Forecasting Photovoltaic Production and Building Energy Demand Using LSTM Networks: A Comparative Study with ANN Models

Christian Schlüter Prieto[1], Farzin Golzar[1]

[1] KTH Royal Institute of Technology, Department of Energy Technology, Stockholm (Sweden)

Abstract

A growing global population demands more efficient and sustainable energy systems. Accurately modeling their production has become essential with the increasing integration of renewable energy sources. Equally important is forecasting energy demand to ensure a stable and optimized energy network. Improved forecasting of both production and consumption enhances energy management and supports the development of robust optimization strategies. This study investigates the use of Long Short-Term Memory (LSTM) networks to forecast photovoltaic (PV) production and building energy demand, comparing their performance with a standard Artificial Neural Network (ANN) model.

Using an hourly data set from 2020 to 201, for PV production, the LSTM achieved a MAE of 5.31 and RMSE of 13.78, compared to 23.41 and 47.97 for the ANN, respectively. Similarly, for energy demand forecasting, the LSTM reduced the MAE from 2.52 to 1.60 and the RMSE from 3.69 to 2.33, indicating a significant improvement in predictive accuracy. In terms of the coefficient of determination (R^2), the LSTM achieved 0.86 for PV prediction and 0.74 for demand, compared to 0.93 and 0.69 for the ANN. The findings highlight the superior performance of LSTM networks in capturing temporal dependencies and improving forecast precision for both PV generation and energy demand.

Keywords: ANN, LSTM, forecasting, PV production, electricity demand

Introduction

In recent decades, climate change and environmental degradation, driven by the consumption of fossil fuels and the resulting greenhouse gas emissions, have been a major issue for social environmental security worldwide (Zouboulis & Tolkou, 2015). To mitigate these effects and counteract climate change, society strives toward more sustainable and resilient energy systems. A significant increase in the proportion of renewable energy sources can be seen worldwide (PALZ, 1994).

In Sweden, particularly, the transition into renewables has become a reality and could reach a 100% share within 20 years (Zhong, Bollen, & Rönnberg, 2021). Sweden has gradually implemented renewable energy over the years. Sweden's abundant access to water resources and utilization of hydropower (Jain, 2023), coupled with sustainable decision-making by the Swedish Government (Regeringen och Regeringskansliet, 2018), has played an important role in shaping its energy landscape. Additionally, the integration of wind and solar power is becoming an ever-increasing source for the Swedish energy mix. The former represents 19,1%, according to the International Energy Agency, as of 2021 (IEA, 2018). The fact that the latter is a relatively small percentage of the total mix means it still has room to develop and expand. Compared to other countries or regions with similar amounts of sunlight, Sweden's photovoltaics still fall behind in terms of installed capacity. Northern Germany is an example of that (Stridh, Yard, Larsson, & Karlsson, 2014).

To incentivize society to buy and use PV, technical advancements and policy adjustments need to be made. An example of these technical advancements is an optimization model for PV coupled with energy storage, such as a battery (Flor Lopes, Golzar, & Lundmark, 2022). The optimization model can significantly reduce costs. Moreover, the effectiveness of the linear optimization depended greatly on the accuracy of the forecasted values of production and demand. Therefore, achieving more precise predicted values can have a high impact resulting in a reduction in energy costs.

10.18086/swc.2025.04.23 Available at http://proceedings.ises.org

(Ağbulut, Gürel, & Biçen, 2021) used four different models to predict daily global solar radiation. The study used the same input dataset for all four machine learning algorithms (SVM, k-NN, DL, and ANN) to ensure a fair comparison. The input variables included daily minimum temperature (Tmin), daily maximum temperature (Tmax), cloud cover, daily extraterrestrial solar radiation, and day length. The last two variables were calculated using standard solar radiation equations. The dataset was randomly divided into 70% training data and 30% testing data using shuffled sampling, and the same dataset and labels were used for all algorithms. The results varied significantly. Thus, it becomes imperative to investigate many types of machine learning models to find the most accurate results. Adding to that, the inherent randomness of machine learning gives motivation to examine different models. Finding the optimal type of model can be a difficult task. It is also why it is important to research the methods used and use more than one for a specific task.

In this study, a Long-Short-Term-Memory (LSTM) model was used to make predictions. The proposed LSTM model has been compared to an artificial neural network (ANN). With these two models taken into consideration, several questions will be answered. The prediction models use weather forecast data (temperature and solar irradiance/GHI) and historical electricity demand data as input variables to forecast PV production and building electricity consumption.

