SolarPACES 2023, 29th International Conference on Concentrating Solar Power, Thermal, and Chemical Energy Systems

Operations, Maintenance, and Component Reliability

https://doi.org/10.52825/solarpaces.v2i.802

© Authors. This work is licensed under [a Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Published: 15 Oct. 2024

Development of a DNI Forecast Tool Based on Machine Learning for the Smart Operation of a Parabolic Trough Collector System with Thermal Energy Storage

Theory and Results

Johannes Christoph [Sa](https://orcid.org/0009-0003-0270-8205)ttler¹ [,](https://orcid.org/0000-0001-7380-8172)Siddharth Dutta² D, Spiros Alexop[oulo](https://orcid.org/0000-0002-6938-0860)s¹ D, Adnan Kawam¹ D. Cristiano Teixeira Boura¹ D. Ulf Herrmann¹ D. and Ioannis Kioutsioukis³th

¹ Solar-Institut Jülich (SIJ) of the FH Aachen University of Applied Sciences, Germany

² Protarget AG, Zeissstrasse 5, 50859 Cologne, Germany

³ University of Patras, Laboratory of Atmospheric Physics, 26504 Rio, Patras, Greece

*Correspondence: Johannes Christoph Sattler, sattler@sij.fh-aachen.de

Abstract. In the research project Smart Solar System (S3), the Solar-Institut Jülich (SIJ) developed forecast tools to predict the direct normal irradiance (DNI) in hourly resolution for the current day. The aim of the daily DNI forecast is to use it as input to enable a smart operation of a parabolic trough collector (PTC) system with a concrete thermal energy storage (C-TES) located at the company KEAN Soft Drinks Ltd in Limassol, Cyprus. The main focus in this work is on the development of a DNI forecast tool based on long short-term memory (LSTM), which is a recurrent neural network (RNN), and its comparison with a DNI forecast tool based on analytical algorithms (non-machine learning). Only the non-machine learning DNI forecast tool with hourly update was tested in real-life PTC plant operation since end of 2022. The comparisons between the three DNI forecast tools show that the potential of using machine learning is very high. Different comparisons were made including an evaluation of the accuracy of the tool (i.e. comparison of the DNI forecast data with DNI measurement data). The DNI forecast based on the LSTM network proved to be more accurate than the nonmachine learning DNI forecasts when considering errors greater than ± 100 W/m². The error of the LSTM network compared to DNI measurement data was as follows: 43.6 % of the data was within ±100 W/m², 74.8 % was within ±200 W/m², 89.8 % was within ±300 W/m² and 95.8 % was within ± 400 W/m².

Keywords: Direct Normal Irradiance (DNI) Forecasting, Long Short-Term Memory (LSTM), Recurrent Neural Network, RNN, Parabolic Trough Collector, PTC, Thermal Energy Storage, **TES**

1. Introduction

In the research project Smart Solar System (S3), funded by national and regional funding organisations in the European network SOLAR-ERA.NET, the Solar-Institut Jülich (SIJ) developed forecast tools to predict the direct normal irradiance (DNI) in hourly resolution for

the current day. The aim of the daily DNI forecast is to use it as input to enable a smart operation of a parabolic trough collector (PTC) system with a concrete thermal energy storage (C-TES) located at the company KEAN Soft Drinks Ltd in Limassol, Cyprus. More details on the PTC system and C-TES are described in [1], [2] and [3]. First results for a non-machine learning DNI forecasting tool developed by the SIJ were published by [4], hereafter referred to as method 1. In the present work, a DNI forecast tool based on a long short-term memory (LSTM) network, which is a recurrent neural network (RNN), was developed using the Keras library in Python (method 2). An RNN is a type of artificial neural network (ANN). An RNN network can process time series data or sequences [5], which is why it is suitable for DNI forecasting. Also, method 1 was developed further.

