

## Article

# Power Equipment Defects Prediction Based on the Joint Solution of Classification and Regression Problems Using Machine Learning Methods

Ivan Shcherbatov <sup>1</sup>, Evgeny Lisin <sup>2,\*</sup>, Andrey Rogalev <sup>1</sup>, Grigory Tsurikov <sup>1</sup>, Marek Dvořák <sup>3</sup>   
and Wadim Strielkowski <sup>3,\*</sup> 

<sup>1</sup> Department of Innovative Technologies for High-Tech Industries, National Research University “Moscow Power Engineering Institute”, Krasnokazarmennaya st. 14, 111250 Moscow, Russia; shcherbatovia@mpei.ru (I.S.); rogalevan@mpei.ru (A.R.); tsurikovgn@mpei.ru (G.T.)

<sup>2</sup> Department of Economics in Power Engineering and Industry, National Research University “Moscow Power Engineering Institute”, Krasnokazarmennaya st. 14, 111250 Moscow, Russia

<sup>3</sup> Department of Trade and Finance, Faculty of Economics and Management, Czech University of Life Sciences Prague, Kamýčká 129, Prague 6, 165 00 Prague, Czech Republic; dvorakmarek@pef.czu.cz

\* Correspondence: lisinym@mpei.ru (E.L.); strielkowski@pef.czu.cz (W.S.)

**Abstract:** Our paper proposes a method for constructing a system for predicting defects and failures of power equipment and the time of their occurrence based on the joint solution of regression and classification problems using machine learning methods. A distinctive feature of this method is the use of the equipment’s technical condition index as an informative parameter. The results of calculating and visualizing the technical condition index in relation to the electro-hydraulic automatic control system of hydropower turbine when predicting the defect “clogging of drainage channels” showed that its determination both for an equipment and for a group of its functional units allows one to quickly and with the required accuracy assess the arising technological disturbances in the operation of power equipment. In order to predict the behavior of the technical condition index of the automatic control system of the turbine, the optimal tuning of the LSTM model of the recurrent neural network was developed and carried out. The result of the application of the model was the forecast of the technical condition index achievement and the limiting characteristic according to the current time data on its values. The developed model accurately predicted the behavior of the technical condition index at time intervals of 3 and 10 h, which made it possible to draw a conclusion about its applicability for early identification of the investigated defect in the automatic control system of the turbine. Thus, we can conclude that the joint solution of regression and classification problems using an information parameter in the form of a technical condition index allows one to develop systems for predicting defects, one significant advantage of which is the ability to early determine the development of degradation phenomena in power equipment.

**Keywords:** defect forecasting system; power equipment; technical condition index; machine learning; neural network; SCADA; logistic regression; machine learning algorithms



**Citation:** Shcherbatov, I.; Lisin, E.; Rogalev, A.; Tsurikov, G.; Dvořák, M.; Strielkowski, W. Power Equipment Defects Prediction Based on the Joint Solution of Classification and Regression Problems Using Machine Learning Methods. *Electronics* **2021**, *10*, 3145. <https://doi.org/10.3390/electronics10243145>

Academic Editor: Zehui Mao

Received: 28 October 2021

Accepted: 14 December 2021

Published: 17 December 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

The production assets of the electric power industry include many units of power equipment interconnected by a single technological process for the production and conversion of energy at power plants. In order to ensure the reliable and continuous operation of a power plant, it is necessary to organize the planning of operation processes, repairs, and modernization of power equipment. Ensuring the implementation of these processes is entrusted to the maintenance and repair management information system.

One of the key problems in building an effective maintenance and repair management system is the high level of capital intensity of power equipment of power plants associated

with the specific nature of the process of energy production and supply to consumers. Thus, carrying out maintenance, repairs and implementation of programs for modernization and technical re-equipment of the power plant requires attracting large investments. Incorrect distribution of limited investment funds in programs for technical re-equipment, reconstruction, and repairs significantly reduces the level of competitiveness of the power plant when operating in the wholesale electricity market. This leads to an increase in the cost of energy products for consumers and creates a threat of disruption to stable energy supply.

In Russia, the use by power enterprises of various scientific and methodological approaches to the organization of information systems for the management of maintenance and repairs has led to a large difference in the technical condition and efficiency of using the main power equipment at power plants. Due to the fact that more than 68% of the installed capacity of power plants operates within the framework of the unified energy system of the country, a decrease in the efficiency of operation of the power equipment of one power plant has a significant impact on the cost and capacity and electric power balances of other energy sources.

Within the context of limited investment into the Russian electric power industry, the task of maintaining the operating condition of the currently operating power plants, with some of their power equipment being outdated, would require a considerable volume of current and major repairs. Maintenance of obsolete equipment leads to an increase in the cost of its operation, and the downtime of power units increases, which reduces their annual electricity production. In addition, the equipment operates at the limit parameters, and the predicted service life is not restored even after a major overhaul. Thus, there is a lot of focus in the industry on improving the efficiency of maintenance systems, which motivated the study.

One of the most common maintenance systems in the energy industry is through making repairs based on the technical conditions of the equipment. At the same time, the specifics of the activities of energy enterprises do not allow the full use of the classical theory of production assets management, which necessitates the development on the basis of special approaches and tools for managing the operation of the power equipment. All of this is due to the fact that control according to the criterion of ensuring a given level of reliability of power equipment has a number of limitations caused by the complexity of statistical data analysis. Statistics on failures and changes in the parameters of the state of a single equipment cannot be obtained, since it assumes the presence of an array of the equipment belonging to the same type which is operated under the same conditions. At the same time, the main production equipment at many power plants is unique and designed for various operating modes. This feature of power plants limits the use of statistical methods of the theory of reliability [1–4].

Further development of the statistical approach to the management of production assets is associated with the development of a methodology for diagnosing the resource of each piece of equipment at a power plant, based on progressive methods for measuring the main technological parameters of a production asset which determine the reliability and efficiency of its operation and also affect the economy of the power plant [5–7].

In recent years, the concept of managing production assets based on the technical condition of equipment used at energy enterprises has grown into a system for organizing the activities of an energy enterprise as a whole [8–10]. Thus, in order to solve the problem of planning maintenance and repairs at power plants, it is merely not enough to assess the current technical conditions of the equipment. It is also of a great importance to predict the residual resource of this equipment which consists of determining the time of the onset of the limiting values of its technological parameters that determine the occurrence of possible defects and failures.

The presence of this large task facing energy enterprises has led to a broad discussion within the scientific community focused on various methods and approaches that could be applied to the construction of systems for predicting defects and failures of power equipment. The most promising from the point of view of the implementation of the

mathematical apparatus of forecasting are the methods and algorithms of machine learning used to solve problems of classification, regression and clustering [11–15].

In accordance with the previous research [16–21], the construction of a modern predictive analytics system that predicts defects and failures of power equipment requires the implementation of the following sequence of stages:

Formation of a data warehouse coming from a supervisory control and data acquisition (SCADA) system and representing unstructured data on technological parameters, events, alarms, and locks;

Data preprocessing, searching for data outliers and anomalies, filling in gaps and conducting data to a single structure;

Solution of the classification problem based on machine learning algorithms and predictive models. In this case, the output of the predictive models is not a discrete result (1 is equipment failure, 0 is no failure), but the probability of a given technical state of equipment belonging to the “failure” state (to class “1”);

Setting the optimal classification threshold. When the threshold is exceeded, the technical condition of the equipment is characterized as a condition close to failure, damage, or defect.

The considered approach to the construction of predictive analytics systems cannot be characterized as complete in the case of solving problems of predicting the technical state of equipment. This is due to the fact that the indicator of the probability showing whether the current technical condition of the equipment belongs to the class of defects is not sufficiently informative for making a decision on the maintenance required for preventing an emergency shutdown of the equipment. In other words, it does not give an idea of when the defect will occur and how long it takes to analyze the developing degradation situation.

The aim of the study is to propose an approach to the construction of a system for predicting defects and failures of power equipment based on machine learning methods which makes it possible to analyze the developing degradation situation in a span of time. Thus, the research contribution lies in the development of predictive analytics tools to solve the problem of early notification of the need for equipment maintenance.

