Input Feature Selection and Optimization for ANN Models Predicting Daylight in Buildings

C.-L. Lorenz[1], M. Packianather[1], C. Bleil De Souza[1], A. B. Spaeth[1], T. I. Lorenz[2]
[1]Cardiff University, Wales, U.K.,
[2]Leiden University, The Netherlands
lorenzc4@cardiff.ac.uk

Abstract. Artificial Neural Networks (ANNs) were used as prediction models to explore design solutions for the atrium design of a school building. To this end, a solution space of 165 design variants was generated via parametric modeling. This paper details the process of extracting, selecting and optimising the input features required for ANN training in order to predict the Daylight Autonomy (DA) metric. The feature selection undertaken in this study mainly consisted of two steps: Firstly, a computationally less expensive machine learning model was used to rank the input features according to their relevance in predicting daylight levels. Secondly, ANNs were trained applying sequential forward selection (SFS) as a method of optimization. A novel aspect of this work is found in the use of machine learning technique to justify the predictors and application of forward sequential selection for optimizing the ANN models in the field of daylighting. The proposed method was investigated in terms of achievable improvements to prediction accuracy, reduceable training time and the feasibility of the method. Although the prediction accuracy improvement was marginal, the biggest contribution of the proposed approach was in reducing the training time by 57% by discarding up to 38% of training features, which were found to be redundant.

1. Introduction

Access and exposure to daylight not only regulates the circadian rhythm, but has been shown to reduce fatigue (Figueiro *et al.*, 2017), improve well-being, and positively impact work and school performance (Maesano and Annesi-Maesano, 2012; Veitch, Christoffersen and Galasiu, 2013). In recent years, the Daylight Factor (DF), a standard metric for assessing daylight in buildings, came under scrutiny for its unsuitability as a design driver, given that it considers neither orientation nor climate (Reinhart, Mardaljevic and Rogers, 2013). Instead, climate-based metrics were introduced (Reinhart and Walkenhorst, 2001) and have only recently been adopted into European and British building standards (EN 17037:2018). Of particular interest in this paper was the Daylight Autonomy (DA) metric, which quantifies the percentage of occupied hours in a year in which a threshold of 300 lux can be met with daylight. However, the annual simulation of climate-based metrics is more time consuming than point-in time daylight factor calculations, making it inherently more difficult to assess design options as an integral part of the design process (Jones and Reinhart, 2019).

Research has introduced Artificial Neural Networks (ANNs) as surrogate energy simulation software in order to improve the feasibility of performance-driven design investigations (Nguyen, Reiter and Rigo, 2014). The benefits of such ANN-based models include their instantaneous response rate and high accuracy, if trained with sufficient performance data, an appropriate model selection and parameter settings (Zhao and Magoulès, 2012). Such models have been found to be particularly useful for automated building control, where an immediate response is required (Hu and Olbina, 2011). Their use further extends to the modeling of large design solution spaces in order to solve design problems with numerous design variables (Machairas, Tsangrassoulis and Axarli, 2014). In general, ANNs present a promising avenue in research as a mature technology that works well with noisy data (i.e. simulation data) and can operate with various types of input and output (i.e. real, discrete, Boolean), the relationship between which need not be known (Mckee, Schulz and Caruana, 2006). Such research has laid the groundwork of this paper, as we applied ANN-based modeling to efficiently predict daylight for the solution space of a building design. More specifically, this study trained ANNs to predict climate-based metrics to explore design variants for the central atrium of a school building, focusing on the selection and optimization of ANN input features.

1.1 ANN-based Daylight Predictions

In the field of daylight design, ANNs have been used to predict luminous efficacy (Lopez and Gueymard, 2007), sky luminance and sky irradiance (Pattanasethanon, Lertsatitthanakorn and Atthajariyakul, 2008; Janjai and Plaon, 2011), as well as horizontal internal illuminance levels (Kazanasmaz, Gunaydin and Binol, 2009). These ANN models mainly relied on historic data as training data – a practice which can be challenging in terms of the time and the effort required to generate them. The need for data to be collected over a longer period of time, e.g. around 3 months as in Kazanasmaz et al. (2009), undermines the feasibility of ANNs when intended as a method of rapid feedback.

One line of research that has circumvented this problem applied ANNs to optimization problems and extracted the training data from a fraction of simulations required as part of the optimization process. Magnier and Haghihat (2010) showed that simulation-based ANNs could be used within the fitness function of a Genetic Algorithm to considerably reduce simulation time. Specific to daylight design, ANNs have been used as emulators for electric energy consumption and visual comfort predictions (Wong, Wan and Lam, 2010; Kim, Jeon and Kim, 2016). Regarding climate-based daylight performance metrics, Zhou and Liu (2015) were able to predict the specific hourly illuminance range of the Useful Daylight Illuminance (UDI) metric.

