A comparison of machine learning functions for time-series prediction in buildings

Blair Birdsell[1], Ralph Evins[2]
[1]Energy in Cities group, Victoria, BC, Canada
[2]Energy in Cities group, Victoria, BC, Canada

Abstract

This work undertakes a comprehensive comparison of RNNs and CNN-based ResNets for both multivariable day-ahead and annual predictions with specific focus on their application in buildings. This varied comparison of two types of nets under two types of prediction conditions documents and describes methods that will be impactful to building designers and operators. The comparison shows that both CNN-based ResNets and RNNs are suitable for short-term forecasting, however, the study establishes RNN's strength in forecasting sequences. This has implications when trying to predict the short-term behavior of building systems.

When the annual performance was estimated, CNN-ResNets showed a distinct advantage in accurately predicting the expected monthly minimum and maximum values. This result can contribute to better decision making in the design and planning process.

Introduction

The goal of this project is to inform best practice in regards to predicting building performance using machine learning. Like many papers on this topic, it is fundamentally motivated by the fact that in North America buildings are responsible for consuming 30-40% of all energy produced (U.S. Department of Energy 2012, Statistics Canada 2016). Therefore, improving building operations has the potential to make a significant impact on reducing energy consumption and greenhouse gas emissions.

In recognition of the increasing amount of data a monitored building can produce, and the rapidly evolving ecosystem of machine learning algorithms, building owners and operators are left with a lot of building data and few answers about how best to use it. Instead of modifying an existing neural net architecture until it's considered novel, adding to the confusing landscape of choices, this project instead undertakes a detailed study of well-known and well-established network architectures and compares their accuracy and computational performance. The research looks at two major classes of neural networks, CNN-based residual networks and recurrent neural networks, under two types of prediction conditions, day-ahead forecasting and long-term predictions. It's a body of work that can contribute to the decision-making process in several of the applications described below.

Applications

Smart Buildings

With the decreasing costs of computation and sensors deployment, it's becoming more common for buildings to implement building management systems, of which model predictive control is a major piece. Intelligent management systems have two goals: Firstly, achieving a high degree of comfort for users concerning the thermal, air quality and visual performance of the space. And secondly, achieving high energy efficiency while minimizing operational costs (Serale et al. 2018).

Model predictive control with neural nets can address several criteria that traditional control methods based on simulation cannot. MPC has the ability to handle uncertainties and interdependence among the parameters, track slow-moving processes that might be time delayed within the evolving system, capture nonlinear dynamics, and, finally, predict multiple objectives at the same time. The underlying thrust of the benefits of MPC is that it has the ability to use the system for anticipatory control rather than corrective control (Afram 2014).

Retrofits

Building owners and operators have started to generate large volumes of data regarding the performance and behaviour of their assets over time (Gonzalez-Vidal et al. 2017). Previous work with RNNs and multivariate time series have looked at the classification of buildings (Baasch 2019). Deb (2021) looked at machine-learning applications for cost-optimal building retrofits. The types of long-term estimates discussed in this paper aggregate characteristics of the building's performance over time so they can to contribute to the decision making process for the existing building stock.

IoT Building Sensors

The deployment of sensors and actuators within buildings allows them to both perceive the world as it is and interact with it. The web of sensors and controllers colloquially known as the Internet of Things presents challenges to

developers, designers, and engineers trying to integrate them for the purpose of gathering building intelligence (Carli et al. 2020). This opens the door for opportunities to optimize performance when looking ahead at expected demand, predicting variables not recorded by sensors, or when the relationships among a collection of sensors is unknown. For example, if the relationship between zone temperature and HVAC units is unknown, ANNs provide an opportunity to statistically undercover and predict these relationships through computation.

Prediction & Forecasting

Seeking a robust comparison of the algorithms in question, both day-ahead and annual predictions are considered. This brings up a subtle and informal distinction that should be explicitly noted between these two types of prediction and why there were chosen. Nate Silver in his 2012 book *The Signal and the Noise* states that in many fields the terms "forecast" and "prediction" are used interchangeably. In this paper they are not used interchangeably. *Forecasting* will be taken to mean a short-term estimate, while *prediction* will refer to a longer term aggregated estimate. Using energy consumption as an example, an energy consumption forecast would look at the past week of data and estimate the next day's usage in high-resolution. Prediction, on the other hand, would estimate the total energy consumed for the day, month, or year.