The novelty of this paper is to present non-hardware based and low-cost options for DNI forecasting using ordinary weather forecast data from a typical weather website with the focus on deep learning (long short-term memory (LSTM) network) for the specific solar application in Cyprus. To show the potential of deep learning for DNI forecasts, the results are compared with analytical algorithms. For the validation, DNI measurement data was used from a weather station located at the company KEAN Soft Drinks Ltd. Generally, the DNI forecast tools developed by the SIJ use freely available weather forecast data as input. The weather forecast data includes the following variables: cloud covers at high, medium and low altitude, wind speed, wind direction, relative humidity, ambient temperature, ambient pressure, fog and dew point, but not all variables were used. Additionally, whole-day DNI reference data of sunny, cloudless days of every month of the year were used as input data. The weather data was automatically downloaded and saved in a database on an hourly basis. This allows, on the one hand, an observation of the change in weather prediction throughout the day compared to the weather forecast for the next 24 hours at midnight. On the other hand, live DNI forecasts can be carried out flexibly, e.g. once a day with weather data downloaded at midnight or on an hourly basis. In this work, the below three DNI forecast tools, here referred to as methods, are evaluated and compared to each other.

- Method 1: DNI forecasting tool with analytical algorithm (24 hour forecast in hourly resolution carried out once at night) – as previously presented in [4]
- Method 2: DNI forecasting tool based on a long short-term memory (LSTM) network (24 hour forecast in hourly resolution. The model was trained and tested with historical hourly weather forecast data from the database that had been downloaded at midnight of each day.)
- Method 3: DNI forecasting tool with analytical algorithm (DNI forecast updated hourly during the day, tested in real-life PTC plant operation since end of 2022)

In the case of method 2 with the LSTM network, live forecasts could not carried out due to time constraints. Method 2 was also not tested with the hourly updated weather data from the weather database. The basics of a neural network are described in chapter 2. The results of the reliability and accuracy of the forecast methods are shown in chapter 3. A discussion is presented in chapter 4 and finally the conclusion and outlook in chapter 5.

Regarding state-of-the-art solar irradiance forecast, ANN methods for predicting daily values of GHI and DNI have been explored, for example, by [6] and [7], respectively. Deep learning and machine learning methods for hourly predictions of GHI were investigated in depth as shown, for example, in the reference list of [8]. Deep learning models for predicting DNI in hourly resolution were developed and explored by [8] using data from the NSRDB database for the period from 2017 through to 2019 in hourly resolution for a specific region in Morocco. Compared to a nowcasting method using all sky images such as the one presented by Nouri et al. (2018) in [9] which predicts the DNI for the next 15 minutes, the above described lowcost method 3 for predicting DNI has a lower temporal resolution for the forecast update (i.e. hourly) and also a far lower accuracy.

2. Development of DNI forecast tools and smart operation

Using DNI forecast data for the smart operation of a PTC system has several advantages. With a smart decision about when a steam boiler shall be operated and when a TES shall be charged or discharged based on the expected solar irradiance for the current operation day, the solar energy can be used more efficiently. Another very important advantage is that even during the rainy winter period when the PTC system in Limassol is normally shut down in the months of December and January, the PTC system can now be efficiently operated if the forecast predicts sufficiently high DNI irradiance during a day. In general, the energy yield during a year's operation (especially for the months of January, February, March, April, November and December, which have frequent cloud passages) can be increased if a reliable DNI forecast is available compared to a non-smart PTC operation. For smaller-sized PTC systems such as the one in Cyprus presented here, a non-hardware based DNI forecast using the methods 1 to 3 can be an interesting option. This could be the case when the aim is to reduce the complexity of the operating system (less hardware and less data processing) while also having the option to plan in advance the operation of the PTC system for a time span of one up to several hours (e.g. for more effective thermal storage charging and discharging decisions).

2.1 Working principle of analytical algorithm methods 1 & 3

The DNI forecast method 1 is based on the development of own algorithms for predicting the DNI. Regarding the freely available weather data for the DNI forecast input, predicted cloud covers at high, medium and low altitude, wind speed, wind direction and relative humidity were used. As the details for method 1 are given in [4], further details are not given here. The same applies to method 3, which is essentially the same as method 1 with the difference that the DNI forecast is being updated each hour of the day.

