INTERNATIONAL JOURNAL OF ENERGY RESEARCH
Int. J. Energy Res. 2007; **31**:412–421
Published online 1 August 2006 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/er.1232

SHORT COMMUNICATION

An energy benchmarking model based on artificial neural network method utilizing US Commercial Buildings Energy Consumption Survey (CBECS) database

Melek Yalcintas[1,*,†] and U. Aytun Ozturk[2]

[1] *AMEL Technologies, Inc. 1164 Bishop St., Suite 124-302, Honolulu, HI 96813, U.S.A.*
[2] *Hawai'i Pacific University, College of Business Administration, 1164 Bishop Street, Suite 912G, Honolulu, HI 96813, U.S.A.*

SUMMARY

This study focuses on development of an energy benchmarking model utilizing U.S. Commercial Buildings Energy Consumption Survey (CBECS) Database. An artificial neural networks (ANN) method based approach was used in the study. Office type buildings in the CBECS database were used in the benchmarking model development and weighted energy use intensity (EUI) was selected as the benchmarking index. The benchmarking model included input variables describing building's physical properties, occupancy and climate. Yearly electricity consumption per square meter, or EUI, was estimated by the ANN model. The correlation coefficient for each census division benchmarking model varied between 0.45 and 0.73, and mean squared error (MSE) varied between 9.60 and 15.25. It was observed that when the data set for a census division was grouped by different climate zones, ANN benchmarking model provided more accurate predictions. It was also observed that ANN model provides more accurate estimations when compared with predictions obtained with multi-linear regression models. For comparison, the MSE values varied between 10.24 and 40.43. Overall, the ANN model proved itself a better prediction model for energy benchmarking. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: energy; model; artificial neural network; commercial buildings

*Correspondence to: Melek Yalcintas, AMEL Technologies, Inc. 1164 Bishop St., Suite 124-302, Honolulu, HI 96813, U.S.A.
† E-mail: melekyalcintas@hawaii.rr.com

Received 25 April 2005
Revised 19 May 2006
Accepted 9 June 2006

1. INTRODUCTION

Energy benchmarking is a method that is typically utilized to compare average energy performance of similar buildings in the same climate zone. If the benchmarking indicates that a building has high energy consumption, compared to buildings with similar characteristics, it becomes necessary to further investigate that building to identify the Energy Conservation Opportunities (ECOs).

Traditionally, energy benchmarking studies use two general approaches in analysing energy related statistical data: (i) benchmarking based on energy use intensity (EUI), which is defined as energy use per square meter of the building floor area ($\mathrm{kWh\,m^{-2}}$) (Kinney and Piette, 2002), (ii) benchmarking based on other variables in addition to building area, such as building envelope, occupancy rate, plug load density, operation hours, etc. (Hick and Von Neida, 2003; Sharp, 1996; Sharp, 1998; Matson and Piette, 2005; Yalcintas, 2006). In the first benchmarking method, a building's EUI is compared against the statistical distribution of the database, and its EUI percentage ranking within the entire group of buildings is identified. The second benchmarking method takes into account other building variables that contribute to the EUI such as building operation hours; lighting, heating and cooling equipment properties; floor percentage that is provided with heating, lighting and cooling; equipment density; building envelope conditions, including construction material, window to wall ratio, window shading and heat transfer properties, etc.

Other energy benchmarking studies employed are model-based benchmarking (Federspiel *et al*., 2002; Mathew *et al*., 2003), high performance building benchmarking (New Buildings Institute, 2003), and point based benchmarking rating system (LEED, 2005). Matson and Piette (2005) provide summaries and comments on the earlier benchmarking studies.

Availability of sufficient data is usually the main concern in developing an energy benchmarking model. Such energy related data may be collected by utility companies, performance contracting/energy analysis companies, building management companies, or government organizations dealing with energy. The Commercial Buildings Energy Consumption Survey (CBECS) database is one such energy database, developed by Energy Information Administration of the U.S. Department of Energy. The database consists of statistical information collected periodically on energy consumption, energy expenditures, and energy-related characteristics of commercial buildings in the United States. In the CBECS database, energy consumption is listed under the following categories: electricity, natural gas, fuel oil, and main fuel (which represents the sum of the former three categories).

