Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

Article

Wavelet Long Short-Term Memory to Fault Forecasting in Electrical Power Grids

Nathielle Waldrigues Branco [1]*, Mariana Santos Matos Cavalca [1], Stefano Frizzo Stefenon [2,3], and Valderi Reis Quietinho Leithardt [4,5]

[1] Department of Electrical Engineering, Santa Catarina State University, R. Paulo Malschitzki 200, 89219-710 Joinville, Brazil; (email: mariana.cavalca@udesc.br)
[2] Fondazione Bruno Kessler. Via Sommarive 18, 38123 Trento, Italy; (email: sfrizzostefenon@fbk.eu)
[3] Department of Mathematics, Informatics and Physical Sciences, University of Udine. Via delle Scienze 206, 33100 Udine, Italy
[4] COPELABS, Lusófona University of Humanities and Technologies. Campo Grande 376, 1749-024 Lisboa, Portugal; (email: valderi@ipportalegre.pt)
[5] VALORIZA, Research Center for Endogenous Resources Valorization. Instituto Politécnico de Portalegre, 7300-555 Portalegre, Portugal
* Correspondence: nathi@uniplaclages.edu.br

Abstract: The electric power distribution utility is responsible for providing energy to consumers in a continuous and stable way, failures in the electrical power system reduce the reliability indexes of the grid, directly harming its performance. For this reason, there is a need for failure prediction to reestablish power in the shortest possible time. Considering an evaluation of the number of failures over time, this paper proposes to perform a failure prediction during the first year of the pandemic in Brazil (2020) to verify the feasibility of using time series forecasting models for fault prediction. The Long Short-Term Memory (LSTM) model will be evaluated to obtain a forecast result that can be used by the electric power utility to organize the maintenance teams. The Wavelet transform shows to be promising in improving the predictive ability of the LSTM, making the Wavelet LSTM model suitable for the study at hand. The results show that the proposed approach has better results regarding the evaluation of the error in prediction and has robustness when a statistical analysis is performed.

Keywords: Electrical Power Grids, Fault Forecasting, Long Short-Term Memory, Time Series Forecasting, Wavelet Transform.

1. Introduction

For the electricity reach the consumers in a stable and continuous way, the electrical power grid must be working independently of the weather conditions [1]. To keep the electrical distribution system running, it is necessary to evaluate the performance of the electrical system's equipment through simulation, then disturbance conditions present in the electrical power grid can be identified [2]. Disturbances that occur in the electrical power system can significantly affect the power supply, variations in voltage level, increased surface conductivity, or contact of conductors with the ground can result in faults, which affect power quality [3].

Time series forecasting can be used to identify the possibility of a failure occurring, which is a promising way to assist the decision-making process for maintenance teams in an electric power utility [4]. As the increase in failures has a strong relationship with weather conditions, in rainy seasons there is a greater chance of a failure occurring, so the study of this variation in relation to a time series is an important aspect in this context [5].

The use of Wavelet transform for noise reduction is an approach that is effective when there is high nonlinearity in the time series [6], using high-frequency bandwidth filters there may be a loss of information considering that a high frequency might be related to the occurrence of a failure. Considering that the Wavelet transform evaluates the signal energy, high frequencies are not totally eliminated, thus maintaining the main signal characteristics [7]. Thus a hybrid approach that combines a deep learning model with Wavelet transform can be an interesting approach [8].

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

Long Short-Term Memory (LSTM) is a model applied in deep learning that has been widely used by researchers for time series forecasting [9–11], its units solve the vanishing gradient problem partially since the LSTM units allow the gradients to flow unchanged [12]. Based on the advantages of the Wavelet transform and the promising capabilities of LSTM [13–16], this work proposes to use a combination of those techniques in a method named Wavelet LSTM. For this purpose, a study will be conducted using the alarm data obtained from a re-closer of a power utility company in the *Serrana* region of Santa Catarina, Brazil.

The main contributions of this research are:

- Proposal of a hybrid Wavelet LSTM model, which has higher predictive capacity than the standard LSTM model. The Wavelet LSTM shows to be a more stable model for time series prediction that can be used in several applications.
- Evaluate a time series regarding the variation of the number of failures in distribution networks with bare cables, due to the presence of contamination and contact of foreign materials with the grid, resulting in disruptive discharges in the power grid.
- Presentation of a solution for evaluating the failure history based on a time series that can be used in other works in which it is necessary to evaluate the number of failures over time.

The continuation of this paper is organized as follows: Section 2 presents a review of related works and the used data. Section 3 presents the proposed method. In Section 4 the results are analyzed. Section 5 presents a conclusion and a discussion of possible future works.

2. Related Work

In electrical distribution systems, an electrical fault is defined as an anomaly in a particular equipment causing a forced interruption in the operation of the electrical power grid [17]. There are two classes of faults: transient and permanent. Transient faults are anomalies of short duration that disappear soon after the action of protective devices, having as common causes atmospheric discharges, momentary contacts between conductors and ground, opening of an electric arc and materials without adequate insulation. Permanent faults are faults that continue to exist until it is possible to replace the defective component or equipment [18].

Through fault diagnosis it is possible to detect where the fault occurred, its size, duration, and impact on the electrical power system [19]. Among the most current fault diagnosis methods are Bayesian Networks [20], Fuzzy Logic [21], Kalman Filter [22], and other mathematical models based on artificial intelligence. The use of artificial intelligence techniques for fault identification has been growing over the years, becoming nowadays a hot topic, especially for the electric power system [23]. Deep learning models have being increasingly used to improve the ability to identify faults in the electrical grid [24–26]. However, as these models have a large number of layers, they require more computational effort, making the choice of the appropriate model a challenge [27]. From the image processing of failed components, it is possible to identify patterns and thus improve their identification in the field [28]. Several researchers are using object detection and image classification based on convolutional neural networks (CNNs) models [29–31]. The CNNs can be specially applied to improve the ability to identify faulty components, as shown by Liu *et al.* [32] and Sadykova *et al.* [33] using You Only Look Once (YOLO), Li *et al.* [34] with an improved Faster R-CNN, and Wen *et al.* [35] using Exact R-CNN. As presented by Sadykova *et.* [33] the YOLO model is a promising alternative to identify insulators during power grid inspections, being able to handle large datasets, where data augmentation techniques can also be applied to avoid early overfitting. In this context, CNN super-resolution can perform the reconstruction of the blurred images to perform the expansion of the dataset [36].

The YOLO model has been updated and variations in its structure can result in significant performance improvements. According to Liu *et al.* [32] the YOLOv3-dense

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

model proposed by them reaches up to 94.47% for insulator identification using varied image backgrounds, in comparison for the same dataset YOLOv3 reaches 90.31% accuracy. Previous versions such as YOLOv2 reach a maximum of 83.43%, making it clear that in many situations the use of non-standard models can be a promising alternative.