2.2 Working principle of DNI forecast with LSTM network (method 2)

As previously described, an LSTM network was developed (method 2) for DNI forecasting. The basic working principle of an artificial neural network, which is modelled as a so-called perceptron that mimics the biological neuron, is shown in [Figure 1](#page-3-0) below and can be described as follows. The user's input consists of a set of time series of variables *xi* . A transfer function multiplies variables x_i with weighting factors w_i and may add a bias *b* whereby the sum of all is the result *α*. The sum *α* is used as parameter for the subsequent activation function *φ*(*α*). There are several activation functions why may be used such as the identity function, Heaviside function, sigmoid, tanh(), rectified linear unit, SoftSign and SoftMax. The result of the activation function is used to adjust the weighting factors ^w*i*. [10] The final result is the output *y* [10], which is, in the case of this work, the DNI forecast data. The LSTM code itself requires various so-called hyperparameter settings in the code such as the number of neurons to be used, the number of hidden layers, number of epochs, the batch size and a dropout factor. Due to the many settings options in the LSTM code, it is therefore necessary to optimise the code such that it is suited for the problem it is used for. The machine learning forecast tool was tested with various input variables whereby the best result was achieved when using high, medium and low altitude cloud cover, DNI reference data, wind speed, wind direction and relative humidity and additionally the ambient temperature.

Figure 1. Schematic composition of an artificial neural network modelled as a perceptron net (according to [10]).

2.3 Data and methodology for developing DNI forecast

For developing forecast method 1 (and 3), historical measurement data (from an own weather station) as well as historical weather forecast data was analysed. The analysis included the identification of dependencies between variables as well as influences among different variables depending on gradients and absolute values (e.g. wind direction, wind speed and humidity). Therefore, for developing own algorithms, essentially a manual data training for a human took place for understanding the data in more detail. The advantage of this method is that the dependencies between variables are identified and understood such that algorithms can be developed based on experience of the human. The disadvantage is that the DNI forecast method 1 (and 3) algorithms are very likely to be lower in quality and accuracy compared to the algorithms developed by a recurrent neural network such as a LSTM network. This is because the LSTM model automatically finds patterns and simply tries out an enormous number of possibilities of combining the available input data set. In the case of this work, it was desired to obtain knowledge in both, i.e. to develop of own algorithms and understanding of the data as well as to develop a recurrent neural network. Generally, own observations of cloud movements and cloud types at the site in Limassol were also helpful in the course of the work with respect to knowing the limitations of the forecasts. An example are predictions of frequent cloud passages in Limassol in the afternoon. These clouds appear to be cumulus or altocumulus clouds, which move as patches of clouds leading to a high fluctuation of DNI reaching the ground. The problem with applying an artificial or recurrent neural network is that the user will not be given any information about how the neural networks algorithm works (i.e. there are no details given about the dependencies between the variables). Rather, an artificial or recurrent neural network can be understood as a blackbox, which must be optimised in terms of tuning hyperparameters for obtaining the best results. Hence, LSTM also requires a thorough optimisation, which, however, is not very time consuming compared to developing own sets of algorithms if previous knowledge in LSTM exists.

3. Results

The reliability of the three DNI forecast methods was evaluated by comparing the daily DNI forecast data with DNI measurement data and by sorting the forecast days into one of three criteria in order to rate the reliability of the forecasts. The term "reliability" in the context of this work refers to three pre-defined criteria (categories): (1) DNI forecast is principally correct with respect to sufficiently high DNI $($ > 400 W/m²) for steam production from morning until the afternoon and/or charging of C-TES in the afternoon for more than 4 hours possible, (2) a correct prediction that the DNI is mostly below the defined DNI threshold, (3) often high overestimation or high underestimation of predicted DNI. The extended details of the criteria are given in [4]. Although the accuracy of the DNI forecast is, at times, very inaccurate, the importance is that the accuracy is sufficient for the DNI forecast to be reliable for the specific

