7th International Building Physics Conference

IBPC2018

Proceedings

SYRACUSE, NY, USA

September 23 - 26, 2018

Healthy, Intelligent and Resilient
Buildings and Urban Environments

ibpc2018.org | #ibpc2018

Neural networks to predict the hygrothermal response of building components in a probabilistic framework

Astrid Tijskens[1,*], Staf Roels[1] and Hans Janssen[1]

[1]KU Leuven, Department of Civil Engineering, Building Physics Section, Kasteelpark Arenberg 40, 3001 Leuven, Belgium

[*]*Corresponding email: Astrid.Tijskens@kuleuven.be*

ABSTRACT
In recent years, probabilistic assessment of hygrothermal performance of building components has received increasing attention. Given the many uncertainties involved in the hygrothermal behaviour of building components, a probabilistic assessment enables to assess the damage risk more reliably. However, this typically involves thousands of simulations, which easily becomes computationally inhibitive. To overcome this time-efficiency issue, this paper proposes the use of much faster metamodels. This paper focusses on neural networks, as they have proven to be successful in other non-linear and non-stationary research applications. Two types of networks are considered: the traditional multilayer perceptron (with and without a time window) and memory neural networks (LSTM, GRU). Both are used for predicting the hygrothermal behaviour of a massive wall. The results showed that all networks are capable to predict the temperature profiles accurately, but only the LSTM and GRU networks could predict the slow responses of relative humidity and moisture content. Furthermore, the LSTM and GRU network were found to have almost equal predicting accuracy, though the GRU converged faster.

KEYWORDS
Internal insulation, hygrothermal performance, metamodel, neural networks, LSTM

INTRODUCTION
In recent years, traditional deterministic assessments in building physics have evolved towards a probabilistic framework (Annex 55, 2015; Vereecken et al. 2015). When evaluating the hygrothermal behaviour of a building component, there are many inherently uncertain parameters, such as the exterior climate, geometry, material… A probabilistic simulation enables taking into account these uncertainties, which allows evaluating the hygrothermal behaviour and the related damage risks more reliably. However, this often involves thousands of simulations, which easily becomes computationally inhibitive, especially when analysing more-dimensional component connection details. To overcome this time-efficiency issue, this paper proposes the use of metamodels, which aim at imitating the original hygrothermal model with a strongly reduced calculation time. Many different metamodelling strategies exist, of which multiple linear regression (MLR) or polynomial linear regression might be the most frequently used. MLR attempts to model the relationship between multiple input variables and an output variable by fitting a linear equation. It often performs well when predicting aggregated values such as the total heat loss (Van Gelder et al. 2014). On the other hand, many damage criteria, such as wood decay or mould growth, require evaluation over time, as such damage often has a long incubation time, whereafter it accumulates. Hence, more advanced metamodelling strategies that can handle time series prediction are needed. Furthermore, the model must be able to capture the highly non-linear and non-stationary pattern of the hygrothermal response of building components. Hence, this paper focusses on neural networks, as they have proven successful in other non-linear and non-stationary research applications.

The multilayer perceptron (MLP) is the most known artificial neural network. It has a feed-forward structure with one input layer, one output layer, and at least one hidden layer in between. Because the MLP can perform a non-linear mapping, it is a widely used (meta-)modelling method. In time-series prediction, it is sometimes used for predicting the next time step based on the current step (Soleimani-Mohseni et al. 2006). However, the MLP is static and has no memory of past time steps; hence, it cannot model input-output relations that span multiple time steps. When these temporal dependencies are short-term, often a time window is added (MLP TW), i.e. the current time step as well as a number of past time steps are used as input to the network (Kemajou et al. 2012). However, the MLP TW fails to capture patterns outside of this time window. Since the time window needs to be determined beforehand, a considerable number of experiments is required to identify the optimum time window.
Recurrent neural networks (RNN) overcome this problem by introducing memory. A simple RNN has a cyclic structure that feeds the output from previous time steps into the current time step as input. Hence, RNNs can model temporal contextual information along time series data. The simple RNN easily suffers from the vanishing or exploding gradient problem though, which makes it difficult for the network to learn correlations between temporally distant events. To deal with this problem, the long-short memory network (LSTM) was proposed by Hochreiter (1998). The LSTM model changes the structure of the hidden units to memory cells with gates. Via these gates, the LSTM unit is able to decide what information to keep from its existing memory, while the simple recurrent unit overwrites its content at each time-step. Hence, if the LSTM unit detects an important feature from an input sequence at early stage, it easily carries this information over a long distance, thus capturing potential long-term dependencies (Chung et al. 2014). Consequently, the LSTM has been widely used for many time-series forecasting and sequence-to-sequence modelling tasks. More recently, the gated recurrent unit (GRU) was proposed by Cho et al. (2014) to make each recurrent unit adaptively capture dependencies of different time scales. Similarly to the LSTM, the GRU has gating units that modulate the flow of information inside the unit, but without having separate memory cells. Both the LSTM unit and the GRU keep their existing memory and add the new content on top of it.