The benchmarking model developed in our study uses electricity consumption as the energy medium and utilizes 1999 CBECS survey results. We developed a benchmarking method that utilizes artificial neural networks (ANN) and showed that the ANN method is not only more accurate in its benchmarking prediction but also more flexible due to its computational structure when compared to less sophisticated methods. For instance, linear regression, which is widely used in benchmarking today, requires manual determination of empirical constant coefficients each time the database is modified. A recent study by Yalcintas (2006) presents an energy benchmarking model based on ANN approach that utilizes a tropical data covering a small city. In contrast, we focused on developing a benchmarking model that utilizes a database covering a wide geographic area (United States). The ANN model coefficients, for each of the nine census divisions classified in CBECS database, were determined and the model performance for wider geographic areas was evaluated in this study. It should be noted that the modelling tool developed can be used for similar databases for different geographic areas, e.g. Canada, Europe, etc.

DOI: 10.1002/er

2. ANN ENERGY BENCHMARKING MODEL USING CBECS DATABASE

An ANN model was developed for each of the nine Census Divisions in the CBECS database. Once an ANN based benchmarking model that utilized data from a wider geographic area was developed, its performance was compared with other benchmarking models. It should be noted that only the office type buildings were modelled in the CBECS database. For ANN computations MATLAB Neural Network Toolbox was used.

The independent variables for the model were identified as follows: First, the office buildings data in one census division were selected as the initial modelling data. Ninth census division office data were most suitable for this purpose, since its climate zones included samples from all climates (i.e. Climate Zone 1 through 5 by CBECS definition). Next, a group of building-parameters in the data set were selected as the initial input variables for the ANN model. The initial input variables considered in the ANN model includes building-floor area, age, operation hours, number of workers, number of computers, number of floors, building-area cooling percentage, heating percentage, lighting percentage, cooling degree days (CDD), heating degree days (HDD), main cooling equipment, main heating equipment, presence of economizer, presence of energy management system, owner occupies, electricity used for heating, presence of electronic ballasts, percent lit by electronic ballasts, percent lit by fluorescent lighting, type of window, and tinted or reflective glass.

By repeated trials of different input variables and ANN configuration, a final set of input variables and ANN architecture was identified. Eight of the input variables listed above were included in the final set. The final input variables include building-operation hours, age category, building-area per worker category, building-area per computer, cooling percentage category, lighting percentage category, cooling degree days and number of floors category. Some of the final input variables were either in ordinal or in broadly readjusted forms. Among those input variables, some were already categorized in the CBECS database, and others were categorized by the authors. The output of the model, the electricity use per square feet of building space per year, EUI, was also categorized. This helped to reduce the influence of very high and very low EUI's on the model. The building age category, operation hours and cooling degree days in the CBECS database, listed in Table I, required no reformatting. Square feet area per worker category and square feet area per PC category were not originally available at the CBECS. They were calculated by the authors as the ratio of the building square feet to the number of workers and number of PCs in the building, respectively. Table I lists the categorized or broadly adjusted input data and range for each variable. The same data processing and categorizing approach was applied to the output EUI. Table II lists the output electricity use index in the categorized form.

In addition to categorizing several input variables, some of the lighting and cooling percentage data were also re-categorized in an effort to simplify their processing in the ANN model. The purpose here was to present the ANN algorithm with repeated input values. Therefore, less repeating percentage values were approximated to a closest more repeating value. Finally, 'number of floors' input variable was re-adjusted. In the original CBECS database, the number of floors higher than 14 were represented arbitrarily by 991 or 992, where 991 corresponded to a building with 15–24 floors, and 992 represented a building that has 25 or higher number of floors. In order to provide a logical flow of the numbers in the ANN model, 991 was replaced by 20 and 992 was replaced by 40, to better represent the corresponding floor range.

Table I. Final input data category.