The studies undertaken typically predicted point-in-time or hourly daylight levels, none addressed annual climate-based daylight predictions. The potential for predicting the annual Daylight Autonomy (DA) metric was shown in a simple model in the authors' previous work (Lorenz *et al.*, 2018), in which input features were empirically selected. A complex model is explored in the current paper, showing a robust, automated and replicable method for improving the selection of ANN input features.

1.2 Input Feature Selection and Optimization Found in the Literature

Input features, also known as predictor variables, make up the training data which is passed to ANN models in order to facilitate supervised, unsupervised and reinforcement type learning. Input feature selection refers to the reduction of the dimensionality of training data by selecting a subset of input features. An objective function is commonly used as selection criterium to minimize the predictive error and thereby identify an optimal or suboptimal subset of input features, improving prediction accuracy and identification of the minimum number of features needed (Marcano-Cedeño *et al.*, 2010; Özşen, 2013).

As exhaustive search methods are computationally expensive, two common methods of input feature selection are Sequential Forward Selection (SFS) and Sequential Backward Selection (SBS). SFS is a bottom-up approach that starts with an empty set of features to which features are iteratively added (Whitney, 1971). Its counterpart, SBS, starts from the complete set from which features are iteratively removed (Marill and M. Green, 1963). Inevitably, both methods inhibit a nesting problem, whereby potentially important features, once removed, cannot be re-introduced. To ensure that important features would be kept alive, mixed method approaches were introduced (Pudil, Novovi and Kittler, 1994). A comparative evaluation of methods is given by Zongker and Jain (1996). As forward-based methods have been shown to be faster than backward-based methods, we propose an SFS approach in combination with a machine learning algorithm to determine the sequence of input features.

2. Proposed Research Methodology

The different steps to developing, optimizing and validating ANN models are illustrated in Figure 1. Performance data was collected from selected design variants of a proposed complex design solution space (A). This data was extracted from daylight simulations and recorded features describing the design changes and the corresponding daylight results. While a majority of data was used for training the ANN models, a part of it was retained for validation of the models (D).