The research hypothesis is the possibility of achieving the research goal by jointly solving the following problems using machine learning methods:

- The problem of predicting the technical condition index of power equipment and determining the time to reach its limiting state;
- The problem of determining the probability that the current technical condition of power equipment belongs to the state of defects.
- This hypothesis determined the following structure and content of the paper:
- Analysis of approaches to building a system for predicting defects in power equipment;
- Predicting the time of occurrence of defects in power equipment based on the determination of the technical condition index;
- Predicting the probability that the current technical state of power equipment belongs to a class of defects.
- Defect forecasting method development based on the joint solution of regression and classification problems;
- Approbation of the compiled forecast model in accordance with the proposed method.

The rest of this paper is organized as follows. Section 2 focuses on the analysis of approaches to building a system for predicting defects in power equipment. Section 3 describes the defect forecasting method development based on the joint solution of regression and classification problems. Section 4 reports the results and provides their discussion. Finally, Section 5 provides overall conclusions and lists possible implications.

The following major contributions can be highlighted:

- An algorithm is proposed for constructing a modern system for predicting defects and failures of power equipment based on a study of the technical condition index of equipment and determining the time it will reach its limit state;

- An approach has been developed to determine the technical condition index of power equipment and its functional units according to their current technological parameters;
- An approach is proposed to determine the probability that the current technical state of power equipment belongs to a class of defects based on ensembles of machine learning methods;
- The LSTM model of a recurrent neural network has been developed and tuned to solve the problem of predicting the behavior of the technical condition index of power equipment.

A distinctive feature of the developments is the use of the equipment technical condition index as an information parameter. If compared with existing methods, in which the probability of the current technical condition of equipment belonging to the state of defects is usually used as a predicted parameter, the technical condition index is a more visual and informative parameter. Also, its use makes it possible to determine the time before the onset of the limiting state of the equipment and to predict the development of degradation phenomena at an early stage, which is discussed in Section 4.

## 2. Analysis of Approaches to Building a System for Predicting Defects in Power Equipment

### 2.1. Predicting the Time of Occurrence of Defects in Power Equipment Based on the Determination of the Technical Condition Index

As it has been noted above, the process of predicting the time before the onset of defects and failures of power equipment consists of constructing a forecast (solving the regression problem) for the key parameters that directly or indirectly characterize its current technical state. In this case, the predicted time is calculated as the time until the occurrence of the limiting value (corresponding to the state of defect or failure) of these parameters in the future. Since such predicted parameter characterizes the current technical condition, the probability of the current technical condition belonging to the state of defects can be used [16,17].

However, a more visual and informative parameter is represented by the technical condition index (TCI) of equipment [22–26] which makes it possible to assess and predict not only the technical condition of power equipment as a whole but also its individual elements.

There are various approaches to determining the technical condition index of equipment [24,27,28]. One of them is based on comparing the current measured technological parameters of the equipment with the values required for normal operation. With such a comparison in hand, each technological parameter describing the technical condition of the equipment in accordance with the generated knowledge base is assigned its respective estimate on the basis of which the calculation of the TCI is carried out. The disadvantage of this approach for the definition of TCI is the subjectivity of expert assessments entered into the knowledge base, as well as the lack of a complete list of technological parameters required for further correct prediction of many possible defects and equipment failures.

An effective approach to determining the TCI of equipment from the point of view of automation of calculations and ease of interpretation is to assess the degree of deviation of the current values of technological parameters characterizing the technical condition of the equipment from the values of these parameters corresponding to warning alarms [29,30]. This approach is reflected in Equation (1) that follows:

$$HI_t = \begin{cases} \frac{y_t - y_{\min}}{y_{\text{mean}} - \Delta y - y_{\min}} \text{ if } |y_t - y_{\min}| < |y_{\max} - y_t| \\ \frac{y_{\max} - y_t}{y_{\max} - y_{\text{mean}} - \Delta y} \text{ if } |y_t - y_{\min}| \geq |y_{\max} - y_t| \end{cases} \quad (1)$$

where  $HI_t \in [0; 1]$  is the TCI value for a given technological parameter at each moment of time;  $y_t$  is the value of the technological parameter describing the technical condition of the equipment at each moment of time;  $y_{\min}, y_{\max}$  is the minimum and maximum technological parameter values  $y_t$  corresponding to the alarm level;  $y_{\text{mean}} = \frac{y_{\max} - y_{\min}}{2}$  is the average

value of technological parameter, ideal in relation to alarm levels; and  $\Delta y = \delta \cdot y_{mean}$  is the deviation from mean value.

The value of TCI for a group of technological parameters characterizing the state of the equipment is determined as the harmonic mean of all individual TCI given by the following Formula (2).

$$HI_t^{group} = \frac{N}{\frac{1}{HI_{1t}} + \frac{1}{HI_{2t}} + \dots + \frac{1}{HI_{Nt}}} \quad (2)$$

where  $N$  is the number of technological parameters describing the current technical state of the considered energy facility.

The determination of the technical condition index in accordance with Formulas (1) and (2), both for a unit of technological equipment and for a group of its functional units on the basis of technological parameters characterizing the technical state of the equipment, makes it possible to assess the arising technological disturbances in the operation of power equipment and, as a result of this, to identify the development of a manufacturing defect or the occurrence of a failure.

## 2.2. Predicting the Probability That the Current Technical State of Power Equipment Belongs to a Class of Defects

This approach is based on the determination of the probability that the current technical condition of the equipment belongs to the class of defects. The arising classification problem is solved using machine learning methods.

Historical data on technological parameters characterizing the technical condition of the equipment (in the form of an attribute space  $X$ ), as well as data on defects and failures in the form of a response space  $Y \in [0; 1]$  (0 is a class characterizing the normal operation of the equipment, and 1 is a class characterizing the presence of a defect) are sent as input for the machine learning models with the help of which the problem of determining the probabilities of belonging of each element of the feature space to the class "1" is then solved.

Furthermore, after building and tuning the optimal machine learning models, the optimal classification threshold is determined, above which a developing defect is identified. A notification signal is generated and sent to the operator who then decides about the need for the equipment maintenance.

The following machine learning algorithms can be used here: logistic regression [31–34], random forest [35–37], gradient boosting of decision trees [38–40] and ensembles of the presented algorithms [41–43]. In the process of training the listed algorithms, the data is divided into training, validation (using the strategy of stratified K-fold cross-validation [44,45]), and test sets.

The hyperparameters of each algorithm are adjusted in such a way as to minimize the error functional-logistic losses (log-loss, cross-entropy) when training these algorithms, which is shown in Equation (3).

$$Q(a(x, hyperparam), x) = \sum_{i=1}^l \ln(1 + e^{-y_i a(x_i)}) \rightarrow \min_{hyperparam} \quad (3)$$

where  $a(x, hyperparam)$  is the predictive response of a machine learning algorithm;  $hyperparam$  is the algorithm hyperparameter;  $Q$  is the functional error; and  $i = \overline{1, l}$  is the training sample element.

The comparison of the performance of algorithms on a test sample is carried out using the metrics of completeness (recall) and accuracy (precision) which characterize such indicators of the algorithm as the number of false alarms of the algorithm (false positive, FP), false classifications of the defect and false omissions of the algorithm (false negative, FN). In order to prevent false notifications about eventual defects in power equipment, the selection of such classification threshold is carried out, which makes it possible to minimize

the number of false omissions of the algorithm (FN) and observe the least number of false alarms (FP).

If the probability of belonging to a class of defects is greater than the optimal classification threshold, a warning message “state close to failure” flashes in the predictive analytics system.

### 3. Defect Forecasting Method Development Based on the Joint Solution of Regression and Classification Problems

The concept of a new method for constructing an effective forecasting system is based on the combination of the two previously presented approaches for predicting defects and failures of power equipment. The task of classifying the current technical state of equipment is complemented by the task of regression (forecasting time series), which is possible to solve not only for determining the presence of a developing emergency situation based on historical events but also for predicting the dynamics of its development and the time of occurrence of the limiting state.

The problem of predicting the probability of belonging to the current technical state to a class of defects is solved on the basis of ensembles of “risky” (the method of gradient boosting of decision trees) and “cautious” (method of logistic regression) methods of machine learning, followed by a visually understandable and interpretable definition and forecasting of TCI and the time to reach it limiting value. As a mathematical apparatus for solving the problem of forecasting a time series (TCI behavior over time), we suggest using machine learning methods that are usually applied when working with the data sequences in regression tasks. One of these methods is based on the construction of a long short-term memory model of a recurrent neural network.