Input data category	
Operation hours	Cooling degree days
Available in CBECS database	Available in CBECS database
Building age category	Number of floors category
Available in CBECS database	Available in CBECS database, up to 14 floors is the same. Above 14 floors:
	20 = 991 (15–24 floors)
	40 = 992 (25 and higher floors)
Building-area per worker category	Cooling percentage category
(Generated by authors using existing data in the CBECS database)	(Generated by authors using existing data in the CBECS database)
01 = 2.0–13.9 m^2 (20–150 ft^2)	0 = 0–20%
02 = 13.91–18.6 m^2 (150.1–200 ft^2)	30 = 20.1–40%
03 = 18.61–23.2 m^2 (200.1–250 ft^2)	45 = 40.1–52%
04 = 23.21–27.9 m^2 (250.1–300 ft^2)	65 = 52.1–70%
05 = 27.91–32.5 m^2 (300.1–350 ft^2)	75 = 70.1–75%
06 = 32.51–37.2 m^2 (350.1–400 ft^2)	80 = 75.1–84%
07 = 37.21–41.8 m^2 (400.1–450 ft^2)	90 = 84.1–90%
08 = 41.81–46.5 m^2 (450.1–500 ft^2)	95 = 90.1–97%
09 = 46.51–55.7 m^2 (500.1–600 ft^2)	100 = 97.1–100%
10 = 55.71–65.0 m^2 (600.1–700 ft^2)	
11 = 65.01–79.0 m^2 (700.1–850 ft^2)	
12 = 79.01–92.9 m^2 (850.1–1000 ft^2)	
13 = 92.91–139.4 m^2 (1000.1–1500 ft^2)	
14 = 139.41 m^2 and higher (1500.1 ft^2 and higher)	
Building-area per PC category	Lighting percentage category
(Generated by authors using existing data in the CBECS database)	(Generated by authors using existing data in the CBECS database)
01 = 2.0–13.9 m^2 (20–150 ft^2)	20 = 0–40%
02 = 13.91–18.6 m^2 (150.1–200 ft^2)	55 = 40.1–60%
03 = 18.61–23.2 m^2 (200.1–250 ft^2)	75 = 60.1–75%
04 = 23.21–27.9 m^2 (250.1–300 ft^2)	80 = 75.1–80%
05 = 27.91–32.5 m^2 (300.1–350 ft^2)	90 = 80.1–90%
06 = 32.51–37.2 m^2 (350.1–400 ft^2)	95 = 90.1–97%
07 = 37.21–41.8 m^2 (400.1–450 ft^2)	100 = 97.1–100%
08 = 41.81–46.5 m^2 (450.1–500 ft^2)	
09 = 46.51–55.7 m^2 (500.1–600 ft^2)	
10 = 55.71–65.0 m^2 (600.1–700 ft^2)	
11 = 65.01–79.0 m^2 (700.1–850 ft^2)	
12 = 79.01–92.9 m^2 (850.1–1000 ft^2)	
13 = 92.91–139.4 m^2 (1000.1–1500 ft^2)	
14 = 139.41 m^2 and higher (1500.1 ft^2 and higher)	

The final form of the ANN model used in this study consists of a three layer feed-forward type configuration: an input layer, a hidden layer and an output layer. The total number of input variables used in the model was eight. Depending on the census division, the hidden layer contained 7 or 3 neurons. Levenberg–Marquardt back-propagation algorithm was employed.

DOI: 10.1002/er

Table II. Final output category.

Output category
Electricity use per square meter per year (square feet per year)
(Generated by Authors using existing data in the CBECS database)
01 = 5.4–32.3 k Wh m^{-2} yr^{-1} (0.5–3 k Wh ft^{-2} yr^{-1})
02 = 32.4–53.8 k Wh m^{-2} yr^{-1} (3.01–5 k Wh ft^{-2} yr^{-1})
03 = 53.9–75.4 k Wh m^{-2} yr^{-1} (5.01–7 k Wh ft^{-2} yr^{-1})
04 = 75.5–96.9 k Wh m^{-2} yr^{-1} (7.01–9 k Wh ft^{-2} yr^{-1})
05 = 97.0–118.4 k Wh m^{-2} yr^{-1} (9.01–11 k Wh ft^{-2} yr^{-1})
06 = 118.5–139.9 k Wh m^{-2} yr^{-1} (11.01–13 k Wh ft^{-2} yr^{-1})
07 = 140.0–161.5 k Wh m^{-2} yr^{-1} (13.01–15 k Wh ft^{-2} yr^{-1})
08 = 161.6–183.0 k Wh m^{-2} yr^{-1} (15.01–17 k Wh ft^{-2} yr^{-1})
09 = 183.1–204.5 k Wh m^{-2} yr^{-1} (17.01–19 k Wh ft^{-2} yr^{-1})
10 = 204.6–226.0 k Wh m^{-2} yr^{-1} (19.01–21 k Wh ft^{-2} yr^{-1})
11 = 226.1–247.6 k Wh m^{-2} yr^{-1} (21.01–23 k Wh ft^{-2} yr^{-1})
12 = 247.7–279.9 k Wh m^{-2} yr^{-1} (23.01–26 k Wh ft^{-2} yr^{-1})
13 = 280.0–322.9 k Wh m^{-2} yr^{-1} (26.01–30 k Wh ft^{-2} yr^{-1})
14 = 323.0–430.6 k Wh m^{-2} yr^{-1} (30.01–40 k Wh ft^{-2} yr^{-1})
15 = 430.7 k Wh m^{-2} yr^{-1} and up (40.01 k Wh ft^{-2} yr^{-1} and up)