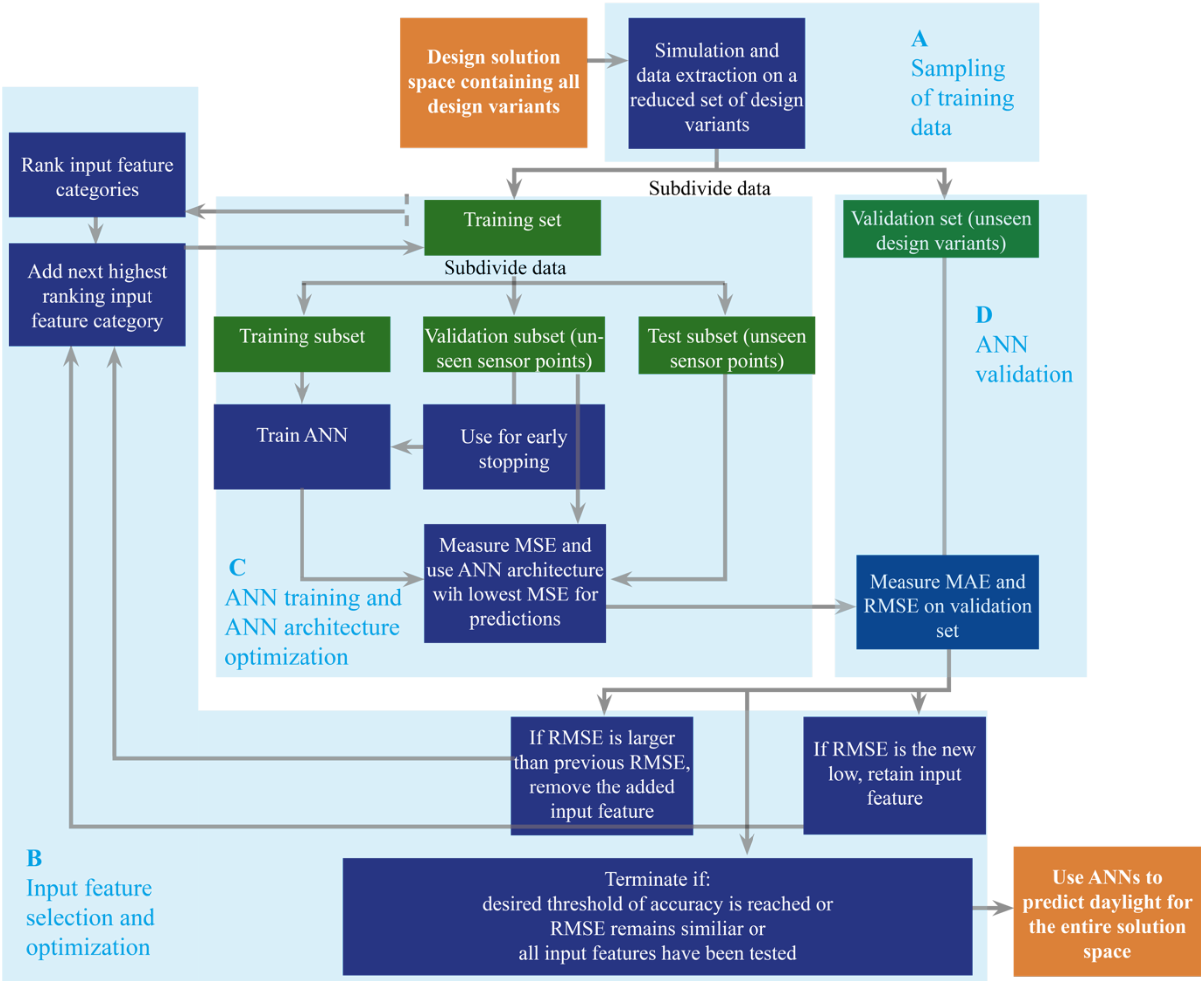

Figure 1: ANN model development and validation

After generating a training set, bagged decision trees, a machine learning technique, were used to determine a sequence for the SFS of features (B). The sequence of input features was established according to their relevance in predicting daylight. Thus, when using SFS, the ANN models were first trained with the input features that showed the highest impact as predictors.

The ANN training data set was subdivided into a training subset, a validation subset and a test subset at the ratio of 65:25:15. The training subset was used to measure the training accuracy, the validation subset was used to avoid over-fitting, and the test subset was used to determine its generalisation capability (C). The data was then passed to a three-layered ANN model with a varying number of neurons in the hidden layer. Each network architecture was trained ten times with a different initial weight setting and distribution of training samples into the training, validation and test subsets. The mean squared error (MSE) was observed on all subsets and the ANN architecture with the lowest error was used for predictions (C). The prediction of ten ANNs was then averaged to improve generalization (as seen in Lorenz *et al.*, 2018).

The validation subset was used for early stopping to avoid overfitting and the test subset was used to estimate prediction accuracies on new cases. Both subsets were also used to assist in the optimization of the network architecture. Each network architecture was trained ten times with the initial weight settings and distribution of samples across the subsets varying in each training run. The network architecture with the lowest error across ten training runs was used for predictions, and the output of all ten networks was averaged to improve generalization.

The ANN predictions were validated against simulation results withheld from training (D). In the case of insufficient accuracy, another training feature was added to the training set (B) and the optimization of network architecture (C), and validation process (D) were repeated. Three options for terminating this cycle were identified: once a desired threshold of accuracy was reached, once the accuracy converged for a certain number of training cycles, or once all input features had been tested. We chose the latter to concurrently assess the method. The other two options are useful to identify the minimum number of input features required and to reduce training time.

2.1 Design Solution Space and Design Variables

The exploration of design solutions was done for the central atrium of a school-building located in Hamburg (Figure 2a). As a first design variable, the atrium base dimension was reduced from 225m^2 to 56,25m^2, thereby splaying the atrium well walls in acute angles between 90 and 79° (Figure 2b). With a second design variable, the atrium well was slanted in northward and southward orientation (Figure 2c), resulting in 9 possible solutions and additionally modifying atrium well splay angles between 113 and 58°. This resulted in a 9 by 6 matrix of 54 possible design variants for the atrium well geometry. As the third design variable, 3 possibilities were specified for the window-to-wall ratios (WWR) across floor levels of the 6-storey building. The WWR was reduced from the ground to higher floors (Figure 2d), increasing the reflected light in the atrium well and daylight levels on the ground floor (Samant, 2017). The entire solution space thereby contained 162 design variants.

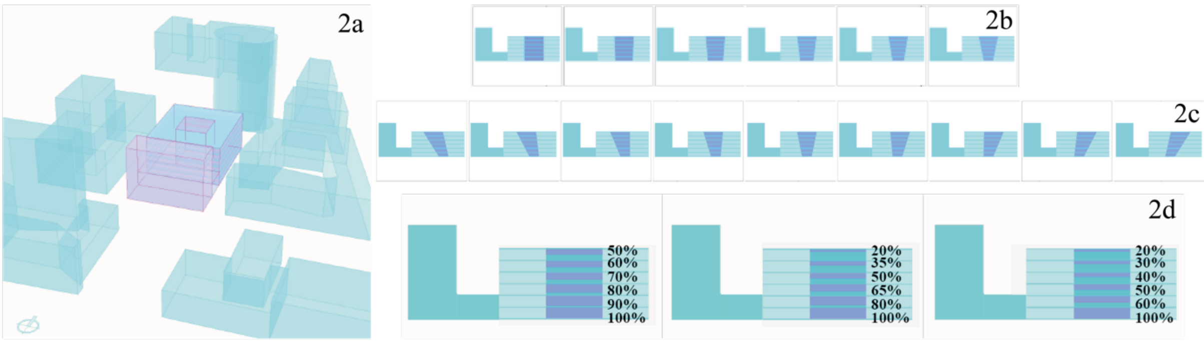

Figure 2a: The school building and surrounding buildings. 2b: Six variations resulting from a change to the atrium base dimension (the atrium well is highlighted in dark blue). 2c: Nine variations for atrium well orientation. 2d: Three variations for the WWR distribution across floor levels

