Electric Power Systems Research 189 (2020) 106823

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

# An LSTM model for power grid loss prediction

Jarkko Tulensalo[a,b], Janne Seppänen[*,b], Alexander Ilin[a]

[a] Department of Computer Science, Aalto University, Espoo, Finland
[b] Fingrid Oyj, Helsinki, Finland

ARTICLE INFO

*Keywords:*
Power grid loss prediction
Recurrent neural networks
Long short-term memory
Deep neural networks

ABSTRACT

With the changes that renewable energy sources bring to the electricity markets in all over the world, prediction of grid losses gets more complex as current methods have limited capability to take local weather conditions into account. This paper suggests a Long Short-Term Memory (LSTM) recurrent neural network model for power grid loss prediction. The model learns long-term relations of hourly time series data from electricity markets, local weather and calendar. We apply the model to predict the total transmission grid losses in Finland. We find that the proposed model outperforms both the reference method currently used in industry and linear regression proposed by previous studies.

## 1. Introduction

Power grid losses account for 2–70 percentage of the total electricity consumption in different countries in the world [1]. When power balance responsible parties acquire the energy consumed by grid losses from the electricity markets, inaccurate grid loss prediction may lead to economical losses or non-optimal grid operation. This paper proposes to use a neural network model to predict future grid losses based on past observations and short-term forecasts of weather and electricity market variables.

The difficulty of power grid loss estimation lies in the complexity of the electrical power system and nonlinear correlations related to grid losses. The power system is also constantly changing, which makes the prediction problem even more difficult. With an increasing share of intermittent, distributed renewable sources in the electricity power systems all over the world, the relationship between power grid losses and weather conditions increase. This makes it more difficult to predict the losses with traditional rule-based models and methods.

The current standard approach in the industry is based on historical reference value, previous knowledge and available prognosis data. This was mentioned in [2] and agrees well with our experience. The industry-standard methods work well for a stable system in which all components do not change with time. However, most power systems are complex dynamical systems affected by many external factors, which makes it difficult to design heuristics that work well in different situations. Using a data-driven (machine learning) approach can be beneficial in this task.

In previous studies, power grid losses in the network have been predicted using various methods of calculating loss rates for each individual line in the system as a function of the transmitted power. The total loss was then calculated as the sum of predicted losses in all lines in the network. The proposed models of line loss calculations included heuristic methods [3], numerical methods [4], regression [5,6] and neural networks [7,8]. Although this approach can account for possible changes in the network configuration, estimating line loss rate for each of the thousands or tens of thousands of lines in a system requires very detailed and accurate information on incoming and outgoing energy flows and other local conditions [2]. Acquiring this information in practice can be tedious and expensive.

Sahlin et al. [2] proposed a method that estimates the line loss rate for a total (aggregate) grid system, which decreases the amount of detailed modelling work. They proposed a linear regression model which was designed using the domain knowledge of the line loss calculations. One potential problem with that approach is that it does not directly take local weather conditions into consideration. However, temperature, wind speed and solar radiation has a large (and complex) effect on local power generation, consumption and the dynamic behaviour of each line. Taking into account those variables can be beneficial in practice.

Neural networks, especially Long Short-Term Memory (LSTM) recurrent neural networks, have successfully been applied to predict electricity demand [9,10], electricity generation of renewable energy sources [11,12] and electricity prices [13]. The studies show that LSTM can outperform more traditional machine learning methods. However, previously there have not been studies on applying LSTM neural networks to power-grid loss predictions.

* Corresponding author.
*E-mail addresses:* Jarkko.Tulensalo@aalto.fi (J. Tulensalo), janne.seppanen@fingrid.fi (J. Seppänen), Alexander.Ilin@aalto.fi (A. Ilin).

https://doi.org/10.1016/j.epsr.2020.106823
Received 5 March 2020; Received in revised form 4 July 2020; Accepted 9 August 2020
Available online 01 September 2020

In this paper, we show that Long Short-Term Memory recurrent neural network can be applied to predict power grid losses. The proposed model can take the nonlinear effect of the local temperature, wind speed and solar radiation into account. We show that the model outperforms both the method currently used for grid loss predictions at Fingrid Oyj (Finnish transmission system operator) and linear regression method proposed by literature [2].

The rest of the paper is organised as follows. We present the data used in the study in Section 2. The proposed model as well as the methods used for comparison are explained in Section 3. We present the results in Section 4 and conclude the study in Section 5.