The neural network is a universal approximator based on Kolmogorov’s theorem [46,47] and Hornik’s approximation theorem [48,49]. The graphical representation of this mathematical model may be different depending on the modification of its mathematical apparatus [50].

One of the modifications of neural networks is a recurrent neural network (RNN) [51,52]. This network model is commonly used for solving regression and classification problems when working with data sequences. The structure and operating principle of this model is fundamentally different from the direct neural network model (Figure 1).

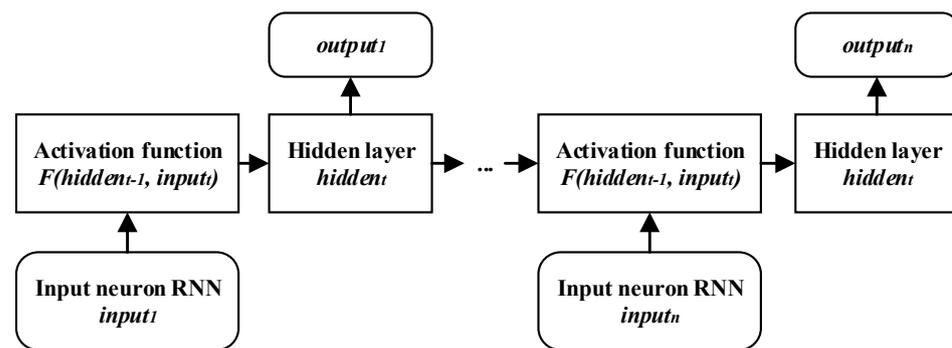


Figure 1. Schematic diagram of signal propagation in a recurrent neural network (many to many type).

It can be noted based on the schematic diagram of signal propagation in the RNN that, in contrast with the direct propagation model, the input data is sent to the model sequentially at each moment of time  $t$ . In this case, at each step of signal propagation, based on the current input data and the previous calculated state, the current state ( $output_t$ ) is calculated. This process is repeated by  $n$  steps until the required output of the model (predicted value) is determined or until the input data ( $input_t$ ) of the model is exhausted.

Signal propagation in the recurrent neural network model is represented through the values of each hidden state ( $hidden_t$ ) (4), calculated based on the previous hidden state ( $hidden_{t-1}$ ) and the current input data ( $input_t$ ).

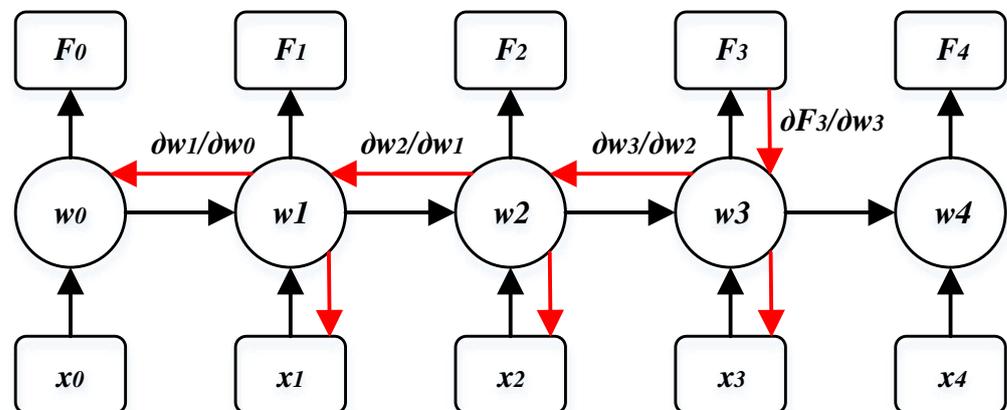
$$hidden_t = \sigma(\langle w_{hidden}, hidden_{t-1} \rangle + \langle w_{input}, input_t \rangle) \quad (4)$$

where  $\sigma()$  is the activation function (sigmoid function, hyperbolic tangent, linear rectifier ReLu);  $w_{hidden}$ ,  $w_{input}$  are the weights for hidden and input states, respectively.

The output value at each calculation step ( $output_t$ ) (5) is computed as the dot product of the weights at the output state by the values at the hidden state, similar to the regression equation.

$$output_t = \langle w_{output}, hidden_t \rangle \quad (5)$$

The training of a recurrent neural network can be carried out by the error backpropagation method [53], the schematic diagram of which in relation to RNN is shown in Figure 2.



**Figure 2.** Schematic diagram of the error back propagation method for a recurrent neural network model.

Thus, during training, after calculating the output signal (the initial stage), the error functional (6) is determined (in regression problems, the root of the standard deviation between the answers RNN  $output_t$  and the values from the response space  $y_t$  is used).

$$F(w_{output}, w_{hidden}, w_{input}) = \sqrt{\sum_{t=1}^n \left( \frac{(output_t - y_t)^2}{n} \right)} \rightarrow \min_w \quad (6)$$

Based on the chain rule for calculating the gradient, the gradient of the error functional is determined. The weight coefficients  $w_{ij}$  are adjusted in the direction of decreasing of the given functional until it takes the minimum value or training iterations reach the set limit. It is worth noting that the weights for the hidden state of the recurrent neural network  $w_{hidden}$  remain the same after the backpropagation of the error from each output  $output_t$ , while the coefficients  $w_{output}$  and  $w_{input}$  change at each step of the gradient descent.

The use of the error backpropagation method in recurrent neural network models for solving approximation problems for large data sequences determines the main disadvantage of these models. As the sequence size increases, the number of inputs also increases, and so does the number of hidden states. When using a chain rule to compute a gradient, the number of computations increases. At the same time, it was proved [51] that the value of the gradient after many iterations of gradient descent can exponentially decrease, which prevents the updating of the weight coefficients or exponentially increases, and leads to an unstable solution of the approximation problem. With regard to the above, the models of recurrent neural networks are called “neural networks with short-term memory” [51,54] since they demonstrate good quality only on small data sequences.

In order to improve the quality of models on large data sequences, we suggest using the long short-term memory (LSTM) method for constructing recurrent neural networks [55–57]. A schematic diagram of signal propagation using this method is shown in Figure 3.

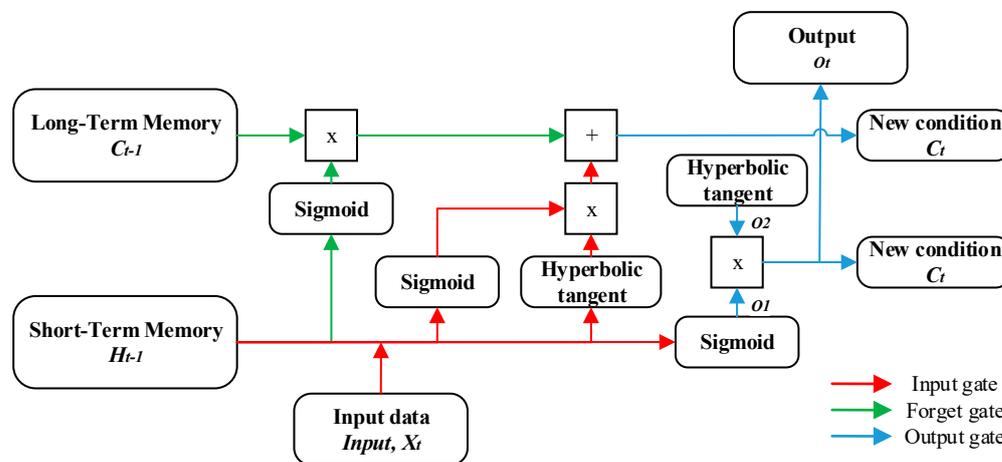


Figure 3. Schematic diagram of signal propagation for the LSTM model of a recurrent neural network.