In this paper, the LSTM and GRU, as well as the MLP and MLP TW, are applied for predicting the hygrothermal behaviour of a massive masonry wall, as an explorative study. All network types are compared based on their prediction performance and training time.

DATA AND METHODS

Data description

As an example of predicting more complicated damage patterns, the hygrothermal performance of a massive masonry wall is evaluated for frost damage, wood decay of embedded wooden beam ends and mould growth. To estimate whether these damage patterns will occur, the temperature (T), relative humidity (RH) and moisture content (MC) are monitored at the associated positions for a period of 6 years (see Table 1). In this explorative study, the probabilistic aspect of the influencing parameters was not yet taken into account fully, as this allows for a more efficient exploring of several network architectures on a smaller dataset. The data was obtained via hygrothermal simulations in Delphin 5.8. The used input parameters are shown in Table 2; the brick material properties are given in Table 3. To account for variability in boundary conditions, different years of the external climate were used, as well as different wall orientations. Since the interior climate is calculated based on the exterior climate, this variability is also included. In total, 24 samples were simulated, of which 18 samples were used for training and 6 for testing. The neural networks are trained to predict the T, RH and MC time-series, given the time-series of the external temperature and relative humidity, the wind-driven-rain load, the short-wave radiation and the internal temperature and relative humidity.

Table 1. The monitored quantities for the damage patterns at different positions in the wall.

Damage pattern	Position	Quantity
Frost damage	0.5 cm from exterior surface	T, RH, MC
Decay of wooden beam ends	5 cm from interior brick surface	T, RH
Mould growth	Interior surface	T, RH

Table 2. Used input parameters for hygrothermal simulations of brick wall.

Input parameter	Value
Brick wall thickness	360 mm
External climate	Gaasbeek, Belgium
Internal climate	cfr. EN 15026 A
Wall orientation	U(0,360)
Rain exposure factor	1
Solar absorption	0.4

* U(a,b): uniform distribution between a and b

Table 3. Properties of the used brick type.

Material property	Value
Dry thermal conductivity (W/m²K)	0.87
Dry vapour resistance factor (-)	14
Capillary absorption coefficient (kg/m²s$^{0.5}$)	0.277
Capillary moisture content (m³/m³)	0.25
Saturation moisture content (m³/m³)	0.35

Network architecture

Several hyper-parameters should be pre-set before building and training the networks, including the number of hidden layers and the number of neurons in these hidden layers. In this paper, all constructed networks have a single hidden layer, as comparative experiments showed no benefits using multiple hidden layers. Furthermore, networks with 32, 64, 128 and 256 hidden units were tested. In case of the MLP TW, a time window of 24 hours was explored. A larger time window, which would be required to capture long-term dependencies, resulted in an extensive input dataset which became too memory intensive. The networks are trained by minimising the mean squared error (MSE) via backpropagation (MLP and MPL TW) or backpropagation-though-time (LSTM, GRU). Based on the results of comparative experiments, the LSTM and GRU networks were trained using the RMSprop learning algorithm (Hinton, 2012) with a learning rate of 0.002. The MLP and MLP TW networks were trained using the Adam learning algorithm (Kingma and Ba, 2015) with a learning rate of 0.001. In general, before presenting data to the network, the data is standardised (zero mean, unit variance) to overcome influences from parameter units. In this paper, both the input and output data are standardised, as this was found to improve training speed and accuracy. The network's accuracy is evaluated by three indicators: the normalised root mean square error (NRMSE), normalised mean absolute error (NMAE), and coefficient of determination (R2), formulated as follows:

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum(y-y^*)^2}}{y_{max}-y_{min}} \qquad NMAE = \frac{\frac{1}{n}\sum|y-y^*|}{y_{max}-y_{min}} \qquad R2 = 1-\frac{\sum(y-y^*)^2}{\sum(y-\bar{y})^2} \tag{1}$$

where y is the true output, y^* is the predicted output, $\bar{y}$ is the mean of the true output and n is the total number of time steps. The RMSE and the MAE are normalised to remove the scale differences between the different outputs. Although the networks are trained to predict all outputs simultaneously, these performance indicators are calculated for each output separately. This allows assessing which outputs are more difficult to predict and which ones are easy.

RESULTS AND DISCUSSION

Figure 1 presents the three performance indicators (rows) for all outputs (columns) and each network type. This graph shows that some outputs are easier to predict than others; all networks are capable to predict the temperature accurately, as well as the interior surface relative humidity. Since the wall temperature and surface relative humidity respond almost immediately to a change in boundary conditions, not much memory is needed to accurately predict these profiles; for these outputs, the MLP with a time window of 24 hours is about as accurate as the more complicated LSTM or GRU. The relative humidity and moisture content at 0.5 cm from the exterior surface (frost damage) and the relative humidity at the wooden beam ends (wood decay) appear less evident to model. As moisture is transported inwards only slowly, there is a large delay between a change in boundary conditions (e.g. a heavy rain shower) and the response in relative humidity in the wall. At the wooden beam ends, this response delay can go up to several months. Hence, the MLP and MLP TW, which have no or only limited memory, are incapable to capture these long-term temporal dependencies and perform poorly. The LSTM and GRU, on the other hand, are able to capture these complex long-term patterns because of their connections to information from long-past time-steps. Figure 2 shows the temperature and relative humidity at the wooden beam ends predicted by the GRU network with 256 hidden units, compared with the true value. The residuals ε_t show that that the error is very small.

