Author Accepted Manuscript (AAM)

© 2025 IEEE. Personal use of this material is permitted.
Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, or reuse of any copyrighted component of this work in other works.

This is the **Author Accepted Manuscript** of the following paper:

Tancredi Testasecca, Patrizia Ribino, Giacomo Potenza, Cristina Baglivo, Marco Beccali and Marina Bonomolo

A comparison between indoor daylight prediction methods:
Daysim simulation and LSTM Models

Published in the proceedings of:
2025 IEEE International Conference on Environment and Electrical Engineering and 2025 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)

The final published version of this paper is available in **IEEE Xplore** at:

https://ieeexplore.ieee.org/document/11169147/

https://doi.org/10.1109/EEEIC/ICPSEurope64998.2025.11169147

This version is the **accepted manuscript**, prior to IEEE copy-editing, typesetting, and publication, and **is not the Version of Record**.

A comparison between indoor daylight prediction methods: Daysim simulation and LSTM Models

Tancredi Testasecca
Department of Engineering
University of Palermo
Palermo, Italy
tancredi.testasecca@unipa.it

Patrizia Ribino
ICAR
National Research Council
Palermo, Italy
patrizia.ribino@icar.cnr.it

Giacomo Potenza
ICAR
National Research Council
Palermo, Italy
giacomo.potenza@icar.cnr.it

Cristina Baglivo
Department of Engineering
University of Salento
Palermo, Italy
cristina.baglivo@unisalento.it

Marco Beccali
Department of Engineering
University of Palermo
Palermo, Italy
marco.beccali@unipa.it

Marina Bonomolo
Department of Engineering
University of Palermo
Palermo, Italy
marina.bonomolo@unipa.it

***Abstract*—The simulation software is widely used in many research and professional fields. Certainly, they are very useful tools, but it could be a difference between the expected results and the simulated ones. This paper aims to validate methods for calculating natural lighting contributions comparing the results of simulated data calculated by using the climatic file, the results of simulated data calculated by using real climate data, and results calculated by using machine learning methods.**

The study was applied for a case study located at the Campus of the University of Palermo. The measurement campaign ran from April to September, collecting more than 24,000 measurements of solar radiation and illuminance. The results showed that the machine learning approach provided the most accurate predictions, with the LSTM model achieving an R^2 of 0.99, validated using approximately 700 measurements. In contrast, the Daysim simulation software reached an R^2 of 0.78 over the same validation period as LSTM. However, machine learning methods require extensive datasets for training, unlike Daysim. Nevertheless, they can be a valuable tool for estimating indoor daylight distributions using only external solar radiation data.