2.2 Daylight Simulation and Data Extraction

From the solution space of 162 variants, a reduced set of 36 variants was selected to provide training data and another 21 variants were selected for validation of ANN accuracy. The architectural models were built in Grasshopper, and daylight simulation on the selected variants were run in Diva – a radiance-based and validated software (Mohsenin and Hu, 2015). DA was calculated for sensor points at a work-plane height of 0.8 m above floor level and the sensor points were spaced in 0.6 m distance from each other. The input features were extracted for every sensor point and passed to the input layer of the network model (Figure 3). The corresponding DA levels were passed to the output layer of the network for it to undergo supervised training. As shown in Figure 3, 26 input features in total were extracted from the simulation model and grouped according to categories (rectangles in Figure 3). In a next step, these categories were ranked according to their impact on predictions.

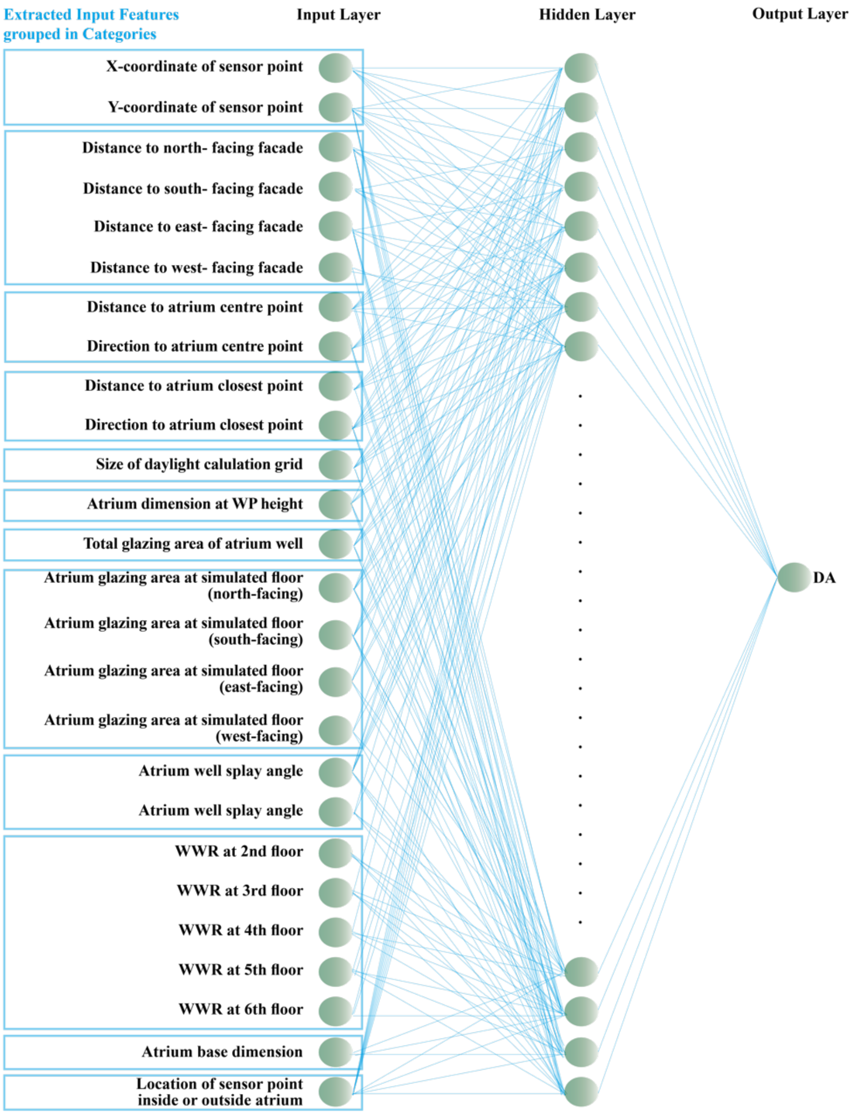

Figure 3: Representation of ANN model as a construct of neurons with an input layer, hidden layer and output layer. The extracted input features were passed to the input layer of the model and the daylight performance data (in DA) to the output layer.