Methodology

A schematic representation of the methodology for this project is presented in Figure 1. The two models are treated separately but with similar steps. A 24-hour forecast will be provided by implementing LSTM for both PV production and building energy demand. Instead of using these predictions to optimize the battery, we want to utilize the results to compare them with the ANN model's results. To efficiently compare the models, similar metrics are used.

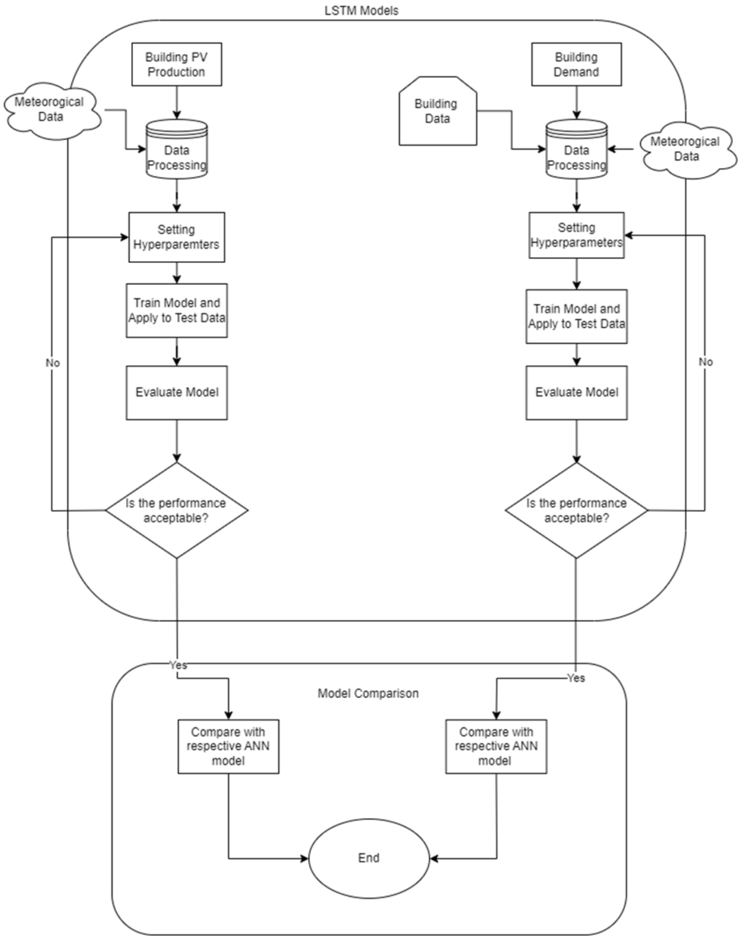

Figure 1: schematic representation of the project

2.1. Hyperparameter-Tuning

The structure of the LSTM model, including the size and quantity of the layers, greatly impacts the model's results. Similarly to an ANN model, it is therefore imperative to carefully design the model's architectural structure and explore different configurations of hyperparameters. Thus, the choice of hyperparameter has been made using a "trial and error" method and by analyzing similar projects. Similar in the sense of the size of data or the number of features roughly being the same. Different hyperparameter configurations are further tested to evaluate the most promising one and compare the computational time required to train the model.

The main hyperparameters for an LSTM network, as well as the parameters investigated in this study, include the configuration of layers, time window (or time steps), learning rate, and epochs. Beyond these parameters, additional ones can be changed, such as the activation functions and optimizers. These were not deeply analyzed for this project due to their requiring a more refined knowledge about neural networks and A.I. models. Also, if they were included in a trial-and-error method for hyperparameter tuning, the number of configurations would increase substantially.

Regarding the configuration of the layers, this includes the number of LSTM layers stacked on top of each other. For both models, a maximum of 3 LSTM layers were used. Additionally, it includes the size of the layers, indicating the number of LSTM units or cells in the layer. Lastly, dropout layers are considered. These are added between the LSTM layers to prevent overfitting, and they randomly drop a fraction of the LSTM units during training.

2.2. Training, Validation and Test

The training for LSTM networks is supervised, meaning the correct output for a given timeframe is known. In this case, 80% of the dataset was used to train the models for both solar and demand prediction, while the remaining 20% was split equally into validation and test data. The process for training the weights of the network follows the same procedure as for most neural networks. By minimizing the loss function, it adjusts the network weights to find the closest output with respect to the given training data set. This process is also known as backpropagation.

