

Data-Fusion for Increasing Temporal Resolution of Building Energy Management System Data

Dumidu Wijayasekara, Milos Manic

Department of Computer Science
Virginia Commonwealth University
Richmond, Virginia, USA

dumidu.wijayasekara@gmail.com, miskko@ieee.org

Abstract—Buildings are known to be significant energy consumers throughout the world. Thus, improving the energy efficiency of buildings is a key research goal. However, maintaining occupant comfort while improving energy efficiency in buildings requires close monitoring of the building environment and immediate control actions taken when sub-optimal behavior is identified. Such monitoring requires high frequency data from sensors. Therefore, increasing the data collection rate or the temporal resolution of sensors can lead to improved building control and state-awareness. This paper presents an on-line learning, data-fusion based methodology that uses Artificial Neural Networks (ANNs) to increase temporal resolution of building sensor data. The presented method utilizes sensor information from different sensors in the building to predict higher temporal resolution data of specific sensors. Furthermore, the presented method is capable of learning changing building behavior for improved long-term accuracy. The presented method was applied to a real-world building dataset and was shown to be able to predict high temporal resolution data with a higher accuracy compared to classical methods. Furthermore, the on-line learning was shown to increase the prediction accuracy in long-term operation.

Keywords—*building energy management systems; data-fusion; on-line learning; artificial neural networks*

I. INTRODUCTION

According to several studies, 20-40% of the world energy production is consumed by buildings [1], [2], [3]. In industrialized countries such as the U.S. and the Euro region countries, the energy consumption of buildings is closer to the 40% mark and surpasses other areas such as industry and transportation [1], [4]. Such energy consumption also means that buildings account for significant portion of greenhouse gas production [1]. It has been estimated that buildings are responsible for over 30% of greenhouse gas production worldwide [5]. Furthermore, economic and population growth among other factors has led to an increasing trend in building energy usage in recent years [1], [4], [6].

Thus, increasing energy efficiency of buildings is gaining significant interest [2], [7], [8]. It has been identified that the largest consumer of building energy is Heating Ventilation and Air Conditioning (HVAC) systems, which are responsible for 30-50% of total energy consumed in the buildings [2], [5], [8], [9]. Thus, improving HVAC performance is a significant factor

that will yield better building energy performance [3], [4], [10], [11].

By implementing very low cost building management strategies alone, it has been shown that the energy efficiency of modern HVAC systems can be improved by more than 5% [2], [5]. Furthermore, it has been shown that the energy efficiency of HVAC systems can be improved by up to 40% by close monitoring and advanced control [12].

Another key factor in HVAC control is occupant comfort [13], [14]. As humans spend significant amount of time indoors, a comfortable environment ensures improved productivity and health [15].

Thus, modern buildings utilize Building Energy Management Systems (BEMS) for controlling building HVAC and lighting to improve energy efficiency while maintaining a comfortable indoor environment [3], [16]. Modern BEMS systems are highly complex systems that utilize sensor data from various locations inside and outside the building to achieve optimal control [17], [18], [19]. Furthermore, these systems act as a state-awareness tool for building managers who are responsible for maintaining building comfort levels and operation of HVAC [3], [4], [20]. To achieve close monitoring of building environment and thereby achieving the energy efficiency goals and maintaining high levels of state awareness, sensor readings with a high temporal resolution is required [21], [22]. However, because of the high cost and time involved in installing new or additional sensors, increasing the temporal resolution of building sensors is a difficult task [23], [24].

Therefore, this paper presents a novel, on-line learning data-fusion based methodology for increasing temporal resolution of building sensors. The data-fusion techniques enable combining information from multiple inter-related sources to generate most accurate predictions [25], [26]. Using data-fusion techniques has been shown to be effective in predicting behavior more accurately [26], [27]. The presented method utilizes Artificial Neural Networks (ANNs) for predicting high temporal resolution sensor data using multiple sensor information from throughout the building. Furthermore, the presented method is capable of learning new building behavior on-line, through known sensor data, at normal measurement intervals. This enables the presented method to consistently produce accurate predictions in the long term. The presented method was applied