2.3 Ranking of Input Feature Categories

The training data set was passed to machine learning (ML) models to evaluate the relevance of the extracted input features as predictors. Although ANNs can be trained to identify significant input features, an alternative ML model was used to reduce computation time. Tested ML techniques included: linear regression models, fine, medium and coarse trees, boosted tree ensembles, liner, quadratic, cubic, fine Gaussian support vector machines and Gaussian process regression models. Bagged decision trees showed the lowest root mean squared error (RMSE) during fitness approximation, and therefore superior performance compared to the other models. Additionally, the computation time was low, with approximation taking around four minutes. Hence, bagged decision trees were selected to identify the feature selection sequence.

The training data set was passed to the bagged decision tree model (as shown in Figure 1). Every input feature category was then individually removed from the data set and the corresponding variance in RMSE was measured. Subsequently, the features were ranked according to the variance they inflicted on the error, with features causing the largest variance ranked highest. The RMSE during approximation on the training data set and resulting sequence of input feature categories is given in Table 1.

Table 1: Ranking of input feature categories according to variance in prediction error

Input Feature Category	Sequence Order	RMSE	Input Feature Category	Sequence Order	RMSE
Distance and direction to atrium closest point (2 features)	1	1.19	X-, Y-coordinates of sensor points (2 features)	7	.87
Distance to façade (4 features)	2	.99	Glazing area at simulated floor level (5 features)	8	.85
Distance and direction to atrium centre point (2 features)	3	.96	Glazing area across all floors (1 feature)	9	.84
WWR (5 features)	4	.90	Atrium base dimension (1 feature)	10	.84
Location of sensor point inside or outside atrium (1 feature)	5	.90	Atrium dimension at WP height (1 feature)	11	.84
Atrium well spay angles (2 features)	6	.89	Daylight calculation grid size (1 feature)	12	.84

2.4 Sequential Feature Selection, Optimization and Validation

The ANN models were trained starting from a data set with one input feature category. Once trained, the ANN model was used to predict the DA metric for the 21 retained variants (13% of data) and predictions were compared to the simulated DA. Two measures were used: the mean absolute error (MAE) and the root mean square error (RMSE). The MAE, which gives the absolute difference between the simulated and predicted values, was chosen for its ease of interpretation. The RMSE was selected as it weighs larger errors more heavily and is a commonly used measure of accuracy. The RMSE was observed for each added input feature and its minimization was used as the objective function to optimize the selection. As long as the RMSE reached a new low, one additional input feature category was added to the training data set. If the RMSE worsened, the last added feature was removed before adding the next feature in the sequence. In this way, all input features were added to the training data set at some point in a forward-sequential manner. Results discuss the observed errors and accuracy for every added input feature category.

2.5 ANN Training Settings and Optimization of ANN Architecture

Back-propagation feedforward ANN models were employed in conjunction with the Lavenberg-Marquardt algorithm. The training parameters were set to an initial mu of 1, a mu decrease factor of 0.8 and mu increase factor of 1.5. The training was run for 200 epochs, during which the connection strengths between neurons were adjusted to minimise the mean squared error (MSE) on the training data set. The maximum number of validation failures was set to 6. Ten network architectures with 38 to 40 neurons in the hidden layer were trained and tested with said settings and the network architecture with the lowest MSE on the training, validation and test subset was used for predictions. This was done to ensure that a network with an optimized architecture was used to evaluate the input features.

3. Results and Discussion

The MSE on the training data set (36/162 simulations with 150.706 sensor point data samples) and the RMSE and MAE on the validation set (21/162 simulations with 87.850 sensor point data samples) were recorded at each training stage of the sequential search. Figure 4 shows the MSE on the training data during SFS (the sequence refers to Table 1). Figure 5 shows the MAE and Figure 6 the RMSE obtained on the validation set. The minimum achieved error on the validation set, which indicates the optimum set of input features, has been highlighted in red. The sequences at which an added input feature category was found to be redundant and was therefore removed from the training data have been highlighted in yellow.

