Machine Learning in Design Exploration:

An Investigation of the Sensitivities of ANN-based Daylight Predictions

Clara-Larissa Lorenz[1], A. Benjamin Spaeth[1], Clarice Bleil De Souza[1], Michael Packianather[2]

[1]Cardiff University, Welsh School of Architecture
[2]Cardiff University, School of Engineering

Abstract. The use of Artificial Neural Networks (ANNs) promises greater efficiency in the assessment of daylight situations than simulations. With the daylight factor under scrutiny and the recent adaptation of climate-based daylight metrics in British and European buildings standards, ANNs provide a possibility for instantaneous feedback on otherwise time-consuming performance metrices. This study demonstrates the application of ANNs as prediction systems in design exploration. A specific focus of the research is the flexibility of ANNs, their reliability and sensitivity to changes.

Keywords: Artificial neural networks, atria, climate-based daylight modeling, daylight autonomy, daylight performance, parametric design

1 Introduction

Exposure to daylight is a key determinant of human well-being and health [1]. As a measure of ensuring daylight quality in buildings, climate-based metrics have been introduced in building design standards (BS EN 17037:2018). Two climate-based metrices that have been developed are the daylight autonomy metric (DA) and the spatial daylight autonomy metric (sDA). DA was developed for energy considerations in sustainable building design and defines the number of occupied hours in a year in which an illuminance threshold (of typically 300 lux) can be maintained by daylight alone [2]. sDA was developed as a threshold to ensure occupant satisfaction and well-being and defines the percentage of space that can meet a DA threshold of 50% [3]. With the impetus towards annual simulations of daylight however, design exploration has become much more time consuming. Research into emulators and surrogate models has addressed the challenges of computational burden by proposing various models to mimic and simplify time-consuming processes. One of the proposed models is Artificial Neural Networks (ANNs), which can be used to replicate the behavior of simulation programs without degrading their accuracy [4].

ANNs are often formulated as a construct of neurons comprising an input layer, a hidden layer and an output layer (Fig. 1). The input layer is provided with data describing a problem and the output layer with the solution. The data that characterizes the problem is referred to as training data and the data describing the solution is referred to as training target. Typically using gradient-descent and Newton-based algorithms between all neurons, ANN models are trained to form relationships between the provided data in such a way that the model can solve new problems [5]. Trained ANNs have been shown to successfully model complex functions and non-linear problems [6].

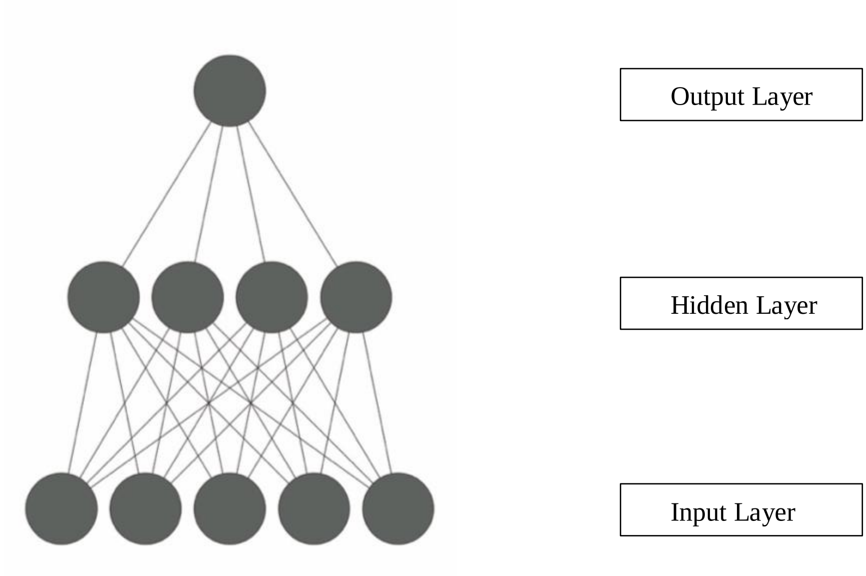

Fig. 1**.** Representation of a network architecture.

Research has also looked into using ANNs to model daylight. Kazanasmaz et al. [7] for instance used an ANN prediction model to determine daylight illuminances in office buildings. The results were compared to field measurements, ultimately showing that the model could successfully predict daylight. Daylight illuminances have also been predicted using ANNs for the automation of split blinds [8]. Other studies were able to use ANNs to predict energy savings as a result of daylighting [9], model sky luminance and luminous efficacy, as well as irradiance data [10–12].

