Means	Stds	Means	Stds	Means	Stds	Means	Stds	Means	Stds
10	0.9088	0.0141	0.8732	0.0124	0.6652	0.0330	0.2717	0.0502	0.1093	0.0299
20	0.9110	0.0127	0.8799	0.0096	0.6700	0.0319	0.2747	0.0444	0.1086	0.0342
30	0.9125	0.0117	0.8820	0.0107	0.6692	0.0292	0.2720	0.0440	0.1086	0.0317
40	0.9134	0.0115	0.8820	0.0084	0.6692	0.0283	0.2779	0.0475	0.1077	0.0337
50	0.9116	0.0105	0.8827	0.0096	0.6702	0.0322	0.2778	0.0460	0.1105	0.0325
60	0.9121	0.0117	0.8849	0.0108	0.6701	0.0308	0.2734	0.0508	0.1070	0.0334
70	0.9099	0.0136	0.8842	0.0106	0.6713	0.0322	0.2776	0.0501	0.1094	0.0326
80	0.9136	0.0113	0.8850	0.0111	0.6702	0.0304	0.2678	0.0493	0.1085	0.0291
90	0.9112	0.0107	0.8832	0.0093	0.6719	0.0324	0.2725	0.0446	0.1045	0.0310
100	0.9096	0.0126	0.8849	0.0103	0.6673	0.0316	0.2766	0.0519	0.1055	0.0333
Learning rate	Means	Stds	Means	Stds	Means	Stds	Means	Stds	Means	Stds
0.0001	0.8748	0.0220	0.8484	0.0189	0.5732	0.0537	0.2444	0.0375	0.1035	0.0300
0.001	0.9100	0.0118	0.8828	0.0109	0.6704	0.0341	0.2754	0.0482	0.1069	0.0311
0.01	0.8841	0.0158	0.8547	0.0221	0.5203	0.0650	0.1602	0.0756	0.1029	0.2175
Number of Epochs	Means	Stds	Means	Stds	Means	Stds	Means	Stds	Means	Stds
250	0.9113	0.0118	0.8843	0.0099	0.6706	0.0338	0.2704	0.0495	0.1054	0.0298
500	0.9130	0.0104	0.8855	0.0126	0.6703	0.0329	0.2719	0.0503	0.0867	0.0423
750	0.9106	0.0108	0.8869	0.0128	0.6688	0.0361	0.2350	0.0817	0.0471	0.0821
1000	0.9109	0.0124	0.8861	0.0122	0.6700	0.0308	0.2105	0.0919	0.0653	0.0843
L2	Means	Stds	Means	Stds	Means	Stds	Means	Stds	Means	Stds
0.0001	0.9102	0.0124	0.8835	0.0106	0.6693	0.0304	0.2716	0.0532	0.1053	0.0340
0.001	0.9106	0.0121	0.8848	0.0111	0.6710	0.0311	0.2787	0.0505	0.1116	0.0334
0.01	0.9097	0.0105	0.8840	0.0113	0.6728	0.0289	0.2777	0.0527	0.1082	0.0318
Decrease const	Means	Stds	Means	Stds	Means	Stds	Means	Stds	Means	Stds
0.00001	0.9144	0.0109	0.8820	0.0083	0.6722	0.0297	0.2701	0.0453	0.1096	0.0313
0.0001	0.9111	0.0106	0.8867	0.0102	0.6709	0.0325	0.2815	0.0509	0.1032	0.0337
0.001	0.8808	0.0230	0.8515	0.0245	0.6623	0.0350	0.2614	0.0399	0.1074	0.0288
Alpha	Means	Stds	Means	Stds	Means	Stds	Means	Stds	Means	Stds
0.0001	0.9105	0.0120	0.8855	0.0136	0.6699	0.0302	0.2658	0.0497	0.1062	0.0312
0.001	0.9113	0.0104	0.8839	0.0114	0.6716	0.0314	0.2655	0.0460	0.1054	0.0329
0.01	0.9109	0.0103	0.8846	0.0107	0.6699	0.0338	0.2733	0.0451	0.1034	0.0304

Table A8. The results of the selection of the most suitable subset of training variables for the prediction of the daily minimum temperature.