Variations of the YOLO model have proven to be very efficient for locating insulators on transmission lines. According to Liu *et al.* [37], MTI-YOLO has a higher average precision than YOLO-tiny and YOLO-v2. Liu *et al.* [38] proposed an improved YOLOv3 model that is better than YOLOv3 and YOLOv3-dense models. Hu and Zhou [39] show that YOLOv4 can reach an accuracy of 96.2% for insulator defect detection, also using YOLOv4 Xing and Chen [40] had a precision of 97.78% for insulator identification.

When the distribution power system does not have insulation on the medium voltage conductors, trees might touch the conductors resulting in discharges to the ground [41]. This type of fault is common in rural power grids that are close to wooded areas. To prevent these faults, the electric power utility performs pruning of trees that are close to the network, thus reducing the chance of discharges to the ground [42].

Insulators installed outdoors are exposed to environmental variations, such as dust accumulation on their surface [43]. When contamination accumulates on insulators their surface conductivity increases, generating leakage current until a discharge occurs [44], when there is high humidity in the air the conductivity increases even more, consequently increasing the chance of faults in the grid [45]. One type of contamination that has a significant impact on the conductivity of insulators is salt contamination, which can be measured by the equivalent salt deposit density [46].

Considering all these kind of possible faults [47], time series forecasting comes as an alternative to prepare maintenance teams in advance to an event based the historical knowledge of data variation over time [48]. A forecast with many steps ahead is challenging as each step ahead contains the accumulated forecast error of the previous step [49], so time series forecasting needs to take into account how many steps ahead can be considered to obtain acceptable assertiveness [50].

Among algorithms for time series forecasting, ensemble learning models in general have high performance and lower computational effort [51], and may be promising approaches for failure prediction. Various ways of combining the weak learners can be used to create a model that has greater capacity, such as bagging, boosting, and stacking [52]. Further optimized models, such as the bayesian optimization-based dynamic ensemble proposed by Du *et al.* [53] can be used, and are even applied with nonlinear data [54].

Many variations of ensemble models for time series forecasting can be found such as efficient bootstrap stacking presented by Ribeiro *et al.* [55], extreme gradient boosting proposed by Sauer *et al.* [56]. Especially for power system failure prediction, the Wavelet transform combined with ensemble models becomes a superior approach to well-established models such as the adaptive neuro-fuzzy inference system [57]. Therefore, ensemble models are successful approaches for multi-step forward prediction [58], which is equivalent to what is being evaluated in this paper.

Due to the existing features in the structure of the LSTM, it is one of the best qualified model to handle chaotic time series, since it has the ability to remember distant values and interpret order of dependencies, which are essential characteristics for a prediction models. Abbasimehr and Paki [59] used the attention mechanism to have a enhanced LSTM model. Related to the power system using LSTM, Guo *et al.* [60] and Ko *et al.* [61] presented a research about wind power forecasting. Specially for fault prediction, Guo *et al.* [62] proposed an modified LSTM version to improve safe and reliable operation of mechanical equipments.

2.1. Faults

The vast majority of faults that occur in the electrical power distribution system with naked cable are caused by direct contact with the network, this occurs mainly when the weather conditions are bad (rain and intense wind), increasing the likelihood of the contact

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

of trees with the power grid. Another failure that can occur frequently is when insulators lose their insulating capacity due to contamination or when the insulators are damaged.

To perform the time series evaluation all failures that occurred on the same day are added up, to obtain a daily failure rate over time and thus evaluate the influence of the change of season in relation to the increase of failures in the electrical power grid. These failures are evaluated in relation to the alarms registered by the electric power utility company during the evaluated period, some examples of alarms are presented in Table 1.

Table 1. Example of alarms that have been registered in the considered period.

Day	Time	Failure Record
06/01/2020	11:09:41	Current Phase B
06/01/2020	11:09:50	Current Phase A
06/01/2020	17:10:32	Current Phase C
26/01/2020	13:57:37	Recloser Communic. Failure
06/04/2020	10:04:23	Relay 50/51 (Neutral)
01/06/2020	17:24:51	Current Phase A
30/06/2020	14:11:56	Phase Voltage C
27/08/2020	10:00:54	Neutral Protection
27/08/2020	11:58:48	Current Phase C
11/09/2020	03:06:12	Current Phase A
29/12/2020	13:56:32	Relay 50/51 (Phase A)

Since failures generally occur in a non-linear pattern, this evaluation is based on a statistical analysis, and it is not possible to determine exactly when a failure will occur, however it is possible to evaluate in which period of the year there is a greater chance of the highest number of failures occurring.

In this paper, the evaluation of the history of recorded faults is in relation to the year 2020 (from January 1 to December 31), this history corresponds to the sum of all the faults of the distribution branches in the Lages region. In total there are 366 days recorded considering that the year 2020 was a leap year, Figure 1 presents the sum of the alarms regarding faults per day in this period.

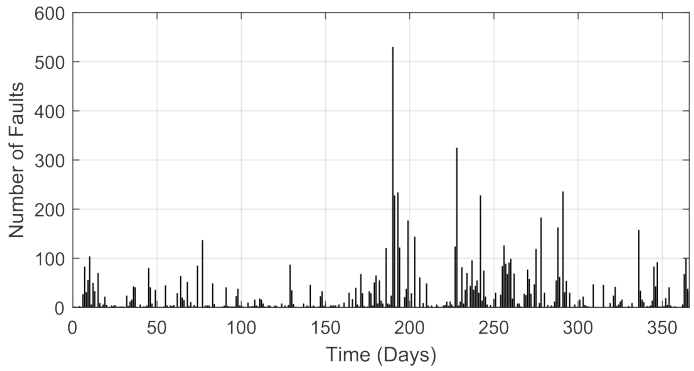

Figure 1. Failures registered in the power grid in 2020 (Lages region), data provided by *Centrais Elétricas de Santa Catarina* (CELESC).

3. Wavelet LSTM

The Wavelet LSTM method is a combination of the Wavelet transform and the Long Short-Term Memory. This approach has been widely used for fault diagnosis, as presented in the work of Sabir *et al.* [63] and Jalayer, Orsenigo, and Vercellis [64] for electrical machines, and specially for rolling bearing in the work of Tan *et al.* [65].

To apply the Wavelet LSTM here, initially the time series passes through the Wavelet filter to reduce noise and non-linearities, after the signal is decomposed and reconstructed, the LSTM receives the filtered signal and performs the prediction. The complete structure

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

of this approach is presented in Figure 2 and will be explained in this section. The values of the original signal (shown in Figure 1) are transformed into an initial time series, that is used as input to the model.