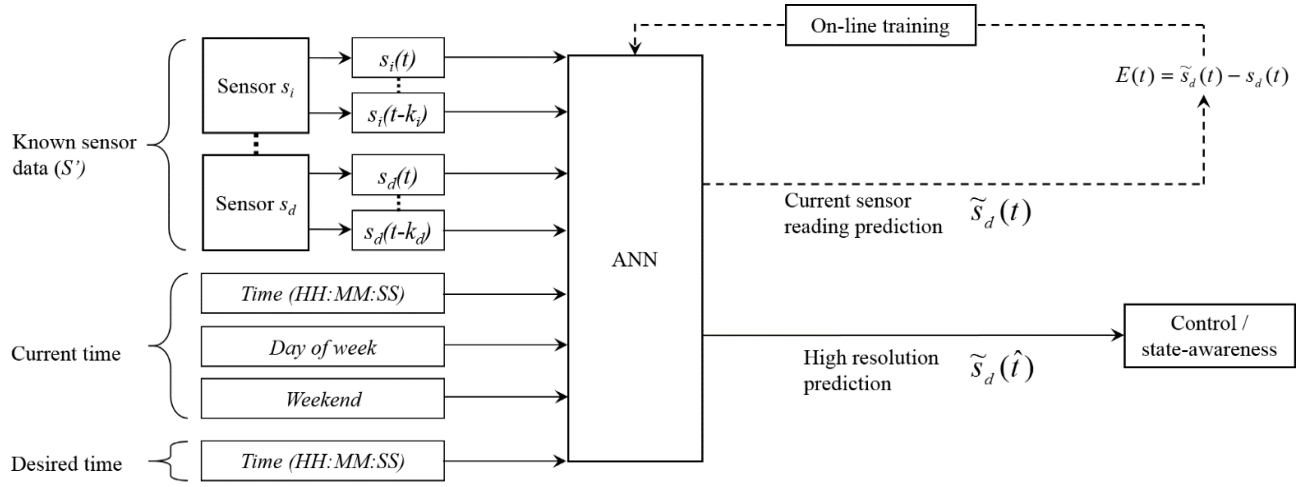


Fig. 1 The overall architecture of the presented methodology

to a real-world dataset and was shown to be more accurate than classical interpolation based and polynomial regression based prediction. Furthermore, the presented method was shown to be adaptive to changing building behavior through on-line learning, and was shown to improve the prediction accuracy in the long-term.

This paper is organized as follows; section II details each module of the presented method for increasing temporal resolution of sensor data. Details about the implementation and the dataset used is presented in section III. Section IV provides the experimental setup and results. Finally, section V presents final conclusions and future research directions.

II. ON-LINE LEARNING DATA-FUSION FOR INCREASING TEMPORAL RESOLUTION OF SENSOR DATA

This section first describes the presented data-fusion based method for increasing temporal resolution of data. Then the on-line learning method presented for improved long-term accuracy will be discussed. Finally the overall presented methodology is briefly described.

A. Increasing Temporal Resolution of Building Sensor Data

The goal of the presented method is to increase the temporal resolution of sensor data using only information available from low temporal resolution data.

A BEMS contains a large number of sensors collecting data at different time intervals. The low temporal resolution, known data reading from a sensor $s_i \in S$ at time t can be represented as $s_i(t)$, where S is the set of sensors in the building, $t \in T$ and $T = \{t_0, (t_0 + \Delta), (t_0 + 2\Delta), \dots\}$ is the constant interval sensor time series at which the data will be collected by the sensor s_i , $\Delta > 0$ being the interval, and t_0 the data collection start time. Similarly, the high resolution data for the same sensor that will be predicted can be represented as $s_i(\hat{t})$, where $\hat{t} \in \hat{T}$ and $\hat{T} = \{t_0, (t_0 + \delta), (t_0 + 2\delta), \dots\}$ is the higher temporal

resolution prediction times. Furthermore, $0 < \delta < \Delta$ yielding the desired higher temporal resolution data.

At a given time t , the higher temporal resolution data for sensor s_d , $s_d(\hat{t})$ where $t < \hat{t} < (t + \Delta)$ needs to be predicted. The simplest and the most common method for predicting $s_d(\hat{t})$ is using either linear interpolation of the previous points or polynomial regression that minimizes sum of error squares of training data [28]. However, because of complex occupancy patterns, weather patterns, and other factors, the resulting BEMS behavior is highly non-linear and complex [3], [29]. This leads to sub-optimal results with interpolation and polynomial regression models. Therefore, in this paper, an ANN based data-fusion model is presented that utilizes information from multiple sensors in the building to provide more accurate results. The ANN is capable of “learning” the inter-relationships between sensors and therefore provide a more holistic approach to model the sensor data.