Keywords—Daylight, machine learning, Daysim, simulation tools

I. INTRODUCTION

The assessment of indoor lighting availability can be useful to evaluate the visual and thermal comfort conditions in indoor space, to design Daylight-Linked control systems, and to estimate the energy consumption related to the lighting systems [1]. Many studies applied different methods to do this. In [2] authors reported a review of estimation tools for potential energy saving from integrating daylight with artificial lighting and divided them in four groups: software simulation, which is fast but need validation; field measurement, which is more accurate but takes longer time and not cost effective; empirical formulae from previous research and standards; and estimation algorithms considering various influential factors in details such as lighting control properties and illuminance contributed from surrounding artificial lightings, etc. Of course, early methods of daylight estimation, including diagrams [3][4], protractors [5][6], rules-of-thumb [7] and analytical formulae [8][9], suffer from over-simplification, limiting their practical and wide application [10]. Over the past years, developers of building performance simulations have been pioneering a variety of daylight simulation tools (DSTs) of growing sophistication and accuracy [11], ensuring their reliability for practitioners and research community as well. DSTs became well-incorporated into the architectural design process, as new family of tools is being developed [12][13] that depends on Daylight Coefficient [14], Climate-Based Daylight Modelling (CBDM) [15][16], 3- and 5-Phase Methods [17][18][19]. The selection of the methods and of the software to be used depends on the accuracy of the data, the availability of the machines (e.g. computer and server), and the time. Regarding the accuracy, the validation of the software is important in order to calculate the errors of simulation and the difference between the simulated data and the expected ones [20]. Obviously, as for many fields, in last decades many researches applied Machine Learning to estimated daylight distribution. In [3] the authors performed a review of the current use of machine learning tools for daylighting design and control and found that 22 of the studies used supervised neural network algorithms to optimize daylight availability, while the remaining 3 studies used multiple linear regression or density-based clustering algorithm. Among the 22 studies utilising Artificial Neural Network (ANN), eighteen used ANN only, and the remaining four conducted a comparative analysis of several ML methods, including ANN. The number of hidden layers and the training algorithms for ANN was different from study to study. It was found that most studies used feed-forward multi-layer perception, and, regarding the training method, most studies adopted a gradient-descent training approach. Arbab et al. [6] compared the performance of RF, DT, Support Vector Regression and ANN models in predicting daylighting and energy performance for office space in Tehran, Iran. It was reported that ANN models produced more accurate results in predicting illuminance. On the other hand, only few studies considered both physics-based and data-driven approach and presented the best-performing algorithm. It is worth noting that presenting a comparative analysis of various algorithms should be viewed as the best practice mainly because it increases the possibility of comparing models developed across multiple studies subsequently increasing the chances of identifying the best possible algorithms for daylight tasks. Of course, ANNs are one of the most applied technics that have been used to predict illuminance values [10] [11] [12] [13] [14] [15] [16] [17] [18] [21] [22] [23], but as well other parameters related to the visual conditions: Useful Daylight Illuminance [24] [25-27], Daylight Autonomy (DA) [4] [28-29], Daylight Glare Probability [30] [31], Annual Sunlight Exposure [32], Daylight Factor [33], etc. Among the various machine learning techniques currently available, Long Short-Term Memory (LSTM) networks have emerged as a particularly effective approach for time-series prediction, due to their capacity to model long-range temporal dependencies in sequential data. Their architecture makes them especially suitable for applications involving complex and non-linear temporal patterns. In this study, LSTM networks are employed for the prediction of illuminance levels, aiming to enhance the accuracy and reliability of lighting forecasts by leveraging the strengths of deep learning in handling dynamic environmental data. The objective is to compare the performance of this advanced machine learning approach with traditional physics-based simulation methods, implemented via the Daysim [34] software, to evaluate their respective accuracy and practical applicability in lighting analysis. The accuracy of both methods is assessed using over six months of measured data collected from a real-world case study. An additional contribution of this study is the comparison between Daysim simulations based on standard weather datasets available in the literature and those using real, site-specific weather measurements. This study will mainly focus on identifying the most accurate method for simulating indoor illuminance at multiple spatial points. These outcomes are particularly relevant for applications such as Digital Twins, where accurate, real-time illuminance estimation based solely on solar radiation data could enable more efficient and adaptive building control strategies.

II. Methodology

In this section the methodology applied is explained. As is possible to see in Table 1, the methodology is based on the comparison of four different set datasets. The results are related to the illuminance values calculated on a workplane.

TABLE I. Dataset Description

Tool	ID	Input	Output
Daysim	D_{wea}	Climate data file radiation	Simulated illuminance
Daysim	D_m	Real climatic data	Simulated illuminance
ML	ML	Real climatic data	Real illuminance (learning)
			Predicted illuminance
Real	R	Real climatic data	Real illuminance values

The first set of illuminance simulated by using the software Daysim. In this case the climate data file has been used. The second set of simulated data has been generated by using as input, the real climate data. The third database includes data generated thanks to the application of Machine Learning algorithms. These latter consider real radiation data as input and real illuminance values as output. Finally, the all the dataset has been compared with real data.

A. Daysim modelling

Daysim is a simulation tool based on Radiance, designed to calculate dynamic indicators of daylight such as DA, Useful Daylight Illuminance, Annual Sunlight Exposure, Spatial Daylight Autonomy, and illuminance values. In the case presented in this manuscript, authors used Sketchup software to develop geometrical models and to fix the point where the indices values were calculated. Unlike static simulations, Daysim considers local weather conditions and the temporal behavior of daylight throughout the year. The required inputs are a geometric model of the building in Radiance-compatible format, surface reflectance data, and the geographic location with its corresponding weather file (usually TMY .wea or .epw format). Additionally, the specifications for openings such as windows and skylights should be included, along with the possible presence of light control devices like curtains or sunshades.

B. Long Short-Term Memory Networks

LSTM networks are a specialized type of Recurrent Neural Network (RNN) designed to effectively model and predict sequential data with temporal dependencies. Unlike traditional RNNs, which often struggle with vanishing or exploding gradients during training, LSTMs incorporate a gating mechanism—comprising input, output, and forget gates—that enables them to retain and selectively update relevant information over extended time periods. This capability makes LSTM networks particularly well-suited for time-series forecasting tasks. In the context of illuminance prediction, where light levels vary dynamically in response to environmental factors such as solar position, weather conditions, and building orientation, capturing these temporal variations is essential for accurate forecasting. Differently to traditional physics-based simulation methods, such as those based on Radiance's daylighting algorithms, LSTM networks offer a data-driven alternative that can learn complex, non-linear patterns from historical measurements. As seen in Fig. 1, in its core, an LSTM cell has a memory, called the cell state, which carries information along the sequence. This memory gets updated at each time step, and the cell decides what to keep, what to add, and what to discard. At each time step, an LSTM unit receives three main inputs: the current input, the output at the previous point in time, known as the previous hidden state, and the last state cell Ct-1(i.e., the long term memory of the network). The LSTM unit processes this information through three gates. **Forget Gate** – This gate looks at the current input and the past state and decides what information is no longer useful. **Input Gate** – This gate decides what new information should be added to the memory. **Output Gate** – After the memory is updated, the output gate decides what part of the memory should influence the output (and the next step in the sequence).