Variable	Training MSE	Testing MSE
Mean daily temperature	6.5819	6.937
Minimum daily temperature	9.7589	9.6867
Maximum daily temperature	10.0932	11.2083
Min. daily temp. at ground level	10.873	10.8093
Mean daily atm. pressure	34.7073	31.1968
Daily dew duration	34.9254	33.0771
Mean daily relative humidity	41.4679	37.7907
Mean daily general cloud cover	43.2027	38.159
Daily sum of precipitation	44.1642	42.2249
Mean daily wind speed	44.4846	38.8985

Table A9. The results of the selection of the most suitable subset of training variables for the prediction of the mean daily atmospheric pressure.

Variable	Training MSE	Testing MSE
Mean daily atm. pressure	25.2469	29.3162
Daily sum of precipitation	65.7232	79.602
Mean daily wind speed	69.265	83.9249
Mean daily general cloud cover	74.6802	83.7544
Min. daily temp. at ground level	74.8499	86.6531
Mean daily temperature	75.2118	86.9998
Minimum daily temperature	75.2322	87.6592
Maximum daily temperature	75.9025	88.0402
Daily dew duration	76.2899	86.9606
Mean daily relative humidity	76.5645	87.7578

Table A10. The results of the selection of the most suitable subset of training variables for the prediction of the mean daily wind speed.

Variable	Training MSE	Testing MSE
Mean daily wind speed	1.5095	1.3262
Mean daily temperature	2.0238	1.5537
Daily sum of precipitation	2.0519	1.6556
Mean daily atm. pressure	2.0599	1.6657
Daily dew duration	2.0703	1.6514
Minimum daily temperature	2.0718	1.6297
Min. daily temp. at ground level	2.0823	1.6414
Maximum daily temperature	2.0855	1.5734
Mean daily general cloud cover	2.0984	1.616
Mean daily relative humidity	2.1058	1.6688

Table A11. The results of the selection of the most suitable subset of training variables for the prediction of the daily sum of precipitation.

Variable	Training MSE	Testing MSE
Daily sum of precipitation	16.2719	10.5237
Mean daily atm. pressure	16.4217	10.8412
Min. daily temp. at ground level	16.4605	11.2146
Minimum daily temperature	16.5323	11.2197
Mean daily temperature	16.5921	11.2773
Maximum daily temperature	16.6668	11.2979
Daily dew duration	16.7754	10.9771
Mean daily general cloud cover	16.8164	10.9645
Mean daily relative humidity	16.8345	11.1375
Mean daily wind speed	16.8671	11.0588

Table A12. The results of the sensitivity analysis for the prediction of the minimum daily temperature.

Inputs Out	Training MSE	Testing MSE
Mean daily temperature	6.7997	7.1771
Minimum daily temperature	5.5283	5.7243
meanhumidity	5.4248	5.6432
Maximum daily temperature	5.3333	5.5946
Min. daily temp. at ground level	5.2886	5.3792
Daily dew duration	5.1688	5.4749
Mean daily atm. pressure	5.1317	5.5430

Table A13. The results of the sensitivity analysis for the prediction of the mean daily atmospheric pressure.

Inputs Out	Training MSE	Testing MSE
Mean daily atm. pressure	58.6317	77.9652
Mean daily temperature	24.3695	29.3873
Mean daily general cloud coverage	24.2859	29.2418
Mean daily wind speed	24.1769	28.8816
Minimum daily temperature	24.1667	29.1359
Daily sum of precipitation	24.1326	29.2514
Min. daily temp. at ground level	23.9672	28.9715

Table A14. The results of the sensitivity analysis for the prediction of the mean daily wind speed.

Inputs Out	Training MSE	Testing MSE
Mean daily wind speed	1.7597	1.5639
Daily sum of precipitation	1.4691	1.2879
Mean daily temperature	1.4595	1.2810
Daily dew duration	1.4467	1.2846
Mean daily atm. pressure	1.4360	1.2863
Minimum daily temperature	1.4160	1.2974
Min. daily temp. at ground level	1.3459	1.3423

Table A15. The results of the sensitivity analysis for the prediction of the daily sum of precipitation.