The original data sets were composed of 57 (Census Division 6) to 221 (Census Division 9) sets of data from all nine census divisions each with 8 input parameters (x_1, x_2, x_3 ..., x_8), and 1 output parameter (y). No outlier data was eliminated from the input data sets in the ANN modelling processes. In order to avoid over-fitting of the data and to provide a good generalization capability to the ANN building energy prediction algorithm, an automated regularization method, which divides the data into two subsets, was used. The first subset is the training set, which is used in computing network weights and biases. The second subset is the testing set. In training process, the training continues until the sum of squared errors and sum of square weights in the Levenberg–Marquardt algorithm reaches a steady level. When a steady level is reached the training stops, and the model is tested with the test set. The test set error was not used during the training process; it was used only to compare different models. However, a low error on the testing network is a good measure of an acceptable ANN model. In this study, for any CBECS Census Division data, three fourths of the data was used for training and remaining one fourth was used for testing.

MATLAB Neural Network Toolbox was used in development of the ANN model. A built in MATLAB normalization function was used to normalize the input and output values of the training set. The MATLAB normalization function 'prestd' produces zero mean and unity standard deviation for the training set input and target output variables. The input and target output data in the testing set is rescaled by the numeric normalization scale determined from the training data. After the ANN model training and testing, another MATLAB function 'poststd' transfers the model predictions to the original scale.

3. MULTIPLE LINEAR REGRESSION (MLR) APPROACH

To be able to demonstrate the effectiveness of the ANN algorithm a regression analysis was conducted for all 9 census divisions. The independent variables listed earlier were evaluated for

Table III. Resulting independent variables used in MLR approach for each census division.

Census division	Independent variables
1	Operation Hours, Heating Degree Days, Floor Area, Lighting Percentage, Number of PC, Main Heating Equipment, Main Cooling Equipment
2	Floor Area, Number of PC
3	Floor Area, Number of PC, Owner Occupied
4	Operation Hours, Heating Degree Days, Floor Area, Heating Percentage, Cooling Percentage, Number of PC, Main Heating Equipment, Main Cooling Equipment, Tinted or Reflective Glass Windows
5	Operation Hours, Number of Workers, Heating Percentage, Main Heating Equipment, Main Cooling Equipment, Number of Floors
6	Heating Degree Days, Cooling Degree Days, Building Age, Owner Occupied, Presence of Economic Ballasts, Main Heating Equipment, Main Cooling Equipment
7	Floor Area, Heating Percentage, Number of PC, Main Heating Equipment, Main Cooling Equipment, Number of Floors
8	Floor Area, Number of PC, Main Cooling Equipment
9	Operation Hours , Floor Area, Lighting Percentage, Number of PC, Owner Occupied, Main Cooling Equipment

predicting energy usage. The box plot and a histogram of each variable for each division were analyzed to apply the appropriate transformation such as a logarithm or exponential. A backward elimination regression method was then used on the transformed data set where the significance level was at 0.05. The resulting independent variables are listed in Table III for each division.

It was observed that the floor area and number of PCs appeared as dominant variables in all divisions, which is not counterintuitive. The rest of the variation in the variables may be attributed to geographic and climatic differences as well as the quantity and quality of the data collected. It should be noted that even though MLR model did not use ordinal or adjusted variables as ANN model did, MLR model performed worse than the ANN model in all of the census divisions except one.

4. RESULTS AND DISCUSSION

ANN model developed for each Census Division predicted EUI's-electricity use indexes (kWh m^{-2} yr^{-1}). The following Figures 1–3 presents the ANN model predictions for census divisions 1, 6 and 9. The ANN model along with the original EUIs, for Census Division 1 are shown in the figures. In these figures, 'A' corresponds to the predicted EUI and 'T' corresponds to the target EUI (or the original output data). The line represents the linear regression relating the predicted EUI to the target EUI. R in the Figures represents the correlation coefficient between the network predicted output and target (or actual outputs). In addition, Table IV lists the coefficient of correlations and mean square errors of the ANN models for each of the census divisions. Both the coefficient of correlations and mean square errors are in acceptable ranges, indicating a good performance of the ANN in modelling the energy benchmarking.