III. Case Study

A. Solarlab description and experimental setup

SolarLab is a laboratory located on the third floor of Building 9 at the University of Palermo. It has previously been used in studies focused on visual comfort [35] and thermal comfort [36]. However, to ensure this study remains self-contained, a brief description of the room is provided below.

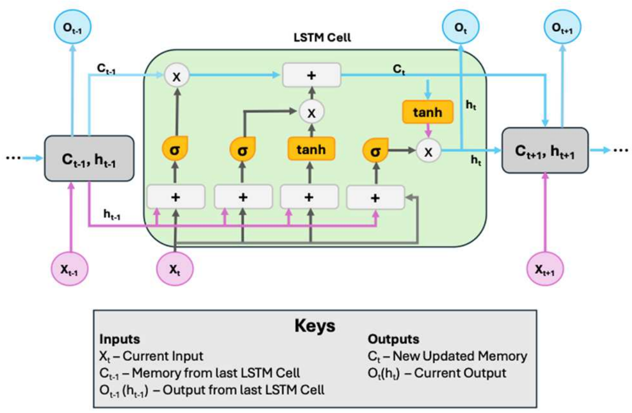

Fig. 1. Functional diagram of an LSTM cell

The room has a C-shaped layout and is bordered by two non-air-conditioned rooms and one air-conditioned space. The only glazed façade is on the south-east side where four double-glazed windows are located (2.40 m in width and 2.90 m in height). Additionally, a solar shading structure measuring 19 m in width and 2.7 m in depth is installed on this façade, covering the exposed south-east surface. Despite the presence of the solar shading system, which effectively blocks direct solar radiation, episodes of high illuminance were observed during morning hours. The experimental setup includes sensors designed to measure illuminance, luminance, and electrical consumption under controlled conditions. These sensors are connected to a Yokogawa data acquisition system, which transfers the recorded signals to a Windows-based PC. Data collection and visualization were carried out using the LabVIEW platform [37]. More details of the experimental setup are available in previous works [38]. The data used in this study were collected using LabVIEW every 5 minutes and include illuminance measurements on the desk and ceiling. The luxmeter was positioned as shown in Fig. 2, (the red circle).

B. *SolarLab Daysim model*

The Daysim model of the SolarLab is shown in Fig. 2 (the ceiling part of the model is set to *"hidden"* to visualize the virtual sensor mesh). The measurement of illuminance used to conduct this study are simulated in the point circled in red (the same position of the sensor used to measure the real data).

Regarding material properties, details on the reflection coefficient assumed for each surface are reported in Table 2.

TABLE II. SURFACES REFLECTANCE COEFFICIENTS ASSUMED

Surface	Material	Reflectance
Interior Wall	White paint	0.8
Exterior Façade	Light yellow paint	0.5
Solar Shelter	Black plastic	0.2
Floor	Marble	0.3
Ceiling	White paint	0.8

C. *Dataset description*

The measurement campaign (R dataset) was carried out between April and September 2024. The illuminance values on the workplane and global horizontal irradiance (GHI) data were collected at five-minute intervals, resulting in over 24,000 measurements, despite the presence of some gaps in the dataset.