## 2. Data

The data was obtained from the Finnish transmission system operator Fingrid Oyj. The data consists of historical, hourly electricity market data, local weather data, calendar data and data specific to the studied power grid system. In total, we selected 54 variables with a long enough availability period and sufficient data quality.

The electricity market data includes electricity demand, wind power generation and physical flows from Sweden price area SE1, Sweden price area SE3, Estonia and Russia [14]. The weather data consists of historical measurements of the air temperature, wind speed and solar radiation in different locations in Finland: Vantaa, Peipohja, Maaninka, Sodankylä and Siikajärvi for the air temperature, Vantaa and Sodankylä for the wind speed, and Vantaa, Peipohja, Sodankylä and Siikajärvi for the solar radiation.

The calendar data consists of one-hot coded indicators of weekdays, public holidays and years and a two-dimensional sine-cosine representation of the daily cycle.

The power-grid data consists of the aggregate grid loss computed with a 24-h lag, a moving 24-h average of the aggregate grid loss also computed with a delay of 24 h and the electricity flow between northern and southern parts of Finland with a lag of 2 h. These variables were chosen based on the data availability in real-time operation.

The total span of the data is the period of 2011–2019. We split the dataset into train/validation/test periods of length 6 years/1 year/1 year respectively. The model inputs were pre-processed with a quantile transformation that creates features which follow a uniform distribution on the interval $[0, 1]$ [15]. The target values for the model outputs were scaled to have zero mean and unit variance.

## 3. Model

### 3.1. Proposed model

The Long Short-Term Memory [16] recurrent neural network is commonly applied to text and language processing but also to time-series modeling. The LSTM neural network was chosen in our application because it can process sequential data and it can capture non-linear time-variate dependencies.

The proposed model is designed for forecasting purposes of total grid losses of a power system in the intra-day market such as Elbas market. The gate of intra-day market closes one hour in advance of the traded hour which means that there is a lead time between forecasting moment and forecasting horizon of two hours. Therefore, we propose a model which can forecast the grid losses two hours ahead with a forecasting horizon of one hour. Consequently, the model can be run every hour with up-to-date weather forecasts to produce a forecast for the market two hours ahead.

The computational graph of our model is shown in Fig. 1. The recurrent neural network (RNN) sequentially processes a sequence that contains 70 past observations $(x_{t-69}, x_{t-68}..., x_t)$ and one- and two-hour ahead forecasts $\hat{x}_{t+1}, \hat{x}_{t+2}$ of the 54 variables selected as model inputs.[1]

At each processing step, the RNN updates its hidden states $h$ which are passed as inputs to the next iteration. At the very final step processing step, the hidden states are linearly combined to compute the model output which is desired to be close to the predicted grid loss $\hat{y}_{t+2}$ two hours ahead. We used mean absolute error (MAE) as the loss function:

$$MAE = \frac{1}{T}\sum_{\tau=1}^{T} |y_\tau - \hat{y}_\tau|, \tag{1}$$

where $\hat{y}_\tau$ is the predicted and $y_\tau$ is the true value of the grid loss at time $\tau$. We chose the mean absolute error because it is less sensitive to possible outliers.

### 3.2. Comparison methods

We compare the proposed model against two methods. The first comparison method is the current method used in Fingrid Oyj. It is based on historical reference values, previous knowledge and available prognosis data. We call this method a *reference* method.

The second comparison method is a linear regression model (as used by Sahlin et al. [2])

$$\hat{y}_{t+2} = A\hat{x}_{t+2} + \text{noise}$$

which uses the last measurements $\hat{x}_{t+2}$ (or the corresponding forecasts) of the 54 chosen variables as inputs to produce a two-hour ahead forecast of the power grid loss.

## 4. Experiments

### 4.1. Hyperparameter tuning

We optimize the hyperparameters of the LSTM model using random search [18] in which each hyperparameter value is sampled from a fixed set of pre-defined values. We observed that different weight initializations resulted in 5–10% variability of the MAE score. We tune each individual LSTM model using stochastic gradient descent and the Adam optimiser with a learning rate of 0.001. The optimized hyperparameters, the set of their pre-defined values and and the optimal hyperparameter values found with the random search are shown in Table 1.

### 4.2. Main results

Table 2 shows the obtained MAE scores for the tuned LSTM model as well as for the comparison methods. The proposed LSTM model outperforms both the reference method by 40% when predicting the total power grid losses of the case study and the linear regression model by 30% which was presented in the literature. Fig. 2 shows that it was possible to beat the reference methods with minimal hyperparameter tuning.