Related Work

Approximating Functions

Before continuing to look in depth at the strengths of the proposed network architectures, it's worth stepping back and describing the main statistical approach to prediction prior to the development machine learning techniques. Statistics-based, classical prediction can be summarized in one historical touch point; the introduction of Box and Jenkins' 1970 book Time series Analysis; Forecasting and Control where the authors first describe how to model time series using regression, differences, and moving averages of the residual error. Autoregressive Moving Average techniques can also be further generalized to multivariable cases through Vector Autoregression. Under this method, interaction among the variables to be predicted are taken into consideration (Lütkepohl 1991, Goel et al. 2017).

As the computational efficiency of computers increased over the last number of decades, it came to highlight the shortcomings of classical methods. Time dependence between past and future values are vital for a system's evolution. Causal interrelationships among coupled variables might occur in a non-uniform manner, but statistics-based methods assume linear dependence over time making them ill-suited for modeling long-term dependencies (Goel 2017, Koutlis 2020). Scalability is also an issue. Selection of the model's order and other parameters to achieve good results is nontrivial (Goel 2017, Koutlis 2020). This means in practice the model's parameters are often adjusted in an ad hoc manner (Gonzalez-Vidal 2017).

Artificial Neural Networks are black-boxes and – as used in this paper – do not rely on modeling or simulating the underlying building physics. The scenario is modeled using only data. Gonzalez-Vidal (2017) concludes these types of black-box models are more cost effective because of their ability to integrate new parameters easily compared to white-box models that are physical simulations of the building. The target variable only relies on current indicators such as temperature, humidity and other features (Gonzalez-Vidal 2017). Divina et al. (2019) and Cai et al. (2019) each published studies reporting ANN techniques beating traditional ARIMA in regards to accuracy.

Convolutional Based Residual Networks

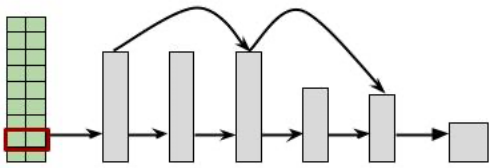

Figure 1: Schematic of CNN-based ResNet.

Convolutional layers in a residual neural network configuration is the first type of neural network architecture to be considered in this study. The basic idea of a convolution is to calculate the product of a function (the kernel) applied to another function (the time series). This product can reveal latent structure in the data. Time series are not often thought of as containing spatial information, but indeed, under certain assumptions, that is exactly one way they can be interpreted. The growth of the method in 20th and 21st century follows from the increased availability of computational power where the kernel can be applied over and over again to update arrays of weights to approximate complex behaviour.

In this project, convolutional layers are packaged into blocks allowing for skipped connections between layers that have been shown to improve accuracy on time series prediction and classification (He et al. 2016, Wang et al. 2019). While a linear network tries to learn the output function $H(x)$, the layers in a residual block aim to learn the residual as well: $H(x)=F(x)-x$ (Choi 2018, Benhaddi 2021). This structure allows the net to extract more characteristic information from the data.

Recurrent Neural Networks

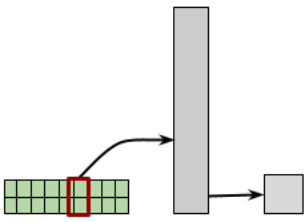

Figure 2: Schematic of RNN

Recurrent neural networks are used widely for time series modeling and are currently considered state-of-the-art for natural language processing. RNNs are seen in application in speech recognition (Graves et al. 2013), text generation (Graves 2013), and language translation (Sutskever 2014). The connection to time series becomes immediately clear when thinking of both language and building performance time series as both fundamentally sequences of values.

Their construction is difficult to describe and computationally intensive but their properties allow them to exhibit temporally dynamic behaviour. RNNs contain the expected arrays of weights similar to linear regression layers, but also have different configurations of state memory depending on the specific type of RNN layer used. The addition of state memory help the model capture and remember dependencies among consecutive input features. As the sequence is fed into the net during training, one by one the layer weights are updated toward the target value (Cheng 2006). Because of RNNs success in the temporal domain it has been studied in the field of building performance and energy consumption by Gao (2019) and Sehovac (2019).