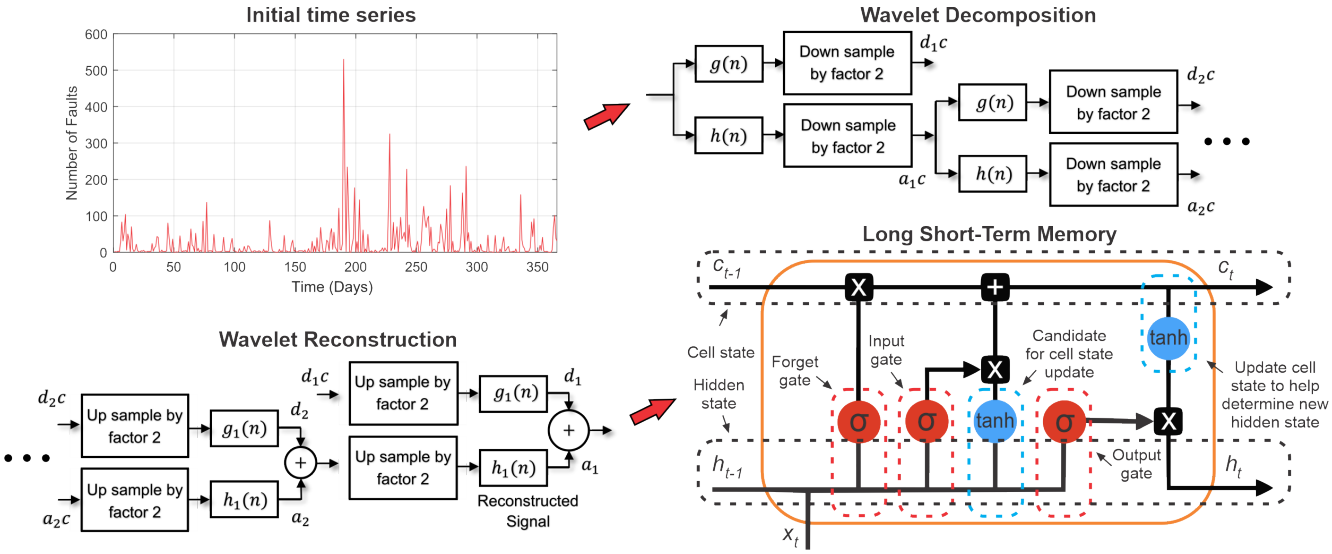

Figure 2. Structure of the Wavelet Long Short-Term Memory model.

To use the Wavelet transform first the signal is decomposed using the Wavelet packet method. The decomposition may be denoted by:

$$W_{\Psi,x}(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\Psi * \left(\frac{t-b}{a}\right) dt,\ a \neq 0 \tag{1}$$

where $x(t)$ is the signal to be decomposed, Ψ is the time-based function (mother Wavelet), a and b are the scale and the displacement parameters respectively [66]. Given a discretization the high-pass filter $g(n)$ is:

$$g(n) = h(2N - 1 - n). \tag{2}$$

where $h(n)$ is the low-pass filter. Thereby, the mother Wavelet and the scaling function (Φ) are given by:

$$\Psi(n) = \sum_{i=0}^{N-1} g(i)\Phi(2n - i), \tag{3}$$

$$\Phi(n) = \sum_{i=0}^{N-1} h(i)\Phi(2n - i). \tag{4}$$

After reconstructing the filtered signal, a time series is obtained that is used for LSTM forecast evaluation.

LSTM is a recurrent neural network, that has feedback allowing the model to remember distant values. For the time series forecasting starting from D samples,

$$x(t - (D-1)\Delta), \ldots, x(t - \Delta), x(t) \tag{5}$$

to predict future value,

$$x(t + P), \tag{6}$$

where P are the steps forward, and Δ is the period of the samples. In this paper, Δ is equal to one day, where all faults of the same day were summed.

The LSTM is capable of understanding order dependence in problems that require sequence prediction, making it promising for time series forecasting [67]. In an LSTM algorithm, each cell is divided into three gates, the input (i_t), output (o_t) and forgetting

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

(f_t) gates [68]. The f_t controls how much information will be forgotten and how much will be remembered. The useful information for the states is added through the i_t, and o_t determines how much of the current state must be assigned to the output [69]. The LSTM can be defined by the following equations:

$$\begin{aligned} i_t &= \sigma_g(W_i x_t + R_i h_{t-1} + b_i), \\ f_t &= \sigma_g(W_f x_t + R_f h_{t-1} + b_f), \\ o_t &= \sigma_g(W_o x_t + R_o h_{t-1} + b_o). \end{aligned} \tag{7}$$

in which b is the polarization matrix, R and W are earnings matrices, and σ_g is the activation function [70]. The LSTM has input activation function G and output activation function H, that are used to update the cell and the hidden state, as given in the equations:

$$\begin{aligned} \tilde{c}_t &= G(W_c x_t + R_c h_{t-1} + b_c), \\ c_t &= f_t \circ c_{t-1} + i_t \circ \tilde{c}, \\ h_t &= o_t \circ H(c_t). \end{aligned} \tag{8}$$

To perform the predicted values of future time steps, the training responses sequences are shifted by a time step. Thus, at each time-step in the input sequence, the net learns to forecast the following time-step value.

3.1. Considered Measures

In this paper the root-mean-square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) were considered, given by:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{9}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}, \tag{11}$$

where y_i is the observed value, $\hat{y}_i$ is the predicted output, and $\bar{y}_i$ is the average of the observed value [71]. The final statistical evaluation is performed with 50 runs using the same parameters configuration, where the mean, median, and standard deviation were evaluated. The simulations were computed using an Intel Core I5-7400, 20 GB of RAM, with MATLAB software.

4. Analysis of Results

The first evaluation is regarding the analysis of the time series forecast performed in relation to the percentage of data used for training and testing of the neural network. The evaluation of this parameter is important because it can be used to define the minimum amount of data needed for training the model. The evaluation results are shown in Table 2, considering a training ratio from 50 to 90 percent. The test uses the difference equivalent percentage to complete the data set, in which the validation stage is not considered.

Table 2. Evaluating the influence of the training/testing relationship.

Train / Test	RMSE	MAE	R^2	Time (s)
50 / 50	8.02×10^{-3}	3.03×10^{-3}	0.1516	17.21
60 / 40	6.24×10^{-3}	7.47×10^{-4}	0.1681	18.53
70 / 30	4.73×10^{-3}	3.59×10^{-4}	0.0325	18.64
80 / 20	3.60×10^{-3}	1.18×10^{-3}	0.2779	20.47
90 / 10	4.29×10^{-3}	5.49×10^{-5}	0.0884	19.68

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

Using 80% of the data for training and 20% of the data for testing gives the best RMSE and R^2 values, hence this ratio was used in the further analysis. As can be seen in Figure 3, there is a major difficulty in predicting the data due to the nonlinearities in the time series, considering that in some cases there were several failures in a short period of time.