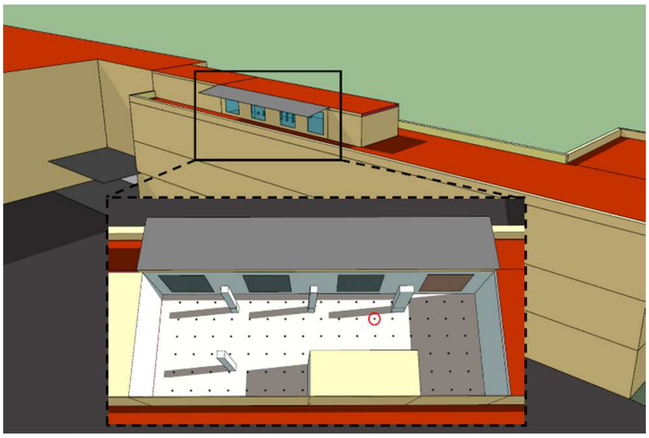

Fig. 2. Daysim model of the SolarLab

The input features used for the LSTM model consist of a set of time-series data derived from both direct measurements and physically-based estimations relevant to solar irradiance and geometry. The only directly measured variable is the GHI, recorded using a pyranometer installed at SolarLab. All other features are obtained through physical and mathematical models based on solar geometry. In particular, pvlib library [39] as used. These include estimates of the direct and diffuse components of solar radiation derived from GHI using two different modeling approaches, identified by the suffixes "disc" and "dirint, (i.e., DNI_{disc}, DNI_{dirint}, DHI_{disc}, DHI_{dirint}, BHI_{disc}, BHI_{dirint}). Specifically, the features include Direct Normal Irradiance (DNI), which represents the direct radiation on a surface perpendicular to the sun's rays; Diffuse Horizontal Irradiance (DHI), which captures the diffuse component on a horizontal plane; and Beam Horizontal Irradiance, referring to the direct radiation incident on a horizontal surface. Additionally, the solar zenith and solar elevation angles are included, as they are fundamental descriptors of solar position and are essential for accurately modeling irradiance dynamics throughout the day. Daysim simulations required a full year of DNI and diffuse DHI or alternatively, beam horizontal irradiance (BHI). Related to the Dm dataset, DHI was estimated using the Erbs method [40] implemented via the Python pvlib library [39]. It is considered one of the most reliable methods for diffuse radiation modelling. BHI was then calculated as the difference between GHI and DNI. Since the experimental dataset did not span a complete year, missing data were supplemented using the Open-Meteo API [41]. For standard simulations requiring a .wea weather file, D_{wea} dataset, the TMY 2007–2021 dataset for Palermo Boccadifalco was selected as it represents the most recent and geographically relevant dataset for the case study. Finally, model outputs were filtered to align with the time series corresponding to the measured data.

D. *LSTM Model for daylight Illuminance prediction*

In this paper, we implemented a stacked LSTM with two hidden LSTM layers, each containing multiple memory cells, and an output layer for daylight illuminance prediction as shown in Fig. 3.

To assess the predictive power of the input features, an ablation study was conducted. This involved systematically evaluating the contribution of each feature—or group of features—to the overall performance of the LSTM model. By selectively removing individual inputs and observing the corresponding changes in predictive accuracy, the study aimed to identify which features are most influential in enhancing the model's ability to forecast illuminance levels. The evaluation was based on two commonly used performance metrics: the coefficient of determination (R2) and the Root Mean Square Error (RMSE), which together provide a comprehensive assessment of both the accuracy and reliability of the predictions.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 1, 64)	20,480
lstm_1 (LSTM)	(None, 64)	33,024
dense (Dense)	(None, 1)	65
activation (Activation)	(None, 1)	0

Fig. 3. Architecture of the LSTM network

In the experimental setup, the LSTM model was trained using a dataset comprising 24,000 time-series observations, with an additional 700 observations reserved for validation. The training process was conducted over 100 epochs with a batch size of 64, allowing the model to iteratively optimize its parameters while monitoring validation performance to prevent overfitting.

IV. Results and Discussion

In this section, the results of each modelling approach are presented in detail, followed by a comparative discussion of their respective accuracy and limitations.

A. *Daysim Results*

As a result, the simulation outputs, shown in Fig. 4, differed notably from the measured dataset (R) at the workplane (Plane 1*). During the week in April, for instance, both Daysim simulations (D_m and D_{wea}) underestimated illuminance on the measurement plane. On 14th April, the maximum simulated illuminance ranged between 1000 lx (D_{wea}) and 1120 lx (D_m), whereas the real measurements (R) exceeded 1500 lx. Nonetheless, D_m produced early morning peaks that more closely aligned with R. On 4th of July, D_m successfully reproduced real illuminance levels during the morning hours. However, in the afternoon, it overestimated illuminance by up to 1.5 times compared to R.

This discrepancy may be attributed to two main factors: (1) an overestimation of the contribution from diffuse radiation within the room, and (2) inherent assumptions in the Daysim modelling methodology, as this error was observed in both D_m and D_{wea}. Neither model was able to reproduce the sharp peak of 2500 lx recorded at 05:45 on 6 July in the R dataset. Similar trends were observed in September.