In terms of climate-based metrics, ANNs have been used to classify the UDI (Useful Daylight Illuminances) metric, showing a better performance than support vector machines in predicting daylight [13]. The predictions were done for hourly data which meant that weather data had to be included in the training data. The necessity of including weather data was bypassed in two studies that trained ANN models directly to predict annual simulation results [14-15]. These papers serve as foundations for the current work. This study predicted the DA and sDA metric to explore design options for a central atrium. Specifically, we evaluated the sensitivities of ANN prediction accuracies to different daylight simulation setting (i.e., ambient bounces, sensor-point spacing) and the number of ANN training samples. By identifying a minimum number of training samples, we also investigated the efficacy of these models in terms of the overall reduceable simulation time. The research thus aims to provide guidance on the application and integration of ANNs and serves as a display of the efficiency of ANN based daylight modeling.

2 Methodology

ANNs were used to predict daylight performance in a building design. The corresponding ANN-integrated process is shown in Fig. 2. First, a space of possible design solutions had to be identified. For this purpose, a parametric model was built in Grasshopper, where the design variables were selected and ranges specified in order to generate the necessary design alternatives for a base case design scenario. The respective design variables of the solution space will be specified in the following sections. In a second step, samples from the solution space were randomly identified. Daylight simulations were run on the samples and data was extracted as training data for our ANN models. The training data describes the design changes and the resulting daylight levels. Third, ANN models were developed by passing the data to a network architecture. The topology that achieved the lowest Mean Squared Error (MSE) was selected for predicting daylight autonomy. Before being used to predict daylight for all solutions in the solution space, the ANN models were validated against unseen simulation results. In the case of insufficient accuracies, additional training samples were taken from the design solution space and the networks were retrained.

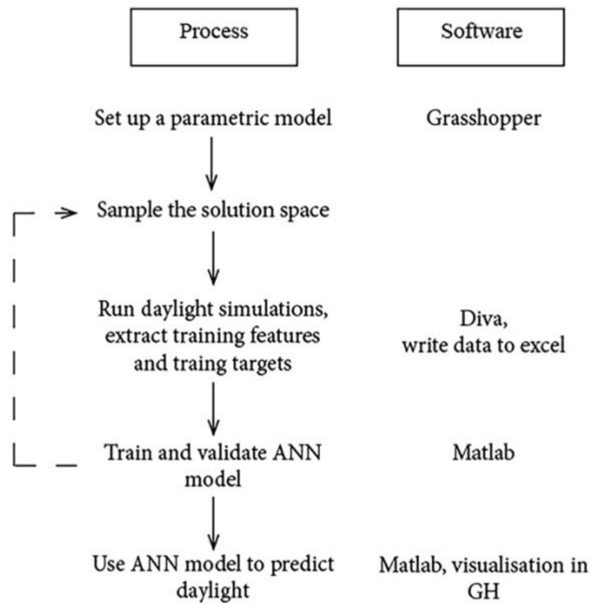

Fig. 2. ANN-integrated workflow

The research was divided into three parts, each with a distinct setup and focus for exploring design solutions for a central atrium. The first part focused on evaluating the sensitivities of prediction accuracies to daylight simulation settings, specifically Radiance Parameter Settings, the inclusion of surrounding buildings and the spacing of sensor points in the daylight simulation model. The second part focused on sampling a design solution space for training data and validating the ANNs for DA and sDA predictions. In the third part, the solution space was extended by an additional design variable. The sampling of training data and resulting accuracies were therefore reevaluated for this solution space. The following section details the design and ANN setup for each part.

2.1 Experimental Setup: ANN Sensitivities to Daylight Simulation Settings

Design Setup: The Katharinen School in Hamburg was used as a base case to explore design solutions for a central atrium. The dimension of the atrium base was selected as a design variable and scaled in order to generate 'V'-shaped atria with varying splay angles (Fig. 3). The atrium base dimension (in m^2) was scaled by a factor of .05 to 1 in increments of .05, resulting in 20 possible solutions.

Simulation Settings: Daylight simulations were undertaken in Diva. The simulation settings were varied, producing different results to train the neural networks with. A calculation plane was set at a height of .8m above the floor, and sensor points for the calculation of daylight levels were distributed with .6m and 1.2m spacing across the plane. Diva calculates daylight through backward ray-tracing and the ambient bounces (ab) determine the number of bounces undertaken to trace sources of light. The settings were set to 2 ab and 4 ab. Simulations were run twice, once with and once without surrounding buildings in the model. This was done so as to investigate whether including surrounding buildings would negatively affect the accuracy of ANN predictions.