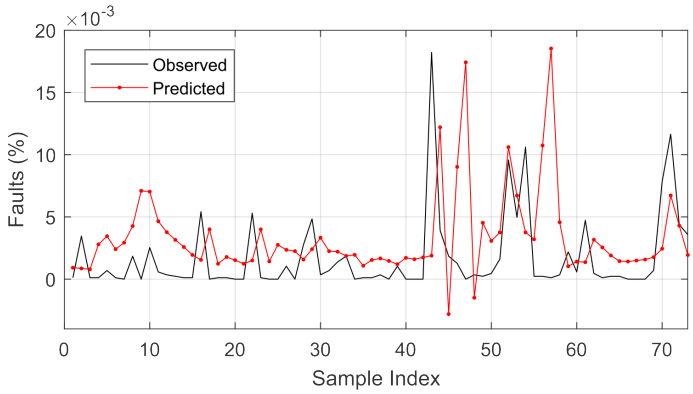

Figure 3. Preliminary analysis of fault prediction capability.

The failures that occurred after the middle of the year were due to the rainy season that starts after winter in the southern hemisphere. The greater presence of bad weather conditions favors the development of faults in the electrical power distribution system. In the following analysis, the optimizer and the number of hidden units are evaluated (see Table 3).

Table 3. Assessment of the number of hidden units (HU) using different optimizers.

Optimizer	HU	RMSE	MAE	R^2	Time (s)
SGDM	50	3.49×10^{-3}	9.12×10^{-4}	0.1957	17.74
	100	3.50×10^{-3}	9.12×10^{-4}	0.2035	18.29
	200	3.51×10^{-3}	1.17×10^{-3}	0.2089	19.58
	500	3.44×10^{-3}	9.18×10^{-4}	0.1624	25.06
	1000	3.44×10^{-3}	9.37×10^{-4}	0.1623	36.66
ADAM	50	4.79×10^{-3}	1.77×10^{-3}	-	17.94
	100	7.69×10^{-3}	4.58×10^{-3}	-	19.94
	200	3.98×10^{-3}	6.51×10^{-4}	0.5554	19.21
	500	3.96×10^{-3}	8.29×10^{-4}	0.5469	25.55
	1000	3.94×10^{-3}	1.52×10^{-3}	0.5242	35.22
RMSprop	50	6.22×10^{-3}	3.01×10^{-3}	-	21.58
	100	5.65×10^{-3}	1.22×10^{-3}	-	19.14
	200	3.93×10^{-3}	6.64×10^{-4}	0.5212	21.64
	500	3.35×10^{-3}	1.00×10^{-3}	0.1074	27.51
	1000	3.16×10^{-3}	2.96×10^{-5}	0.0190	36.50

In this evaluation, the SGDM optimizer had more stable results for the coefficient of determination, presenting a smaller variance in relation to the change of hidden units. Comparing all the models the best results occurred using 200 hidden units considering the coefficient of determination. In some cases, it was not possible to measure the coefficient of determination due to the high intensity of variation in the prediction using the ADAM and RMSprop optimizers. Considering that there was a large variation in the values, a statistical analysis was performed and will be presented using 200 hidden units.

The inclusion of the Wavelet transform for noise reduction is added for the following analysis, this transform should be used with caution as it can result in a loss of features of the signal. Figure 4 shows the result of the Wavelet transform relative to the original signal using 1 node, and Figure 5 shows this comparison using 2 nodes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

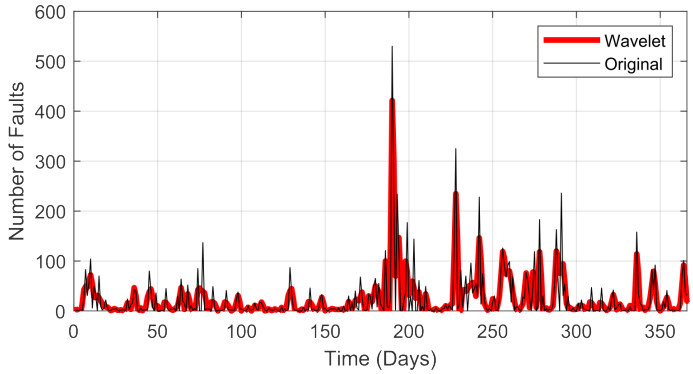

Figure 4. Evaluation of the Wavelet transform with 1 node.

The use of 2 nodes considerably alters the response of the transform, hindering a practical application. When 3 nodes or more are used, the signal loses its characteristics and is not considered in this paper.

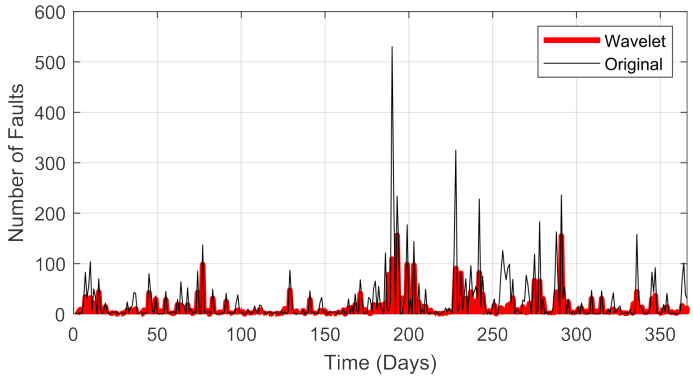

Figure 5. Evaluation of the Wavelet transform with 2 nodes.

The complete analysis of the Wavelet transform depth variation is presented in Table 3. Considering that because there is an error value that makes the prediction not suitable for analysis, the use of 2 nodes is disregarded after this evaluation, as mentioned earlier this can also be observed when the Wavelet transform was compared in this configuration in relation to the original signal (see Figure 5).

Table 4. Assessment of depth using different 1 and 2 nodes.