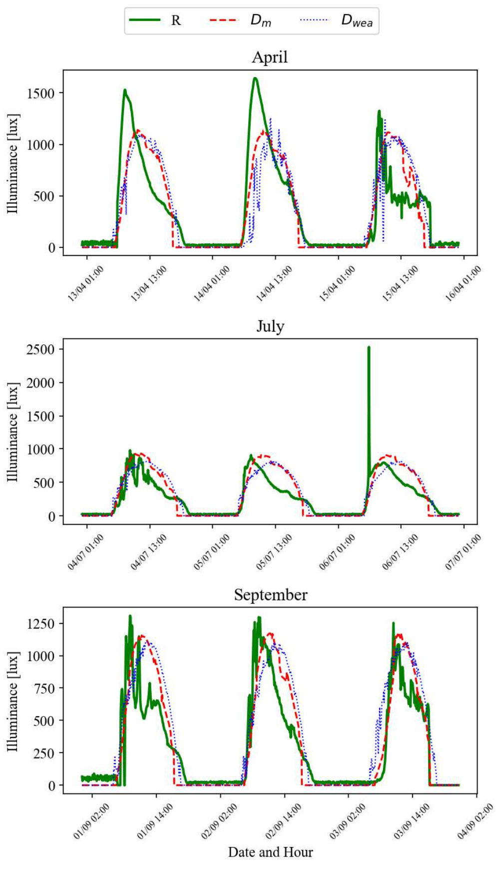

Fig. 4. Daysim model results for different months

The D_m results were again more consistent with the measured values (R) in the morning. It is necessary to remind that, as common, although the geometry of the model accurately replicates the real case study, some uncertainties could exist (e.g. related to the surface reflection coefficients to be assumed).

B. *LSTM Ablation Study Results*

The ablation study consisted of 511 tests and revealed that 88 input features configurations yielded a coefficient of determination $R^2 \geq 0.99$ and an RMSE = 0.002, indicating consistently high predictive performance across different feature sets. These results indicate an exceptionally high level of predictive accuracy, with the model able to explain 99% of the variance in the observed illuminance data. The remarkably low RMSE further confirms that the model's predictions are not only consistent with the true values but also exhibit minimal absolute error. Notably, among these high-performing configurations, the combination of only two features DNI_{dirint} and BHI_{disc} matches the performance of more complex feature sets. This result is particularly significant, as it demonstrates that a minimal, yet physically meaningful subset of input variables can be sufficient for accurate illuminance prediction. The inclusion of DNI_{dirint} captures the direct component of solar radiation as estimated by one of the physical models, while BHI_{disc} represents the direct radiation on a horizontal surface, derived from an alternative model. The strong performance of this reduced feature set suggests that these two variables together encapsulate the most critical information required for modeling daylight illuminance, potentially reducing model complexity and computational cost without sacrificing accuracy. Moreover, from the perspective of deploying the LSTM model in a real-time lighting control system, the configuration with the minimal number of input features is the most suitable. This minimal setup not only reduces computational overhead, making the model less time-consuming to evaluate, but also decreases the dependency on extensive input data, which is advantageous in scenarios where sensor availability or data acquisition may be limited. Therefore, this configuration offers a highly practical and efficient solution for real-time applications without compromising predictive performance. As shown in Fig. 5, the temporal trends of the actual daylight illuminance and the values predicted by the LSTM model are closely aligned. The near-complete overlap of the two curves visually confirms the high predictive accuracy of the model. In contrast, the configuration that yielded the lowest performance in the ablation study consisted of the features DHI_{disc}, DNI_{dirint}, DHI_{dirint}, and solar zenith, which resulted in a R^2 of 0.8869 and RSME of 0.009. While still moderately accurate, this value is significantly lower than those achieved by other configurations, indicating reduced predictive capability as it is also visually confirmed by the trend reported in Fig. 6. This diminished performance may be attributed to the absence of features that directly represent the global or beam components of irradiance on horizontal surfaces, which are particularly relevant to modeling indoor daylight illuminance. Additionally, the presence of redundant or less informative variables, such as both DHI_{disc} and DHI_{dirint}, which provide similar information about diffuse irradiance, may contribute to reduced model generalization, possibly introducing noise into the input data. Although the solar zenith is a critical geometric parameter, it alone may not

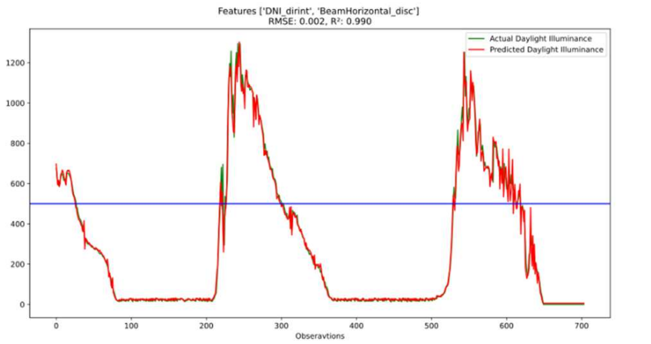

Fig. 5. Trend of actual and predicted daylight illuminance for the best performing configuration.