The left subplot of Fig. 3 presents the weekly average MAE for the proposed LSTM model and the reference method. We can see that the proposed model achieves lower average of mean absolute error than the reference method on a weekly level. Since the summer season is usually more stable and more predictable for electricity markets, both methods have a lower prediction error during the summer period. Winter periods are subject to high variations in the electricity demand and renewable

[1] In this study, we assumed perfect foresight of the one- and two-hour ahead forecasts, since we did not have historical forecasts of these variables in our dataset. We assume that the forecasts of weather and electricity market variables can be easily obtained from corresponding service providers. We conducted tests on robustness of the proposed model with respect to forecasting errors of the input variables by inserting gaussian noise to all forecasted inputs with 1%–10% variance which was relevant scale in terms of the current context [17]. The results showed that the model is robust for the forecast errors and the performance of the proposed model did not notably change.

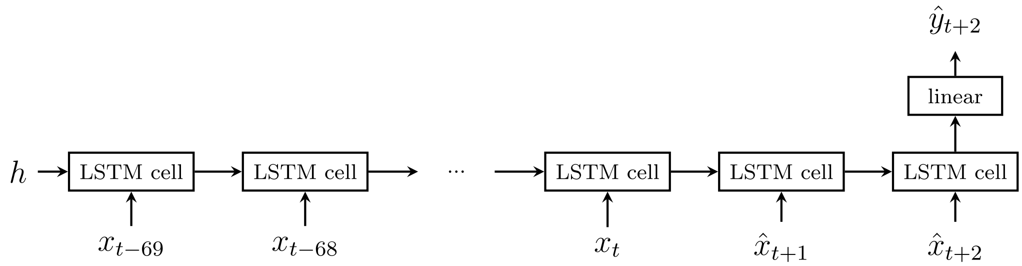


**Fig. 1.** Network architecture with input feature sequence length of 72. The RNN sequentially processes of last 70 observations and the next two forecasts of the 54 variables selected as model inputs. The model output $\hat{y}_{t+2}$ is the power grid loss forecast for two hours ahead.

**Table 1**
Hyperparameters tuned for the LSTM model.

| Hyperparameter | Experimented values | Chosen value |
|---|---|---|
| Data sequence length | 48, 72, 96, 168, 334 | 72 |
| LSTM number of layers | 1, 2, 3 | 2 |
| LSTM hidden size | 54, 216, 432, 864, 1296 | 432 |
| LSTM dropout | 0.0, 0.1, 0.2, 0.3 | 0.1 |

**Table 2**
The mean absolute error on the test set for the proposed LSTM model compared to the reference and linear regression methods.

| Model | Mean absolute error (MWh/h) | Mean absolute percentage error (%) | Relative performance (%) |
|---|---|---|---|
| Reference | 19.1 | 14.4 | - |
| Linear regression | 16.5 | 12.5 | 14 |
| LSTM | 11.5 | 8.7 | 40 |

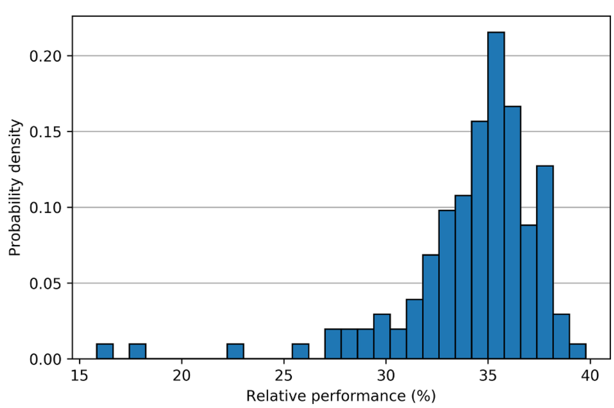


**Fig. 2.** Histogram of the test performance of the LSTM model relative to the reference model for different hyperparameter values used in the random search.

generation and thus more difficult to predict which we can notice by spikes in mean absolute error with both methods.

The right subplot of Fig. 3 shows the relative difference in performance of the LSTM model compared to the reference method in weekly average mean absolute error (MWh/h). Positive values correspond to the weeks when the proposed LSTM model performed better than the reference method. We can see that the proposed LSTM model outperformed the reference method in a total of 50 weeks and showed weaker relative performance only during two weeks.

The weeks where the LSTM model performed worse or similarly to the reference method are mainly during the winter period except for one week in the summer. We can observe from the right subplot in Fig. 4 that the weaker relative performance for LSTM model occurred during week 27. However, looking at the same week in the left subplot shows that the forecast error of both the LSTM model and the reference method is low during the week. The lower relative performance is more due to well performing reference prediction during the specific week as the proposed model seems to achieve low prediction error as well.