Nodes	Depth	RMSE	MAE	R^2	Time (s)
1	1	2.19×10^{-3}	4.54×10^{-4}	0.3375	23.14
	2	2.16×10^{-3}	4.69×10^{-4}	0.3547	22.13
	3	2.22×10^{-3}	5.28×10^{-4}	0.3132	23.61
	4	2.16×10^{-3}	3.53×10^{-4}	0.3509	22.76
	5	2.17×10^{-3}	3.70×10^{-4}	0.3493	21.71
2	1	2.22×10^{10}	1.24×10^{9}	0.6888	21.34
	2	2.13×10^{10}	1.24×10^{9}	0.7155	22.75
	3	2.45×10^{10}	6.79×10^{9}	0.6214	18.05
	4	2.32×10^{10}	1.90×10^{9}	0.6601	18.74
	5	2.44×10^{10}	6.58×10^{9}	0.6244	18.69

The best coefficient of determination was reached using a depth equal to 2 in the Wavelet transform, getting close to the best MAE value that happened using the depth equal to 4. Considering these results, the depth equal to 2 was used for statistical analysis which is presented in Table 5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

Table 5. Statistical evaluation.

Model	Optimizer	Mean	Median	Std Dev.
LSTM	SGDM	3.50×10^{-3}	3.49×10^{-3}	2.62×10^{-5}
	ADAM	7.87×10^{-3}	8.04×10^{-3}	2.63×10^{-3}
	RMSprop	4.90×10^{-3}	4.95×10^{-3}	6.41×10^{-4}
Wavelet LSTM	SGDM	2.19×10^{-3}	2.19×10^{-3}	2.76×10^{-5}
	ADAM	1.79×10^{-3}	1.27×10^{-3}	2.27×10^{-3}
	RMSprop	1.67×10^{-3}	1.54×10^{-3}	7.15×10^{-4}

The Wavelet LSTM model was superior in all comparative analyses to the LSTM model with respect to RMSE. Even varying the optimizer, the Wavelet LSTM showed promise for the analysis in question. The best average RMSE result was obtained using the RMSprop optimizer in the Wavelet LSTM model. The comparison between the prediction result and the original signal is presented in Figure 6.

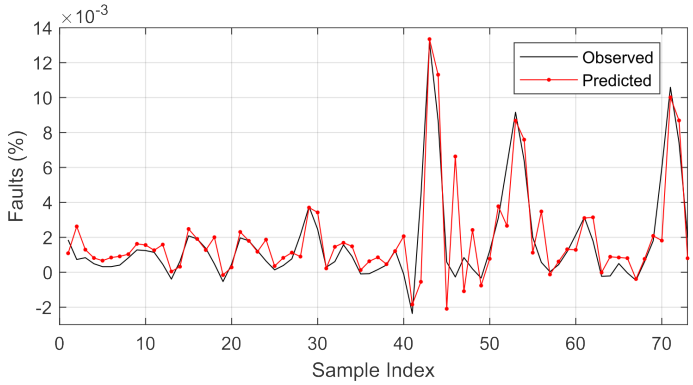

Figure 6. Comparison of the prediction using the Wavelet LSTM model to the observed values.

The prediction results using the Wavelet LSTM in comparison to the original signal showed that this model is suitable for the application in question. As can be seen, after the highest accumulated value of the number of failures there was an oscillation in the prediction, something that is expected due to this abrupt variation in the time series.

5. Conclusion

The prediction of faults in the electrical distribution system is necessary to ensure the operation of the power grid. Analyzing the variation of a time series it is possible to verify the presence of a higher number of failures during a certain time and thus define a more effective correction strategy. Based on a time series forecasting, the electric utility may know when there are higher chances to have faults before it happened, and so have a better define a strategy to deal with that.

It is noticeable that there is a difficulty in this prediction due to the large variation in the number of failures in some seasons of the year, mainly related to the rainy season. Using traditional models the forecast results are ineffective, so it is necessary to combine algorithms to create a hybrid model to meet the needs of the problem.

The Wavelet LSTM model showed better results in all analyses compared to the standard LSTM model, including better results in the statistical analysis, being an appropriate model for the evaluation presented in this paper. Using this model it is possible to have failure prediction indicators that can help the organization of maintenance teams, thus reducing the response time when a disruptive failure occurs.

Future work can be done regarding the type of failure, the failures can vary, for instance, as a result of direct contact with the grid and/or leakage current. Specific analysis on which type of failure that occurs more frequently and how to avoid this type of failure is promising work to be done in the future.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

Author Contributions: Writing—original draft, methodology, Nathielle Waldrigues Branco; Writing—review and editing, supervision, Mariana Santos Matos Cavalca; Writing—review and editing, software, formal analysis, Stefano Frizzo Stefenon; Project administration, supervision, Valderi Reis Quietinho Leithardt.

Funding: This work was supported by the National Funds through the Fundação para a Ciência e a Tecnologia, I.P. (Portuguese Foundation for Science and Technology) by the Project "VALORIZA—Research Centre for Endogenous Resource Valorization" under Grant UIDB/05064/2020 and Grant UIDB/04111/2020, and in part by the Instituto Lusófono de Investigação e Desenvolvimento (ILIND) under Project COFAC/ILIND/COPELABS/3/2020.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this paper were provided by *Centrais Elétricas de Santa Catarina*, regarding the alarms of the power distribution grids in Lages region, Brazil, from January 1 to December 31 of 2020. These records are available at: https://github.com/SFStefenon/FailuresPowerGrid2020 (Accessed on September 21, 2021).