The ANN uses currently available known data from multiple sensors to predict the desired value. A subset of sensors $S' \subseteq S$ is first selected such that the desired sensor s_d is in $s_d \in S'$. From each of the sensors selected, a number of previous time steps $T' = \{t, (t - \Delta), (t - 2\Delta), \dots, (t - k_i\Delta)\}$ is selected as inputs to the ANN, where time t is the immediate preceding data collection time to the desired high resolution time interval \hat{t} , i.e. $t < \hat{t} < (t + \Delta)$. The number of time steps selected, k_i , from each sensor i may vary according to the importance of that sensor and correlation of that sensor to the desired sensor s_d .

In addition to the known sensor values, current and desired time information is presented to the ANN as inputs in terms of time of day, day of week and a weekend flag. The output produced by the ANN will be the sensor value for the given desired time.

Therefore, given the previously known sensor data and the current and desired time information as the inputs, the ANN

TABLE I. SELECTED INPUT VARIABLES TO THE ANN

Input Type	Description	Unit
Known sensor data (S')	Zone temperature	°C
	Outside air temperature	°C
	Chiller temperature	°C
	Mixed air temperature	°C
	Return air temperature	°C
	CO ₂ concentration	PPM
	Supply fan load	%
Current time	Hour	Integer (1-24)
	Minute	Integer (0-59)
	Day of week	Integer (1-7)
	Weekend flag	Binary
Desired time	Hour	Integer (1-24)
	Minute	Integer (0-59)

readings of BEMS sensors and wireless sensors in zones Z5 and Z8 were deemed significantly different. Therefore, these zones were not used for the experimental phase. Thus, a total of 7 zones (Z1 to Z9, except Z5 and Z8 in Fig. 2) were selected for the experiment.

The installed temporary wireless sensors are capable of recording sensor values at a high temporal resolution. For the experimentation purposes the data collection rate of these wireless temperature sensors were set to 5 minute intervals, i.e. $\delta = 5 \text{ mins}$ of the higher temporal resolution prediction time series \hat{T} . The wireless sensor data with high temporal resolution is necessary to validate the presented method. While the availability of high temporal resolution data will improve performance, the presented method was also shown to produce good results even without using the high temporal resolution data from wireless sensors for training (see Section IV.C)

For experimental purposes 6 weeks of data collected during the months of August and September was selected. The first two weeks of data was selected as the training data and the remaining as the testing and validation data.

While a large number of sensor readings are available for each floor that monitor the air quality and the state of Air Handling Units (AHUs) along with outside conditions, only a subset of these sensors were selected for the data-fusion (S'). While having a large number of sensor data strengthens the holistic approach of modelling the system, it may lead to over-complex an ANN which will increase the computation time and may yield poor results. Therefore, 7 sensors were selected as S' and are shown in Table I. These sensors were selected because they are directly related to zone temperature. As mentioned in section 2, along with the sensor data, current and desired time information is also used as an input to the ANN. The selected time variables used are also shown in Table I. For time information, the seconds were not used in this example because the resolution of the predicted data is 5 minutes.

For all the sensors, the number of previous time steps used was set to 2, i.e. $k_i = 2$. Therefore, the set of time steps used for the input was $T' = \{t, (t-45), (t-90)\}$. Thus, from each sensor 3 values were used as inputs. Therefore the total number

of sensor inputs to the ANN was 21. Along with the 6 time information, the total number of inputs to the ANN was 27.

A feed-forward ANN with soft bipolar transfer functions for all layers was used. The architecture of the ANN selected for experimentation contained one input layer, two hidden layers and one output layer. The input layer contained 27 neurons and the output layer contained 1 neuron, representing the number of inputs and outputs. The hidden layers contained 15 and 7 neurons each. This architecture was selected after experimenting with several different architectures. The training was done using Error Backpropagation, with Levenberg-Marquardt optimization.

The selection of the set of sensor S' and the time periods for each sensor T' , along with the most appropriate architecture of the ANN can be improved by using techniques such as cross-validation. Using such techniques will increase the prediction accuracy of the presented methodology. However, the current selection was deemed sufficient for the demonstration purposes of this paper.

IV. EXPERIMENTAL RESULTS

The presented method was implemented using the dataset and ANN architecture mentioned in section III. Thus, 7 different ANNs were trained (one for each selected zone). As mentioned, the first 2 weeks of data was used as the training data and the remaining 4 weeks was used for testing the trained models.

The results of the tested models, which is the high temporal resolution temperature data was validated against the measurements from the temporary wireless sensor data. The results presented are the averaged prediction errors of all selected zones. The errors are presented in terms of mean and standard deviations of absolute errors in °C.