Winter periods are generally more difficult for grid loss prediction because local weather is subject to higher variability and corona effects (short-term high peaks of a local power grid loss) are more probable. The left subplot in Fig. 4 demonstrates that high peaks in power grid losses can be difficult to predict: both methods perform poorly on December 29th when there is a significant increase of the total grid loss. In order to improve the prediction of the peaks caused, for example, by the corona effect, one has to collect more detailed information about local weather. It is known that the corona effect only occurs during specific weather and voltage conditions along a transmission line. Unfortunately this type of information was not available in the used dataset.

The right subplot of Fig. 4 gives a closer look at another week during the winter. We can see that both methods perform poorly during the day just before Christmas. That result suggests that a prediction model can benefit from using better features representing holiday periods.

## 5. Conclusion

In this paper, we proposed a recurrent LSTM neural network model to predict power grid losses of a power system network. The proposed model can take complex effects of local weather conditions on the power grid loss into account, which the previous methods have not been able to do. The proposed LSTM model works without prior knowledge of the power system network, which is a clear advantage compared to the previous methods in the literature. For this reason, the proposed method is directly applicable to other power systems without any time-consuming network modelling. We showed that the proposed model outperforms the reference method used in the industry and linear regression method presented in literature.

The accuracy of the proposed model can be further improved in a number of ways. We observed in our experiments that increasing the size of the training data generally has a positive effect on the model performance. Thus increasing the size of the training data by spanning a larger period of time and using more detailed local information is one way to improve the model. We could also see that the model performs poorly during days close to public holidays. This suggests that the model can benefit from better handling of holiday periods, for example, through using better features. Poor performance on public holidays also indicates overfitting related to the small size of the used dataset. We also observed that the model has difficulties in predicting grid losses caused by corona. This can be improved by using more detailed weather data and other types of local measurements.

## CRediT authorship contribution statement

**Jarkko Tulensalo:** Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Janne Seppänen:** Writing - original draft, Writing - review & editing, Visualization. **Alexander Ilin:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Visualization, Supervision.

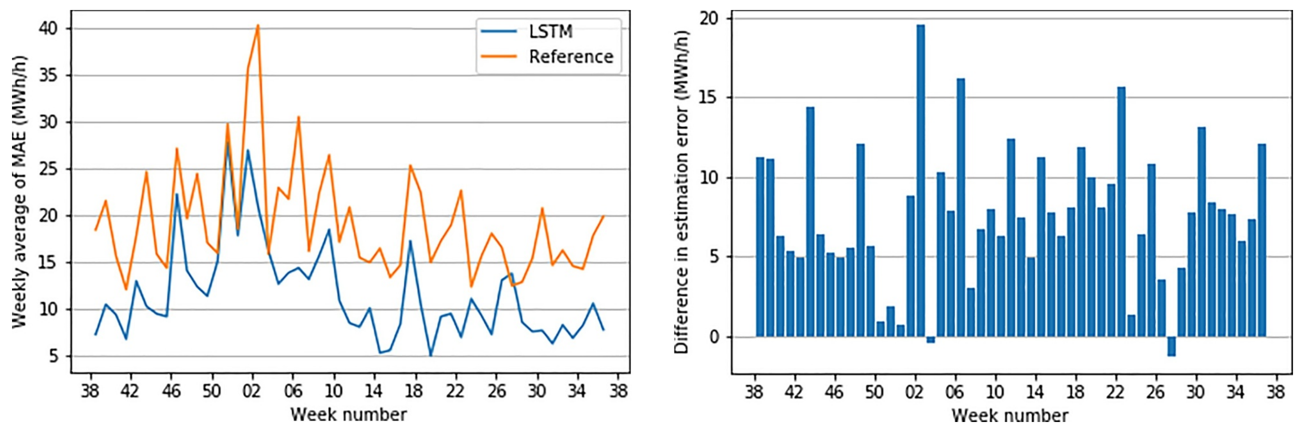


**Fig. 3.** Left: Weekly average mean absolute error for the proposed LSTM model and the reference method. Right: The relative difference in performance of the LSTM model compared to the reference method in weekly average mean absolute error (MWh/h). For example, a value of 10 means that the proposed LSTM model had a lower estimation error by an average of 10 MWh/h during the week compared to the reference method.