Inputs Out	Training MSE	Testing MSE
Daily sum of precipitation	15.9476	10.9150
Mean daily atm. pressure	15.8352	10.5831
Daily dew duration	15.8150	10.5845
Min. daily temp. at ground level	15.7052	10.5052
Mean daily temperature	15.6931	10.5876
Minimum daily temperature	15.6718	10.5391
Maximum daily temperature	15.6270	10.5490

Table A16. Examples of the weather service forecast for Szczecin city from 1 January 2018 to 31 December 2020.

Date (Year-Month-Day)	Max. temp. [°C]	Min. temp. [°C]	Pressure [hPa]	Wind [m/s]	Precip [mm]
2018-01-01	5	2	1010.7	6.81	0
2018-01-02	5	3	1016.5	4.03	0
2018-01-03	5	4	994.74	7.57	3
2018-01-04	6	4	1002.31	5.49	3
2018-01-05	6	4	1004.5	7.71	0

Table A17. Comparison of the MLP prediction and weather service for the maximum temperature.

Date (Year-Month-Day)	Observed Value [°C]	MLP Prediction [°C]	Weather Service Forecast [°C]
2018-01-01	11.7	11.3	5
2018-01-02	5.4	9.7	5
2018-01-03	5.9	5.7	5
2018-01-04	6.1	6.2	6
2018-01-05	8.1	5.7	6

Table A18. Comparison of the MLP and weather service forecast for the minimum temperature.

Date (Year-Month-Day)	Observed Value [°C]	MLP Prediction [°C]	Weather Service Forecast [°C]
2018-01-01	4.8	5.5	2
2018-01-02	2.6	4.3	3
2018-01-03	1.2	1.0	4
2018-01-04	4.3	1.7	4
2018-01-05	4.7	2.8	4

Table A19. Comparison of MLP forecast and weather service for the mean daily atmospheric pressure.

Date (Year-Month-Day)	Observed Value [hPa]	MLP Prediction [hPa]	Weather Service Forecast [hPa]
2018-01-01	998	1002.07	1010.7
2018-01-02	1004.1	1000.53	1016.5
2018-01-03	989.3	1008.97	994.74
2018-01-04	989.4	991.62	1002.31
2018-01-05	992.4	999.02	1004.5

Table A20. Comparison of MLP forecast and weather service for the mean daily wind speed.

Date (Year-Month-Day)	Observed Value [m/s]	MLP Prediction [m/s]	Weather Service Forecast [m/s]
2018-01-01	5.5	4.45	6.81
2018-01-02	2.4	4.76	4.03
2018-01-03	6.9	3.22	7.57
2018-01-04	5.4	5.73	5.49
2018-01-05	5.6	5.90	7.71

Table A21. Comparison of the MLP and weather service forecast for the daily precipitation sum.

Date (Year-Month-Day)	Observed Value [mm]	MLP Prediction [mm]	Weather Service Forecast [mm]
2018-01-01	0.9	2.2	0
2018-01-02	4.3	2.5	0
2018-01-03	16.7	2.2	3
2018-01-04	2.6	3.7	3
2018-01-05	0.5	3.6	0