In this schematic diagram (see Figure 3), several signal propagation streams can be observed. In the first stream (*input, input gate*), filtering (selection) of data arriving in the long-term memory of the model takes place. The stream involves new data  $X_t$  arriving at the input of the model and data  $H_{t-1}$  stored in the short-term memory of the model. With the help of the sigmoid function (7), the incoming data is converted to the limits [0,1], where 0 is assigned to data that are not useful for making a forecast and 1 is assigned to data useful for further training of the model.

$$i_1 = \sigma(\langle w_{i_1}, H_{t-1} \rangle + \langle u_{i_1}, X_t \rangle + bias_1) \tag{7}$$

In the process of training the neural network using the backpropagation method, the weight coefficients  $w_{i_1}$  and  $u_{i_1}$  are updated in such a way as to discard data that noises the forecast while missing useful ones.

With the activation function (hyperbolic tangent) (8), the same input data  $X_t$  and data from short-term memory are adjusted within the range [−1; 1]:

$$i_2 = \tanh(\langle w_{i_2}, H_{t-1} \rangle + \langle u_{i_2}, X_t \rangle + bias_2) \tag{8}$$

Next, the scalar multiplication of the first  $i_1$  and second  $i_2$  layers of the considered signal propagation flow is determined and, thereby, the information entering the long-term memory cell  $i_{input\_gate} = \langle i_1, i_2 \rangle$  is determined.

In the second stream of signal propagation (*forget gate*), it is determined which data from long-term memory at the calculated step  $t - 1$  should be added or removed when new information appears at the calculated step  $t$ . Similarly, as in the case of the first stream (*input gate*) (7), using the same sigmoid function with different weight coefficients  $w_{forget\_gate}$ , useful information (among the input data  $X_t$  and data coming from the short-term memory  $H_{t-1}$ ) is determined by Equation (9).

$$f = \sigma(\langle w_{forget\_gate}, H_{t-1} \rangle + \langle u_{forget\_gate}, X_t \rangle + bias_{forget\_gate}) \tag{9}$$

The new state of long-term memory (data set)  $C_t$  (10) is formed taking into account (9) and the data received on the input stream.

$$C_t = C_{t-1}f + i_{input\_gate} \tag{10}$$

On the third (output) stream of information dissemination, based on the input data for the current calculated step  $X_t$ , coming from short-term memory at the previous calculated step  $H_{t-1}$  and the new state of long-term memory  $C_t$  (10), the value of the current output of the neural network model  $O_t$  and the new state of short-term memory (new hidden state)  $H_t$  are formed by Equations (11)–(13). At the same time, a filter implemented in the form of a sigmoid function is also applied to the data  $X_t$  and  $H_{t-1}$ , and the new state of long-term memory  $C_t$  is processed using the hyperbolic tangent.

$$O_1 = \sigma(\langle w_{o_1}, H_{t-1} \rangle + \langle u_{o_1}, X_t \rangle + bias_{O_1}) \quad (11)$$

$$O_2 = \tanh(w_{o_2} C_t + bias_{O_1}) \quad (12)$$

$$H_t, O_t = \langle O_1, O_2 \rangle \quad (13)$$

Thus, when determined at the current calculated step  $t$  and corresponding to each new input, the new states at each next step take the index  $t - 1$ , and the calculations within the described signal propagation streams are repeated. The weights are updated using the backpropagation method. In contrast to the standard model of a recurrent neural network, the values of the gradients during gradient descent operations remain stable due to the structure of the LSTM of recurrent neural networks [53].

In order to develop and train the LSTM model, the PyTorch library, which is one of the most commonly used Python libraries for deep learning, is recommended. For preliminary processing and analysis of data obtained from the SCADA system of the power plant, the most convenient method is the use of the statistical software package Statgraphics.

The application of the LSTM model of recurrent neural networks, in practice, results in high quality in solving problems of forecasting data sequences and time series. In the case of a time series forecast, as input data  $X_t$ , the neural network model cannot be fed data from the feature space  $X$  (as it happens in standard regression problems solved by machine learning methods) but instead data from the response space  $Y$  shifted by the time required to make the forecast. With regard to the problem of predicting the technical condition index of power equipment, the indicated offset (time lag) can be used for the current temporal data on the values of the index. That is, the input of the neural network model is supplied with the values of the index shifted by time, which makes it possible to predict whether the index would reach the limit value. In this case, the bias is a hyperparameter of the neural network model.

#### 4. Approbation of the Results and Discussion of Results

In recent decades, Russia focused on the development of renewable energy, first of all, on the use of the country's potential in hydro power. The leader in this area is PJSC RusHydro, one of the largest hydro-generating companies in the world which generates electric power from renewable sources.

In addition to its high-capacity HPPs, PJSC RusHydro is actively developing small HPPs focused on supplying power to isolated and hard-to-reach areas of the country. Reducing the rated capacity and placing energy sources closer to the local consumer is currently a global trend.

Small hydroelectric power plants represent a low-carbon energy source and have a number of advantages over other RES-generation facilities (above all in terms of the average cost of electricity production throughout the entire life cycle of operation). In this case, the systems of maintenance and repair of equipment take on key importance, which allow for maintaining of the main power equipment in working order based on SCADA.

The most common SCADA system in the power industry is ClearSCADA which supports various standards for information transfer in telecommunication systems. The support for the third-party software and hardware through the use of open standards and communication protocols allows ClearSCADA to work, in particular, with controllers Siemens, Schneider, Yokogawa, Control-Logic, and Omron. In 2009, a series of tests was conducted in order to confirm the compliance of the ClearSCADA functionality with

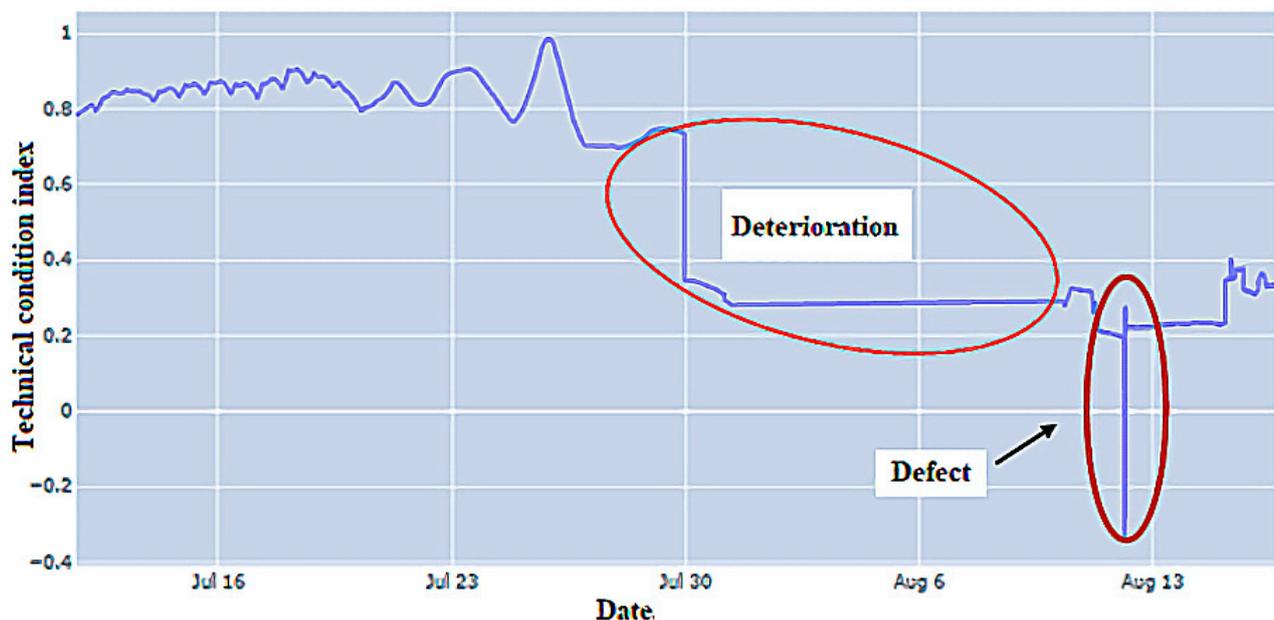
the requirements of power enterprises. Since then, ClearSCADA has been used for the purposes of automated process control system (APCS) at geographically distributed power plants, such as small HPPs.

The object of the study was the defect “clogging of drainage channels” of the electro-hydraulic automatic control system of the hydroelectric turbine. Equipment failure data were collected from the plant’s SCADA system during the summer period of 2017. Table 1 shows the set values of the lower and upper limits of the warning alarms that are most correlated with the defect under investigation.