Challenges

The challenges of working with building time series data fall into two categories.

Data Quality

The data that comes directly from sensors or mechanical units has several drawbacks that if left unaddressed could lead to substandard results for any machine learning application. One requirement for most traditional neural nets is that the input data is required to have uniform time steps. Gaps in data collection of varying lengths, from hours to months, can also occur when a sensor and mechanical unit fails or is down for maintenance etc. Unpredictable transient events are also common in building data. These events can be caused an influx of users at the site, such as at a special event, or by powering up large mechanical units. To a certain degree, depending on the requirements of the project, these sorts of artifacts can be mitigated by pre-processing the data.

Data Structure

Characteristics inherent in the data itself also make time series prediction challenging. These are important characteristics to understand because they are often tied to the structure of the data the model is trying to learn and make predictions from. Notably, these challenges are similar to the characteristics with which traditional statistical methods struggled.

Included in long-term time series are patterns of different length such as seasonal, weekly and daily variations. Certain high resolution building time series, taken in 15 minutes or shorter time steps, can contain high frequency noise that present a difficult environment to model. Trends in the data can also represent non-linear relationships. Systems can evolve in ways that go beyond linearly proportional effects. Lastly, the building data used for this project is multivariate, where each variable's future value rely on its own past values as well as the past values of other variables. The possibility of interactions among the variables is a challenge to undercover and predict.

Methodology

Data Source

In this study, Building 59 at Lawrence Berkeley National Laboratory in Berkeley, California was used for both annual and day-ahead predictions. It is a 112 000 square foot office building containing two office floors, one mechanical equipment floor, and one floor dedicated to a data and computing center. Only the datasets for the office floors were released. Building data was collected from January 2017 to Dec 2019. It included HVAC and lighting systems states, building-level energy consumption, indoor environmental conditions, and on-site weather (U.S. Department of Energy 2021). The recorded time resolution depends on the data type, and ranges from 1 minute to 30 minute ticks. The dataset was analyzed and prepared by LBNL and addressed data file aggregation and categorization, diagnostic gap filling, and outlier modification. Detailed information about the raw data's post-processing can be found in the dataset's accompanying documentation (Lawrence Berkeley National Laboratory 2021).

Data Processing

Two further steps were taken to prepare the dataset for the training and testing undertaken in this experiment. After selecting which features would be used and targeted, the data was down sampled to 30 minute time steps. This created a dataset of 35041 rows of 70 features, two of which were types of time specification and not used for training.

The second step was to standardize the input features. This was done by subtracting the mean of the feature from each value and dividing it by the standard deviation of the feature. No modification was done to the prediction targets. When training for the annual predictions the dataset was further down sampled to a 1 hour resolution to ease the computational burden of the training phase.

Use Case

The dataset was used to predict the flow rate of the building's 4 roof top units. The prediction model has no underlying knowledge of the physical system, nor is the flow rate ever used as an input for training. During the training process the required flow rate is deduced from the provided data set, with the approximating function created by the neural nets used to map the inputs to the flow rate at $t+1$. The model is never given the relationship between temperature zones and RTUs, and therefore must deduce that information itself. Predicting the required flow rate of the RTUs in the future given the system's previous state is

representative of what needs to be predicted and controlled by the model in a building management system.

Training

64 features were used for training including the building zone temperatures, light electricity loads, miscellaneous electricity loads, heat pump status, site weather, business day bool, and calendar month.

A 72-hour window was used on 5 weeks of training data. The output of each window at 30 min timesteps was the next 30 minutes of RTU flow rate values. The annual training data was segmented into 24-hour periods with the previous 7 days of data. This makes the target an array of 24 values for each of the 4 rooftop units.

Training was conducted on a 4 Ghz Quad-core iMac with 32GB of RAM operating macOS BigSur. All coding and neural net construction was completed in Wolfram Mathematica 13.