References

1. Amani, F.A.; Fadlalla, A.M. Data mining applications in accounting: A review of the literature and organizing framework. *Int. J. Account. Inf. Syst.* **2017**, *24*, 32–58. [\[CrossRef\]](#)
2. Siregar, S.P.; Wanto, A. Analysis of Artificial Neural Network Accuracy Using Backpropagation Algorithm In Predicting Process (Forecasting). *IJISTECH Int. J. Inf. Syst. Technol.* **2017**, *1*, 34–42. [\[CrossRef\]](#)
3. Waziri, B.S.; Bala, K.; Bustani, S.A. Artificial Neural Networks in construction engineering and management. *Int. J. Archit. Eng. Constr.* **2017**, *6*, 50–60. [\[CrossRef\]](#)
4. Nagapurkar, P.; Smith, J.D. Techno-economic optimization and social costs assessment of microgrid-conventional grid integration using genetic algorithm and Artificial Neural Networks: A case study for two US cities. *J. Clean. Prod.* **2019**, *229*, 552–569. [\[CrossRef\]](#)
5. Hu, H.; Tang, L.; Zhang, S.; Wang, H. Predicting the direction of stock markets using optimized neural networks with Google Trends. *Neurocomputing* **2018**, *285*, 188–195. [\[CrossRef\]](#)
6. Galeshchuk, S. Neural networks performance in exchange rate prediction. *Neurocomputing* **2016**, *172*, 446–452. [\[CrossRef\]](#)
7. Benedetti, M.; Cesarotti, V.; Introna, V.; Serranti, J. Energy consumption control automation using Artificial Neural Networks and adaptive algorithms: Proposal of a new methodology and case study. *Appl. Energy* **2016**, *165*, 60–71. [\[CrossRef\]](#)
8. Guyot, D.; Giraud, F.; Simon, F.; Corgier, D.; Marvillet, C.; Tremeac, B. Overview of the use of Artificial Neural Networks for energy-related applications in the building sector. *Int. J. Energy Res.* **2019**, *43*, 6680–6720. [\[CrossRef\]](#)
9. Runge, J.; Zmeureanu, R. Forecasting energy use in buildings using Artificial Neural Networks: A review. *Energies* **2019**, *12*, 3254. [\[CrossRef\]](#)