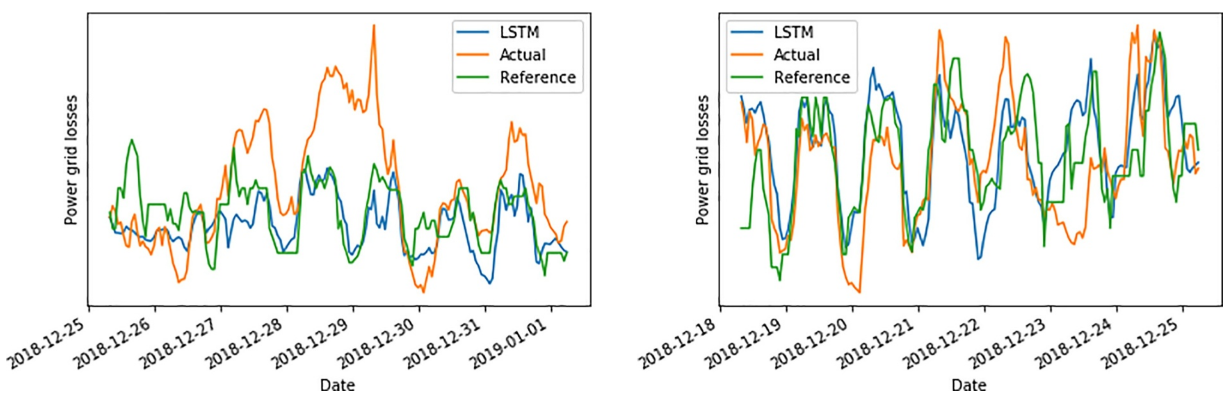


**Fig. 4.** Hourly grid loss values and their predictions. Left: Larger errors are more often in winter (week 52). Right: Large errors can be observed in the day just before Christmas (week 51).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] Electric power transmission and distribution losses, 2018, Accessed: 2020-03-05. Available at https://data.worldbank.org/.
[2] J. Sahlin, R. Eriksson, T. Ali, M. Ghandhari, Transmission line loss prediction based on linear regression and exchange flow modelling, 2017 IEEE Manchester PowerTech, (2017), pp. 1–6.
[3] L. Tian, Q. Wang, A. Cao, Research on SVM line loss rate prediction based on heuristic algorithm, Appl. Mech. Mat. 291–294 (2013) 2164–2168.
[4] P. Nallagownden, T. Hong, Development of a new loss prediction method in a deregulated power market using proportional sharing, 2011 5th International Power Engineering and Optimization Conference, PEOCO 2011 - Program and Abstracts, (2011).
[5] W. Liu, T. Zeng, T. Jun, L. Guan, B. Li, Research on electric power with development and application of line loss rate forecasting software based on MLRM-GM, Adv. Mat. Res. 977 (2014) 182–185.
[6] S. Appalasamy, H.-S. Gan, O. Jones, N. Moin, C.S. Tan, Transmission loss modelling and analysis with multiple linear regression, in Proceedings of the 20th International Congress on Modelling and Simulation. Adelaide, Australia: MODSIM2013, December 2013.
[7] Y. Ren, X. Zhang, X. Huang, Study on the prediction of line loss rate based on the improved RBF neural network, Adv. Mat. Res. 915–916 (2014) 1292–1295.
[8] F. Xie, B. Zhou, Q. Zhang, L. Jiang, Line loss rate forecasting based on grey model and combination of neural network, Adv. Mat. Res. 621 (2012) 340–343.
[9] S. Bouktif, A. Fiaz, A. Ouni, M. Serhani, Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches, Energies 11 (2018) 1636.
[10] R. Agrawal, F. Muchahary, M. Tripathi, Long term load forecasting with hourly predictions based on long-short-term-memory networks, 2018 IEEE Texas Power and Energy Conference (TPEC), 2018, pp. 1–6.
[11] Q. Xiaoyun, K. Xiaoning, Z. Chao, J. Shuai, M. Xiuda, Short-term prediction of wind power based on deep long short-term memory, 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), (2016), pp. 1148–1152.
[12] X. Qing, Y. Niu, Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, Energy 148 (2018).
[13] U. Ugurlu, I. Oksuz, O. taş, Electricity price forecasting using recurrent neural networks, Energies 11 (2018) 1255.
[14] Nordpool historical market data, 2019, (Available at https://www.nordpoolgroup.com/historical-market-data/). Partly publicly available. Accessed: 2019-11-30.
[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[16] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.
[17] A. Heikura, Ennusteiden laadun vaikutukset sähköjärjestelmän käyttötoimintaan, Master's thesis. Aalto University, School of Electrical Engineering, Department of Electrical Engineering and Automation. (2020).
[18] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, J. Mach. Learn. Res. 13 (Feb) (2012) 281–305.