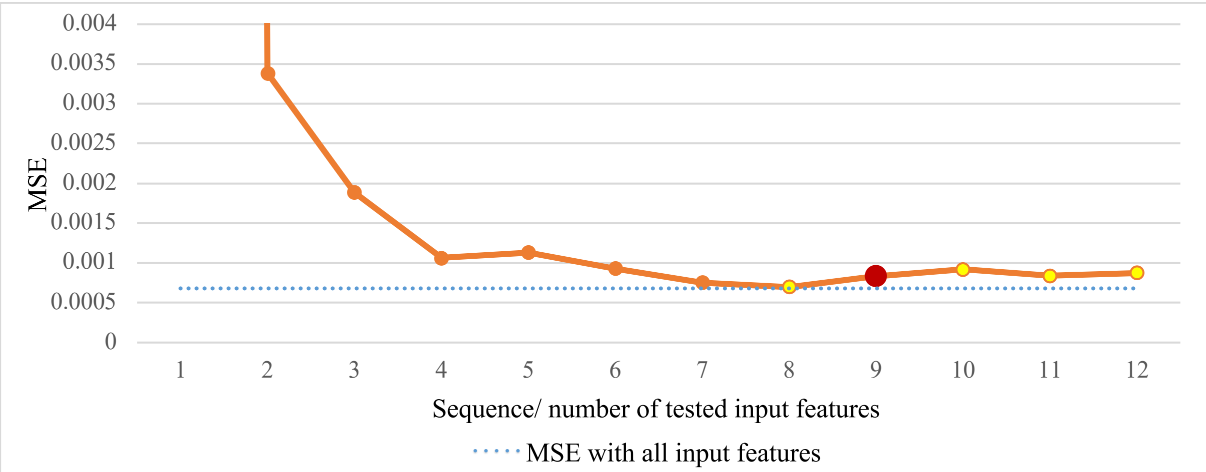

Figure 4: MSE on the training data set after every added input feature category

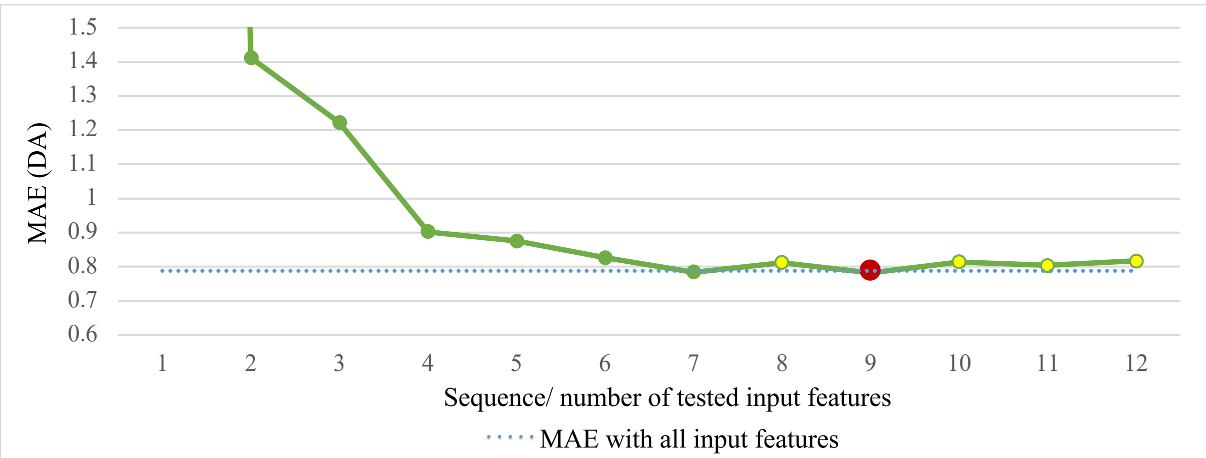

Figure 5: MAE on the validation data set after every added input feature category

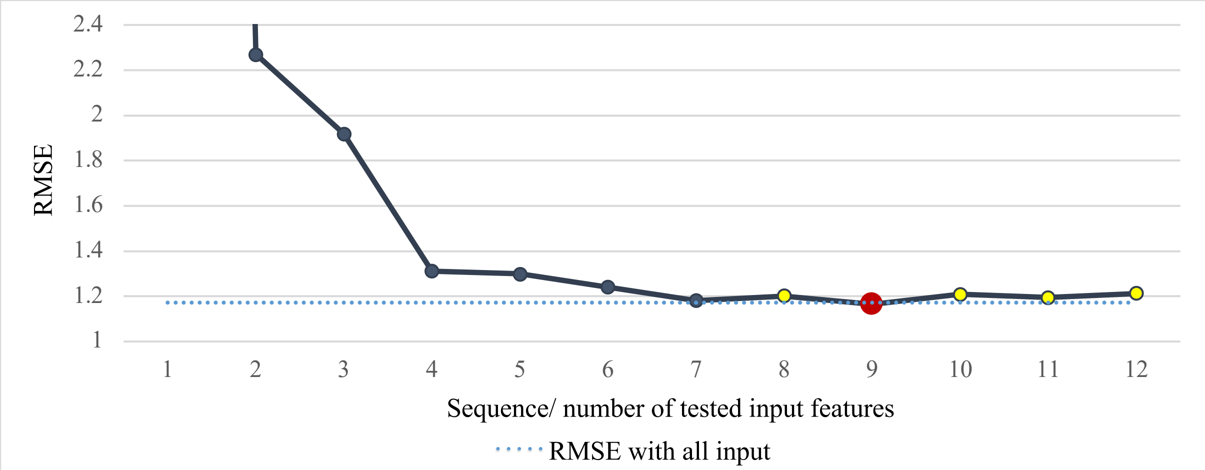

Figure 6: RMSE on the validation data set after every added input feature category

The MSE dropped below 0.01 after including six input feature categories and remained below this threshold in consecutive training runs. The MAE and RMSE on the predictions remained consistently low after the 7th input feature category was added to the training data. This suggests that the first 7 added input feature categories (see Table 1) suffice in predicting DA for the given design solution space.