Inputs	Outputs
Business day boolean	Roof top unit 1 (CFM)
Month + fraction of month	Roof top unit 2 (CFM)
Air temperature set 1 & 2 (°C)	Roof top unit 3 (CFM)
Dew point temperature (°C)	Roof top unit 4 (CFM)
Relative humidity (%)	
Solar radiation (W/m²)	
Zone #16 to #72 temperature (°C)	
Heat pump temperature (°C)	
Miscellaneous electrical load north & south (W)	
HVAC demand north & south (W)	
Lighting (W)	

Table 1. Table of model parameters.

Performance Metrics

To frame the day-ahead results this project uses the Mean Absolute Error to represent the magnitude of error compared to the historically recorded performance of the RTUs. Secondly, the Mean Absolute Percentage Error is calculated and presented for comparison which has been used in a wide number of building performance studies. (Edwards et al. 2012, Fan et al. 2014, Gonzalez-Vidal 2017). In the literature reviewed for this study, the MAPE reported by CNNs and RNNs ranged between 5-10% (Gonzalez-Vidal 2017, Koutlis 2020).

In regards to annual predictions, the well-known ASHRAE Guideline 14 lets annual predictions vary by up to 20%, but recent authors have noted this may not be sufficient or meaningful in practice (A. S. Committee et al., 2002, Garrett 2014, Gonzalez-Vidal 2017). In this paper the monthly min and max values are calculated and compared between the estimated and ground truth values.

Results

Description of results

Day-ahead Forecast

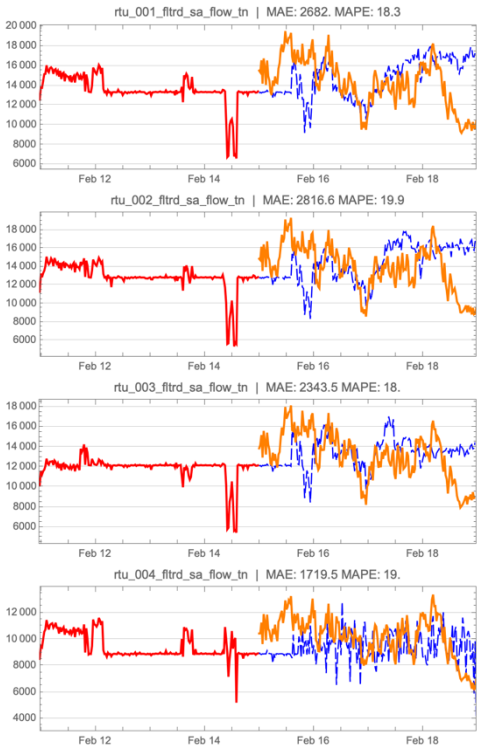

Figure 3. Day ahead ResNet cubic feet per minute results (y-axis).