ANN training: To train the ANN model, a list of features describing the design changes was extracted. The features included the coordinates of all sensor points, their distance to the facade and the atrium, the direction of the atrium (from the sensor point), the atrium dimension and the glazing area of the atrium well. All data was normalized between -1 and 1 before it was passed to a back-propagation feedforward network. The Levenberg-Marquardt training algorithm was employed [16] and the training epochs were set to 150. A custom script was used to optimize the network architecture between 1-20 neurons in the hidden layer. The MSE was used to measure how well the network was able to fit the provided input features with the provided DA results and predict daylight for individual unseen sensor points. The MSE was calculated for the training data set, which was subdivided into a training, validation and test subset at the ratio of 65:20:15. An ensemble of ten networks was trained for every network architecture and the networks with the lowest MSE were selected for predicting daylight.

ANN training: The training data was extracted from 18 out of the 20 solutions. The ANN models were used to predict daylight for the 2 remaining scenarios. The accuracy of prediction was measured using the Mean Absolute Error (MAE), which describes the absolute difference between the simulated and predict DA values.

$$\mathrm{MAE} = \frac{\sum_{t=1}^{n}|P_t - T_t|}{n} \quad (1)$$

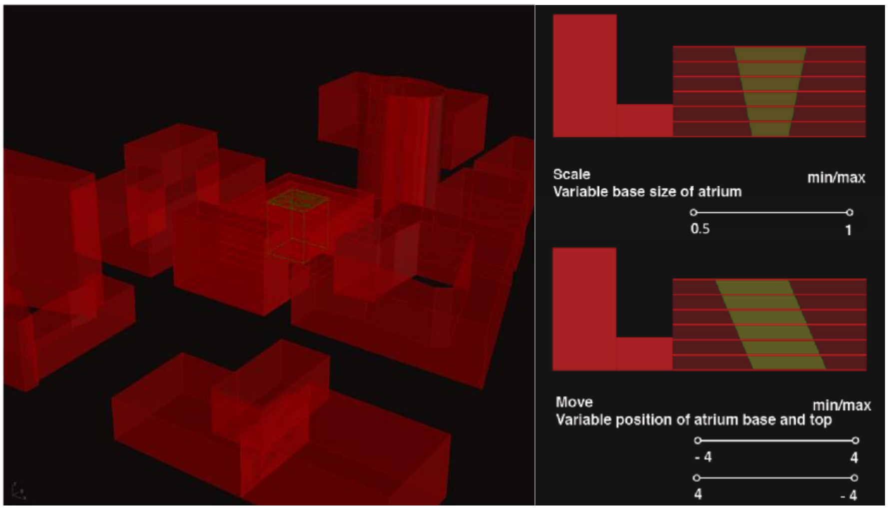

Fig. 3. The Katharinen School. The central atrium of the building is highlighted in green.

2.2 Experimental Setup: Sampling of Design Solution Spaces

Design Setup: For this part of the study, ANNs were integrated as a prediction system in a larger solution space of 54 possible solutions. The design variables are shown in Fig. 3. The atrium base was scaled by a factor of .5 to 1, in increments of .1, generating 6 possible solutions with dimensions ranging from 56.25 to 225m^2. Additionally, the central atrium was slanted by moving the atrium base and atrium top in opposite direction, 1 unit at a time. This generated another 9 possible solutions, resulting in a 6 by 9 matrix of 54 combinations.

Simulation Settings: Diva simulations were settings were set to 6 ambient bounces and the sensor point spacing to .6m. The daylight model included the surrounding buildings.

ANN training: Additional training features were extracted for this solution space. Optimizing the selection of input features reduced the bias of certain features, improved prediction accuracy, and reduced the overall training time. A detailed description of the selection process is beyond the scope of this paper and will therefore not be discussed any further. ANN prediction accuracies were investigated for four training data sets comprising different sample sizes. The simulations from which the training data was extracted are highlighted in Fig. 4 and 5. Training data set A consisted of data from 18 simulations and 71.167 samples, training data set B of 12 simulations and 50.049 samples, training data set C of 9 simulations and 33.464 samples and training data set D of 6 simulations and 25.088 samples. ANN architectures with up to 40 neurons in the hidden layer were tested and the training epochs were increased to 200.