The lowest MSE on the training data was reached after including the 8th input feature category (describing glazing areas on the simulated floor level). The MAE and RMSE on the validation data however increased, resulting in the feature being removed from the feature set. The minimum MAE and RMSE were reached at .78 MAE and 1.16 RMSE during the 9th sequence of the SFS after including 8, and removing 1 input feature categories. This was therefore considered the ‘optimum’ set of input features. In comparison to the complete set of 12 extracted input feature categories however, the improvement in accuracy was only marginal, as the original feature set showed errors of .79 MAE and 1.17 RMSE (indicated by the dotted blue line in Figures 4 to 6). Although this shows that the input feature selection did not result in a significant improvement of accuracies, not all input features were needed, and overall training time could be reduced.

The training time of ANN models (including optimization of ANN architecture) was measured on a 2.6 GHz Intel Core i9 processor. ANN training of the model that included all 12 input features (26 individual features) had a duration of 05:55 (hh:mm). In comparison, training of the ‘optimal’ feature set with 8 input feature categories (19 individual features) took 03:23 hours. When further limiting the data to 7 added feature categories (16 individual features), training time was 02:32 hours. Predictions on the validation set were made in less than 1 second.

The ANN models were able to predict DA with an accuracy that converged around .8 DA MAE, meaning that this was, on average, the absolute difference between the simulated and predicted DA. The simulated DA range from 0 to 88%, with 1% referring to 1% of occupied hours in a year. A difference of .8 DA can therefore hardly be interpreted and constitutes a negligible error. The RMSE, which converged around 1.2%, supports this finding and shows a high accuracy of ANNs in predicting the DA metric.

In order to achieve the above-mentioned accuracies, data was extracted from a total of 57/162 simulations for training and validation. As daylight simulations took approximately 3 hours per design variant, a total of 315 hours of simulations were saved through ANN predictions. Compared to a conventional brute-methods approach, integrating ANNs reduced simulation time by 65%.

4. Conclusion and Recommendations for Future Research

Overall, the following conclusions for ANN-based daylight predictions can be drawn:

a) On average, prediction accuracies of around .8 DA mean absolute error (MAE) were achieved. These accuracies could be maintained as the number of input features increased.

b) The MSE on the training data set did not directly correlate with the MAE and RMSE on the validation data set, as the prediction accuracies could improve (lower MAE and RMSE) even though the ability of the network to fit the data decreased (higher MSE).

c) Having superfluous input features did not significantly lower accuracies (e.g. calculation grid size and atrium dimension at WP height). It did however increase training time. On the other hand, too few input features (e.g. before sequence 4), though requiring less training time, compromised prediction accuracies.

d) Using the proposed method, accuracies marginally improved compared to the empirically selected full set of input features. However, the training time for the ANN models could be reduced by 43% with the optimum number of training features. After further input feature reduction, training time was reduced by 57% without significantly compromising the accuracies.

e.) In the input feature set with the optimum prediction performance, 7 input features (27%) were discarded. After further reduction, three additional input features (38%) could be removed without significantly compromising accuracies.

The paper proposes a viable method for selecting input features useful for predicting daylight in buildings, which can be applied in other fields as well. It investigated how far improvements in terms of accuracy and time-savings could be achieved. ANN training time was greatly reduced, although accuracy improvements were marginal. The original, empirically selected input feature series performed well in this study. In more complex design scenarios however, where the number of design variables is greater, it may become more difficult to empirically identify the appropriate training features. In such cases, the proposed selection method may prove to be especially valuable. A downside of the method is that it remains computationally expensive, as it requires multiple training runs with already computationally demanding ANN models. It would therefore be useful to evaluate the results of feature selection using smaller and more feasible network architectures, or completely rely on computationally less expensive ML models for feature selection. A comparison to alternative feature selection methods is recommended, as well as the application of the proposed method on more complex solution spaces.

Acknowledgements

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

References

Figueiro, M. G., Steverson, B., Heerwagen, J., Kampschroer, K., Hunter, C. M., Gonzales, K., Plitnick, B. and Rea, M. S. (2017) 'The impact of daytime light exposures on sleep and mood in office workers', *Sleep Health*. National Sleep Foundation., 3(3), pp. 204–215.

Hu, J. and Olbina, S. (2011) 'Illuminance-based slat angle selection model for automated control of split blinds', *Building and Environment*. Elsevier Ltd, 46(3), pp. 786–796.

Janjai, S. and Plaon, P. (2011) 'Estimation of sky luminance in the tropics using artificial neural networks: Modeling and performance comparison with the CIE model', *Applied Energy*. Elsevier Ltd, 88(3), pp. 840–847.

Jones, N. L. and Reinhart, C. F. (2019) 'Effects of real-time simulation feedback on design for visual comfort', *Journal of Building Performance Simulation*. Taylor & Francis, 12(3), pp. 343–361.