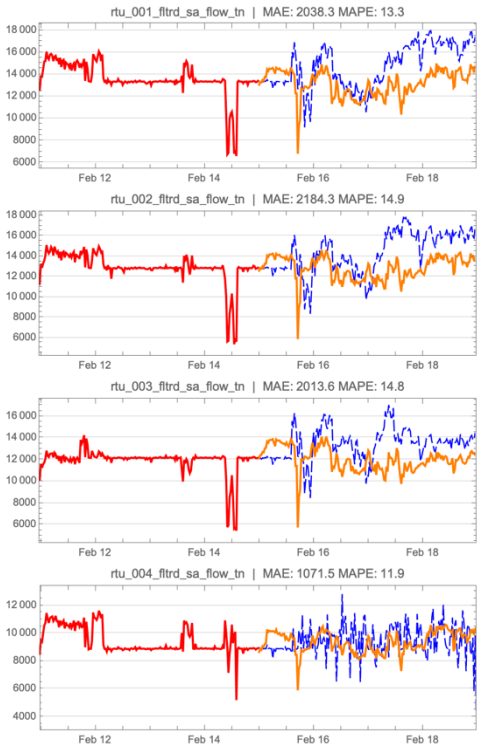

Figure 4. Day Ahead RNN cubic feet per minute

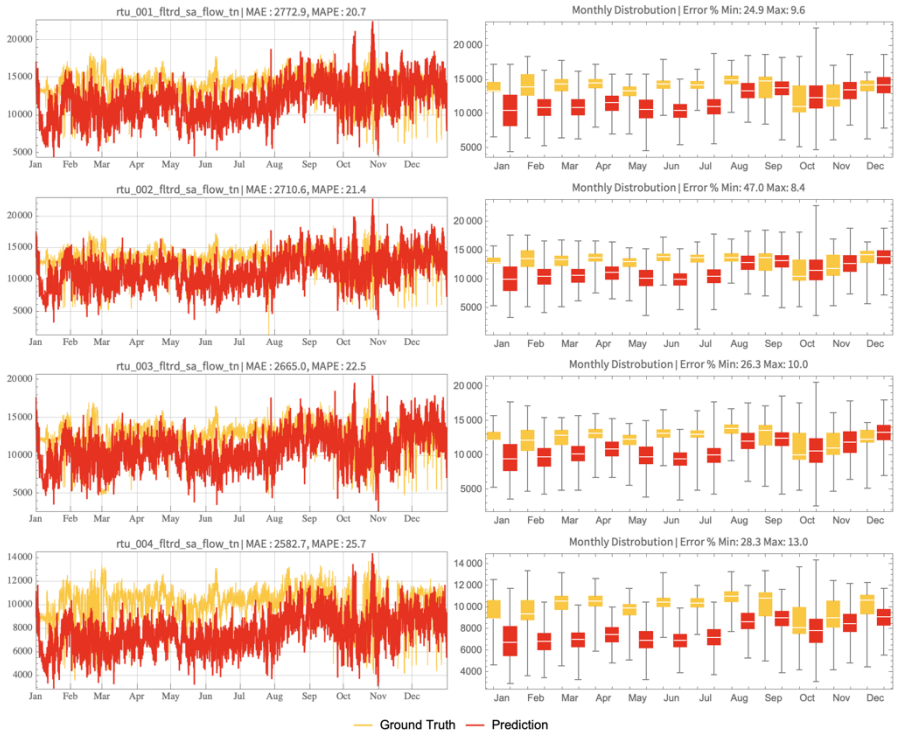

Figure 5. Annual ResNet cubic feet per minute prediction (y-axis) time series and monthly distribution of values as Box-whisker diagram.

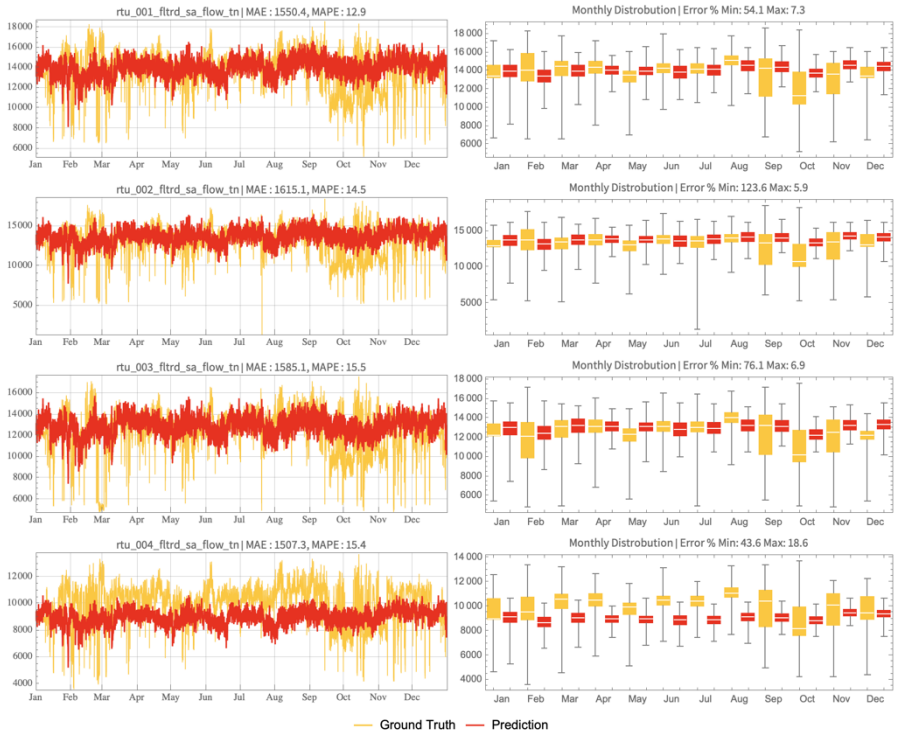

Figure 6. Annual RNN cubic feet per minute prediction (y-axis) time series and monthly distribution of values as Box-whisker diagram.

results (y-axis).