10. Katsatos, A.; Moustris, K. Application of Artificial Neuron Networks as energy consumption forecasting tool in the building of Regulatory Authority of Energy, Athens, Greece. *Energy Procedia* **2019**, *157*, 851–861. [CrossRef]
11. Rodríguez, F.; Fleetwood, A.; Galarza, A.; Fontán, L. Predicting solar energy generation through Artificial Neural Networks using weather forecasts for microgrid control. *Renew. Energy* **2018**, *126*, 855–864. [CrossRef]
12. Nguyen, H.; Bui, X.N. Predicting blast-induced air overpressure: A robust artificial intelligence system based on Artificial Neural Networks and random forest. *Nat. Resour. Res.* **2019**, *28*, 893–907. [CrossRef]
13. Annarumma, M.; Withey, S.J.; Bakewell, R.J.; Pesce, E.; Goh, V.; Montana, G. Automated triaging of adult chest radiographs with deep artificial neural networks. *Radiology* **2019**, *291*, 196–202. [CrossRef] [PubMed]
14. Shahid, N.; Rappon, T.; Berta, W. Applications of Artificial Neural Networks in health care organizational decision-making: A scoping review. *PLoS ONE* **2019**, *14*, e0212356. [CrossRef]
15. Hayati, M.; Mohebi, Z. Temperature forecasting based on neural network approach. *World Appl. Sci. J.* **2007**, *2*, 613–620.
16. Hayati, M.; Mohebi, Z. Application of Artificial Neural Networks for temperature forecasting. *World Acad. Sci. Eng. Technol.* **2007**, *28*, 275–279.
17. Scher, S.; Messori, G. Predicting weather forecast uncertainty with machine learning. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 2830–2841. [CrossRef]
18. Schultz, M.; Betancourt, C.; Gong, B.; Kleinert, F.; Langguth, M.; Leufen, L.; Mozaffari, A.; Stadler, S. Can deep learning beat numerical weather prediction? *Philos. Trans. R. Soc. A* **2021**, *379*, 20200097. [CrossRef]
19. Lynch, P. The origins of computer weather prediction and climate modeling. *J. Comput. Phys.* **2008**, *227*, 3431–3444. [CrossRef]
20. Baek, S. A revised radiation package of G-packed McICA and two-stream approximation: Performance evaluation in a global weather forecasting model. *J. Adv. Model. Earth Syst.* **2017**, *9*, 1628–1640. [CrossRef]
21. Biffis, E.; Chavez, E. Satellite data and machine learning for weather risk management and food security. *Risk Anal.* **2017**, *37*, 1508–1521. [CrossRef]
22. Štulec, I. Effectiveness of weather derivatives as a risk management tool in food retail: The case of Croatia. *Int. J. Financ. Stud.* **2017**, *5*, 2. [CrossRef]
23. Xu, L.; Wang, S.; Tang, R. Probabilistic load forecasting for buildings considering weather forecasting uncertainty and uncertain peak load. *Appl. Energy* **2019**, *237*, 180–195. [CrossRef]
24. Van Gerven, M.; Bohte, S. Artificial neural networks as models of neural information processing. *Frontiers* **2017**, *11*, 114. [CrossRef]
25. Schuld, M.; Sinayskiy, I.; Petruccione, F. The quest for a quantum neural network. *Quantum Inf. Process.* **2014**, *13*, 2567–2586. [CrossRef]