**Table 1.** Values of warning alarms for the parameters of the technical state of the automatic turbine control system.

Signal Name (Feature)	Ymin	Ymax
Water level in the turbine cover, mm	550	750
Hydroelectric head, m	18	25
Oil temperature in the drain tank of the high-pressure oil unit, °C	30	42
Oil pressure in the boiler of the high-pressure oil unit (sensor 1), kgf/cm <sup>2</sup>	0	140
Oil pressure in the boiler of the high-pressure oil unit (sensor 2), kgf/cm <sup>2</sup>	0	140
Oil level in the drain tank of the high-pressure oil unit, mm	190	230
Oil pressure in the boiler of the low-pressure oil unit, kgf/cm <sup>2</sup>	0	28
Oil level in the boiler of the low-pressure oil unit, mm	−1500	800
Oil level in the drain tank of the low-pressure oil unit, mm	−800	270

In accordance with the above formulas (see Formulas (1) and (2)), a study of the behavior of the technical condition index of the electro-hydraulic automatic control system of the hydroelectric turbine was carried out (Figure 4).



**Figure 4.** Behavior of the technical condition index of the automatic control system of the hydroelectric turbine.

From the analysis shown on the graph, it can be seen that the generalized TCI of the automatic control system of the turbine 15 days before the detection of the defect “clogging of drainage channels” decreased from 75% to 35%. At the same time, on the day of the defect detection, the calculated TCI value became negative. This behavior of the TCI is a consequence of going beyond the upper set level of the warning alarm for the technological parameter “water level in the turbine cover” (see Figure 5) which was the cause of the defect.

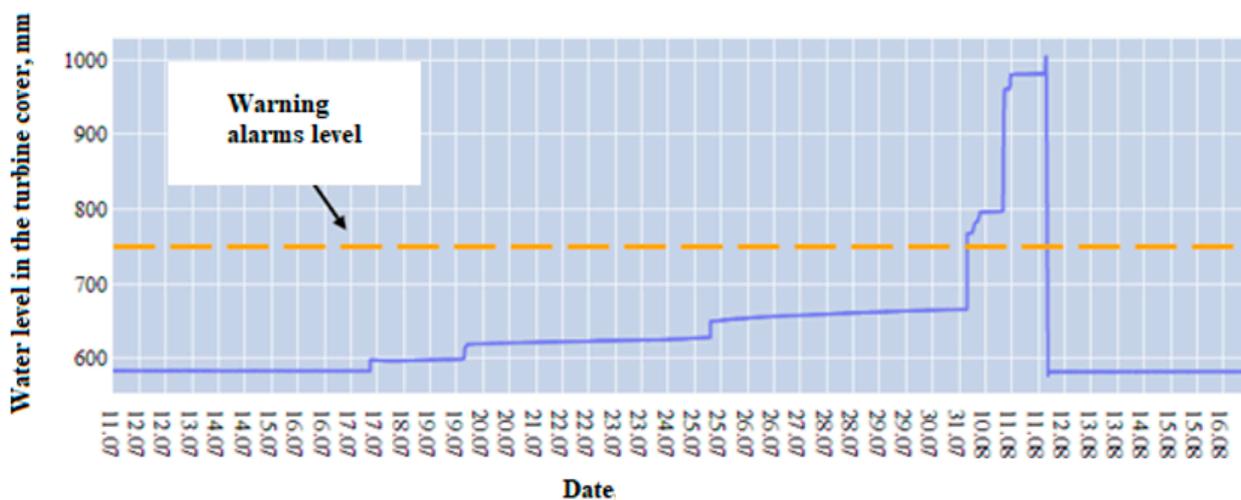


Figure 5. Behavior of the parameter “water level in the turbine cover”.

Thus, the determination of the technical condition index both for a piece of equipment and for a group of its functional units on the basis of technological parameters characterizing the technical state of the equipment makes it possible to quickly, and with the required accuracy, assess the arising technological disturbances in the operation of power equipment. As a result, one can identify the development a manufacturing defect or failure and take a timely decision to start the maintenance of the equipment in order to prevent an emergency shutdown.

As noted earlier, determining the probability that the current technical state of the equipment belongs to the class of defects “clogged drainage channels” is a classification task that is solved by machine learning methods. Here, we considered the technological parameters presented in Table 1 and discrete signals for switching on/off the pumps of oil pressure units as a feature space.

In the course of research and tuning of optimal machine learning models, it was determined that the most qualitative model for solving this problem is a combination of a “riskier” [15] random forest algorithm and a more “careful” logistic regression. At the same time, the approach to building machine learning models using the ensemble technique (the weighted voting technique was applied) was as follows: each basic algorithm  $b_N(x)$  participating in the construction of the composition was assigned a weight coefficient  $\beta$ , and then, for each element of the sample, by voting these basic algorithms, the answer  $a(x)$  (14) was chosen that optimizes the quality metrics of the machine learning model.

$$a(x) = \max_{\beta_1, \dots, \beta_N} \{\beta_1 b_1(x), \dots, \beta_N b_N(x)\} \quad (14)$$

The optimal hyperparameters for each tuned machine learning model participating in the developed ensemble of algorithms, the metrics of the data quality of the models, as well as the optimal weights  $\beta$  that were used in the weighted voting of these models, are presented in Table 2.

Table 2 shows the presence of machine learning errors (FP type) which means that the model can mistakenly predict the state of equipment close to a defect.

**Table 2.** Hyperparameters of tuned machine learning models and quality metrics.

Hyperparameter/Metric Name	Value
<b>I. (a) Logistic Regression: Hyperparameters</b>	
Regularization coefficient L2 (Ridge regularization)	100
<b>I. (b) Logistic Regression: Quality metrics</b>	
Cross-entropy on the training set	0.0006270288
Cross-entropy on the test set	0.0011196107
Completeness on the test set (after determining the optimal threshold)	1.0
Accuracy on the test set (after determining the optimal threshold)	0.999523719
Optimal classification threshold	0.613599
Number of false skips of a defect (FN)	0
Number of false positives of the algorithm on the test set (FP)	5
<b>II. (a) Random Forest: Hyperparameters</b>	
Number of basic algorithms (number of decision trees)	15
Depth of basic algorithms	12
Minimum subset of elements in tree leaves	1
Minimum subset of elements in tree tops	2
Value of the share of the number of features required for training decision trees	36.3%
<b>II. (b) Random Forest: Metrics</b>	
Cross-entropy on the training set	
Cross-entropy on the test set	0.000364
Completeness on the test set (after determining the optimal threshold)	1.0
Accuracy on the test set (after determining the optimal threshold)	0.999523719
Optimal classification threshold	0.0665
Number of false skips of a defect (FN)	0
Number of incorrectly identified defects (false positives) in equipment on the test set (FP)	5
<b>III. (a) Random Forest and Logistic Regression Ensemble: Hyperparameters</b>	
Logistic Regression weighting factor	2.5
Random Forest weighting factor	8.0
<b>III. (b) Random Forest and Logistic Regression Ensemble: Metrics</b>	
Number of false skips of a defect (FN)	0
Number of incorrectly identified defects (false positives) in equipment on the test set (FP)	2
Optimal classification threshold	0.2

In order to test the proposed method for predicting a defect based on the study of the behavior of the technical condition index and reaching its limit values, the LSTM model of the recurrent network was tuned. The root of the standard deviation is used as the error functional. To minimize the error functional by directed enumeration, the following optimal model hyperparameters were determined:

- Percentage ratio of training and validation set volumes;
- Time shift of input data (predicted time of TCI change);
- Number of latent state neurons in one layer of the recurrent neural network;
- Number of layers of the recurrent neural network;
- Size of long-term memory, that is, the amount of data needed to be stored in memory;
- Model regularization factor;
- Learning rate when using gradient descent in the backpropagation method;

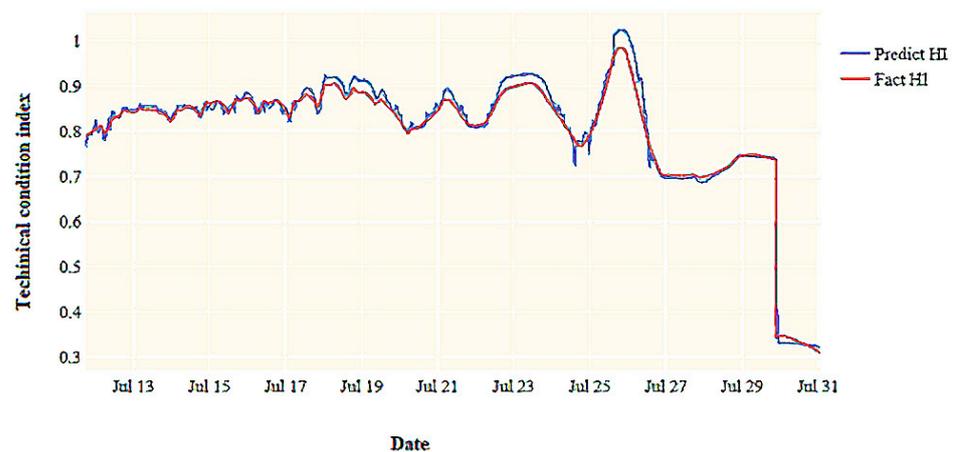
Number of steps of the backpropagation algorithm (number of training eras).