Three different experimental cases were performed, each of which are described below.

A. Experimental Case: 1

The goal of this experiment was to compare the prediction accuracy of the presented method with typically used techniques. Therefore, in this case, the presented method is tested against a simple interpolation using the previous two points and a polynomial approximation which minimizes squared error. The polynomial model was selected using 10 fold cross validation and the 12th degree polynomial was shown to be the most accurate. Thus, the results of the 12th degree polynomial is presented.

For this experiment, the data from the first 2 weeks was used as the training data. Furthermore, the high temporal resolution temperature data from the temporary wireless sensors was also used for training each model. For this case, the results of the weekends were ignored since the polynomial model cannot model such behavior.

The on-line learning was not used for this experiment as it may increase the prediction accuracy of the presented system, giving it an unfair advantage.

The prediction error of 5 minute interval temperature data for this case is shown in Table II. The polynomial model performed the worse since the highly complex behavior of the

TABLE II. EXPERIMENTAL RESULTS CASE 1

Method	Absolute Error (°C)			
	Training		Testing	
	Mean	SD	Mean	SD
12 th order polynomial	1.837	1.985	2.241	2.143
Linear interpolation	0.941	0.576	0.921	0.687
Presented ANN	0.272	0.206	0.545	0.210

TABLE III. EXPERIMENTAL RESULTS CASE 2

Method	Absolute Error (°C)			
	Training		Testing	
	Mean	SD	Mean	SD
Presented ANN without on-line learning	0.272	0.206	0.630	0.290
Presented ANN with on-line learning	0.272	0.206	0.391	0.228

TABLE IV. EXPERIMENTAL RESULTS CASE 3

Method	Absolute Error (°C)			
	Training		Testing	
	Mean	SD	Mean	SD
Presented ANN with all data	0.272	0.206	0.399	0.228
Presented ANN with only BEMS data	0.462	0.370	0.567	1.034

building cannot be captured by it. While the linear interpolation works better compared to the polynomial, it was still worse than the presented ANN based method. Furthermore, it was observed that the linear interpolation performs poorly in day-night and cooling on-off transition periods. While this observation is intuitive, it has to be noted that such transitive behavior is the most important for control and state-awareness in buildings for maintaining occupant comfort and increasing energy efficiency.

B. Experimental Case: 2

The goal of this experiment was to exemplify the ability of the on-line learning method to learn new behavior of the building over time. Thus, in this case, the presented method was tested against the same ANN with the on-line learning disabled. As before, the first 2 weeks was used as the training data and the high temporal resolution temperature data was also used for training each model. For this experiment the weekends were included.

The overall results of this case is shown in Table III. It can be seen that with the on-line learning enabled, the overall error is slightly decreased along with the standard deviation. Fig. 3 shows the mean and standard deviation of absolute error of each model at the end of each day for the testing period. It can be observed that with the on-line learning enabled the prediction mean absolute error is almost as good as with the on-line

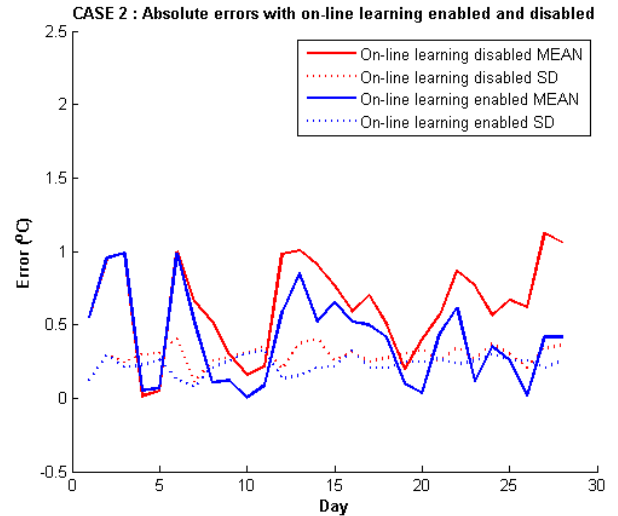


Fig. 3 Daily prediction absolute error curves for experimental case 2 learning disabled. Furthermore, as time progresses the mean absolute error is decreased with the on-line learning enabled.

C. Experimental Case: 3

Identifying whether the presented method can be used without having a temporary sensor network for training was the goal of this experiment. Thus, in this case, two instances of the presented ANN method is compared, where one instance is trained using only the BEMS data and the other instance trained using the wireless sensor data. The on-line learning was enabled in both instances.