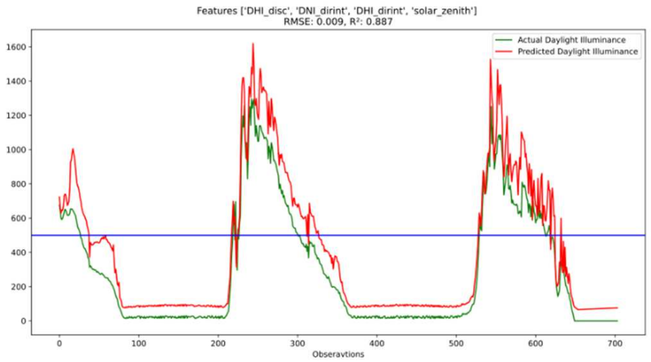

Fig. 6. Trend of actual and predicted daylight illuminance for the worst performing feature configuration.

sufficiently compensate for the lack of surface-incident irradiance data.

These findings underscore the importance of selecting input features not solely based on quantity or variety, but on their physical relevance to the target variable. The performance gap between this configuration and others with fewer but more targeted features highlights the advantage of carefully curated, domain-informed inputs in achieving robust and efficient predictive models.

C. Discussion

Since the LSTM network was only validated on the last 700 measurements, comparing the models on the same time window provides insightful information. As previously said and as represented in Fig. 7, the ML model performed the best, comparing to true measurements (R), with an R^2 of 0.99.

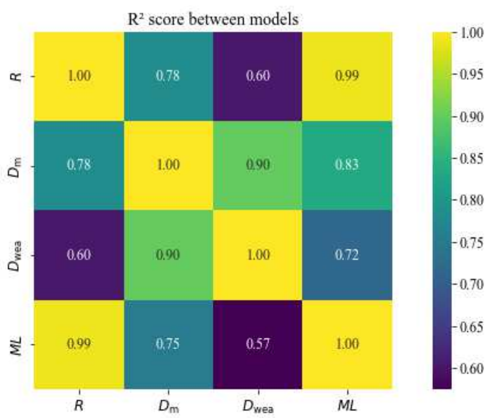

Fig. 7. Heatmap of the R^2 score between models.

Among Daysim-based simulations, simulations using measured climate data (D_m) outperformed baseline climate file simulations (D_m) in having better R^2 (0.78 vs. 0.60) and lower RMSE (165.3 lx vs. 223.5 lx). Additionally, the high similarity between D_m and D_m ($R^2 = 0.90$ and a RMSE = 134.7 lx) suggests the inaccuracy with actual measurements is likely attributable to intrinsic modeling assumptions or calculation. Nevertheless, Daysim remains valuable due to its good performance with a R^2 equal to 0.78. It provides spatial illuminance distribution maps without internal sensor data for the entire year. Relying solely on external solar radiation measurements, it can support DT and smart building applications if few measurements are available. This capability is especially useful for predictive lighting control and comprehensive daylight analyses across entire spaces. One of the advantages to use this approach is that it can be replicated when not-permanent IoT devices (such as sensors) is not available. Indeed, to train ML algorithms it is necessary just to measure the outdoor irradiance and the indoor illuminance for a limited period. Of course, by applying the method used to develop the D_m dataset, only the outdoor radiation measures are necessary.

V. CONCLUSION

This study evaluated and compared two distinct methodologies, physics-based simulation using Daysim and data-driven modeling employing LSTM neural networks, to predict indoor daylight illuminance. The goal was to identify the most accurate and efficient approach suitable for real-time applications, including adaptive lighting control and digital twins. An experimental campaign was conducted at the SolarLab of the University of Palermo. In a three months period, 24,705 illuminance and irradiance measurements were recorded. This dataset was essential for evaluating the Daysim model and training the LSTM network. Additionally, an ablation study on the LSTM model was carried out to assess the best combination of input features. The results demonstrated that the LSTM model achieved exceptional predictive performance ($R^2 > 0.99$) using only two input variables: DNI_dirint and BeamHorizontal_disc. However, this level of accuracy required substantial datasets for training. Daysim simulations, which do not require extensive training data and rely solely on external radiation measurements, produced lower accuracy (R^2 up to 0.46). Nevertheless, Daysim offers the advantage of estimating illuminance distributions throughout an entire spatial area without the necessity of internal sensor data. Future research should focus on employing LSTM and similar deep learning models for real-time illumination control and predictive maintenance in smart buildings. Moreover, hybrid modeling approaches combining physics-based simulations with machine learning corrections represent a promising area for future exploration.

ACKNOWLEDGMENT

This work received funding by European Commission - Next Generation EU - PNRR M4 - C2 - investimento 1.1: Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN) - PRIN 2022 cod. 2022YWW9B8 "Study for a tool for DEsign, COntrol and COmmissioning of Lighting Control systems, Finanziamento PRIN 2022 bando D.D. n. 104 del 02-02- 2022. CUP: B53D23006660006.