When we review Table III, as well as Figures 1–3, we see that the lowest coefficient of correlation was obtained for Census Division 5 (South-East Cost), which is followed by Census

DOI: 10.1002/er

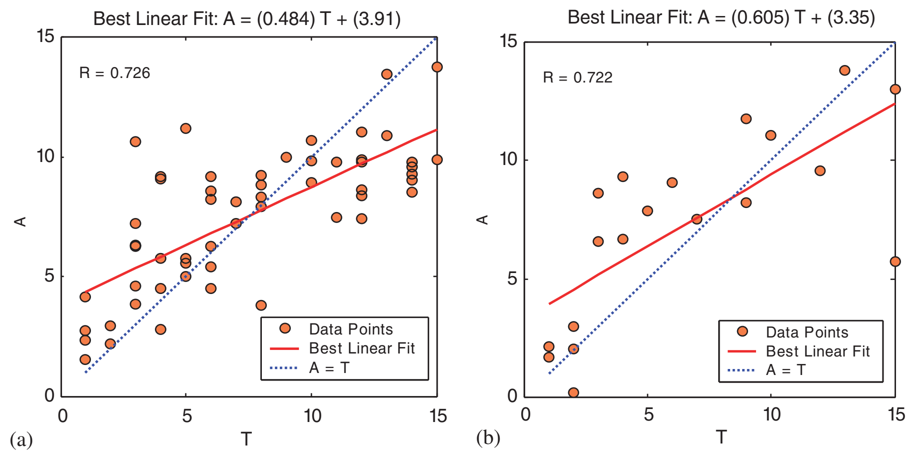

Figure 1. Census Division 1 ANN model estimates: (a) training; and (b) testing.

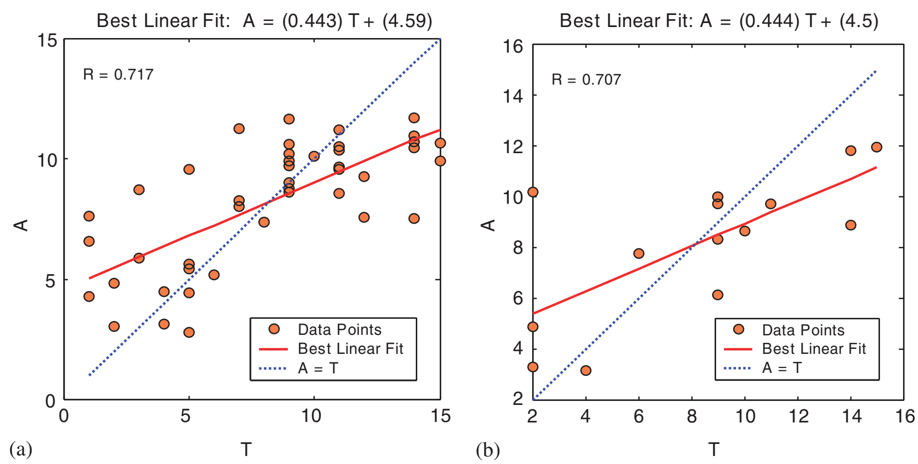

Figure 2. Census Division 6 ANN model estimates: (a) training; and (b) testing.

Division 8 (West Mountain Range). Both divisions cover large geographic areas, and most importantly cover very different climate zones. Division 5 includes cold climate to hot and humid climates, Division 8 covers very cold climate to hot and dry climate. When the buildings in these divisions were benchmarked by the same model, the model accuracy was affected. Census Division 9 (West Coast) is another example with varying climate and geographic properties, and relatively poor ANN prediction. However, when we examine Census Division 1 (New England), representing a dataset with similar climate and closer geographic locations, we see that the coefficient of correlation is highest. Census Division 6 (South East inland) shows similar climate and geographic trend as in Census Division 1, resulting with good prediction capacity.