	MAE	MAPE
Day-ahead ResNet	2390.4	18.8%
Day-ahead RNN	1826.9	13.7%

Table 2. Combined Average Error Results.

When the MAE and MAPE values are averaged among the 4 RTUs, the RNN is shown to be around 5% more accurate. Referring to the plots, the CNN-based ResNet can be seen converging on the next day and replicating the noise of the real signal adequately, but accuracy falls off as the model extends past the 24-hour forecast period. The RNN forecast is more accurate but follows the previous sequence too closely by mimicking the behaviour of the previous day's deep dip in airflow.

Annual Prediction

Both the CNN-ResNet and RNN did an admirable job representing the high-noise environment of the original data. The CNN-ResNet (Figure 5) was better able to predict the distribution of monthly values and the monthly minimum and monthly values with the exception of the anomaly in the October maximum value. The RNN model (Figure 6) tended to constrain the values in an unrealistic manner leading to inaccurate monthly minimum and maximum predicted values when calculated.

Interpretation & analysis of behaviour

Day-ahead Forecast

It's visually noticeable that accuracy decreases as the forecast window extends past the 24 hour estimate and the model begins to encounter more data it's never seen before. This is realistic of real world building applications where time series are continually added to over time and only ever see small periods of new data as the forecast window moves and model updated. Each model was relatively capable at reproducing the high-frequency fluctuations seen in the historic data making them strong candidates to forecast a building's behaviour in high-resolution. Considering the RNN's 5% greater accuracy, that is the author's suggestion of where to start for short-term forecasting of building performance data.

Annual Prediction

While the strengths of RNNs are in predicting sequences, the examples here reveal its shortcomings when aggregating their values as a distribution. The RNN was not able to match the ground truth values' distribution accurately. This makes its performance suspect for use in the decision making process. The CNN-ResNet took longer to converge but standouts as a good candidate to use in the decision making process because of its ability to predict similar minimum and maximum values as the real data. The CNN-ResNet is probably doing a better job at uncovering the deep latent knowledge contained within the annual data.

Estimating annual behaviour from only one previous year is a challenging environment from which to make predictions, and this probably had an effect overall accuracy achieved. Multiple years of data would help the model generalize better to future periods.

Conclusion

Overview

ResNets and RNNs will be widely known by people who do time series modeling in fields such as economics or healthcare. This study outlines their suitability for model predictive control in building management systems. It is valuable record of their performance under both short-term and long-term prediction conditions, and facilitates their adoption in building design and operations. RNN showed it strength in forecasting sequences, with better accuracy in short-term forecasting, and should probably be given more consideration when selecting a model for time series forecasting in the building industry. CNN-ResNets, in both forecasting and prediction, were slow to converge but showed an excellent ability to draw out the latent structure of the data.

A final point worth noting is that the neural networks described in this work were able to achieve a high level of performance without any underlying knowledge of the physical systems that generated the data. In practice however, knowledge of the field should still be included in their application to reduce errors of interpretation or application. Miller et al. (2020) came to a similar conclusion regarding the need of a hybrid approach blending data science and engineering knowledge when modeling building data.

Future Work

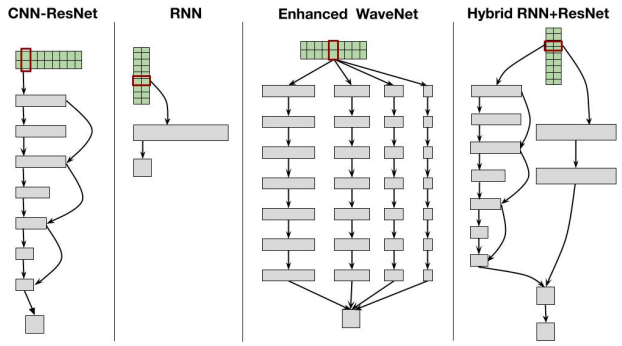

Figure 7. Schematic description of 4 different types of neural net.