Kazanasmaz, T., Gunaydin, M. and Binol, S. (2009) 'Artificial neural networks to predict daylight illuminance in office buildings', *Building and Environment*, 44, pp. 1751–1757.

Kim, W., Jeon, Y. and Kim, Y. (2016) 'Simulation-based optimization of an integrated daylighting and HVAC system using the design of experiments method', *Applied Energy*. Elsevier Ltd, 162, pp. 666–674.

Lopez, G. and Gueymard, C. A. (2007) 'Clear-sky solar luminous efficacy determination using artificial neural networks', *Solar Energy*, 81(7), pp. 929–939.

Lorenz, C.-L., Packianather, M., Spaeth, A. B. and Bleil de Souza, C. (2018) 'Artificial Neural Network-Based Modelling for Daylight Evaluations', in *Symposium on Simulation for Architecture + Urban Design*. Delft, NE.

Machairas, V., Tsangrassoulis, A. and Axarli, K. (2014) 'Algorithms for optimization of building design: A review', *Renewable and Sustainable Energy Reviews*, 31(1364), pp. 101–112.

Maesano, C. and Annesi-Maesano, I. (2012) Impact of Lighting on School Performance in European Classrooms. *Clima2016*, doi 10, pp. 1-14.

Magnier, L. and Haghighat, F. (2010) 'Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network', *Building and Environment*. Elsevier Ltd, 45(3), pp. 739–746.

Marcano-Cedeño, Quintanilla-Domínguez, J., Cortina-Januchs, M. G. and Andina, D. (2010) 'Feature selection using Sequential Forward Selection and Classification applying Artificial Metaplasticity Neural Network', *IECON 36th Annual Conference on IEEE Industrial Electronics Society*, pp. 2845–2850.

Marill, T. and M. Green, D. (1963) 'On the effectiveness of receptors in recognition systems', *Information Theory, IEEE Transactions on*, 9, pp. 11–17.

Mckee, S. A., Schulz, M. and Caruana, R. (2006) 'Efficiently Exploring Architectural Design Spaces via Predictive Modeling', in *Asplos International Conference on Architectural Support for Programming Languages and Operating Systems*. CA, USA, pp. 195–206.

Mohsenin, M. and Hu, J. (2015) 'Assessing daylight performance in atrium buildings by using Climate Based Daylight Modeling', *Solar Energy*, 119, pp. 553–560.

Nguyen, A.-T., Reiter, S. and Rigo, P. (2014) 'A review on simulation-based optimization methods applied to building performance analysis', *Applied Energy*, 113, pp. 1043–1058.

Özşen, S. (2013) 'Classification of sleep stages using class-dependent sequential feature selection and artificial neural network', *Neural Computing and Applications*, 23(5), pp. 1239–1250. doi: 10.1007/s00521-012-1065-4.

Pattanasethanon, S., Lertsatitthanakorn, C. and Atthajariyakul, S. (2008) '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', 49, pp. 1999–2005.

Pudil, P., Novovi, J. and Kittler, J. (1994) 'Floating search methods in feature selection', *Pattern Recognition Letters*, 15(November), pp. 1119–1125.

Reinhart, F, C., Mardaljevic, J. and Rogers, Z. (2013) 'Dynamic Daylight Performance Metrics for Sustainable Building Design', *The journal of the illuminating Engineering Society of North America*, 3(1), pp. 7–31.

Reinhart, C. F. and Walkenhorst, O. (2001) 'Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds', *Energy and Buildings*, 33(7), pp. 683–697.

Samant, S. (2017) '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', *Architectural Science Review*, 54(4).

Veitch, J. A., Christoffersen, J. and Galasiu, A. D. (2013) Daylight and View through Residential Windows : Effects on Well-being, pp. 1-6.

Whitney, A. W. (1971) 'A Direct Method of Nonparametric Measurement Selection', *Computers, IEEE Transactions on*, 20, pp. 1100–1103.

Wong, S. L., Wan, K. K. W. and Lam, T. N. T. (2010) 'Artificial neural networks for energy analysis of office buildings with daylighting', *Applied Energy*. Elsevier Ltd, 87(2), pp. 551–557.

Zhao, H. X. and Magoulès, F. (2012) 'A review on the prediction of building energy consumption', *Renewable and Sustainable Energy Reviews*. Elsevier Ltd, 16(6), pp. 3586–3592.

Zhou, S. and Liu, D. (2015) 'Prediction of Daylighting and Energy Performance Using Artificial Neural Network and Support Vector Machine', *American Journal of Civil Engineering and Architecture*, 3(3A), pp. 1–8.

Zongker, D. and Jain, A. (1996) 'Algorithms for Feature Selection: An Evaluation', in *Proceedings of 13th International Conference on Pattern Recognition*. Vienna, Austria, pp. 18–22.