The results of this case is shown in Table IV. It can be clearly seen that with more training data the prediction error is lower. However, even without high temporal resolution data for training, the ANN is capable of learning the complex behavior of the system and produce good results.

V. CONCLUSION

This paper presented a data-fusion methodology that utilizes Artificial Neural Networks (ANNs) that increases the temporal resolution of sensor data. ANNs are capable of learning complex inter-relationships between various different types of sensors and provide accurate predictions. An on-line learning method is also presented that increases the prediction accuracy in the long-term.

The presented method was applied to a real-world building dataset to increase the temporal resolution of temperature sensors. Experimental results showed that the presented method is capable of predicting high temporal resolution data with high accuracy. Furthermore, the experimental results showed that the on-line learning is capable of improving the performance of the ANN in time. Finally, it was shown that the ANN is capable of learning the building behavior such that by using only low temporal resolution data, higher resolution data can be predicted effectively.

Future work entails applying the presented methodology for a longer period of time and to other sensors that can improve control and state-awareness of the building such as, CO₂, humidity, and air handling unit sensors. Further improvements to the prediction can be made with more complex hybrid machine learning techniques such as combination of Fuzzy and Neural methodologies. The usefulness and feasibility of using the data predicted by the presented system for control should also be explored.

REFERENCES

- [1] L.P. Lombard, J. Ortiz, C. Pout, "A review on buildings energy consumption information," in *Energy and Buildings*, vol. 40, pp. 394–398, 2008.
- [2] B. Sun, P. B. Luh, Q-S. Jia, Z. Jiang, F. Wang, C. Song, "Building Energy Management: Integrated Control of Active and Passive Heating, Cooling, Lighting, Shading, and Ventilation Systems," in *IEEE Trans. on Automation Science and Engineering*, vol. 10, no. 3, pp. 588-602, July 2013.
- [3] D. Wijayasekara, O. Linda, M. Manic, C. Rieger, "Mining Building Energy Management System Data Using Fuzzy Anomaly Detection and Linguistic Descriptions," in *IEEE Trans. on Industrial Informatics*, vol. 10, no. 3, pp. 1829-1840, June 2014.
- [4] D. Wijayasekara, M. Manic, C. Rieger, "Fuzzy Linguistic Knowledge Based Behavior Extraction for Building Energy Management Systems," in *Proc. IEEE Symp. on Resilience Control Systems*, Aug. 2013.
- [5] A. Costa, M. M. Keane, J. I. Torrens, E. Corry, "Building operation and energy performance: Monitoring, analysis and optimisation toolkit," in *Applied Energy*, vol. 101, pp. 310-316, Jan. 2013.
- [6] T. Kalamees, K. Jylhä, H. Tietäväinen, J. Jokisalo, S. Ilomets, R. Hyvönen, S. Saku, "Development of weighting factors for climate variables for selecting the energy reference year according to the EN ISO 15927-4 standard," in *Energy and Buildings*, vol. 47, pp. 53-60, Apr. 2012.
- [7] G. Byeon, T. Yoon, S. Oh, G. Jang, "Energy Management Strategy of the DC Distribution System in Buildings Using the EV Service Model," in *IEEE Trans. on Power Electronics*, vol. 28, no.4, pp. 1544-1554, Apr. 2013.
- [8] N. N. A. Bakar, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, F. Hussin, M. Bandi, "Energy efficiency index as an indicator for measuring building energy performance: A review," in *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 1-11, 2015.
- [9] K. W. Roth, D. Westphalen, J. Dieckmann, S. D. Hamilton, W. Goetzler, "Energy Consumption Characteristics of Commercial Building HVAC Systems: Volume III, Energy Savings Potential," TIAX LLC Report for US Department of Energy Building Technologies Program, 2002.
- [10] O. Linda, D. Wijayasekara, M. Manic, C. Rieger, "Computational Intelligence based Anomaly Detection for Building Energy Management Systems," in *Proc. IEEE Symp. on Resilience Control Systems*, Aug. 2012.