This work lays the foundation to improve the our understanding of ResNets and RNNs by increasing their ability to capture features of time series data at different scales. Oord et al. (2016) introduced WaveNet architecture, a variation of CNN-based ResNets, that has proved to deal well with series containing global and locals trends and high-frequency noise. Rueda et al. (2021) has looked at its short-term load forecasting performance. An Enhanced WaveNet architecture has yet to be robustly compared or applied to building performance data for use in MPC and

long-term predictions. A final major addition to this comparison is a hybrid ResNet-RNN configuration (Goel 2017, Choi 2018, Khan 2020). As summarized in Figure 7, these neural net architectures are not unknown to the world of computer science, but studying and describing their suitability for projects in the building industry is valuable.

References

A. S. Committee et al., (2002). Ashrae guideline 14, measurement of energy and demand savings. Atlanta.

Afram, A. & Janabi-Sharifi, F. (2014). Theory and applications of HVAC control systems – A review of model predictive control (MPC), Building and environment, vol. 72, pp. 343-355.

Baasch, G.M. and Evins, R., (2019). Targeting Buildings for Energy Retrofit Using Recurrent Neural Networks with Multivariate Time Series. In Neural Inf. Process. Syst.

Benhaddi, M. & Ouarzazi, J. (2021) Multivariate Time Series Forecasting with Dilated Residual Convolutional Neural Networks for Urban Air Quality Prediction, Arabian journal for science and engineering (2011), vol. 46, no. 4, pp. 3423-3442.

Cai, M., Pipattanasomporn, M. & Rahman, S. (2019), Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques, Applied energy, vol. 236, pp. 1078-1088.

Carli, R., Cavone, G., Ben Othman, S. & Dotoli, M. (2020). IoT Based Architecture for Model Predictive Control of HVAC Systems in Smart Buildings, Sensors (Basel, Switzerland), vol. 20, no. 3, pp. 781.

Cheng, H., Tan, P., Gao, J. & Scripps, J. (2006). Multistep-Ahead Time Series Prediction. Advances in Knowledge Discovery and Data Mining Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 765-774.

Choi, H., Ryu, S. & Kim, H. (2018), Short-Term Load Forecasting based on ResNet and LSTM, 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, pp. 1.

Deb, C., Dai, Z. & Schlueter, A. (2021), A machine learning-based framework for cost-optimal building retrofit, Applied energy, vol. 294, pp. 116990.

Deb, C., Zhang, F., Yang, J., Lee, S.E. & Shah, K.W. (2017). A review on time series forecasting techniques for building energy consumption, Renewable & sustainable energy reviews, vol. 74, pp. 902-924.

Divina, F., García Torres, M., Goméz Vela, F.A. & Vázquez Noguera, J.L. (2019). A Comparative Study of Time Series Forecasting Methods for Short Term Electric Energy Consumption Prediction in Smart Buildings, Energies (Basel), vol. 12, no. 10, pp. 1934.

Edwards, R.E., New, J., Parker, L.E. (2012). Predicting future hourly residential electrical consumption: A machine learning case study. Energy and buildings, vol. 49, pp. 591-603.

Fan, C., Xiao, F. & Wang, S. (2014). Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. Applied Energy, vol. 127, pp. 1–10.

Gao, Y., Fang, C. & Ruan, Y. (2019). A novel model for the prediction of long-term building energy demand: LSTM with Attention layer, IOP conference series. Earth and environmental science, vol. 294, no. 1, pp. 12033.

Garrett, A. & New, J. (2016). Suitability of ASHRAE guideline 14 metrics for calibration. ASHRAE transactions, vol. 122, no. 1, pp. 469.

George E. P. Box, Jenkins, G. M., & Reinsel, G. C. (2008). Time series analysis: Forecasting and control, fourth edition. Wiley.