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

References

1. Araya, J.; Montaña, J.; Schurch, R. Electric Field Distribution and Leakage Currents in Glass Insulator Under Different Altitudes and Pollutions Conditions using FEM Simulations. *IEEE Latin America Transactions* **2021**, *19*, 1278–1285. doi:10.1109/TLA.2021.9475858.
2. Sun, J.; Yang, Q.; Cui, H.; Ran, J.; Liu, H. Distribution Line Fault Location With Unknown Fault Impedance Based on Electromagnetic Time Reversal. *IEEE Transactions on Electromagnetic Compatibility* **2021**, *63*, 1921–1929. doi:10.1109/TEMC.2021.3097105.
3. Liu, Z.; Chen, H.; Hu, Z.; Li, Y.; Wu, X.; Peng, H. Fault Detection System for 500 kV AC Fault Current Limiter Based on High-Coupled Split Reactor. *IEEE Transactions on Applied Superconductivity* **2021**, *31*, 1–7. doi:10.1109/TASC.2021.3101745.
4. Stefenon, S.F.; Bruns, R.; Sartori, A.; Meyer, L.H.; Ovejero, R.G.; Leithardt, V.R.Q. Analysis of the Ultrasonic Signal in Polymeric Contaminated Insulators Through Ensemble Learning Methods. *IEEE Access* **2022**, *10*, 33980–33991. doi:10.1109/ACCESS.2022.3161506.
5. Medeiros, A.; Sartori, A.; Stefenon, S.F.; Meyer, L.H.; Nied, A. Comparison of artificial intelligence techniques to failure prediction in contaminated insulators based on leakage current. *Journal of Intelligent & Fuzzy Systems* **2021**, *42*, 3285–3298. doi:10.3233/JIFS-211126.
6. Rhif, M.; Ben Abbes, A.; Farah, I.R.; Martínez, B.; Sang, Y. Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review. *Applied Sciences* **2019**, *9*, 1345. doi:10.3390/app9071345.
7. Ameid, T.; Menacer, A.; Talhaoui, H.; Azzoug, Y. Discrete wavelet transform and energy eigen value for rotor bars fault detection in variable speed field-oriented control of induction motor drive. *ISA Transactions* **2018**, *79*, 217–231. doi:10.1016/j.isatra.2018.04.019.
8. Stefenon, S.F.; Kasburg, C.; Nied, A.; Klaar, A.C.R.; Ferreira, F.C.S.; Branco, N.W. Hybrid deep learning for power generation forecasting in active solar trackers. *IET Generation, Transmission & Distribution* **2020**, *14*, 5667–5674. doi:10.1049/iet-gtd.2020.0814.
9. Chandra, R.; Goyal, S.; Gupta, R. Evaluation of Deep Learning Models for Multi-Step Ahead Time Series Prediction. *IEEE Access* **2021**, *9*, 83105–83123. doi:10.1109/ACCESS.2021.3085085.
10. Hu, Y.; Sun, X.; Nie, X.; Li, Y.; Liu, L. An Enhanced LSTM for Trend Following of Time Series. *IEEE Access* **2019**, *7*, 34020–34030. doi:10.1109/ACCESS.2019.2896621.
11. Ma, C.; Dai, G.; Zhou, J. Short-Term Traffic Flow Prediction for Urban Road Sections Based on Time Series Analysis and LSTM BILSTM Method. *IEEE Transactions on Intelligent Transportation Systems* **2022**, *23*, 5615–5624. doi:10.1109/TITS.2021.3055258.
12. Zhang, S.; Wang, Y.; Liu, M.; Bao, Z. Data-Based Line Trip Fault Prediction in Power Systems Using LSTM Networks and SVM. *IEEE Access* **2018**, *6*, 7675–7686. doi:10.1109/ACCESS.2017.2785763.
13. Kim, W.H.; Kim, J.Y.; Chae, W.K.; Kim, G.; Lee, C.K. LSTM-Based Fault Direction Estimation and Protection Coordination for Networked Distribution System. *IEEE Access* **2022**, *10*, 40348–40357. doi:10.1109/ACCESS.2022.3166836.
14. Qiao, M.; Yan, S.; Tang, X.; Xu, C. Deep Convolutional and LSTM Recurrent Neural Networks for Rolling Bearing Fault Diagnosis Under Strong Noises and Variable Loads. *IEEE Access* **2020**, *8*, 66257–66269. doi:10.1109/ACCESS.2020.2985617.
15. Stefenon, S.F.; Freire, R.Z.; Meyer, L.H.; Corso, M.P.; Sartori, A.; Nied, A.; Klaar, A.C.R.; Yow, K.C. Fault detection in insulators based on ultrasonic signal processing using a hybrid deep learning technique. *IET Science, Measurement & Technology* **2020**, *14*, 953–961. doi:10.1049/iet-smt.2020.0083.
16. Ma, Y.; Oslebo, D.; Maqsood, A.; Corzine, K. DC Fault Detection and Pulsed Load Monitoring Using Wavelet Transform-Fed LSTM Autoencoders. *IEEE Journal of Emerging and Selected Topics in Power Electronics* **2021**, *9*, 7078–7087. doi:10.1109/JESTPE.2020.3019382.
17. Furse, C.M.; Kafal, M.; Razzaghi, R.; Shin, Y.J. Fault Diagnosis for Electrical Systems and Power Networks: A Review. *IEEE Sensors Journal* **2021**, *21*, 888–906. doi:10.1109/JSEN.2020.2987321.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