26. da Silva, A.J.; Ludermir, T.B.; de Oliveira, W.R. Quantum perceptron over a field and neural network architecture selection in a quantum computer. *Neural Netw.* **2016**, *76*, 55–64. [CrossRef]
27. Hwang, J.; Orenstein, P.; Cohen, J.; Pfeiffer, K.; Mackey, L. Improving subseasonal forecasting in the western US with machine learning. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019, 2019; pp. 2325–2335.
28. Rasp, S. WeatherBench: A Benchmark Dataset for Data-Driven Weather Forecasting. Available online: <https://raspstephan.github.io/blog/weatherbench/#> (accessed on 2 December 2020).
29. Li, Y.; Lang, J.; Ji, L.; Zhong, J.; Wang, Z.; Guo, Y.; He, S. Weather Forecasting Using Ensemble of Spatial-Temporal Attention Network and Multi-Layer Perceptron. *Asia-Pac. J. Atmos. Sci.* **2020**. [CrossRef]
30. Alotaibi, K.; Ghumman, A.R.; Haider, H.; Ghazaw, Y.M.; Shafiquzzaman, M. Future predictions of rainfall and temperature using GCM and ANN for arid regions: A case study for the Qassim Region, Saudi Arabia. *Water* **2018**, *10*, 1260. [CrossRef]
31. Sheik Mohideen Shah, S.; Meganathan S.; Kamali A. Soft Computing Research For Weather Prediction Using Multilayer Architecture. *Int. J. Eng. Adv. Technol. (IJEAT)* **2019**, *8*, 3779–3783.
32. Maldonado-Correa, J.; Valdiviezo-Condolo, M.; Viñan-Ludeña, M.S.; Samaniego-Ojeda, C.; Rojas-Moncayo, M. Wind power forecasting for the Villonaco wind farm. *Wind Eng.* **2020**. [CrossRef]
33. Chen, T.; Kapron, N.; Chen, J.Y. Using Evolving ANN-Based Algorithm Models for Accurate Meteorological Forecasting Applications in Vietnam. *Math. Probl. Eng.* **2020**.
34. Dupuy, F.; Duine, G.J.; Durand, P.; Hedde, T.; Pardyjak, E.; Roubin, P. Valley Winds at the Local Scale: Correcting Routine Weather Forecast Using Artificial Neural Networks. *Atmosphere* **2021**, *12*, 128. [CrossRef]
35. Shamshirband, S.; Esmailbeiki, F.; Zarehaghi, D.; Neyshabouri, M.; Samadianfard, S.; Ghorbani, M.A.; Mosavi, A.; Nabipour, N.; Chau, K.W. Comparative analysis of hybrid models of firefly optimization algorithm with support vector machines and Multilayer Perceptron for predicting soil temperature at different depths. *Eng. Appl. Comput. Fluid Mech.* **2020**, *14*, 939–953. [CrossRef]
36. Esteves, J.T.; de Souza Rolim, G.; Ferraudo, A.S. Rainfall prediction methodology with binary Multilayer Perceptron neural networks. *Clim. Dyn.* **2019**, *52*, 2319–2331. [CrossRef]
37. Velasco, L.C.P.; Serquiña, R.P.; Zamad, M.S.A.A.; Juanico, B.F.; Lomocoso, J.C. Performance Analysis of Multilayer Perceptron Neural Network Models in Week-Ahead Rainfall Forecasting. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 578–588. [CrossRef]
38. Radhika, Y.; Shashi, M. Atmospheric temperature prediction using support vector machines. *Int. J. Comput. Theory Eng.* **2009**, *1*, 55. [CrossRef]