ANN validation: A full-factorial validation was done by comparing all predicted daylight levels to the simulation results. Prediction errors were calculated for the DA and sDA metrics. With regard to DA, two further measures were included aside from the MAE: the Mean Biased Error (MBE), as a measure for the bias of predictions, and the Root Mean Squared Error (RMSE), as a measure for the robustness of predictions [17].

$$\text{MBE} = \frac{\sum_{t=1}^{n}(P_t - T_t)}{n} \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(P_t - T_t)^2}{n}} \quad (3)$$

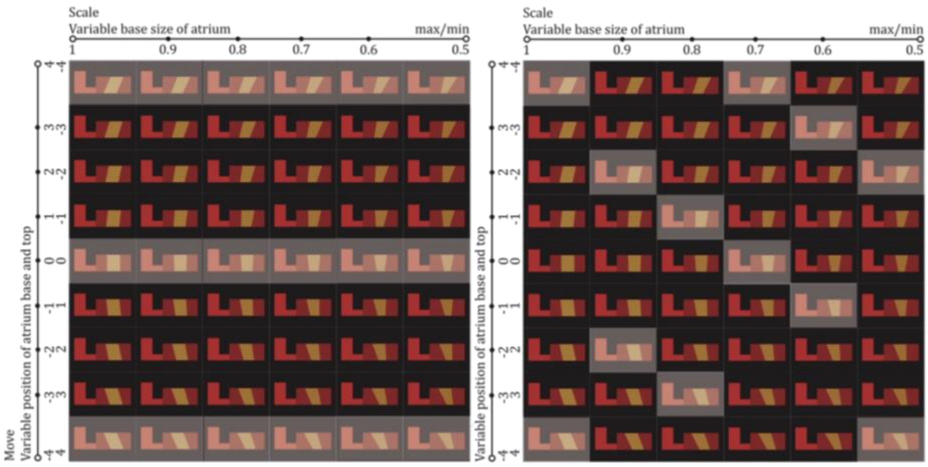

Fig. 4. Training data set A and B (left to right): The matrix represents the solution space of 54 design variants. The simulations from which the training data was extracted have been highlighted. Training data set A comprised 18 simulations – 3 from every column of the matrix. For training data set B, 2 simulations were selected from every column of the matrix.

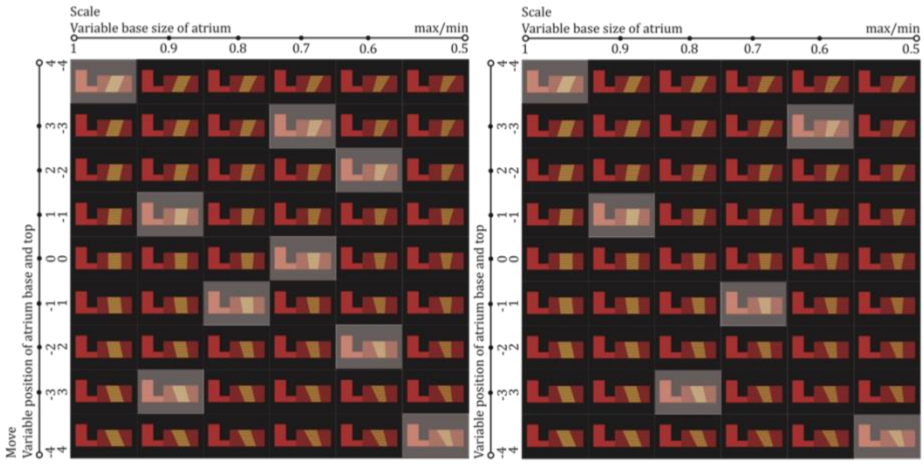

Fig. 5. Training data set C and D (left to right): Training data set C comprised 9 simulations, 1 simulation from every row of the matrix. Training data set D comprised 6 simulations, 1 simulation from every column of the matrix.

Following the validation of DA prediction accuracies, the same was done for the sDA metric. By comparing the simulated and predicted sDA results, we examined how closely the neural networks were able to predict the DA 50 threshold.

2.3 Experimental Setup: Predicting sDA for a Larger Solution Space

Design Setup: Another dimension was added to the solution space by incorporating window-to-wall ratio (WWR) distribution of the atrium well walls as a design variable. The atrium well had been fully glazed in the prior solutions. In order to increase daylight levels on the ground floor, the glazing area was reduced, thus increasing the reflected light within the atrium. Three options were selected for the distribution of window-to-wall ratios, based on recommendations from research in the field [18–20]. The distribution of WWR of the 6-storey building from the 6^{th} to the ground floor were as follows: the first option had a WWR distribution of 50, 60, 70, 80, 90, 100%, the second option had a WWR distribution of 20, 30, 40, 50, 60, 100% and the third option has a WWR distribution of 20, 35, 50, 65, 80, 100%. These options were applied to the previously generated solution space, adding another 3 matrices of 54 solutions each to the design space. Simulation and ANN training settings were kept the same as in the previous section. In terms of testing prediction accuracies, a partial-factorial validation was undertaken with a select number of simulation results.