References

[1] M. Bonomolo, G. Zizzo, M. Beccali, B. Di Pietra, G. Sciumè, and S. Guarino, "Methods to Predict Energy Use for Lighting Systems: An Overview and an Application on a Real Case Study," in *Proc. 2022 Workshop on Blockchain for Renewables Integration (BLORIN)*, 2022, pp. 194–199.

[2] X. Yu and Y. Su, "Daylight availability assessment and its potential energy saving estimation – A literature review," *Renew. Sustain. Energy Rev.*, vol. 52, pp. 494–503, 2015.

[3] J. Ngarambe, I. Adilkhanova, B. Uwiragiye, and G. Y. Yun, "A review on the current usage of machine learning tools for daylighting design and control," *Build. Environ.*, 2022, Art. no. 109507.

[4] C. L. Lorenz, A. B. Spaeth, C. Bleil De Souza, and M. S. Packianather, "Artificial neural networks for parametric daylight design," *Archit. Sci. Rev.*, vol. 63, no. 2, pp. 210–221, 2020.

[5] C. H. Lin and Y. S. Tsay, "A metamodel based on intermediary features for daylight performance prediction of façade design," *Build. Environ.*, vol. 206, Art. no. 108371, 2021.

[6] M. Arbab, M. Rahbar, and M. Arbab, "A Comparative Study of Artificial Intelligence Models for Predicting Interior Illuminance," *Appl. Artif. Intell.*, vol. 35, no. 5, pp. 373–392, 2021.

[7] E. Nault, P. Moonen, E. Rey, and M. Andersen, "Predictive models for assessing the passive solar and daylight potential of neighborhood designs: A comparative proof-of-concept study," *Build. Environ.*, vol. 116, pp. 1–16, 2017.

[8] T. Yu and H. Zhu, "Hyper-parameter optimization: A review of algorithms and applications," *arXiv preprint arXiv:2003.05689*, 2020.

[9] M. Han et al., "A novel reinforcement learning method for improving occupant comfort via window opening and closing," *Sustain. Cities Soc.*, vol. 61, Art. no. 102247, 2020.

[10] Q. He et al., "Predictive models for daylight performance of general floorplans based on CNN and GAN: a proof-of-concept study," *Build. Environ.*, vol. 206, Art. no. 108346, 2021.

[11] Y. Xu et al., "A two-stage multi-objective optimization method for envelope and energy generation systems of primary and secondary school teaching buildings in China," *Build. Environ.*, vol. 204, Art. no. 108142, 2021.

[12] V. Logar, Ž. Kristl, and I. Škrjanc, "Using a fuzzy black-box model to estimate the indoor illuminance in buildings," *Energy Build.*, vol. 70, pp. 343–351, 2014.

[13] S. G. Colaco, A. M. Colaco, C. P. Kurian, and V. I. George, "An Adaptive predictive framework to online prediction of interior daylight illuminance," in *Proc. Int. Conf. Advances in Energy Conversion Technologies (ICAECT)*, 2014, pp. 174–180.

[14] V. N. Katsanou, M. C. Alexiadis, and D. P. Labridis, "An ANN-based model for the prediction of internal lighting conditions and user actions in non-residential buildings," *J. Build. Perform. Simul.*, vol. 12, no. 5, pp. 700–718, 2019.

[15] M. Ahmad, M. Mourshed, J. L. Hippolyte, Y. Rezgui, and H. Li, "Optimising the scheduled operation of window blinds to enhance occupant comfort," 2015.

[16] J. Hu and S. Olbina, "Illuminance-based slat angle selection model for automated control of split blinds," *Build. Environ.*, vol. 46, no. 3, pp. 786–796, 2011.

[17] H. B. Gunay, W. O'Brien, I. Beausoleil-Morrison, and S. Gilani, "Development and implementation of an adaptive lighting and blinds control algorithm," *Build. Environ.*, vol. 113, pp. 185–199, 2017.

[18] E. N. D. Madias, P. A. Kontaxis, and F. V. Topalis, "Application of multi-objective genetic algorithms to interior lighting optimization," *Energy Build.*, vol. 125, pp. 66–74, 2016.

[19] N. K. Kandasamy, G. Karunagaran, C. Spanos, K. J. Tseng, and B. H. Soong, "Smart lighting system using ANN-IMC for personalized lighting control and daylight harvesting," *Build. Environ.*, vol. 139, pp. 170–180, 2018.

[20] L. Bellia, A. Pedace, and F. Fragliasso, "The impact of the software's choice on dynamic daylight simulations' results: A comparison between Daysim and 3ds Max Design®," *Sol. Energy*, vol. 122, pp. 249–263, 2015.