The validation data is used to evaluate the performance during training. For each epoch, the loss of the validation data is seen. Depending on the loss of the validation data, hyperparameters can be tuned accordingly. Lastly, the test data is a set that the model has not seen during training. The performance on the test data provides the clearest indication of how the model will behave when applied to real-time data. Therefore, the metrics most relevant to this project are calculated on the results for the test data set.

2.3. Pre-processing

A key difference between the datasets used in the LSTM model and ANN model is how the data is pre-processed. For ANN, the input for the model is a list of features, and its output for a given timestep is a response. The model does not consider time. It only observes the input values to predict an output. For LSTM, the previous timestep or timesteps are relevant. Thus, the dataset must be set up differently for the LSTM to do its calculations. The dataset is split into time windows, thus creating a tensor. For example, if the objective is to predict the demand for 14:00, then, depending on how big the choice of this time window is, the dataset for this particular timestep would include all the data within this time window as well. If the time window is set to 24 hours, then the data from 24 hours before that timestep would be considered for the calculations. The time window is a hyperparameter for LSTM models and was therefore mentioned earlier in the section about hyperparameter-tuning.

To avoid numerical instabilities and improve the performance of the model, the data is scaled. For this model, a MinMaxScaler was used. This scaler transforms all numbers within the feature into a number between 0 and 1. The x inEq. 1 represents the value that is transforming, while min(x) and max(x) represent the smallest and largest numbers, respectively.

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad \text{(eq. 1)}$$

2.4. Performance Evaluation

The relevant metrics are: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and lastly the Pearson's coefficient R^2. The formulas can be seen below:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \text{(eq. 2)}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \text{(eq. 3)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{(eq. 4)}$$

$$R^2 = 1 - \frac{var(y_i - \hat{y}_i)}{var(\hat{y}_{i)}} \quad \text{(eq. 5)}$$

In all displayed metrics, y_i represents actual output values and $\hat{y}_i$ is the predicted output values.

2.5. Software Tools and Computational Resources

The programming language used for this project is Python 3.11.8, and Jupyter Notebook was used as the integrated development environment (IDE). Many illustrations and graphics have been designed with the help of draw.io. The different configurations of the models and the final ones were run on an ASUS laptop with an AMD Ryzen 7 7300U (2.30 GHz) processor, and 16 GB Memory (RAM) installed.

2.6. Data Description

The prediction models use meteorological forecast data obtained from the SMHI Open Data API. The main input variables include temperature and Global Horizontal Irradiance (GHI), which are commonly used predictors for solar energy and electricity demand forecasting. These meteorological variables are combined with historical electricity demand data, specifically the electricity demand from the preceding day of the week, to capture temporal consumption patterns.

The prediction problem is formulated as a supervised learning task, where the input features consist of the meteorological forecasts and historical demand variables, and the outputs are the day-ahead forecasts of photovoltaic (PV) power generation and building electricity demand. Prior to model training, the data were normalized using a MinMaxScaler, and the dataset was divided into training and testing sets for model development and evaluation.

Results

In this chapter, the results of the LSTM model are presented and compared with the ANN model. Results and comparisons serve as a critical component of this project, addressing the initial question of whether LSTM can more accurately predict the sought-after energy quantities. By analyzing the results, the LSTM model can be effectively compared to the ANN model. Note that the dataset, used to present the results, is the test data, meaning that the networks has not seen this data during training and therefore gives us a more complete understanding of the model's accuracy.

3.1. Suitable hyperparameter configuration

As mentioned earlier, a trial-and-error method was adopted to find a suitable configuration of hyperparameters. To efficiently supervise the process of trying different configurations and documenting their results, an Excel document was created to identify the optimal configuration. The first configuration of hyperparameters started of simple. Simple in the sense that only one single layer of LSTM was used, and the size of the layer was small. Moreover, the time window started with 5 hours, meaning the model only accounted for 5 hours in the past for predicting the next value. With each new configuration, the complexity increased.

Besides identifying the different hyperparameters, the different value options need to be considered. Thus, the following possible values have been assigned to the different hyperparameters:

- Time window = {5, 10, 24, 48}
- Learning rate = {0.001, 0.01, 0.1}

- Number of LSTM layers = {1, 2, 3}
- Number of LSTM units in LSTM layers = {64, 128, 256}
- Epochs = {10, 20, 50}

In addition, the activation functions within the LSTM modules can be changed. For example, one could add the following:

- Activation function = {“ReLU”, “Sigmoid”, “Tanh”}

The first 5 hyperparameters already give us 324 different possible configurations to evaluate for one of the models. This would be a total of 648 configurations for both the PV and demand model. Besides the tweaking of activation functions requiring intricate knowledge about machine learning and deep learning networks, considering different activation functions would triple the number of configurations to analyze. Therefore, the possibility of changing the activation function within the network was not considered for this trial-and-error method. Furthermore, some other configurations were not fully analyzed because it could be concluded that they would not improve the model’s performance or because the computational time to execute the model exceeded a reasonable limit.