3 Results

3.1 Simulation Settings and ANN Prediction Accuracies for DA

The average difference of the simulated and predicted DA for all sensor points of 2 design variants of the scaled atrium are shown in Fig. 6. For simulation settings with 2 ambient bounces in a simulation model without the surrounding buildings, this difference was 2.12 DA. As a reminder, DA values range from 0 to 100% and 1 unit refers to 1% of occupied hours per year. Thus, the result can be considered highly accurate. The results became even more accurate after increasing the number of ambient bounces, in effect reducing the errors to 1.07 MAE. This stands to reason as there were now fewer fluctuations in the daylight simulation results. The result did not worsen after increasing the spacing between sensor points, although this meant that the number of samples used to train the networks were reduced. Aside from the number of ambient bounces, a factor that negatively affected the prediction was the inclusion of surrounding buildings in the simulation models, after which ANN prediction errors increased to 1.8 MAE. Nonetheless, predictions generally remained close to the simulated DA results.

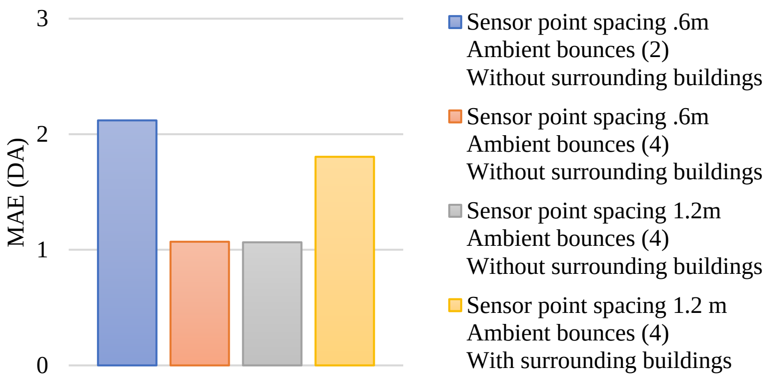

Fig. 6**.** ANN sensitivities to simulation settings

3.2 Data Sampling

Fig. 7 compares the prediction accuracies of ANNs trained on data extracted from 18 (data set A), 12 (data set B), 9 (data set C) and 6 (data set D) daylight simulations. The MSE of the trained networks is shown alongside the prediction errors on the 36 (data set A) to 48 (data set D) unseen cases. Interestingly, while the MSE remained low for all training data sets, the MAE of the predictions increased by 314% from data set B to C (12 to 9 simulations) and by 916% from data set C to D (9 to 6 simulations). The MSE values ranged from .0005 (C) to .0009 (A), whileas the MAE ranged from .66 and .68 (data sets A and B with 18 and 12 simulations) to 2.32 (data set C with 9 simulations) and 6.23 (data set D with 6 simulations).

Although the MSE was low for all training data sets, it should be noted that the MSE is primarily an indicator for the ability of the ANN model to approximate a function for data it has been provided with. As such, the MSE indicates how well the network would be able to predict daylight for sensor points of design solutions which are part of the training data. As can be seen in the results, the MSE does not necessarily give insight into prediction errors on unseen design variants: The models that were trained on data sets C (9 simulations) and D (6 simulations) were able to fit the data very well, but were presumably unable to map daylight performance of the design landscape. This becomes clear when looking at the MAE, MBE and RMSE, which were considerably higher for training sets C (9 simulations) and D (6 simulations), as compared to sets A (18 simulations) and B (12 simulations). Nonetheless, it was yet unclear how the errors were distributed across the solution space and how they affected the overall assessment of daylight performance for each design variant. Therefore, in the next part, we take a look at the sDA accuracies for all predicted design variants individually.