Goel, H., Melnyk, I. & Banerjee, A. (2017). R2N2: Residual Recurrent Neural Networks for Multivariate Time Series Forecasting, arXiv:1709.03159.

Gonzalez-Vidal, A., Ramallo-Gonzalez, A.P., Terroso-Saenz, F. & Skarmeta, A. (2017). Data driven modeling for energy consumption prediction in smart buildings, IEEE, pp. 4562.

Graves, A. (2013). Generating Sequences With Recurrent Neural Networks. arXiv:1308.0850v5

Graves, A., Mohamed, A. & Hinton, G. (2013). Speech recognition with deep recurrent neural networks, 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. pp. 6645-6649.

He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770.

Khan, Z.A., Hussain, T., Ullah, A., Rho, S., Lee, M. & Baik, S.W. (2020). Towards Efficient Electricity Forecasting in Residential and Commercial Buildings: A Novel Hybrid CNN with a LSTM-AE based Framework, Sensors (Basel, Switzerland), vol. 20, no. 5, pp. 1399.

Koutlis, C., Papadopoulos, S., Schinas, M. & Kompatsiaris, I. (2020). LAVARNET: Neural network modeling of causal variable relationships for multivariate time series forecasting. Applied soft computing, vol. 96, pp. 106685.

Lawrence Berkeley National Laboratory (2021). Dataset Description and Data Cleaning for Building 59 Raw Data: The Benchmark Datasets Project.

Lütkepohl, H. (1991). New Introduction to Multiple Time Series Analysis, Springer Berlin Heidelberg, Berlin, Heidelberg.

Miller, C., Arjunan, P., Kathirgamanathan, A., Fu, C., Roth, J., Park, J.Y., Balbach, C., Gowri, K., Nagy, Z., Fontanini, A.D., Haberl, J. & National Renewable Energy Lab. (NREL), Golden, CO (United States) (2020) The ASHRAE Great Energy Predictor III competition: Overview and results. Science & technology for the built environment, vol. 26, no. 10, pp. 1427-1447.

Oord, A.v.d., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. & Kavukcuoglu, K. (2016), WaveNet: A Generative Model for Raw Audio. arXiv:1609.03499

Rueda, F.D., Suárez, J.D. & Alejandro del Real Torres (2021), Short-Term Load Forecasting Using Encoder-Decoder WaveNet: Application to the French Grid, Energies (Basel), vol. 14, no. 2524, pp. 2524.

Sehovac, L., Nesen, C. & Grolinger, K. (2019). Forecasting Building Energy Consumption with Deep Learning: A Sequence to Sequence Approach, IEEE, , pp. 108.

Serale, G., Fiorentini, M., Capozzoli, A., Bernardini, D. & Bemporad, A. (2018). Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities, Energies (Basel), vol. 11, no. 3, pp. 631.

Silver, N., (2012). The Signal and the Noise: Why So Many Predictions Fail—But Some Don't. Penguin.

Statistics Canada. (2016). Survey of Commercial and Institutional Energy Use, 2014. Monthly energy review. The Daily. https://www150.statcan.gc.ca/n1/daily-quotidien/160916/dq160916c-eng.pdf Accessed Feburary 23 2022.

Sutskever, I., Vinyals, O. & Le, Q.V. (2014). Sequence to Sequence Learning with Neural Networks. arXiv:1409.3215.

U.S. Department of Energy (2012). 2011 Buildings Energy Data Book. Buildings Technologies Program. Energy Efficiency and Renewable Energy. https://ieer.org/wp/wp-content/uploads/2012/03/DOE-2011-Buildings-Energy-DataBook-BEDB.pdf. Accessed February 23 2022.

U.S. Department of Energy (2021). The Buildings Data Platform. Office of Energy Efficiency and Renewable Energy's Building Technologies Office. https://bbd.labworks.org/ds/bbd/lbnlbldg59. Accessed February 23 2022.

Wang, K., Li, K., Zhou, L., Hu, Y., Cheng, Z., Liu, J. & Chen, C. (2019). Multiple convolutional neural networks for multivariate time series prediction. Neurocomputing (Amsterdam), vol. 360, pp. 107-119.