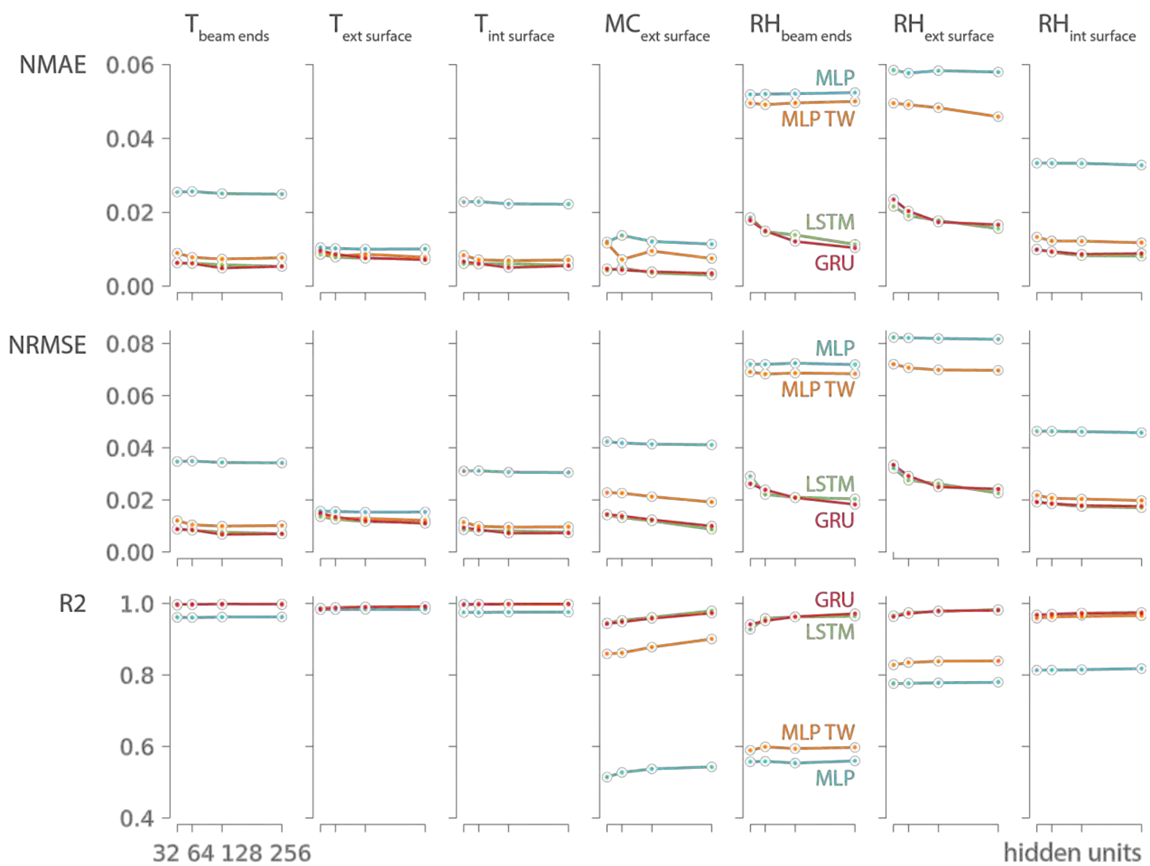

Figure 1. The performance indicators show that the temperature patterns are easy to model, while only the memory networks types (LSTM, GRU) are able to model the moisture content and relative humidity patterns accurately.

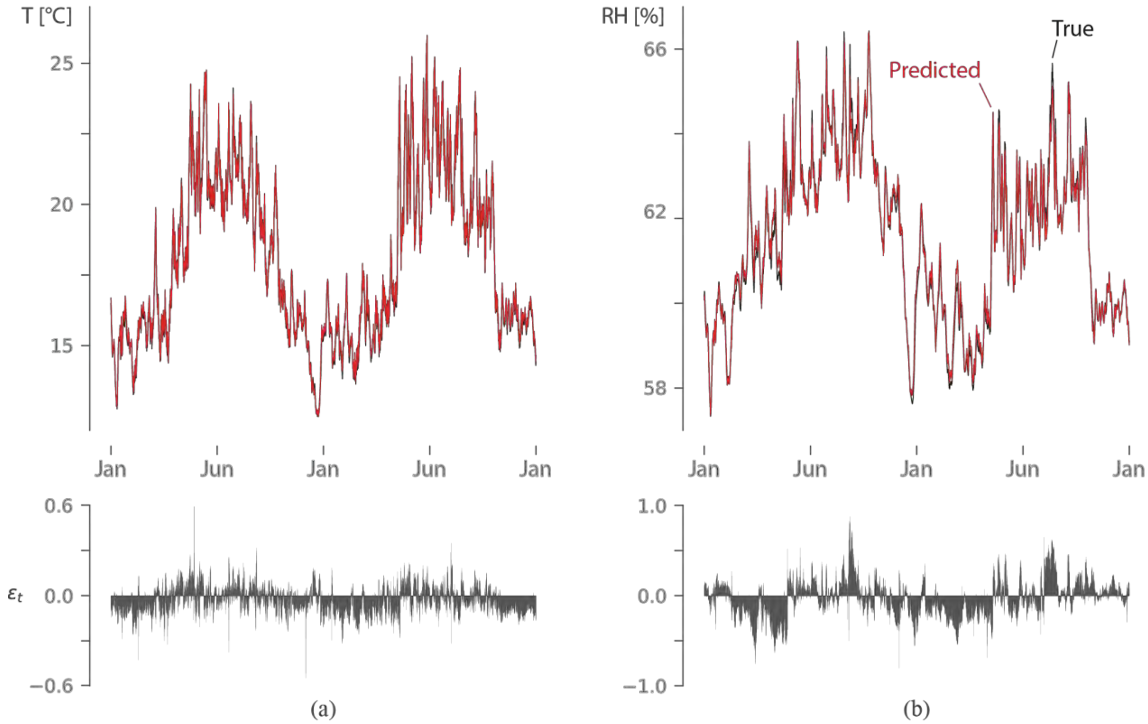

Figure 2. The temperature (a) and relative humidity (b) at the wooden beam ends, predicted by the GRU network with 256 hidden units. The residuals ε_t show that that the error is small.