The optimal values of these parameters obtained during the tuning of the neural network model are presented in Table 3.

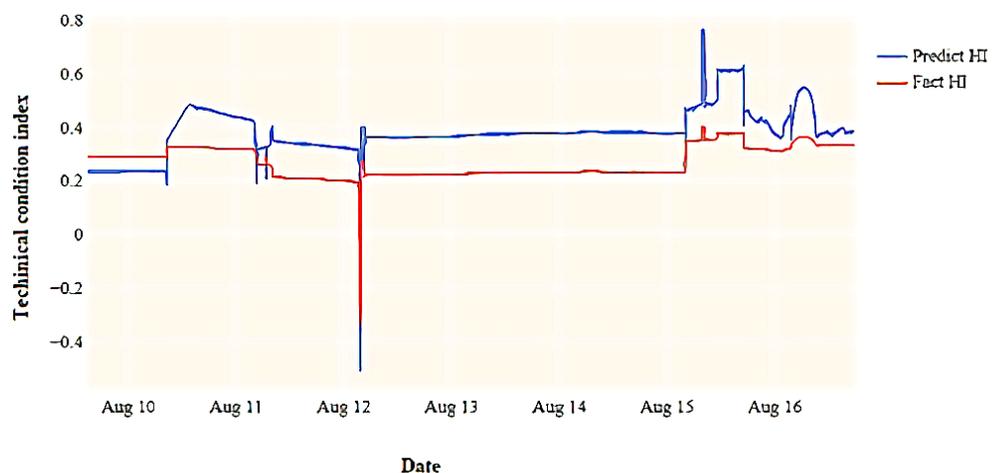
**Table 3.** Values of the optimal hyperparameters of the tuned LSTM model of the recurrent neural network.

Hyperparameter Name	Value
Percentage ratio of training and validation set volumes, %	75/25
Time shift of input data (predicted time of TCI change), hours	3; 10
Number of latent state neurons in one layer of the recurrent neural network	134
Number of layers of the recurrent neural network	2
Size of long-term memory, number of design points	10,132
Learning rate	0.03
Number of training eras	103

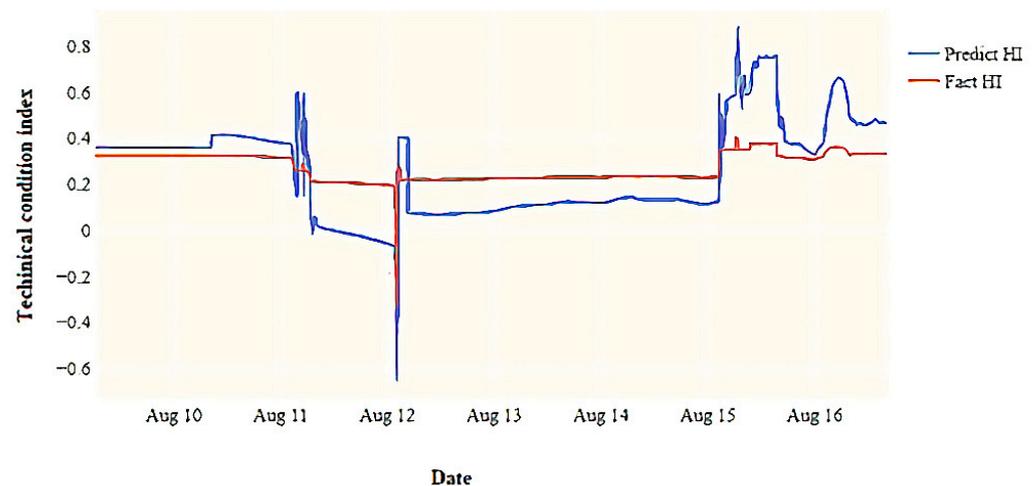
The resulting graphs of changes in the predicted values of the technical condition index on the training and test sets in comparison with real data with time shifts equal to 3 h and 10 h are presented in Figures 6–8.



**Figure 6.** Changes in the predicted values of the technical condition index on the training set in comparison with real data (time shift—3 h).



**Figure 7.** Changes in the predicted values of the technical condition index on the test set in comparison with real data (time shift—3 h).



**Figure 8.** Changes in the predicted values of the technical condition index on the test set in comparison with real data (time shift—10 h).

From the analysis of the resulting graphs, the following outcomes and comments can be made.

With increasing training eras, the calculated value of the error on the training set decreases and takes on a constant value during subsequent iterations of gradient descent and updating the weight coefficients. In this case, the value of the error on the validation set is stochastically decreasing. This behavior of the error on the validation set relative to the error functional on the training set indicates that the tuned LSTM model of the recurrent neural network did not overfit at the specified training interval.

The change in the technical condition index on the training sample with a time offset of 3 h shows that the model quite accurately repeats the real behavior of this index ( $\sqrt{\text{MSD}} = 0.011$ ). However, it is worth noting that due to the existing time offset (the time lag of the data entering the model input), it is possible to observe “run-outs” in the predicted values of the technical condition index relative to the real ones. These “run-outs” can help to accurately determine the time of equipment failure when they are correctly compared with the probability that the current technical condition of the equipment belongs to a defect.

When using a time offset equal to 3 h (Figure 7), predicting the defect under investigation becomes impossible from a practical point of view, since the predicted values of the technical condition index are issued almost simultaneously with the process of deterioration of the equipment condition. Thus, there is practically no time left for an operative analysis of the situation by the operator. However, when using a 10-h time offset, the abnormal behavior of the predicted technical condition index (Figure 8), namely its early sharp decrease, allows the operator to assess the situation in time and take appropriate maintenance measures for the automatic turbine control system unit and not allow the shutdown of equipment.

The use of information on the probability that the current technical state belongs to the condition of the defect “clogged drainage channels”, which for the considered 10 h before the defect changed from 0.873 to 1.000, improves the quality of the operator’s decision-making on the implementation of measures aimed at maintaining the equipment in proper technical condition.

## 5. Conclusions

Overall, the proposed method for constructing a forecasting system for detecting defects and failures of power equipment on test samples demonstrated its effectiveness in terms of early determination of the development of degradation phenomena in equipment. Thus, the hypothesis we put forward about the possibility of studying the developing

degradation situation in time by jointly solving the problems of predicting equipment technical condition index and determining the probability that its current technical state belongs to the state of defects turned out to be correct. This allows us to conclude about the significant contribution of the study to the development of industry predictive analytics tools, which allow one to continue proactively informing others about the need for equipment maintenance.

In the course of this research, the following significant scientific results were obtained.

Based on the analysis of the literature, an algorithm for constructing a modern predictive analytics system that predicts defects and failures of power equipment has been compiled. An approach to its improvement based on the assessment of equipment technical condition index and time determination to reach its limiting state is proposed, which makes it possible to increase the information content of the predictive analytics system.

An approach to determining the index of the technical condition of power equipment by its current technological parameters has been proposed. It has shown that the analysis of the technical condition index both for an equipment and for a group of its functional units makes it possible to quickly and with the required accuracy assess the arising technological disturbances in the operation of power equipment. Determination of the technical condition index of equipment and the time to reach its limit value is carried out by forecasting time series.