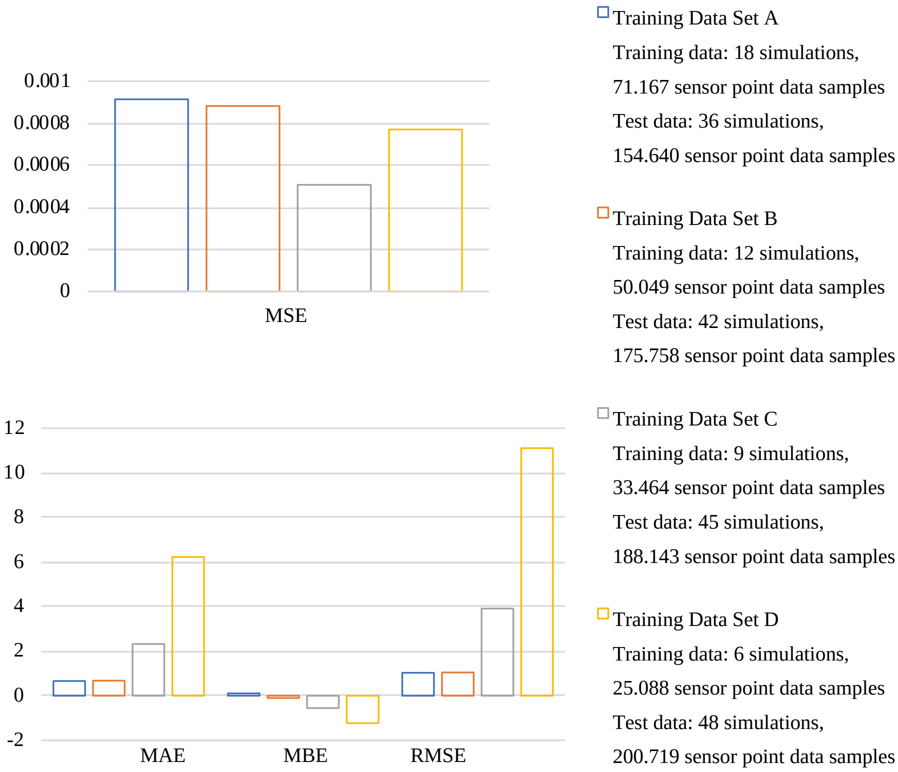

Fig. 7. Mean Squared Errors of the trained networks and resulting prediction accuracies for training data set A to D

3.3 ANN Prediction Accuracies for sDA

As mentioned before, the sDA metric denotes the percentage of space achieving a DA of 50% or more. By assigning a summative value to a space, the metric becomes useful for the assessment of the overall daylight performance of the space. A Grasshopper script was run to convert all DA predictions from the ANN models into sDA results. The predicted sDA results were then compared to the simulated sDA results. The prediction accuracies for the sDA metric are shown in Fig. 8, shedding light on how closely the ANN models were able to predict the DA 50% threshold.

Fig. 8 gives the predicted and simulated sDA results for atrium adjacent spaces on the ground floor for training data sets A to D. The sDA results are shown according to the matrix entries in Fig. 4 and Fig. 5. As seen in the figure, the sDA of the 54 design solutions varied between 20 and 29%. For training data sets A and B with data from 18 and 12 simulations respectively, simulated and predicted values showed considerable and consistent overlap, indicating high prediction accuracies. The ANN models were thus suitable for predicting daylight for the entire design space. For training sets C and

D however, predictions diverged from simulated results, most noticeably so for data set D. The strong discrepancies show that training data extracted from 9 and 6 simulations were insufficient for mapping the daylight performance of the entire solution space of 54 solutions. Data from 12 simulations (data set B) on the other hand were already sufficient to accurately map the performance across the solution space. The achieved time-savings therefore equal the simulation time of 42 simulations less the training time of the ANN models. On a 2.6 GHz Intel Core i7, one daylight simulation took a little over three hours. Using ANNs to predict daylight for 42 design solutions thus reduced the time spent on simulations by approximately 126 hours, or 123 hours after taking into account the time spent on training and optimizing the networks with architectures between 37-40 neurons in the hidden layer.