purpose regarding the smart operation of the PTC system (see [4] for more details). The results for the reliability are shown in [Table 1](#page-4-0) below. For method 1, which was developed first, 87 days could be evaluated (data of the same days for methods 2 and 3 was not available). The second data column compares methods 1 and 2 for a period of 71 days for which data for the same days was available. To compare all three methods with one another, available data of 55 days was evaluated. Depending on the available time periods where data was available, the results for the DNI forecast reliability vary too. When comparing the non-machine learning methods 1 and 3 it is clear that when the DNI forecast is updated hourly, then the reliability is improved too. The reason is that the short-term weather forecast of just a few hours ahead is, on average, more accurate than a forecast which predict e.g. 12 to 24 hours ahead. The reliability of the hourly DNI forecast update (method 3) improves by about 9 % compared to method 1. What is also shown is that method 2 (LSTM network) has a higher reliability compared to method 1, but a slightly lower reliability compared to method 3. However, if method 2 were an hourly updated DNI forecast, then method 2 is likely to have a better reliability compared to method 3. Generally, due to the reason that the DNI forecast is carried out on the basis of using freely available weather forecast data (cloud cover, wind speed etc.), the quality of the DNI forecast is highly dependent on the quality of the weather forecast data.

Table 1. Results of the evaluation of DNI forecast reliability for the three presented DNI forecasting methods. The data used is this work is from the rainy/cloudy months November 2022 to February 2023 (with gaps).

DNI forecast method	Reliability (87 days evaluated	Reliability (71 days evaluated)	Reliability (55 days evaluated)
Method 1	70 1 %	74 6 %	727%
Method 2	N/A	78.9 %	80.0%
Method 3	N/A	N/A	81.8%

[Figure 2](#page-4-1) shows an example of two different days for which a DNI forecast was made. Each graph shows three curves: DNI measurement data is shown as curve "DNI meas", DNI forecast method 2 data is "DNI fc (LSTM)" and DNI forecast with method 1 is "DNI fc". While method 2 is worse than method 1 in the forecast shown in the graph on the left side, the contrary was the result in the graph on the right side.

Figure 2. Example of DNI forecast of LSTM method 2 being worse than method 1 (left), and LSTM method 2 being better than method 1 (right). meas: measurement, fc: forecast.

In a next step, the actual accuracy of the DNI forecast methods was evaluated and the result is shown in [Table 2](#page-6-0) (the evaluation shows two table sections for each method). For the comparison, the identical 55 days, for which data was available for all three methods, were evaluated. Values after sunset and before sunrise were omitted as the DNI forecast and measurement data is always 0 W/m². In total, 567 values of forecast and measurement data were used. Regarding method 1, the left table section shows the number of occurrences (counts) that the DNI forecast is within a predefined error. The right side of the table shows how many data points are within different error ranges. The relevant result for method 1 is that 69.1 % of the data has an error between $0 - 200$ W/m² and 82.4 % of the data has an error between $0 - 300$ W/m². It should be noted that the error was calculated as absolute value (i.e. negative values between -100 to 0 W/m² are counted together with positive values between 0 to 100 W/m²). Therefore, the error shown in the table can be due to an overestimation (positive value) or underestimation (negative value) of the DNI. Considering that the DNI forecast is very basic, the result is very promising. Method 2 has slightly better results, but more interestingly, the data of method 2 was adjusted with an offset of 63 W/m2 because this improved the accuracy even further (the reason being that on average, LSTM slightly underestimated the DNI forecast values). Both results for method 2 and method 2 (adjusted) in [Table 2](#page-6-0) show that the LSTM network leads to a better accuracy compared to method 1. Method 2 (adjusted) and method 2 even have a better accuracy than method 3 when considering errors greater than 100 W/m². However, this should not be confused with the previous statement that the reliability of method 3 is currently still best from all methods because this is another type of comparison. The tendency of the data of [Table 2,](#page-6-0) however, is that the potential of method 2 with LSTM network is very high if the DNI forecast were being updated every hour using updated weather forecast data. A visualisation of the accuracy of method 2 (adjusted) with a positive offset of 63 W/ m^2 as well as method 3 is shown in [Figure 3](#page-7-0) left and right, respectively. The comparison encompasses data of 55 days (with 567 data points, considering only daytime). The two graphs show plots of DNI forecast data vs. DNI measurement data. If a data point is on the thick black line then the forecast is perfectly accurate. Also shown is a black dashed line and a black dotted line, which mark an overestimation by 200 W/m² and an underestimation by 200 W/m², respectively. It can be clearly seen that there are many DNI forecast events where the DNI measurement data had a value of 0 W/m^2 while the forecast predicted a positive DNI value.