3.2. PV Model

The best performing configuration for the PV model consists of a single LSTM layer with a size of 256 cells. The time window used was 24. And other hyperparameters, such as the learning rate and epochs, were set to 0.001 and 10, respectively. Regarding the computational time, it took 241 seconds for the model to adjust the weights.

The results of the LSTM network for the PV model are illustrated in figure 2, figure 3, and figure 4.

In figure 2, the blue lines indicate the forecasted values that the network has calculated. The orange color represents the actual solar irradiance for that time horizon. The time horizon is determined by the test dataset. In this case (and for the demand model), the data’s time frame stretches from end of October 2022 to the end of the year. Visualizing the results in this manner helps us better understand the accuracy of the model and provides a clearer understanding of the predicted data.

The scatter plot in figure 3 depicts an alternative perspective for visualizing the accuracy of the model. Values deviating from the red line stipulate discrepancies between the predicted value and the actual value.

Figure 2: LSTM-testdata predictions vs actual GHI.

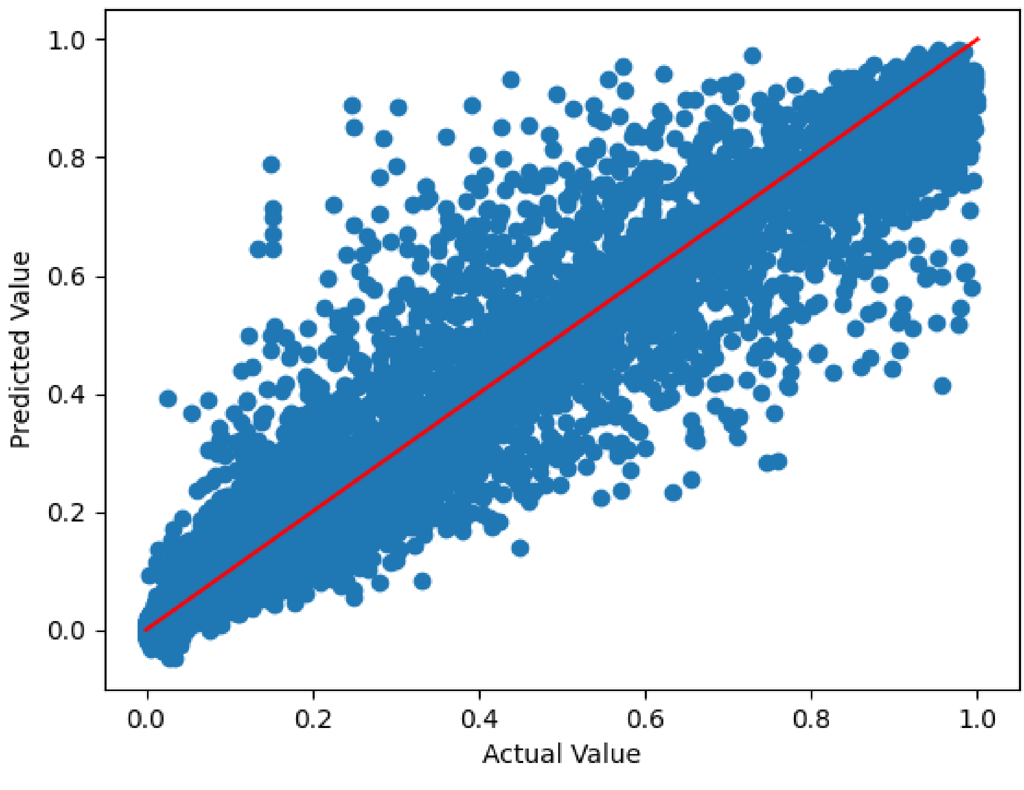

Figure 3: Scatter plot of actual vs predicted solar irradiance.

Moreover, the aim is to forecast the hourly energy production. Additionally, to efficiently compare the model with the ANN model once again, the results are presented similarly. In Figure 4, the PV production for November 2022 is shown. By assessing the figure visually, the two lines seem to be very close to each other, meaning that the model accurately predicted the PV production of the building. We have a substantial spike of energy generated from 8 to 15. This also reflects the time of year and the corresponding solar irradiance for our project. During the winter months in Sweden, the sun usually rises at 8 am and sets around 3 pm.