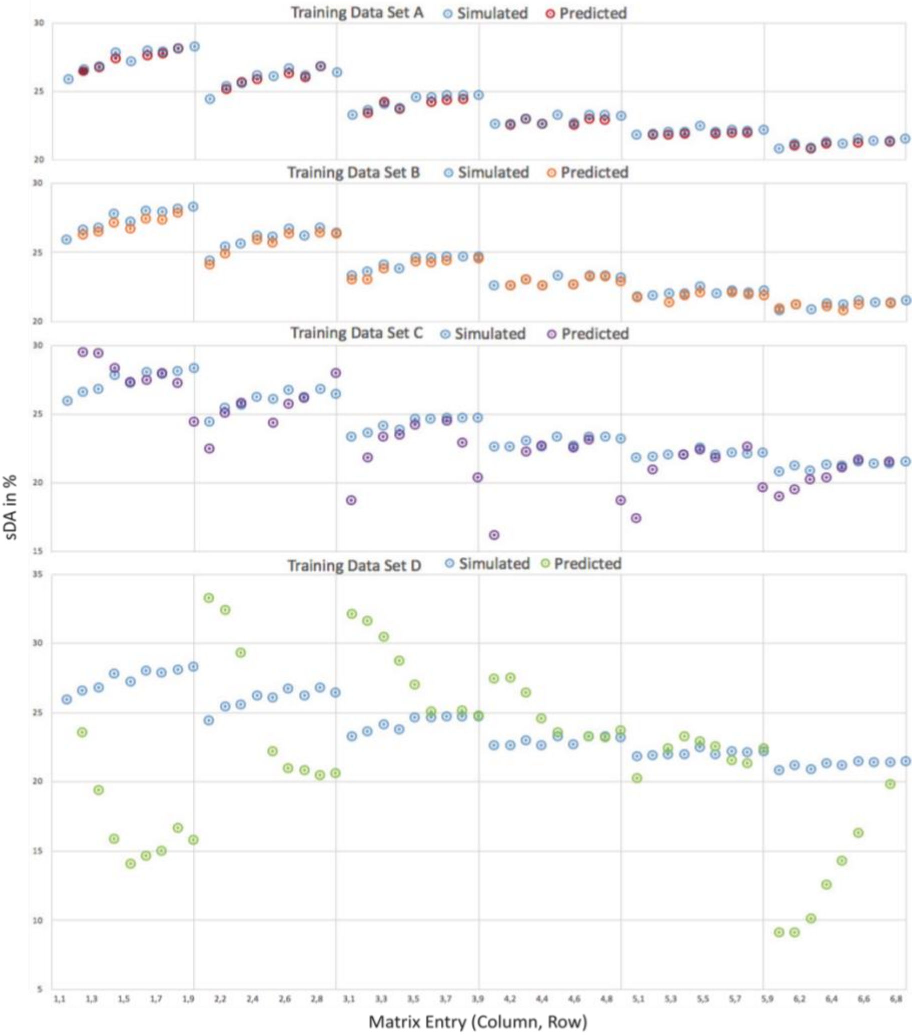

Fig. 8**.** sDA performance of the 54 generated design solutions: Predicted and simulated sDA for training data sets A to D

3.4 Sampling Solution Spaces with Similar Daylight Performance

To increase daylight levels on the ground floor, three WWR distribution options were tested. ANN models were trained on data from 36 randomly selected simulations out of 162. After training, the ANNs were validated against simulation results. The obtained errors of .74 MAE, .01 MBE and 1.12 RMSE for the DA metric were similar to those for the 54 solutions of fully-glazed atrium facades. The sDA metric showed an equally high accuracy. Evaluated against 84 simulations, the highest obtained absolute error for the sDA metric was 1.16. This is in itself a small error, seeing as 1% of floor area receiving more light does not strike as a significant impact.

In the context of choosing the WWR distribution as a design variable, the correct identification of differences in daylight performance between the design options appears to be more crucial than a high prediction accuracy. In line with expectations, a WWR distribution of 20, 30, 40, 50, 60% WWR from the top to the ground floor provided higher daylight levels in atrium adjacent spaces than the other two options. The option with a 50 to 100% WWR distribution from top to ground floor had the weakest daylight performance (Fig. 9). The ANN models were able to identify the trends in the design alternatives, successfully predicting the ranking of options in accordance with the order of the simulated results. Furthermore, the accuracies of the prior solution space of 54 variants could be maintained within a solution space of 162 variants despite an increase in design variables.

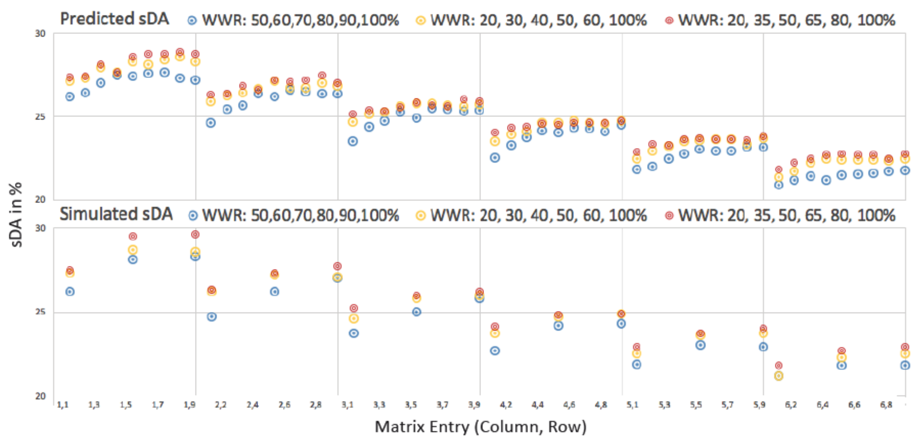

Fig. 9. Predicted and simulated sDA performance for WWR distributions

4 Conclusions

This paper demonstrated the reliability and accuracy of ANN models in mapping solution spaces. A surprisingly small number of 12 and 36 simulations were sufficient to train accurate ANNs to predict daylight for a solution space of 54 and 162 design variants. Importantly, the networks were able to learn the relation between design options with very similar daylight performance. Consequently, the ranking of design solutions according to their predicted performance stayed true to the ranking of design solutions established by simulation. The findings of this paper are limited to the design variables employed in the study. Ongoing research is therefore looking at more complex design variables that generate a greater breadth of design solutions.