Sattler et al. | SolarPACES Conf Proc 2 (2023) "SolarPACES 2023, 29th International Conference on Concentrating Solar Power, Thermal, and Chemical Energy Systems"

Table 2. Accuracy of DNI forecast methods 1, 2 and 3 for 55 days.

Figure 3. Comparison of method 2 (adjusted DNI forecast data) (left) and method 3 with hourly DNI forecast data (right) against DNI measurement data.

4. Discussion

If an LSTM network is to be used for a DNI forecast then it is necessary to have historical weather forecast data as well as historical DNI measurement data available for training the LSTM model. If, for a location, this data is not available then it will need to be generated first which takes up to a few years time in order to gather sufficient amounts of data. This is a big disadvantage with respect to using LSTM but in the long term, LSTM has the potential of generating DNI forecasts with a significantly better accuracy as well as reliability compared to DNI forecast methods with analytical algorithms (i.e. non-machine learning methods).

As shown in [Table 2,](#page-6-0) between 69.1 % and 74.8 % of the data points of all three forecast methods were within an error 200 W/ $m²$ which can be rated as a relatively positive outcome given the simplicity of the approach. It should be noted though that the DNI forecast presented in this work has a lower accuracy than commercially available DNI forecast data. If high forecast data accuracy is not the priority, then the presented low-cost DNI forecast methods of this work may be interesting for operators of solar systems.

5. Conclusion and Outlook

When directly comparing the LSTM method 2 with the analytical method 1, which are both designed to carry out a 24 hour forecast at midnight of each day, an improvement in the accuracy was identified. Notable is also the comparatively small work input needed for the development of an LSTM model compared to the developing of an own forecast algorithm. In this work, the amount of data for testing was relatively little for an LSTM network so it will be interesting to see how much the LSTM network (method 2) can be improved when more data is available for training the model. The LSTM network has significant potential and is expected to further improve the DNI forecast reliability as well as accuracy when the DNI forecast is updated every hour. From own observations in Limassol, Cyprus, it was noticed that patches of clouds lead to a high fluctuation DNI reaching the ground. If the weather forecast data does not predict the correct cloud cover for a specific hour, then this greatly affects the accuracy of a DNI prediction, i.e. large mismatches between predicted and measurement DNI have been observed, which is a major disadvantage when relying on weather forecast data.

In a future project, the use of LSTM for hourly updated DNI forecasts and, perhaps, also global horizontal irradiance (GHI) forecasts shall be explored further.

Data availability statement

The detailed and extensive amount of data supporting the results of this paper is only (and even only in parts) accessible to the consortium members of project S3 within legal restrictions bound by a cooperation agreement. For reasons of maintaining intellectual property, the information and data presented in this paper is limited. Weather forecast data from MET Norway was used for the purpose of exemplarily showing the accuracy of the DNI forecast tool.

Author contributions

J. C. Sattler: Conceptualisation, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software, Validation, Visualisation, Writing – original draft, Writing – review & editing;

S. Dutta: Conceptualisation, Investigation, Methodology, Project administration, Validation, Writing – review & editing;

A. Kawam: Data curation, Formal Analysis, Investigation, Software, Validation, Writing – review & editing;

S. Alexopoulos: Conceptualisation, Formal Analysis, Methodology, Project administration, Supervision, Validation, Visualisation, Writing – original draft, Writing – review & editing;

C. Teixeira Boura: Supervision, Project administration, Writing – review & editing;

U. Herrmann: Supervision, Writing – review & editing;

I. Kioutsioukis: Conceptualisation, Methodology, Validation, Writing – review & editing;

Project consortium: Funding acquisition.

Competing interests

The authors declare no competing interests.

Funding

Project Smart Solar System is supported under the umbrella of SOLAR-ERA.NET Cofund 2 by Projektträger Jülich – Forschungszentrum Jülich GmbH – Energie-Technologie-Nachhaltigkeit (ETN 1) and General Secretariat of Research and Innovation (GSRI). SOLAR-ERA.NET is supported by the European Commission within the EU Framework Programme for Research and Innovation HORIZON 2020 (Cofund ERA-NET Action, N° 786483). Funding from the state of North Rhine-Westphalia on the basis of the directive on the granting of funding from the "Programme for the rational use of energy, regenerative energies and energy saving - pro-gres.nrw - programme area innovation".