Furthermore, it was found that the number of hidden units appears to have limited effect on the prediction performance in case of the MLP and MLP TW, though it increased the training time (Figure 3). In case of the LSTM and GRU on the other hand, more hidden units resulted in an improved prediction performance and a decrease in training time. It appears that the memory networks converge faster when they have more hidden units. Though prediction accuracy is very similar for both the LSTM and the GRU, the latter required less training time as it has fewer network parameters.

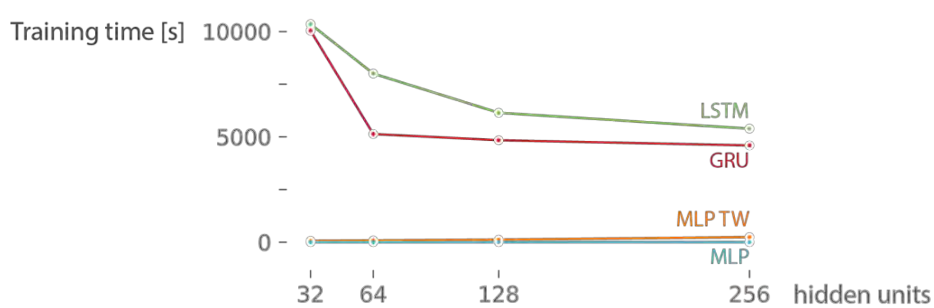

Figure 3. The training time for the memory network types (LSTM, GRU) decreases with increasing number of hidden units. The simulation time for one sample is 1 s for all network types. The reference simulation time for one sample in the original hygrothermal model is about 720 s.

CONCLUSIONS

In this paper, the hygrothermal simulation model for a massive masonry wall is replaced by a much faster metamodel. Four different types of neural networks were considered as metamodel:

a traditional multilayer perceptron (MLP), a multilayer perceptron with a pre-defined time window (MLP TW), a long-short term memory network (LSTM) and a gated recurrent unit network (GRU). Only the last two types have dynamic memory. The MLP TW has only access to the past time steps within the pre-defined time window, and the MLP tries to predict based on the current time step only. It was found that all network types were capable to predict the temperature accurately. Since the temperature responds almost immediately to a change in boundary conditions, not much memory is needed to capture these patterns. By contrast, only the LSTM and the GRU were able to accurately capture the long-term dependencies needed to predict the relative humidity and moisture content, as these respond much slower to a change in boundary conditions. Both types of memory networks were found to have almost equal predicting accuracy, though the GRU converged faster and thus required less training time.

ACKNOWLEDGEMENT
This research is part of the EU project RIBuild, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 637268. This financial support is gratefully acknowledged.

REFERENCES

Cho K., Van Merrienboer B., Bahdanau D. and Bengio Y. 2014. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint* arXiv:1409.1259.

Chung J., Gulcehre C., Cho K. and Bengio Y. 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv preprint* arXiv:1412.3555.

Hinton G., Srivastava N. and Swersky K. 2012. Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. *COURSERA: Neural Networks for Machine Learning.*

Hochreiter S. 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02), 107–116.

Janssen H., Roels S., Van Gelder L. and Das P. 2015. Annex 55 - Reliability of Energy Efficient Building Retrofitting, Probability Assessment of Performance and Cost. Final Report. Chalmers University Of Technology Gothenburg (SE), 152 pages.

Kemajou A., Mba L. and Meukam P. 2012. Application of artificial neural network for predicting the indoor air temperature in modern building in humid region, *British Journal of Applied Science and Technology*, 2(1), 23–34.

Kingma D.P. and Ba L.J. 2015. Adam: A Method for Stochastic Optimization. *In: Proceedings of International Conference on Learning Representations.* arXiv:1412.6980

Soleimani-Mohseni M., Thomas B. and Fahlén P. 2006. Estimation of operative temperature in buildings using artificial neural networks. *Energy and Buildings*, 38(6), 365-640.

Van Gelder L., Das P., Janssen H. and Roels S. 2014. Comparative study of metamodelling techniques in building energy simulation: Guidelines for practitioners. *Simulation Modelling Practice and Theory*, 49, 245–257.

Vereecken E., Van Gelder L., Janssen H., and Roels S. 2015. Interior insulation for wall retrofitting – A probabilistic analysis of energy savings and hygrothermal risks, *Energy and Buildings*, 89, 231–244.