[21] J. Xie and A. O. Sawyer, "Simulation-assisted data-driven method for glare control with automated shading systems in office buildings," *Build. Environ.*, vol. 196, Art. no. 107801, 2021.

[22] J. Y. Park, T. Dougherty, H. Fritz, and Z. Nagy, "LightLearn: An adaptive and occupant centered controller for lighting based on reinforcement learning," *Build. Environ.*, vol. 147, pp. 397–414, 2019.

[23] Z. Cheng et al., "Satisfaction based Q-learning for integrated lighting and blind control," *Energy Build.*, vol. 127, pp. 43–55, 2016.

[24] Y. Han, L. Shen, and C. Sun, "Developing a parametric morphable annual daylight prediction model with improved generalization capability for the early stages of office building design," *Build. Environ.*, vol. 200, Art. no. 107932, 2021.

[25] A. Razmi, M. Rahbar, and M. Bemanian, "PCA-ANN integrated NSGA-III framework for dormitory building design optimization: Energy efficiency, daylight, and thermal comfort," *Appl. Energy*, vol. 305, Art. no. 117828, 2022.

[26] L. Le-Thanh, H. Nguyen-Thi-Viet, J. Lee, and H. Nguyen-Xuan, "Machine learning-based real-time daylight analysis in buildings," *J. Build. Eng.*, vol. 52, Art. no. 104374, 2022.

[27] Y. Zou, Q. Zhan, and K. Xiang, "A comprehensive method for optimizing the design of a regular architectural space to improve building performance," *Energy Rep.*, vol. 7, pp. 981–996, 2021.

[28] I. Chatzikonstantinou and S. Sariyildiz, "Approximation of simulation-derived visual comfort indicators in office spaces: a comparative study in machine learning," *Archit. Sci. Rev.*, vol. 59, no. 4, pp. 307–322, 2016.

[29] D. Uribe, S. Veraand, and W. Bustamante, "Optimization of complex fenestration systems using an artificial neural network," in *Proc. 51st Int. Conf. Architectural Science Association (ANZAScA)*, 2017, pp. 177–185.

[30] S. M. Yacine, Z. Noureddine, B. E. Piga, and E. Morello, "Developing neural networks to investigate relationships between lighting quality and lighting glare indices," *Energy Procedia*, vol. 122, pp. 799–804, 2017.

[31] A. Wagdy et al., "Open-plan Glare Evaluator (OGE): A Demonstration of a New Glare Prediction Approach Using Machine Learning Algorithms," *arXiv preprint arXiv:1910.05594*, 2019.

[32] H. Nourkojouri, Z. S. Zomorodian, M. Tahsildoost, and Z. Shaghaghian, "A machine-learning framework for daylight and visual comfort assessment in early design stages," *arXiv preprint arXiv:2109.06450*, 2021.

[33] S. Wang, Y. K. Yi, and N. Liu, "Multi-objective optimization (MOO) for high-rise residential buildings' layout centered on daylight, visual, and outdoor thermal metrics in China," *Build. Environ.*, vol. 205, Art. no. 108263, 2021.

[34] C. F. Reinhart, "DAYSIM Version 2.1," January 2005. https://www.researchgate.net/publication/44072319_DAYSIM_Version_21, Accessed: May 2025.

[35] M. Beccali, M. Bonomolo, A. Galatioto, M. G. Ippolito, and G. Zizzo, "A laboratory setup for the evaluation of the effects of BACS and TBM systems on lighting," in *Proc. 2015 Int. Conf. Renewable Energy Research and Applications (ICRERA)*, 2015, pp. 1388–1393.

[36] M. Bonomolo, L. Cirrincione, G. Scaccianoce, T. Testasecca, and M. Beccali, "Development of a Methodology for Assessing Indoor Punctual Thermal Comfort Conditions," in *Proc. 2024 IEEE Int. Conf. Environment and Electrical Engineering and 2024 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, 2024, pp. 1–6.

[37] National Instruments, "https://www.ni.com/en.html," Accessed: May 2025.

[38] M. Bonomolo et al., "Analisys on the feasibility to modulate lighting systems load: the case study of Palermo," in *Proc. 2022 IEEE Int. Conf. Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, 2022, pp. 1–6.

[39] K. S. Anderson, C. W. Hansen, W. F. Holmgren, A. R. Jensen, M. A. Mikofski, and A. Driesse, "pvlib python: 2023 project update," *The Journal of Open Source Software*, vol. 8, no. 92, p. 5994, Dec. 2023, doi: 10.21105/joss.05994.

[40] D. G. Erbs, S. A. Klein and J. A. Duffie, Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Solar Energy 28(4), pp 293-302, 1982. Eq. 1

[41] "□Free Open-Source Weather API | Open-Meteo.com." https://open-meteo.com/, Accessedd: May 2025.