In order to present the effect of climate zones on the ANN model performance, Climate 4 and 5 data in the Census Division 9 were isolated, and a new ANN model was developed for this

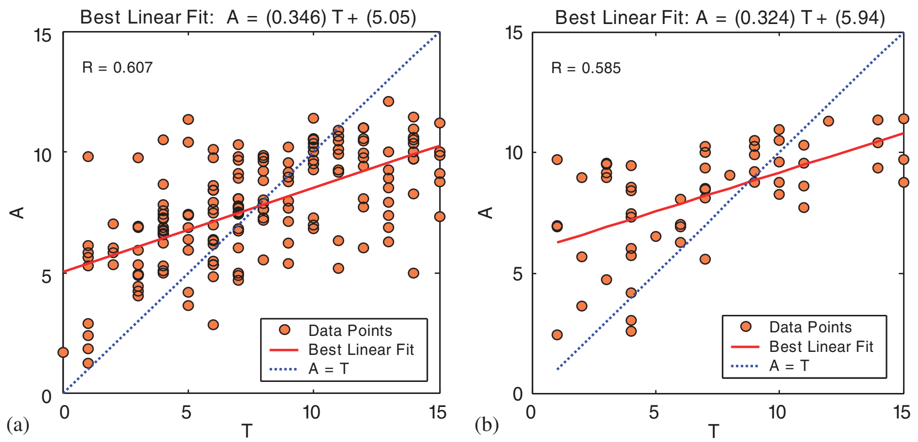

Figure 3. Census Division 9 ANN model estimates: (a) training; and (b) testing.

Table IV. ANN model coefficient of correlations and mean square errors (MSE) for each census divisions.

	Coefficient of correlation		Mean square error (MSE)	
Census division	Model training	Model testing	Model training	Model testing
1	0.726	0.722	8.69	11.04
2	0.651	0.652	9.72	10.47
3	0.622	0.619	10.88	11.20
4	0.602	0.587	11.00	14.80
5	0.468	0.455	12.73	15.25
6	0.717	0.707	8.38	9.60
7	0.615	0.622	10.38	12.05
8	0.589	0.538	12.02	11.70
9	0.607	0.585	9.64	12.03

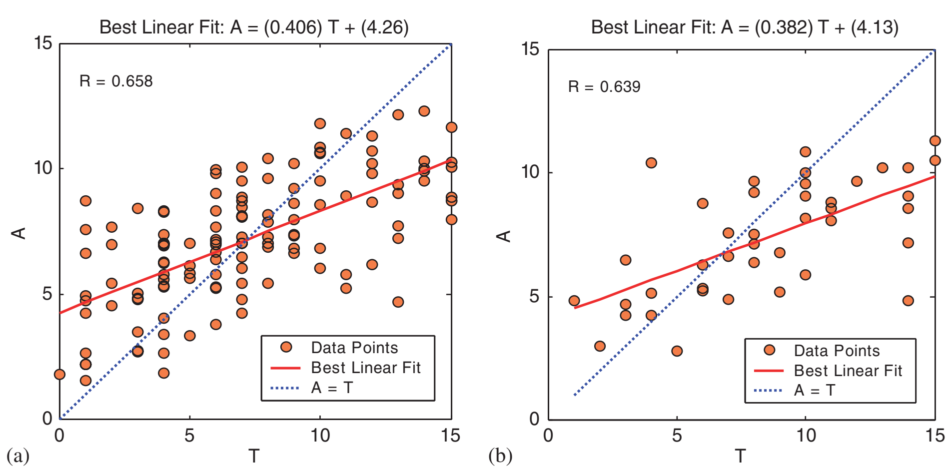

Figure 4. Census Division 1 ANN model estimates: (a) training; and (b) testing.

Table V. Comparison of MSE for both ANN and MLR modelling approaches.

	Mean squared error (MSE)	
Census division	ANN	MLR
1	11.04	37.00
2	10.47	16.00
3	11.20	16.85
4	14.80	29.06
5	15.25	33.57
6	9.60	40.43
7	12.05	21.62
8	11.70	11.41
9	12.03	10.24

condition. As seen in Figure 4, the ANN model benchmark prediction for this case is much better than the model that was developed for the entire Census Division 9, which was illustrated in Figure 3. The MSE of the training data was 9.05 and testing data was 9.93. This MSE is also lower than what was listed in Table IV for the entire Census Division 9.

Table V below provides a comparison of Mean Squared Error (MSE)'s for all divisions. In all cases but one, ANN model outperformed the MLR approach. This is expected since ANN uses a nonlinear functional approximation, whereas MLR assumes linearity of the relationship between dependent and independent variables in the data set analysed.