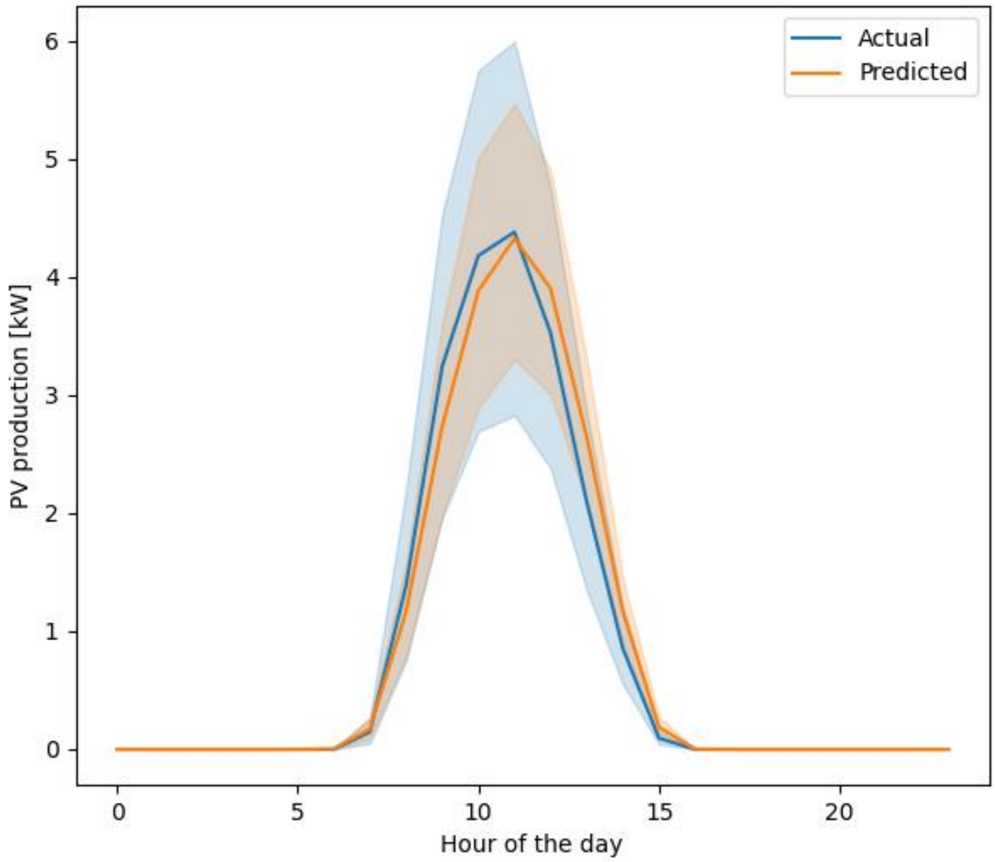

Figure 4: Predicted vs actual PV production for November 2022. The shaded areas illustrate how much the PV production values vary around the average during each hour of the day, showing the range of possible values rather than just the mean trend.

3.3. Load Demand Model

For the demand model, the configuration consists of a double layer. The first layer has 256 modules and the second has 64. The other hyperparameters remained the same as in the configuration for the PV model. A total of 352 seconds was required to render the model.

For the demand model, the results are presented similarly in Figure 5, 6 and 7.

In Figure 5 the actual demand and predicted demand are illustrated. The blue line represents the predictions and the orange line the actual values. The daily spikes in energy consumption are clearly highlighted. There are certain irregularities in the figure, seen around the beginning of November and the middle of December. It is hard to tell how these came to be, but possible explanations include the building or a neighboring building undergoing some reconstruction, or a temporary electricity outage. Altogether, Figure 5 shows good visual accuracy, giving a positive indication of the demand model.

Following the presentation of the PV model and the ANN models, a scatter plot is shown in Figure 6. The red line indicates the line where both the actual values and the predicted values are the same. And lastly, the results for the month of November 2022 are also put forward. The forecasted line in orange appears to follow the actual demand in blue. The predicted values do not seem to spike as much as the actual demand.