Acknowledgement

The project consortium of project S3 would like to sincerely thank the Cyprus University of Technology as well as KEAN Soft Drinks Ltd. for their support given in the project.

References

- [1] J. C. Sattler, S. Alexopoulos, R. A. Chico Caminos, J. Mitchell, V. Ruiz, S. Kalogirou, P. Ktistis, C. Teixeira Boura, U. Herrmann, "Dynamic simulation model of a parabolic trough collector system with concrete thermal energy storage for process steam generation", AIP Conference Proceedings, vol. 2126, p. 150007, Jul., 2019, doi: https://doi.org/10.1063/1.5117663
- [2] J. C. Sattler, R. A. Chico Caminos, N. Ürlings, S. Dutta, V. Ruiz, S. Kalogirou, P. Ktistis, R. Agathokleous, C. Jung, S. Alexopoulos, V. Atti, C. Teixeira Boura, U Herrmann, "Operational experience and behaviour of a parabolic trough collector system with concrete thermal energy storage for process steam generation in Cyprus", AIP Conference Proceedings, vol. 2303, p. 140004, Dec., 2020, doi: https://doi.org/10.1063/5.0029278
- [3] J. C. Sattler, R. A. Chico Caminos, V. Atti, N. Ürlings, S. Dutta, V. Ruiz, S. Kalogirou, P. Ktistis, R. Agathokleous, S. Alexopoulos, C. Teixeira Boura, U. Herrmann, "Dynamic simulation tool for a performance evaluation and sensitivity study of a parabolic trough collector system with concrete thermal energy storage", AIP Conference Proceedings, vol. 2303, p. 160004, Dec., 2020, doi: https://doi.org/10.1063/5.0029277
- [4] J. C. Sattler, V. Atti, S. Alexopoulos, C. Teixeira Boura, U. Herrmann, S. Dutta, and I. Kioutsioukis, "DNI Forecast Tool for the Smart Operation of a Parabolic Trough

Collector System with Concrete Thermal Energy Storage", in: SolarPACES 2022, 26 – 30 September 2022, Albuquerque, New Mexico, USA.

- [5] The MathWorks[®] Inc., https://de.mathworks.com/help/deeplearning/ug/time-seriesforecasting-using-deep-learning.html (last opened on 2 October 2022)
- [6] Z. Bounoua, A. Mechagrane, "Prediction of daily global horizontal solar irradiation using artificial neural networks and commonly measured meteorological parameters", *AIP Conf. Proc.* 20 December 2018, 2056 (1): 020024, doi: https://doi.org/10.1063/1.5084997
- [7] Z. Bounoua, A. Mechaqrane, "Estimation of Daily Direct Normal Solar Irradiation Using Machine-Learning Methods", In: Bennani, S., Lakhrissi, Y., Khaissidi, G., Mansouri, A., Khamlichi, Y. (eds) WITS 2020, 2022, Lecture Notes in Electrical Engineering, vol 745. Springer, Singapore, doi: https://doi.org/10.1007/978-981-33-6893-4_58
- [8] M. Khalifa Boutahir, Y. Farhaoui, M. Azrour, I. Zeroual, A. El Allaoui, "Effect of Feature Selection on the Prediction of Direct Normal Irradiance.", Big data mining and analytics, 5 (2022).:309-317, doi: 10.26599/bdma.2022.9020003
- [9] B. Nouri, P. Kuhn, S. Wilbert, C. Prahl, R. Pitz-Paal, P. Blanc, T. Schmidt, Z. Yasser, L. Ramirez Santigosa, D. Heineman, "Nowcasting of DNI maps for the solar field based on voxel carving and individual 3D cloud objects from all sky images", AIP Conf. Proc., 8 November 2018, 2033 (1): 190011, doi: https://doi.org/10.1063/1.5067196
- [10] M. Deru, A. Ndiaye, "Deep Learning mit TensorFlow, Keras und TensorFlow.js", 2., aktualisierte und erweiterte Auflage 2020, 1. Nachdruck 2023, Rheinwerk Verlag, Bonn 2020, Germany, Ch. 3, p. 71, ISBN: 978-3-8362-7425-8