39. Salcedo-Sanz, S.; Deo, R.; Carro-Calvo, L.; Saavedra-Moreno, B. Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms. *Theor. Appl. Climatol.* **2016**, *125*, 13–25. [CrossRef]
40. Baboo, S.S.; Shereef, I.K. An efficient weather forecasting system using artificial neural network. *Int. J. Environ. Sci. Dev.* **2010**, *1*, 321. [CrossRef]
41. Hossain, M.; Rekabdar, B.; Louis, S.J.; Dascalu, S. Forecasting the weather of Nevada: A deep learning approach. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015; pp. 1–6. doi:10.1109/IJCNN.2015.7280812. [CrossRef]
42. Białobrzewski, I. Porównanie algorytmów uczenia sieci neuronowej jednokierunkowej, z czasowym opóźnieniem, wykorzystanej do predykcji wartości temperatury powietrza atmosferycznego. *Inżynieria Rol.* **2005**, *9*, pp. 7–15.
43. Lai, L.L.; Braun, H.; Zhang, Q.; Wu, Q.; Ma, Y.; Sun, W.; Yang, L. Intelligent weather forecast. In Proceedings of the 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 04EX826), Shanghai, China, 26–29 August 2004; Volume 7, pp. 4216–4221.
44. Khotanzad, A.; Davis, M.H.; Abaye, A.; Maratukulam, D.J. An Artificial Neural Network hourly temperature forecaster with applications in load forecasting. *IEEE Trans. Power Syst.* **1996**, *11*, 870–876. [CrossRef]
45. Abdel-Aal, R.E. Short-term hourly load forecasting using abductive networks. *IEEE Trans. Power Syst.* **2004**, *19*, 164–173. [CrossRef]
46. Devi, C.J.; Reddy, B.S.P.; Kumar, K.V.; Reddy, B.M.; Nayak, N.R. ANN approach for weather prediction using back propagation. *Int. J. Eng. Trends Technol.* **2012**, *3*, 19–23.
47. Seyyedabbasi, A.; Candan, F.; Kiani, F. A Method for Forecasting Weather Condition by Using Artificial Neural Network Algorithm. *ICTACT J. Soft Comput.* **2018**, *8*, 1696–1700.
48. Abhishek, K.; Singh, M.; Ghosh, S.; Anand, A. Weather forecasting model using Artificial Neural Network. *Procedia Technol.* **2012**, *4*, 311–318. [CrossRef]
49. Abdel-Aal, R.; Elhadidy, M. Modeling and forecasting the daily maximum temperature using abductive machine learning. *Weather Forecast.* **1995**, *10*, 310–325. [CrossRef]
50. Abdel-Aal, R.; Elhadidy, M. A machine-learning approach to modelling and forecasting the minimum temperature at Dhahran, Saudi Arabia. *Energy* **1994**, *19*, 739–749. [CrossRef]
51. Holmstrom, M.; Liu, D.; Vo, C. Machine Learning Applied to Weather Forecasting. *Meteorol. Appl.* **2016**, 1–5. Available online: <http://cs229.stanford.edu/proj2016/report/HolmstromLiuVo-MachineLearningAppliedToWeatherForecasting-report.pdf> (accessed on 15 December 2016).
52. Heinermann, J.; Kramer, O. Machine learning ensembles for wind power prediction. *Renew. Energy* **2016**, *89*, 671–679. [CrossRef]
53. Jasiński, T. Zastosowanie sztucznych sieci neuronowych w modelowaniu prędkości wiatru jako jednej z determinantów poboru energii w budynkach. *Fiz. Budowli W Teor. I Prakt.* **2018**, *10*, 9–14.
54. Velo, R.; López, P.; Maseda, F. Wind speed estimation using Multilayer Perceptron. *Energy Convers. Manag.* **2014**, *81*, 1–9. [CrossRef]
55. Liu, J.N.; Hu, Y.; He, Y.; Chan, P.W.; Lai, L. Deep neural network modeling for big data weather forecasting. In *Information Granularity, Big Data, and Computational Intelligence*; International Publishing Switzerland: Cham, Switzerland, 2015; pp. 389–408.
56. Grover, A.; Kapoor, A.; Horvitz, E. A deep hybrid model for weather forecasting. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, 10–13 August 2015; pp. 379–386.
57. Salman, A.G.; Kanigoro, B.; Heryadi, Y. Weather forecasting using deep learning techniques. In Proceedings of the 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 10–11 October 2015; pp. 281–285.
58. Ding, M.; Zhou, H.; Xie, H.; Wu, M.; Nakanishi, Y.; Yokoyama, R. A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting. *Neurocomputing* **2019**, *365*, 54–61. [CrossRef]
59. Schmidhuber, J.; Hochreiter, S. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780.
60. Kwon, B.S.; Park, R.J.; Song, K.B. Short-term load forecasting based on deep neural networks using LSTM layer. *J. Electr. Eng. Technol.* **2020**, *15*, 1501–1509. [CrossRef]
61. Kang, J.; Wang, H.; Yuan, F.; Wang, Z.; Huang, J.; Qiu, T. Prediction of Precipitation Based on Recurrent Neural Networks in Jingdezhen, Jiangxi Province, China. *Atmosphere* **2020**, *11*, 246. [CrossRef]
62. Gad, I.; Hosahalli, D. A comparative study of prediction and classification models on NCDC weather data. *Int. J. Comput. Appl.* **2020**. [CrossRef]
63. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall: Upper Saddle River, NJ, USA, 1999.
64. Lee, S.; Lee, Y.S.; Son, Y. Forecasting daily temperatures with different time interval data using deep neural networks. *Appl. Sci.* **2020**, *10*, 1609. [CrossRef]
65. Shabib Aftab, M.A.; Hameed, N.; Bashir, M.S.; Ali, I.; Nawaz, Z. Rainfall Prediction in Lahore City using Data Mining Techniques. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*. [CrossRef]
66. Baran, Á.; Lerch, S.; Ayari, M.E.; Baran, S. Machine learning for total cloud cover prediction. *arXiv* **2020**, arXiv:2001.05948.
67. Wei, C.C. Development of Stacked Long Short-Term Memory Neural Networks with Numerical Solutions for Wind Velocity Predictions. *Adv. Meteorol.* **2020**, *2020*, 1–18. [CrossRef]