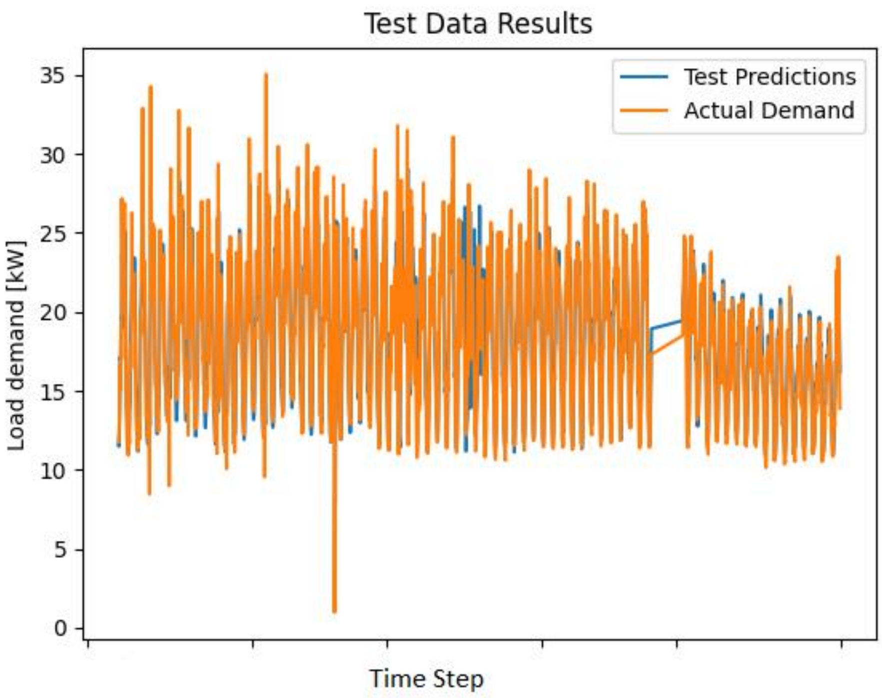

Figure 5: LSTM-testdata predictions vs actual demand.

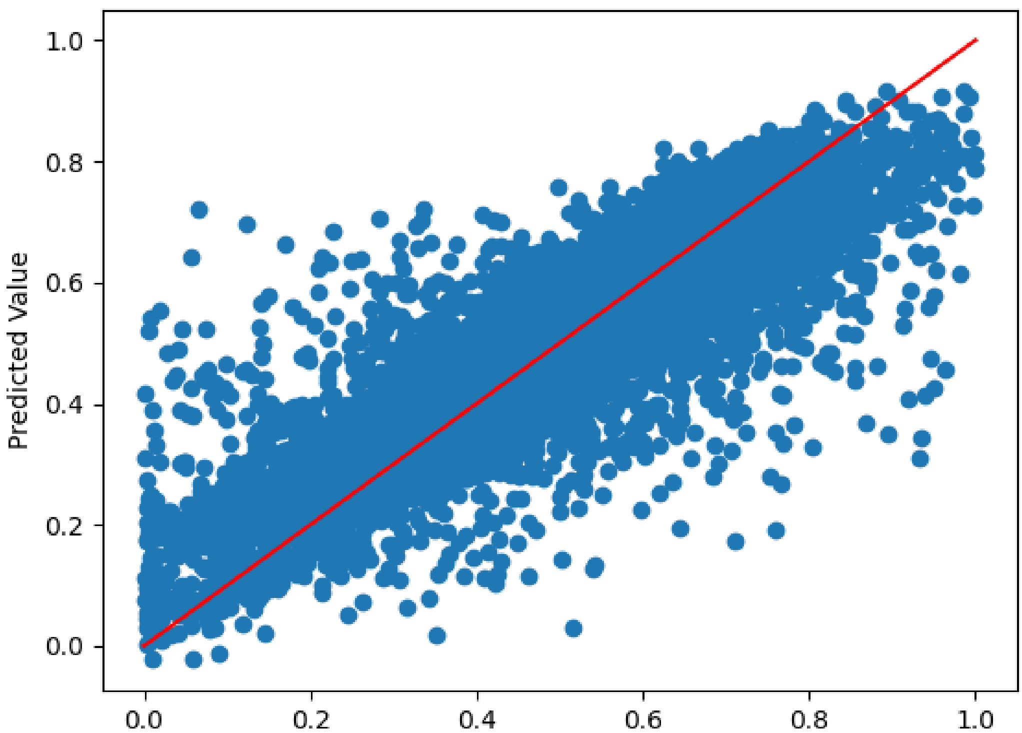

Figure 6: Scatter plot of actual vs predicted demand.