- [11] A. Aswani, N. Master, J. Taneja, D. Culler, C. Tomlin, "Reducing Transient and Steady State Electricity Consumption in HVAC Using Learning-Based Model-Predictive Control," in *Proceedings of the IEEE*, vol. 100, no. 1, pp. 240-253, Jan. 2012.
- [12] K. Whitehouse, J. Ranjan, J. Lu, T. Sookoor, M. Saadat, C. M. Burke, G. Staengl, A. Canfora, H. Haj-Hariri, "Towards Occupancy-Driven Heating and Cooling," in *IEEE Design & Test of Computers*, vol. 29, no. 4, pp. 17- 25, Aug. 2012.
- [13] I. J. Aucamp, L. J. Grobler, "Heating, ventilation and air conditioning management by means of indoor temperature measurements," in *Proc. of Industrial and Commercial Use of Energy Conference*, pp. 1-4, Aug. 2012.
- [14] K. Amarasinghe, D. Wijayasekara, M. Manic, "Neural Network Based Downscaling of Building Energy Management System Data," in *Proc. IEEE International Symp. on Industrial Electronics*, June 2014.
- [15] S. Bhattacharya, S. Sridevi, R. Pitchiah, "Indoor air quality monitoring using wireless sensor network," in *Proc. of 6th International Conf. on Sensing Technology*, pp. 422-427, Dec. 2012.
- [16] H. Mirinejad, K. Welch, L. Spicer, "A review of intelligent control techniques in HVAC systems," in *Proc. of IEEE Energytech*, pp. 1-5, May 2012.
- [17] Z. Raad, K. S. M. S. Homod, A. F. Haider F. Almurib, "Gradient autotuned Takagi–Sugeno Fuzzy Forward control of a HVAC system using predicted mean vote index," in *Energy and Buildings*, vol. 49, pp. 254-267, Jun. 2012.
- [18] Y. Cheng, J. Niu, N. Gao, "Thermal comfort models: A review and numerical investigation," in *Building and Environment*, vol. 47, pp. 13-22, Jan. 2013.
- [19] M. Gangolells, M. Casals, N. Forcada, M. Macarulla, A. Giretti, "Environmental impacts related to the commissioning and usage phase of an intelligent energy management system," in *Applied Energy*, vol. 138, pp. 216-223, 2015.
- [20] N. Moteqi, M. A. Piette, S. K. Kinney, J. Dewey, "Case Studies of Energy Information Systems and Related Technology: Operational Practices, Costs, and Benefits," in *Proc. of ICEBO*, Oct. 2003.
- [21] M. A. Piette, S. K., Kinney, H. Friedman, (2001). EMCS and TimeSeries Energy Data Analysis in a Large Government Office Building (No. LBNL-47699). Lawrence Berkeley National Laboratory.
- [22] R. Gulbinas, A. Khosrowpour, J. Taylor, "Segmentation and Classification of Commercial Building Occupants by Energy-Use Efficiency and Predictability," in *IEEE Trans. On Smart Grid*, vol. 6, no. 3, pp. 1414-1424, May 2015.
- [23] T. Kim, Y. K. Jeong, I. W. Lee, "A Design of Building Group Management Service Framework for On-Going Commissioning," in *Advanced Science and Technology Letters*, vol. 49, pp. 84-88, 2014.
- [24] N. Batra, P. Arjunan, A. Singh, P. Singh, "Experiences with Occupancy based Building Management Systems," in *Proc. IEEE Int. Conf. on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 153-158, Apr. 2013.
- [25] D. L. Hall, J. Llinas, "An introduction to multisensor data fusion," in *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.
- [26] D. Wijayasekara, O. Linda, M. Manic, C. Rieger, "FN-DFE: Fuzzy-Neural Data Fusion Engine for Enhanced Resilient State-Awareness of Hybrid Energy Systems," in *IEEE Trans. on Cybernetics*, vol. 44, no. 11, pp. 2168-2267, Nov. 2014.
- [27] S. L. Sun, "Multisensor optimal information fusion input white noise deconvolution estimators," in *IEEE Trans. on Syst., Man, and Cybern. B, Cybern.*, vol. 34, no. 4, pp. 1886–1893, Aug. 2004.
- [28] H. J. Motulsky, L. A. Ransnas, "Fitting curves to data using nonlinear regression: a practical and nonmathematical review," in *The FASEB journal*, vol. 1, no. 5, pp. 365-374, 1987.
- [29] H. Tianzhen, C. Wen-Kuei, L. Hung-Wen, "A fresh look at weather impact on peak electricity demand and energy use of buildings using 30-year actual weather data," in *Applied Energy*, vol. 111, pp. 333-350, Nov. 2013.
- [30] P. Bermejo, L. Redondo, L. de la Ossa, D. Rodríguez, J. Flores, C. Urea, J. A. Gámez, J. M. Puerta, "Design and simulation of a thermal comfort adaptive system based on fuzzy logic and on-line learning," in *Energy and Buildings*, vol. 49, pp. 367-379, Jun. 2012.