68. Poreba, S.; Ustrnul, Z. Forecasting experiences associated with supercells over South-Western Poland on July 7, 2017. *Atmos. Res.* **2020**, *232*, 104681. [[CrossRef](#)]
69. Matczak, P.; Graczyk, D.; Choryński, A.; Pińskwar, I.; Takacs, V. Temperature Forecast Accuracies of Polish Proverbs. *Weather Clim. Soc.* **2020**, *12*, 405–419. [[CrossRef](#)]
70. Bartoszek, K.; Krzyżewska, A. The atmospheric circulation conditions of the occurrence of heatwaves in Lublin, southeast Poland. *Weather* **2017**, *72*, 176–180. [[CrossRef](#)]
71. Taszarek, M.; Brooks, H.E. Tornado climatology of Poland. *Mon. Weather Rev.* **2015**, *143*, 702–717. [[CrossRef](#)]
72. Kijewska, M.; Pleskacz, K. Niepewność prognoz parametrów wiatru dla Zalewu Szczecińskiego i Zatoki Pomorskiej jako jedno ze źródeł błędów predykcji trasy dryfu rozbitka. *Autobusy Tech. Eksploat. Syst. Transp.* **2017**, *18*, 216–223.
73. Cedro, A.; Cedro, B. Wpływ warunków klimatycznych i zanieczyszczenia powietrza na reakcję przyrostową sosny zwyczajnej (*Pinus sylvestris* L.) rosnącej w Lasach Miejskich Szczecina. *Leśne Pr. Badaw.* **2018**, *79*, 105–112.
74. Raschka, S.; Patterson, J.; Nolet, C. Machine Learning in Python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information* **2020**, *11*, 193. [[CrossRef](#)]
75. Raschka, S.; Mirjalili, V. *Python Machine Learning*; Number 1; Packt Publishing: Birmingham, UK, 2019.
76. Rasp, S.; Dueben, P.D.; Scher, S.; Weyn, J.A.; Mouatadid, S.; Thuerey, N. WeatherBench: A Benchmark Data Set for Data-Driven Weather Forecasting. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2020MS002203. [[CrossRef](#)]
77. Chuchro, M. Analiza danych środowiskowych metodami eksploracji danych. *Stud. Inform.* **2011**, *32*, 96.
78. Yeo, I.K.; Johnson, R.A. A new family of power transformations to improve normality or symmetry. *Biometrika* **2000**, *87*, 954–959. [[CrossRef](#)]
79. Yeo, I.K.; Johnson, R.A.; Deng, X.W. An empirical characteristic function approach to selecting a transformation to normality. *CSAM Commun. Stat. Appl. Methods* **2014**, *21*, 213–224. [[CrossRef](#)]
80. Optis, M.; Kumler, A.; Brodie, J.; Miles, T. Quantifying sensitivity in numerical weather prediction-modeled offshore wind speeds through an ensemble modeling approach. *Wind Energy* **2021**, 1–17. [[CrossRef](#)]
81. De Los Campos, G.; Pérez-Rodríguez, P.; Bogard, M.; Gouache, D.; Crossa, J. A data-driven simulation platform to predict cultivars' performances under uncertain weather conditions. *Nat. Commun.* **2020**, *11*, 1–10. [[CrossRef](#)]
82. Flovik, V. *How (Not) to Use Machine Learning for Time Series Forecasting: Avoiding the Pitfalls*. Available online: <https://towardsdatascience.com/how-not-to-use-machine-learning-for-time-series-forecasting-avoiding-the-pitfalls-19f9d7adf424> (accessed on 3 December 2020).
83. Faris, H.; Alkasassbeh, M.; Rodan, A. Artificial Neural Networks for Surface Ozone Prediction: Models and Analysis. *Pol. J. Environ. Stud.* **2014**, *23*, 341–348.
84. Cateni, S.; Colla, V.; Vannucci, M. General purpose input variables extraction: A genetic algorithm based procedure GIVE a GAP. In Proceedings of the 2009 Ninth International Conference on Intelligent Systems Design and Applications, Pisa, Italy, 30 November–2 December 2009; pp. 1278–1283.
85. Perera, A.; Azamathulla, H.; Rathnayake, U. Comparison of different Artificial Neural Network (ANN) training algorithms to predict atmospheric temperature in Tabuk, Saudi Arabia. *MAUSAM: Q. J. Meteorol. Hydrol. Geophys.* **2020**, *71*, 551–560.
86. De Barcelos Tronto, I.F.; da Silva, J.D.S.; Sant'Anna, N. An investigation of Artificial Neural Networks based prediction systems in software project management. *J. Syst. Softw.* **2008**, *81*, 356–367. [[CrossRef](#)]
87. Lula, P.; Morajda, J.; Paliwoda-Pekosz, G.; Stal, J.; Tadeusiewicz, R.; Wilusz, W. *Komputerowe Metody Analizy i Przetwarzania Danych*; Uniwersytet Ekonomiczny w Krakowie: Cracow, Poland, 2014.
88. Heidari, A.A.; Faris, H.; Aljarah, I.; Mirjalili, S. An efficient hybrid Multilayer Perceptron neural network with grasshopper optimization. *Soft Comput.* **2019**, *23*, 7941–7958. [[CrossRef](#)]
89. White, H.D. Author cocitation analysis and Pearson's r. *J. Am. Soc. Inf. Sci. Technol.* **2003**, *54*, 1250–1259. [[CrossRef](#)]
90. McQuistan, A. Using Machine Learning to Predict the Weather: Part 2. Available online: <https://stackabuse.com/using-machine-learning-to-predict-the-weather-part-2/> (accessed on 1 December 2020).
91. Rahman, M.A.; Yunsheng, L.; Sultana, N. Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman's rho tests and ARIMA model. *Meteorol. Atmos. Phys.* **2017**, *129*, 409–424. [[CrossRef](#)]