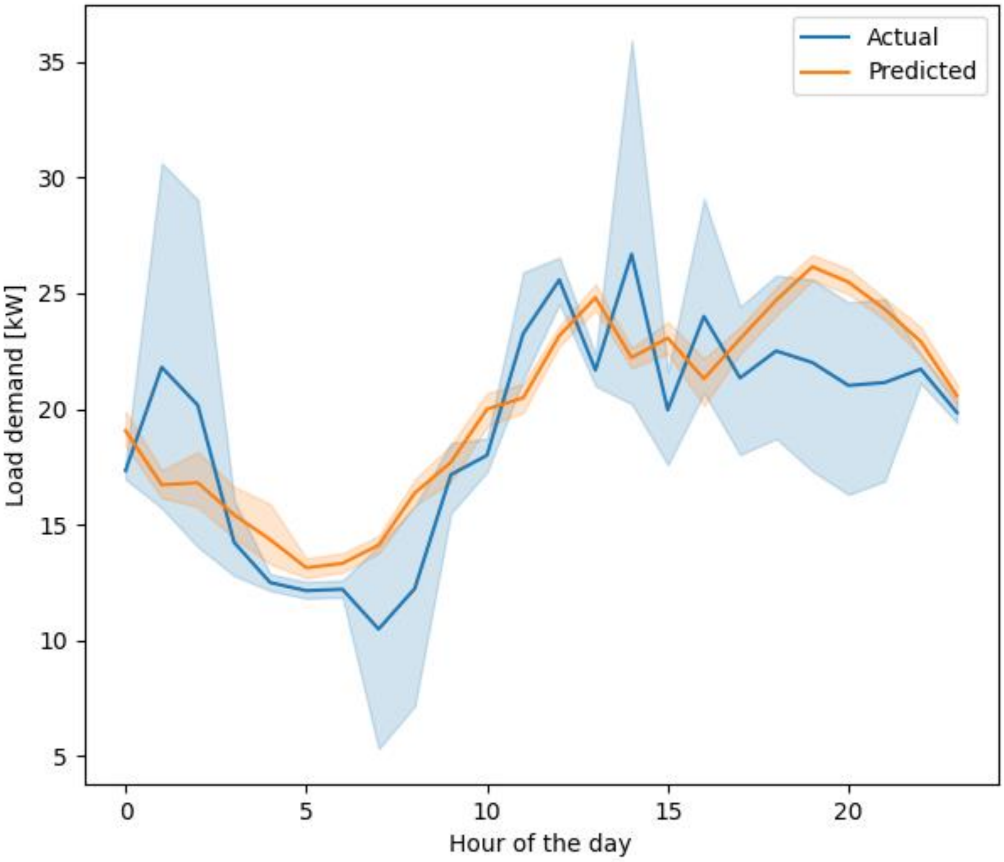

Figure 7: Predicted vs actual PV demand for November 2022. The shaded areas illustrate how much the load demand values vary around the average during each hour of the day, showing the range of possible values rather than just the mean trend.

3.6. Comparison of ANN vs LSTM

This chapter presents a comparative analysis of the LSTM model and ANN model within the context of this project.

Table 1: Metrics Comparison between LSTM and ANN.

	PV Model		Demand Model	
	LSTM	**ANN**	**LSTM**	**ANN**
MAE	5.31	23.40	1.60	2.51
nMAE	0.01	0.03	0.04	0.06
MSE	205.5	2301.09	5.83	13.64
nMSE	0.63	0.00	0.17	0.34
RMSE	13.77	47.96	2.33	3.69
nRMSE	0.04	0.06	0.07	0.09
R^2	0.85	0.93	0.74	0.68

Table 1 presents the results of the test data for both models. Additionally, the n denotes that the metrics are normalized, meaning they are divided by the maximum value minus the minimum value of the dataset. Normalizing the metrics facilitates comparison between models and future models or datasets, helping the reader interpret the metrics as relative values.

In the table above, both the LSTM PV and demand models seems to outperform the respective ANN models in most metrics. One anomaly worth noticing is that the R-squared coefficient for the ANN PV model is 10% greater than for the LSTM model. A possible explanation for this deviation could be overfitting of the ANN model. Besides that, it could also simply be the inherent difference between the two algorithms. LSTM might not capture certain trends, that ANN does.

1. Discussion

The discussion chapter interprets and analyzes the key findings in the context of answering the initial questions. Also, it aims to put the results into a broader literature and understand the significance for the energy field.

An issue with ANN models was the complexity imposed by the summer months due to irregular weather patterns and high levels of solar irradiance. The performance during these months seemed to decrease compared to the rest of the year. Meanwhile, the LSTM manages to maintain a constant accuracy regardless of the time of month. This could be explained by the inherent nature of the network to retain information with the different gates.