5. CONCLUSION

Even though the ANN method outperformed currently used linear regression methods as a benchmarking model, there is still room for improvement. These improvements can be achieved by either changing the structure of the CBECS database, or further perfecting the ANN modeling approach.

The main revision needed in the CBECS database is the capability to classify the data in smaller geographic regions. A quick review of the existing data reveals that when this database is divided into smaller geographic regions, there may not be sufficient number of data points to conduct any type of analysis. For example, Census Division 4 has only 67 data sets for the office buildings. If this data is further divided into two or more geographic regions, the number of data points remaining may not be sufficient to develop an ANN model. Therefore, for some census divisions and geographic regions more survey data needs to be included.

A second point related to the CBECS database is the lack of information on climate. Currently, the database provides heating degree days and cooling degree days as the only climate variables. Other variables that are more crucial in determining the climate effects in the benchmarking model are average summer and winter dry bulb and wet bulb temperatures, and solar radiation. This information is readily available at the National Weather Bureau, and can easily be integrated into the database.

The third item regarding the CBECS database is related to the improvements that can be made to building envelope data. The building envelope related data in the current CBECS

DOI: 10.1002/er

database covers only the main building material, type of window and tinted or reflective glass windows. Further information is needed on the roof and wall insulation properties, wall to window ratios, as well as building orientation.

Furthermore, while building operation hours in the CBECS database indicate average usage of the building, it may not be directly related to the equipment operation schedule in the building. An additional equipment schedule may be useful for the benchmarking model development.

Improvements related to the modeling approach include development of automated data filtering method, search for alternative ANN learning algorithms, and investigation of alternative benchmarking methods based on fuzzy neural networks and/or genetic algorithms.

In conclusion, since benchmarking is becoming an important tool in categorizing energy consumption in various buildings, better benchmarking tools with advanced computational capabilities are required. ANN method is definitely an appropriate candidate to meet these needs. The most effective feature of ANN method is that the same benchmarking algorithm can be applied to different climate zones. Therefore, if the CBECS database is ever divided to cover smaller regions within the census divisions, or states, or even sub-climate zones within a state, generic ANN architecture can be easily trained with the regional data, and can be used to benchmark building-energy usage in that region. Furthermore, the model can update itself when new data is provided. Once the suggested improvements are made to the CBECS database, ANN modelling can be used to develop a highly accurate benchmarking model.

REFERENCES

CBECS (Commercial Building Energy Consumption Surveys) database website: http://www.eia.doe.gov/emeu/cbecs

Federspiel C, Zhang Q, Arens E. 2002. Model-based benchmarking with applications to laboratory buildings. *Energy and Buildings* **34**(3):203–214.

Hicks TW, Von Neida B. 2003. An evaluation of America's first ENERGY STAR buildings: the class of 1999. *Proceedings of the ACEEE 2003 Summer Study on Energy Efficiency in Buildings*, vol. 4, Rye Brook, New York, 177–185.

Kinney S, Piette MA. 2002. Development of California commercial building energy benchmarking database. *Information and Electronic Technologies* **7**:109–120.

LEED Reference Guide. 2005. US Green Building Council.

Mathew P, Sartor D, Van Geet O, Reilly S. 2003. Rating energy efficiency and sustainability in laboratories: results and lessons from the Labs21 program. *Proceedings of the ACEEE 2003 Summer Study on Energy Efficiency in Buildings*, vol. 4, Rye Brook, New York, July 29–August 1 2003; 321–329.

Matson N, Piette MA. 2005. Review of California and national methods for energy-performance benchmarking for commercial buildings. *California Energy Commission, Public Interest Energy Research Program*. LBNL No: 57364.

New Buildings Institute. 2003. High Performance Buildings.

Sharp T. 1996. Energy benchmarking in commercial office buildings. *Proceedings of the ACEEE 1996 Summer Study on Energy Efficiency in Buildings*, vol. 4, 321–329.

Sharp T. 1998. Benchmarking energy use in schools. *Proceedings of the ACEEE 1998 Summer Study on Energy Efficiency in Building*, vol. 3, 305–316.

Yalcintas M. 2006. An energy benchmarking model based on artificial neural network method with a case example for tropical climates. *International Journal of Energy Research*, DOI: 10.1002/er.1212.