An approach to determine the probability that the current technical state of equipment belongs to a class of defects has been provided. It is shown that the arising classification problem can be effectively solved using ensembles of “risky” (random forest) and “cautious” (logistic regression) machine learning methods.

Optimal tuning of the LSTM model of a recurrent neural network for solving the problem of predicting the behavior of the technical condition index of equipment has been developed and carried out. It is shown that if the index values shifted by the time required to make the forecast are fed into the neural network model as input data, then in this way it is possible to predict the index reaching the limit value.

The approbation of scientific results was carried out on the basis of data obtained from the SCADA system of the hydroelectric power plant. The defect “clogging of drainage channels” of the electro-hydraulic automatic control system of the hydro turbine was considered. The discussion part of the study led us to the following conclusions:

The constructed model quite accurately repeats the real behavior of the technical condition index. However, due to the time lag of the data arriving at the input of the model, it is possible to observe “run-outs” in the predicted values of the index relative to the real ones. When correctly compared with the probability that the current technical state belongs to a defect, these “run-outs” can help to accurately determine the time of equipment failure.

Investigation of the defect “clogging of drainage channels” with a time shift of 3 h allowed observing “run-outs” during the period of anomalous behavior of the predictive technical condition index. At the same time, the use of this time shift in predicting the investigated defect loses its practical value, since the predicted values of the technical condition index are issued almost simultaneously with the process of deterioration of the equipment condition.

Investigation of the defect “clogging of drainage channels” at a time offset of 10 h also made it possible to observe “run-outs” during the period of the defect occurrence. In this case, the efficiency of using the forecast data increases. An early sharp decrease in the technical condition index allows the operator to assess the situation in time and take appropriate measures for the maintenance of the unit of the automatic turbine control system and prevent the shutdown of all equipment.

Despite the existing advantages of the developed approach to the design of a system for predicting defects in power equipment based on the determination of the technical condition index and the time to reach its limiting value, the question of calculating the index for equipment consisting of a set of functional units remains open. At the moment, weights are used, reflecting the importance of each unit for the operation of equipment (determined

by an expert method). A further direction of research can focus on the refinement of the methods for calculating the technical condition index for various classes of power equipment and the construction of predictive models for the occurrence of equipment defects in the event of failure of its functional units.

**Author Contributions:** Conceptualization, I.S. and E.L.; methodology, I.S., E.L., A.R., M.D., and W.S.; formal analysis, I.S., M.D., and W.S.; investigation, I.S., G.T., E.L., M.D., and W.S.; resources, I.S., and A.R.; data curation, I.S., A.R., E.L., M.D., and W.S.; writing—original draft preparation, E.L., I.S., M.D., and W.S.; writing—review and editing, E.L., and W.S.; visualization, I.S., E.L., and G.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Ministry of Science and Higher education of the Russian Federation, grant number MD-914.2020.6.

**Acknowledgments:** This work was performed under the grant of RF President for governmental support of young scientists PhD and DSc (MD-914.2020.6).

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

## References

- Babak, V.; Babak, S.; Myslovych, M.; Zaporozhets, A.; Zvaritch, V.M. Principles of construction of systems for diagnosing the energy equipment. In *Diagnostic Systems for Energy Equipments*; Springer: Cham, Switzerland, 2020; pp. 1–22.
- Zaporozhets, A.; Eremenko, V.; Serhiienko, R.; Ivanov, S. Methods and hardware for diagnosing thermal power equipment based on smart grid technology. In Proceedings of the Conference on Computer Science and Information Technologies, Lviv, Ukraine, 11–14 September 2018; pp. 476–489.
- Li, S.; Nie, Y.; Li, J. Condition monitoring and diagnosis of power equipment: Review and prospective. *High Volt.* **2017**, *2*, 82–91. [\[CrossRef\]](#)
- Kabir, S.; Taleb-Berrouane, M.; Papadopoulos, Y. Dynamic reliability assessment of flare systems by combining fault tree analysis and Bayesian networks. *Energy Sources Part A* **2019**, 1–18. [\[CrossRef\]](#)
- Lisin, E.; Rogalev, A.; Strielkowski, W.; Komarov, I. Sustainable modernization of the Russian power utilities industry. *Sustainability* **2015**, *7*, 11378–11400. [\[CrossRef\]](#)
- Lisin, E.; Kurdiukova, G.; Strielkowski, W. Economic prospects of the power-plant industry development in Russia. *J. Int. Stud.* **2016**, *9*, 178–190. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lisin, E.; Shuvalova, D.; Volkova, I.; Strielkowski, W. Sustainable development of regional power systems and the consumption of electric energy. *Sustainability* **2018**, *10*, 1111. [\[CrossRef\]](#)
- Okley, P. Leading prospects for the development of production asset management systems of Russian thermal power plants. In Proceedings of the 3rd International Conference on Social Economic, and Academic Leadership, Prague, Czech Republic, 23–24 March 2019; Volume 318, pp. 361–365.
- Sharovin, I.; Lopatin, V.; Trofimov, V.; Trofimov, A. From Automated Design to Digital Double of APCS for a TPP. *Therm. Eng.* **2021**, *68*, 228–234. [\[CrossRef\]](#)
- Arakelyan, E.; Kosoy, A.; Mezin, S.; Pashchenko, F. Application of the basic principles of “Industry 4.0” in the intellectualization of automated control systems of modern thermal power plants. *Procedia Comput. Sci.* **2021**, *184*, 865–870. [\[CrossRef\]](#)
- Carvalho, T.; Soares, F.; Vita, R.; Francisco, R.; Basto, J.; Alcalá, S. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024. [\[CrossRef\]](#)
- Cinar, Z.; Abdussalam Nuhu, A.; Zeeshan, Q.; Korhan, O.; Asmael, M.; Safaei, B. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability* **2020**, *12*, 8211. [\[CrossRef\]](#)
- Stetco, A.; Dinmohammadi, F.; Zhao, X.; Robu, V.; Flynn, D.; Barnes, M.; Nenadic, G. Machine learning methods for wind turbine condition monitoring: A review. *Renew. Energy* **2019**, *133*, 620–635. [\[CrossRef\]](#)
- Alsina, E.; Chica, M.; Trawinski, K.; Regattieri, A. On the use of machine learning methods to predict component reliability from data-driven industrial case studies. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 2419–2433. [\[CrossRef\]](#)
- Kocsis, G.; Xydis, G. Repair Process Analysis for Wind Turbines Equipped with Hydraulic Pitch Mechanism on the U.S. Market in Focus of Cost Optimization. *Appl. Sci.* **2019**, *9*, 3230. [\[CrossRef\]](#)
- Shcherbatov, I.; Turikov, G. Determination of power engineering equipment’s defects in predictive analytic system using machine learning algorithms. *J. Phys. Conf. Ser.* **2020**, *1683*, 042056. [\[CrossRef\]](#)
- Arakelian, E.; Pashchenko, A.; Shcherbatov, I.; Tsurikov, G.; Titov, F. Creation of Predictive Analytics System for Power Energy Objects. In Proceedings of the Twelfth International Conference “Management of Large-Scale System Development” (MLSD), Moscow, Russia, 1–3 October 2019; pp. 1–5. [\[CrossRef\]](#)
- Wang, J.; Zhang, W.; Shi, Y.; Duan, S.; Liu, J. Industrial big data analytics: Challenges, methodologies, and applications. *arXiv* **2018**, arXiv:1807.01016.