18. Li, B.; Cui, H.; Li, B.; Wen, W.; Dai, D. A permanent fault identification method for single-pole grounding fault of overhead transmission lines in VSC-HVDC grid based on fault line voltage. *International Journal of Electrical Power & Energy Systems* **2020**, *117*, 105603. doi:10.1016/j.ijepes.2019.105603.
19. Wadi, M.; Elmasry, W. An Anomaly-based Technique for Fault Detection in Power System Networks. In Proceedings of the 2021 International Conference on Electric Power Engineering; ICEPE: Palestine, 2021; pp. 1–6. doi:10.1109/ICEPE-P51568.2021.9423479.
20. Wu, J.; Zhang, L.; Bai, Y.; Reniers, G. A safety investment optimization model for power grid enterprises based on System Dynamics and Bayesian network theory. *Reliability Engineering & System Safety* **2022**, *221*, 108331. doi:10.1016/j.ress.2022.108331.
21. Sadi, M.A.H.; AbuHussein, A.; Shoeb, M.A. Transient Performance Improvement of Power Systems Using Fuzzy Logic Controlled Capacitive-Bridge Type Fault Current Limiter. *IEEE Transactions on Power Systems* **2021**, *36*, 323–335. doi:10.1109/TPWRS.2020.3003294.
22. Rigatos, G.; Serpanos, D.; Zervos, N. Detection of Attacks Against Power Grid Sensors Using Kalman Filter and Statistical Decision Making. *IEEE Sensors Journal* **2017**, *17*, 7641–7648. doi:10.1109/JSEN.2017.2661247.
23. Sopelsa Neto, N.F.; Stefenon, S.F.; Meyer, L.H.; Bruns, R.; Nied, A.; Seman, L.O.; Gonzalez, G.V.; Leithardt, V.R.Q.; Yow, K.C. A Study of Multilayer Perceptron Networks Applied to Classification of Ceramic Insulators Using Ultrasound. *Applied Sciences* **2021**, *11*, 1592. doi:10.3390/app11041592.
24. Haj, Y.E.; El-Hag, A.H.; Ghunem, R.A. Application of Deep-Learning via Transfer Learning to Evaluate Silicone Rubber Material Surface Erosion. *IEEE Transactions on Dielectrics and Electrical Insulation* **2021**, *28*, 1465–1467. doi:10.1109/TDEI.2021.009617.
25. Stefenon, S.F.; Singh, G.; Yow, K.C.; Cimatti, A. Semi-ProtoPNet Deep Neural Network for the Classification of Defective Power Grid Distribution Structures. *Sensors* **2022**, *22*, 4859. doi:10.3390/s22134859.
26. Zhao, M.; Barati, M. A Real-Time Fault Localization in Power Distribution Grid for Wildfire Detection Through Deep Convolutional Neural Networks. *IEEE Transactions on Industry Applications* **2021**, *57*, 4316–4326. doi:10.1109/TIA.2021.3083645.
27. Mantach, S.; Lutfi, A.; Moradi Tavasani, H.; Ashraf, A.; El-Hag, A.; Kordi, B. Deep Learning in High Voltage Engineering: A Literature Review. *Energies* **2022**, *15*, 5005. doi:10.3390/en15145005.
28. Corso, M.P.; Perez, F.L.; Stefenon, S.F.; Yow, K.C.; García Ovejero, R.; Leithardt, V.R.Q. Classification of Contaminated Insulators Using k-Nearest Neighbors Based on Computer Vision. *Computers* **2021**, *10*, 112. doi:10.3390/computers10090112.
29. Wu, H.; Hu, Y.; Wang, W.; Mei, X.; Xian, J. Ship Fire Detection Based on an Improved YOLO Algorithm with a Lightweight Convolutional Neural Network Model. *Sensors* **2022**, *22*, 7420. doi:10.3390/s22197420.
30. Vieira, J.C.; Sartori, A.; Stefenon, S.F.; Perez, F.L.; de Jesus, G.S.; Leithardt, V.R.Q. Low-Cost CNN for Automatic Violence Recognition on Embedded System. *IEEE Access* **2022**, *10*, 25190–25202. doi:10.1109/ACCESS.2022.3155123.
31. Hou, L.; Chen, C.; Wang, S.; Wu, Y.; Chen, X. Multi-Object Detection Method in Construction Machinery Swarm Operations Based on the Improved YOLOv4 Model. *Sensors* **2022**, *22*, 7294. doi:10.3390/s22197294.
32. Liu, C.; Wu, Y.; Liu, J.; Sun, Z. Improved YOLOv3 Network for Insulator Detection in Aerial Images with Diverse Background Interference. *Electronics* **2021**, *10*, 771. doi:10.3390/electronics10070771.
33. Sadykova, D.; Pernebayeva, D.; Bagheri, M.; James, A. IN-YOLO: Real-Time Detection of Outdoor High Voltage Insulators Using UAV Imaging. *IEEE Transactions on Power Delivery* **2020**, *35*, 1599–1601. doi:10.1109/TPWRD.2019.2944741.
34. Li, X.; Su, H.; Liu, G. Insulator Defect Recognition Based on Global Detection and Local Segmentation. *IEEE Access* **2020**, *8*, 59934–59946. doi:10.1109/ACCESS.2020.2982288.
35. Wen, Q.; Luo, Z.; Chen, R.; Yang, Y.; Li, G. Deep Learning Approaches on Defect Detection in High Resolution Aerial Images of Insulators. *Sensors* **2021**, *21*, 1033. doi:10.3390/s21041033.
36. Chen, H.; He, Z.; Shi, B.; Zhong, T. Research on Recognition Method of Electrical Components Based on YOLO V3. *IEEE Access* **2019**, *7*, 157818–157829. doi:10.1109/ACCESS.2019.2950053.
37. Liu, C.; Wu, Y.; Liu, J.; Han, J. MTI-YOLO: A Light-Weight and Real-Time Deep Neural Network for Insulator Detection in Complex Aerial Images. *Energies* **2021**, *14*, 1426. doi:10.3390/en14051426.
38. Liu, J.; Liu, C.; Wu, Y.; Xu, H.; Sun, Z. An Improved Method Based on Deep Learning for Insulator Fault Detection in Diverse Aerial Images. *Energies* **2021**, *14*, 4365. doi:10.3390/en14144365.
39. Hu, X.; Zhou, Y. Insulator defect detection in power inspection image using focal loss based on YOLO v4. In Proceedings of the International Conference on Artificial Intelligence, Virtual Reality, and Visualization (AIVRV 2021). SPIE, 2021, Vol. 12153, pp. 90 – 95. doi:10.1117/12.2626694.
40. Xing, Z.; Chen, X. Lightweight algorithm of insulator identification applicable to electric power engineering. *Energy Reports* **2022**, *8*, 353–362. doi:10.1016/j.egyr.2022.01.209.
41. Stefenon, S.F.; Seman, L.O.; Pavan, B.A.; Ovejero, R.G.; Leithardt, V.R.Q. Optimal design of electrical power distribution grid spacers using finite element method. *IET Generation, Transmission & Distribution* **2022**, *16*, 1865–1876. doi:10.1049/gtd2.12425.
42. Yang, C.; Chen, T.; Yang, B.; Zhang, X.; Fan, S. Experimental study of tree ground fault discharge characteristics of 35 kV transmission lines. In Proceedings of the 2021 IEEE Sustainable Power and Energy Conference (iSPEC); , 2021; Vol. 1, pp. 2883–2891. doi:10.1109/iSPEC53008.2021.9735502.
43. Stefenon, S.F.; Neto, C.S.F.; Coelho, T.S.; Nied, A.; Yamaguchi, C.K.; Yow, K.C. Particle swarm optimization for design of insulators of distribution power system based on finite element method. *Electrical Engineering* **2022**, *104*, 615–622. doi:10.1007/s00202-021-01332-3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