Another topic of discussion is the computational time of the models. For LSTM, the time to train the chosen models took somewhere between 4 and 6 minutes. In fact, there were other models with much larger networks and different hyperparameter settings that took up to 10 minutes or more to finish. In contrast to ANN, which requires around 2-3 minutes for the calculations to finish, it appears to be much more computationally demanding. For this project, the time aspect is not as relevant since the primary focus is to compare the models within the energy field. But there might be other projects where the time requirement might be pivotal. Thus, before applying a LSTM model over an ANN model, one should ask if the added computational time is worth the increase in accuracy. Although some might object, arguing that this increase in accuracy cannot be generalized, the literature review suggests otherwise. Regardless, computational time is something to be taken into consideration.

2. Conclusions

As the literature review showed, it is adequate to use LSTM for this project. Learning more about the LSTM model and ANN, additional strength and weaknesses were uncovered. These strengths and weaknesses were further displayed in the results chapter, and some new ones were added. For example, the catch regarding computational time. Lastly, the results clearly indicated a greater accuracy for the LSTM model.

The use of LSTM models for energy systems is thus a very reasonable option for predicting energy quantities within buildings. By achieving greater performance, systems can be more predictable, leading to more efficient use of energy and ultimately saving money. The forecast was used to optimize a battery to save money on electricity. This was done with the help of a reliable, hourly prediction of the production and demand of the building. If these predictions can be done with even greater accuracy, the building can save even more in terms of bills. And there are many other projects where a reliable energy prediction is paramount. This shows that the implementation of LSTM can be a great option.

As mentioned above, the outcomes of this project can be applied to many applications within the field. Battery optimization is one of them. But this could also include things such as electricity markets, microgrids, large energy systems, etc.

Still, there are areas within this project that can be further explored. The first being hyperparameter-tuning. As mentioned, there are many more methods for tuning. Thus, using another method or multiple methods could improve the results for the models and is therefore an interesting area to further explore.

Secondly, there are many more neural networks that could be applied instead of LSTM. Within the category of deep neural networks there is Gated Recurrent Units (GRU), Deep Belief Networks (DBN) and Generative Adversarial Networks (GAN) to name a few. Similar to the methods for hyperparameter-tuning, these could prove advantageous for increasing the accuracy of the predictions.

References

[1]. Ağbulut, Ü., Gürel, A. E., & Biçen, Y. (2021). Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. Renewable and Sustainable Energy Reviews, 135, 110114. https://doi.org/10.1016/j.rser.2020.110114

[2]. Lopes, M. F., Golzar, F., & Lundmark, L. (2022). Forecasting and optimization models for integrated PV-ESS systems: A case study at KTH Live-In Lab.

[3]. International Energy Agency (IEA). (2018). International Energy Agency data and statistics. Retrieved March 5, 2025, from https://www.iea.org/data-and-statistics

[4]. Jain, P. (2023). Hydropower development and its role in Sweden's transition to renewable energy. International Journal of Innovative Science and Research Technology, 8(9), 2456–2165. https://ijisrt.com/assets/upload/files/IJISRT23SEP847.pdf

[5]. Palz, W. (1994). Role of new and renewable energies in future energy systems. International Journal of Solar Energy, 14(3), 127–140. https://doi.org/10.1080/01425919408909805

[6]. Regeringen och Regeringskansliet. (2018). Energipolitikens inriktning. Retrieved March 4, 2025, from https://www.regeringen.se/rattsliga-dokument/proposition/2018/04/prop.-201718228

[7]. Schlüter Prieto, C., & Golzar, F. (2024). Forecasting production and demand for PV-integrated systems with LSTM and comparison with ANN model: A case study at KTH Live-In Lab. Retrieved from https://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1897175

[8]. Stridh, B., Yard, S., Larsson, D., & Karlsson, B. (2014). Profitability of PV electricity in Sweden. IEEE 40th Photovoltaic Specialist Conference (PVSC), 1492–1497. https://doi.org/10.1109/PVSC.2014.6925198

[9]. Zhong, J., Bollen, M., & Rönnberg, S. (2021). Towards a 100% renewable electricity generation system in Sweden. Renewable Energy, 171, 812–824. https://doi.org/10.1016/j.renene.2021.02.153

[10]. Zouboulis, A., & Tolkou, A. (2015). Effect of climate change in wastewater treatment plants: Reviewing the problems and solutions. In S. Shrestha, A. K. Anal, P. A. Salam, & M. van der Valk (Eds.), Managing water resources under climate uncertainty (pp. 197–220). Springer International Publishing. https://doi.org/10.1007/978-3-319-10467-6_10