Acknowledgements. We gratefully acknowledge Funds for Woman Graduates and thank them for their support.

References

[1] M. G. Figueiro *et al.*, "The impact of daytime light exposures on sleep and mood in office workers," *Sleep Heal.*, vol. 3, no. 3, pp. 204–215, 2017.

[2] C. F. Reinhart and O. Walkenhorst, "Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds," *Energy Build.*, vol. 33, no. 7, pp. 683–697, 2001.

[3] Heschong Mahone Group, "Daylight Metrics - PIER Daylighting Plus Research Program Final Report to the California Energy Commission," CA, USA, 2012.

[4] D. Gossard, B. Lartigue, and F. Thellier, "Multi-objective optimization of a building envelope for thermal performance using genetic algorithms and artificial neural network," *Energy Build.*, vol. 67, pp. 253–260, 2013.

[5] Y. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient BackProp," in *Neural Networks: Tricks of the trade*, Berlin, Germany: Springer, 1998.

[6] P. Sajda, "Neural Networks," in *Encyclopedia of the Human Brain Volume 3*, V. S. Ramachandran, Ed. Columbia University, USA: Academic Press, 2002, pp. 373–383.

[7] T. Kazanasmaz, M. Gunaydin, and S. Binol, "Artificial neural networks to predict daylight illuminance in office buildings," *Build. Environ.*, vol. 44, pp. 1751–1757, 2009.

[8] 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.

[9] M. Ehrgott, C. M. Fonseca, X. Gandibleux, J.-K. Hao, and M. Sevaux, Eds., *Evolutionary Multi-Criterion Optimization*, vol. 5467. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[10] G. Lopez and C. A. Gueymard, "Clear-sky solar luminous efficacy determination using artificial neural networks," *Sol. Energy*, vol. 81, no. 7, pp. 929–939, 2007.

[11] S. Pattanasethanon, C. Lertsatitthanakorn, and S. Atthajariyakul, "An accuracy assessment of an empirical sine model , a novel sine model and an artificial neural network model for forecasting illuminance / irradiance on horizontal plane of all sky types at Mahasarakham , Thailand," vol. 49, pp. 1999–2005, 2008.

[12] S. Janjai and P. Plaon, "Estimation of sky luminance in the tropics using artificial neural networks: Modeling and performance comparison with the CIE model," *Appl. Energy*, vol. 88, no. 3, pp. 840–847, 2011.

[13] S. Zhou and D. Liu, "Prediction of Daylighting and Energy Performance Using Artificial Neural Network and Support Vector Machine," *Am. J. Civ. Eng. Archit. Vol. 3, 2015, Pages 1-8*, vol. 3, no. 3A, pp. 1–8, 2015.

[14] C.-L. Lorenz and W. Jabi, "Predicting Daylight Autonomy Metrics Using Machine

Learning," in *International Conference for Sustainable Design of the Built Environment*, 2017.

[15] C.-L. Lorenz, M. Packianather, A. B. Spaeth, and C. Bleil de Souza, "Artificial Neural Network-Based Modelling for Daylight Evaluations," in *Symposium on Simulation for Architecture + Urban Design*, 2018.

[16] D. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *SIAM J. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.

[17] J. M. Twomey and A. E. Smith, "Performance measures, consistency, and power for artificial neural network models," *Math. Comput. Model.*, vol. 21, no. 1/2, pp. 243–258, 1995.

[18] O. Aschehoug, "Daylight in glazed spaces," *Build. Res. Inf.*, vol. 20, no. 4, pp. 242–245, 1992.

[19] S. Samant, "Atrium and its adjoining spaces: a study of the influence of atrium façade design Swinal Samant Atrium and its adjoining spaces: a study of the influence of atrium faç ade design," *Archit. Sci. Rev.*, vol. 54, no. 4, 2017.

[20] R. J. Cole, "The effect of the surfaces adjoining atria on the daylight in adjacent spaces," *Build. Environ.*, vol. 25, no. 1, pp. 37–42, 1990.