19. Maldonado-Correa, J.; Martín-Martinez, S.; Artigao, E.; Gomez-Lazaro, E. Using SCADA data for wind turbine condition monitoring: A systematic literature review. *Energies* **2020**, *13*, 3132. [[CrossRef](#)]
20. Moleda, M.; Momot, A.; Mrozek, D. Predictive maintenance of boiler feed water pumps using SCADA data. *Sensors* **2020**, *20*, 571. [[CrossRef](#)] [[PubMed](#)]
21. Astolfi, D.; Castellani, F.; Lombardi, A.; Terzi, L. Multivariate SCADA Data Analysis Methods for Real-World Wind Turbine Power Curve Monitoring. *Energies* **2021**, *14*, 1105. [[CrossRef](#)]
22. Berge, S.; Lund, B.; Ugarelli, R. Condition monitoring for early failure detection. Frognerparken pumping station as case study. *Procedia Eng.* **2014**, *70*, 162–171. [[CrossRef](#)]
23. Plotnikova, L.; Bainov, A.; Torkunova, Y.; Nadezhdina, M. Digitalizing the Process of Tracking Technical Condition of the Main Equipment of Energy Providing Enterprises. In Proceedings of the 3rd International Scientific Conference on New Industrialization and Digitalization (NID 2020), Ekaterinburg, Russia, 12 December 2020. [[CrossRef](#)]
24. Vianna, E.; Abaide, A.; Canha, L.; Miranda, V. Substations SF6 circuit breakers: Reliability evaluation based on equipment condition. *Electr. Power Syst. Res.* **2017**, *142*, 36–46. [[CrossRef](#)]
25. Vitolina, S. Development of lifetime data management algorithm for power transformers. In Proceedings of the 5th International Conference on Intelligent Systems, Modelling and Simulation, Taichung, China, 10–12 December 2014; pp. 452–457. [[CrossRef](#)]
26. Kavchenkov, V.P.; Nazarov, A. Assessment of the Electric Power System Elements Reconstruction Priority Taking into Account Mode Reliability. In Proceedings of the 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), Vladivostok, Russia, 6–9 October 2020; pp. 1–6. [[CrossRef](#)]
27. Atamuradov, V.; Medjaher, K.; Camci, F.; Zerhouni, N.; Dersin, P.; Lamoureux, B. Machine health indicator construction framework for failure diagnostics and prognostics. *J. Signal Process. Sys.* **2020**, *92*, 591–609. [[CrossRef](#)]
28. Nazarythev, A.; Andreev, D. A Technique for Calculation of Life Limits of Electrical Network Equipment. *Energy Syst. Res.* **2019**, *2*, 73–78. [[CrossRef](#)]
29. Nazarychev, A.; Andreev, D.; Tadjibaev, A.; Vysogorets, S.; Sulynenkov, I. Methods for calculation of the marginal exploitation lifespan of power transformers 35 rV and higher based on the state index. In Proceedings of the Rudenko International Conference “Methodological Problems in Reliability Study of Large Energy Systems” (RSES 2018), Irkutsk, Russia, 2–7 July 2018; Volume 58, p. 02006. [[CrossRef](#)]
30. Plantweb Health Advisor. Product Data Sheet. Available online: <https://www.emerson.com/documents/automation/product-data-sheet-plantweb-health-advisor-plantweb-en-us-176360.pdf> (accessed on 25 October 2021).
31. Bodla, M.; Malik, S.; Rasheed, M.; Numan, M.; Ali, M.; Brima, J. Logistic regression and feature extraction based fault diagnosis of main bearing of wind turbines. In Proceedings of the IEEE 11th Conference on Industrial Electronics and Applications, Hefei, China, 5–7 June 2016; pp. 1628–1633. [[CrossRef](#)]
32. Caesarendra, W.; Widodo, A.; Yang, B. Application of relevance vector machine and logistic regression for machine degradation assessment. *Mech. Syst. Signal Process.* **2010**, *24*, 1161–1171. [[CrossRef](#)]
33. Chen, B.; Chen, X.; Li, B.; He, Z.; Cao, H.; Cai, G. Reliability estimation for cutting tools based on logistic regression model using vibration signals. *Mech. Syst. Signal Process.* **2011**, *25*, 2526–2537. [[CrossRef](#)]
34. Korshikova, A.; Trofimov, A. Model for early detection of emergency conditions in power plant equipment based on machine learning methods. *Therm. Eng.* **2019**, *66*, 189–195. [[CrossRef](#)]
35. Yang, B.; Di, X.; Han, T. Random forests classifier for machine fault diagnosis. *J. Mech. Sci. Technol.* **2008**, *22*, 1716–1725. [[CrossRef](#)]
36. Ullah, I.; Khan, R.; Yang, F.; Wuttisittikulkiij, L. Deep learning image-based defect detection in high voltage electrical equipment. *Energies* **2020**, *13*, 392. [[CrossRef](#)]
37. Wang, H.; Liu, Z. An error recognition method for power equipment defect records based on knowledge graph technology. *Front. Inform. Tech. Electron. Eng.* **2019**, *20*, 1564–1577. [[CrossRef](#)]
38. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the IEEE Proceedings of the 22nd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [[CrossRef](#)]
39. Khalyasmaa, A.; Senyuk, M.; Eroshenko, S. Analysis of the state of high-voltage current transformers based on gradient boosting on decision trees. *IEEE Trans. Power Deliv.* **2020**, *36*, 2154–2163. [[CrossRef](#)]
40. Khalyasmaa, A.; Eroshenko, S.; Shatunova, D.; Larionova, A.; Egorov, A. Digital twin technology as an instrument for increasing electrical equipment reliability. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2020; Volume 836, p. 012005. [[CrossRef](#)]
41. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wires Data Min. Knowl. Discov.* **2018**, *8*, e1249. [[CrossRef](#)]
42. Shi, J.; Yu, T.; Goebel, K.; Wu, D. Remaining useful life prediction of bearings using ensemble learning: The impact of diversity in base learners and features. *J. Comput. Inf. Sci. Eng.* **2021**, *21*, 021004. [[CrossRef](#)]
43. Beretta, M.; Julian, A.; Sepulveda, J.; Cusido, J.; Porro, O. An Ensemble Learning Solution for Predictive Maintenance of Wind Turbines Main Bearing. *Sensors* **2021**, *21*, 1512. [[CrossRef](#)] [[PubMed](#)]
44. Fushiki, T. Estimation of prediction error by using K-fold cross-validation. *Stat. Comput.* **2011**, *21*, 137–146. [[CrossRef](#)]
45. Jiang, G.; Wang, W. Error estimation based on variance analysis of k-fold cross-validation. *Pattern Recogn.* **2017**, *69*, 94–106. [[CrossRef](#)]
46. Gorban, A. Approximation of continuous functions of several variables by an arbitrary nonlinear continuous function of one variable, linear functions, and their superpositions. *Appl. Math. Lett.* **1998**, *11*, 45–49. [[CrossRef](#)]

47. Arnold, V. On the representation of functions of several variables as a superposition of functions of a smaller number of variables. *Collect. Work. Represent. Funct. Celest. Mech. KAM Theory* **2009**, *1*, 25–46. [[CrossRef](#)]
48. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [[CrossRef](#)]
49. Hornik, K. Some new results on neural network approximation. *Neural Netw.* **1993**, *6*, 1069–1072. [[CrossRef](#)]
50. Protalinsky, O.; Shcherbatov, I.; Stepanov, P. Identification of the actual state and entity availability forecasting in power engineering using neural-network technologies. *J. Phys. Conf. Ser.* **2017**, *891*, 012289. [[CrossRef](#)]
51. Haykin, S. *Kalman Filtering and Neural Networks*; John Wiley & Sons: New York, NY, USA, 2004; 284p.
52. Schafer, A.; Zimmermann, H. Recurrent neural networks are universal approximators. *Int. J. Neural Syst.* **2007**, *17*, 253–263. [[CrossRef](#)]
53. Andryushin, A.; Arakelyan, E.; Shcherbatov, I.; Kosoy, A.; Dolbikova, N. Application of neural network technologies in power engineering. *J. Phys. Conf. Ser.* **2019**, *1370*, 012054. [[CrossRef](#)]
54. Bianchi, F.; Maiorino, E.; Kampffmeyer, M.; Rizzi, A.; Jenssen, R. *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*; Springer: Cham, Switzerland, 2017; 72p.
55. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
56. Zheng, S.; Ristovski, K.; Farahat, A.; Gupta, C. Long short-term memory network for remaining useful life estimation. In Proceedings of the IEEE International Conference on Prognostics and Health Management (ICPHM), Dallas, TX, USA, 19–21 June 2017; pp. 88–95. [[CrossRef](#)]
57. Delgado, I.; Fahim, M. Wind Turbine Data Analysis and LSTM-Based Prediction in SCADA System. *Energies* **2021**, *14*, 125. [[CrossRef](#)]