44. Salem, A.A.; Abd-Rahman, R.; Al-Gailani, S.A.; Kamarudin, M.S.; Ahmad, H.; Salam, Z. The Leakage Current Components as a Diagnostic Tool to Estimate Contamination Level on High Voltage Insulators. *IEEE Access* **2020**, *8*, 92514–92528. doi:10.1109/ACCESS.2020.2993630.
45. Stefenon, S.F.; Seman, L.O.; Sopelsa Neto, N.F.; Meyer, L.H.; Nied, A.; Yow, K.C. Echo state network applied for classification of medium voltage insulators. *International Journal of Electrical Power & Energy Systems* **2022**, *134*, 107336. doi:10.1016/j.ijepes.2021.107336.
46. Cao, B.; Wang, L.; Yin, F. A Low-Cost Evaluation and Correction Method for the Soluble Salt Components of the Insulator Contamination Layer. *IEEE Sensors Journal* **2019**, *19*, 5266–5273. doi:10.1109/JSEN.2019.2902192.
47. Stefenon, S.F.; Corso, M.P.; Nied, A.; Perez, F.L.; Yow, K.C.; Gonzalez, G.V.; Leithardt, V.R.Q. Classification of insulators using neural network based on computer vision. *IET Generation, Transmission & Distribution* **2021**, *16*, 1096–1107. doi:10.1049/gtd2.12353.
48. Hou, C.; Wu, J.; Cao, B.; Fan, J. A deep-learning prediction model for imbalanced time series data forecasting. *Big Data Mining and Analytics* **2021**, *4*, 266–278. doi:10.26599/BDMA.2021.9020011.
49. Taieb, S.B.; Atiya, A.F. A Bias and Variance Analysis for Multistep-Ahead Time Series Forecasting. *IEEE Transactions on Neural Networks and Learning Systems* **2016**, *27*, 62–76. doi:10.1109/TNNLS.2015.2411629.
50. Duan, J.; Kashima, H. Learning to Rank for Multi-Step Ahead Time-Series Forecasting. *IEEE Access* **2021**, *9*, 49372–49386. doi:10.1109/ACCESS.2021.3068895.
51. Stefenon, S.F.; Ribeiro, M.H.D.M.; Nied, A.; Yow, K.C.; Mariani, V.C.; Coelho, L.S.; Seman, L.O. Time series forecasting using ensemble learning methods for emergency prevention in hydroelectric power plants with dam. *Electric Power Systems Research* **2022**, *202*, 107584. doi:10.1016/j.epsr.2021.107584.
52. Ribeiro, M.H.D.M.; Coelho, L.S. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Applied Soft Computing* **2020**, *86*, 105837. doi:10.1016/j.asoc.2019.105837.
53. Du, L.; Gao, R.; Suganthan, P.N.; Wang, D.Z. Bayesian optimization based dynamic ensemble for time series forecasting. *Information Sciences* **2022**, *591*, 155–175. doi:10.1016/j.ins.2022.01.010.
54. Kim, D.; Kim, C. Forecasting time series with genetic fuzzy predictor ensemble. *IEEE Transactions on Fuzzy Systems* **1997**, *5*, 523–535. doi:10.1109/91.649903.
55. Ribeiro, M.H.D.M.; da Silva, R.G.; Moreno, S.R.; Mariani, V.C.; Coelho, L.S. Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting. *International Journal of Electrical Power & Energy Systems* **2022**, *136*, 107712. doi:10.1016/j.ijepes.2021.107712.
56. Sauer, J.; Mariani, V.C.; Coelho, L.S.; Ribeiro, M.H.D.M.; Rampazzo, M. Extreme gradient boosting model based on improved Jaya optimizer applied to forecasting energy consumption in residential buildings. *Evolving Systems* **2021**, pp. 1–12.
57. Stefenon, S.F.; Ribeiro, M.H.D.M.; Nied, A.; Mariani, V.C.; Coelho, L.D.S.; Leithardt, V.R.Q.; Silva, L.A.; Seman, L.O. Hybrid Wavelet Stacking Ensemble Model for Insulators Contamination Forecasting. *IEEE Access* **2021**, *9*, 66387–66397. doi:10.1109/ACCESS.2021.3076410.
58. da Silva, R.G.; Ribeiro, M.H.D.M.; Moreno, S.R.; Mariani, V.C.; Coelho, L.S. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. *Energy* **2021**, *216*, 119174. doi:10.1016/j.energy.2020.119174.
59. Abbasimehr, H.; Paki, R. Improving time series forecasting using LSTM and attention models. *Journal of Ambient Intelligence and Humanized Computing* **2022**, *13*, 673–691.
60. Yu, R.; Gao, J.; Yu, M.; Lu, W.; Xu, T.; Zhao, M.; Zhang, J.; Zhang, R.; Zhang, Z. LSTM-EFG for wind power forecasting based on sequential correlation features. *Future Generation Computer Systems* **2019**, *93*, 33–42. doi:10.1016/j.future.2018.09.054.
61. Ko, M.S.; Lee, K.; Kim, J.K.; Hong, C.W.; Dong, Z.Y.; Hur, K. Deep Concatenated Residual Network With Bidirectional LSTM for One-Hour-Ahead Wind Power Forecasting. *IEEE Transactions on Sustainable Energy* **2021**, *12*, 1321–1335. doi:10.1109/TSTE.2020.3043884.
62. Guo, J.; Lao, Z.; Hou, M.; Li, C.; Zhang, S. Mechanical fault time series prediction by using EFMSAE-LSTM neural network. *Measurement* **2021**, *173*, 108566. doi:10.1016/j.measurement.2020.108566.
63. Sabir, R.; Rosato, D.; Hartmann, S.; Guehmann, C. LSTM Based Bearing Fault Diagnosis of Electrical Machines using Motor Current Signal. In Proceedings of the 2019 18th IEEE International Conference On Machine Learning And Applications. ICMLA, 2019, pp. 613–618. doi:10.1109/ICMLA.2019.00113.
64. Jalayer, M.; Orsenigo, C.; Vercellis, C. Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms. *Computers in Industry* **2021**, *125*, 103378. doi:10.1016/j.compind.2020.103378.
65. Tan, W.; Sun, Y.; Qiu, D.; An, Y.; Ren, P. Rolling Bearing Fault Diagnosis Based on Single Gated Unite Recurrent Neural Networks. *Journal of Physics: Conference Series* **2020**, *1601*, 1–9. doi:10.1088/1742-6596/1601/4/042017.
66. Stefenon, S.F.; Kasburg, C.; Freire, R.Z.; Silva Ferreira, F.C.; Bertol, D.W.; Nied, A. Photovoltaic power forecasting using wavelet Neuro-Fuzzy for active solar trackers. *Journal of Intelligent & Fuzzy Systems* **2021**, *40*, 1083–1096. doi:10.3233/JIFS-201279.
67. Casado-Vara, R.; Martin del Rey, A.; Pérez-Palau, D.; de-la Fuente-Valentín, L.; Corchado, J.M. Web Traffic Time Series Forecasting Using LSTM Neural Networks with Distributed Asynchronous Training. *Mathematics* **2021**, *9*, 421. doi:10.3390/math9040421.
68. Fernandes, F.; Stefenon, S.F.; Seman, L.O.; Nied, A.; Ferreira, F.C.S.; Subtil, M.C.M.; Klaar, A.C.R.; Leithardt, V.R.Q. Long short-term memory stacking model to predict the number of cases and deaths caused by COVID-19. *Journal of Intelligent & Fuzzy Systems* **2022**, *6*, 6221–6234. doi:10.3233/JIFS-212788.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2022 doi:10.20944/preprints202210.0004.v1

69. Sagheer, A.; Kotb, M. Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing* **2019**, *323*, 203–213. doi:10.1016/j.neucom.2018.09.082.
70. Liu, Y.; Guan, L.; Hou, C.; Han, H.; Liu, Z.; Sun, Y.; Zheng, M. Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform. *Applied Sciences* **2019**, *9*, 1108. doi:10.3390/app9061108.
71. Sopelsa Neto, N.F.; Stefenon, S.F.; Meyer, L.H.; Ovejero, R.G.; Leithardt, V.R.Q. Fault Prediction Based on Leakage Current in Contaminated Insulators Using Enhanced Time Series Forecasting Models. *Sensors* **2022**, *22*, 6121. doi:10.3390/s22166121.