Available online at www.sciencedirect.com
ScienceDirect
Applied Energy 85 (2008) 271–296
APPLIED ENERGY
www.elsevier.com/locate/apenergy

Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector

Merih Aydinalp-Koksal [a,*], V. Ismet Ugursal [b]

[a] Department of Environmental Engineering, Hacettepe University, Beytepe, Ankara 06532, Turkey
[b] Department of Mechanical Engineering, Dalhousie University, P.O. Box 1000, Halifax, NS, Canada B3J 2X4

Received 8 May 2006; received in revised form 21 August 2006; accepted 29 September 2006
Available online 21 September 2007

Abstract

This paper investigates the use of conditional demand analysis (CDA) method to model the residential end-use energy consumption at the national level. There are several studies where CDA was used to model energy consumption at the regional level; however the CDA method had not been used to model residential energy consumption at the national level. The prediction performance and the ability to characterize the residential end-use energy consumption of the CDA model are compared with those of a neural network (NN) and an engineering based model developed earlier. The comparison of the predictions of the models indicates that CDA is capable of accurately predicting the energy consumption in the residential sector as well as the other two models. The effects of socio-economic factors are estimated using the NN and the CDA models, where possible. Due to the limited number of variables the CDA model can accommodate, its capability to evaluate these effects is found to be lower than the NN model.

Keywords: Residential energy consumption modeling; Conditional demand analysis; Neural networks modeling

Abbreviations: A/C, air conditioning; ALC, appliance, lighting, and cooling; CDA, conditional demand analysis; DHW, domestic hot water; ENG, engineering; GHG, green houses gas; NN, neural networks; NRCan, Natural Resources of Canada; SHEU, survey of household energy use; UEC, unit energy consumption.

* Corresponding author. Fax: +90 312 299 20 53.
E-mail address: merih.aydinalp@gmail.com (M. Aydinalp-Koksal).

0306-2619/$ - see front matter
doi:10.1016/j.apenergy.2006.09.012

Nomenclature

a	regression coefficient
AF_{ij}	features of household i's appliance j
BP	dummy variable: one if the household has a boiler, zero if not
BWTV	number of black and white TVs owned by the household
CAC	dummy variable: one if the household has central A/C unit, zero if not
CLOTH	dummy variable: one if the household has a clothes washer, zero if not
COOK	dummy variable: one if the household has a range, zero if not
CTV	number of color TVs owned by the household
CV	coefficient of variation
DHW	dummy variable: one if the household has DHW heating equipment, zero if not
DISH	dummy variable: one if the household has a dishwasher, zero if not
DRYER	dummy variable: one if the household has a clothes dryer, zero if not
EDC_i	economic and demographic characteristics of household i
e_{ijt}	a random error term for the end-use
FF	dummy variable: one if the household has a furnace, zero if not
FLOU	number of fluorescent lamps owned by the household
FREZ1	dummy variable: one if the household has a freezer, zero if not
FREZ2	dummy variable: one if the household has a secondary freezer, zero if not
HALO	number of halogen lamps owned by the household
HEC_{it}	energy consumption by household i in period t
INCA	number of incandescent lamps owned by the household
LIGHTS	total number of lamps
MC_{it}	market conditions of household i in period t
MICROW	dummy variable: one if the household has a microwave, zero if not
N	number of end-use equipment
POOL	dummy variable: one if the household has a natural gas pool heater, zero if not
R^2	correlation coefficient
REF1	dummy variable: one if the household has a refrigerator, zero if not
REF2	dummy variable: one if the household has a secondary refrigerator, zero if not
SH	dummy variable: one if the household has SH equipment, zero if not
S_{ij}	binary indicator of household i's ownership of appliance j
SSH	dummy variable: one if the household has supplementary space heating equipment, zero if not
$STRUC_i$	structural features of household i
UEC_{ijt}	end-use j unit energy consumption of household i in period t
UP_{ijt}	utilization patterns relating to appliance j
VCR	number of VCRs owned by the household
WAC	dummy variable: one if the household has window A/C unit, zero if not
WC_{it}	weather conditions of household i in period t

1. Introduction

Canada is one of the countries legally bounded to reduce its greenhouse gas (GHG) emissions by 2010. To achieve this commitment Canada should seek all possible ways to reduce its energy consumption and associated GHG emissions. Due to its northerly location and prevalence of single family housing, in 2003 residential sector accounted for 17% of the total energy consumption and 16% of the total GHG emission in Canada [1]. Thus, improving the end-use energy efficiency in the residential sector would play a major role in Canada's commitment to reduce its GHG emissions.

There are many energy efficiency improvements to be considered to reduce the end-use energy consumption in the residential sector. These improvements have complex interrelated effects on the end-use energy consumption of households and detailed mathematical models are required to evaluate their effects. It was recently shown that three models have been used to model residential energy consumption: engineering method [2–4], neural networks (NN) [5–12], and conditional demand analysis (CDA) [13–18].

The engineering method involves developing a housing database representative of the national housing stock and estimating the energy consumption of the houses in the database using a building energy simulation program. Thus, this method requires a database representative of the housing stock with detailed house description data, as well as extensive user expertise and lengthy input data preparation time. A difficulty with this method is the inclusion of consumer behaviour and other socioeconomic variables that have a significant effect on the residential energy use. However, because of the high level of detail and flexibility provided by engineering based models, they can be used to evaluate the impact of a wide range of scenarios for energy conservation on residential energy consumption and GHG emissions [19].

NN are simplified mathematical models of biological neural networks. They are highly suitable for determining causal relationships amongst a large number of parameters such as seen in the energy consumption patterns in the residential sector. The NN approach has been used for prediction problems as a substitute for statistical approaches due to their simplicity of application and accurate estimates. NN models were mainly used for utility load forecasting and to predict energy consumption of individual buildings since they have a high potential to model nonlinear processes such as building energy loads. It has been shown in previous papers [11,12] that while NN can be successfully used to estimate national end-use energy consumption and the impact of socio-economic factors in the residential sector, they are not flexible in evaluating the impact of energy conservation measures.

CDA is a regression-based method. Compared to the engineering and NN based models, the CDA based models are easier to develop and use, and do not require detailed data as the engineering method. However, since these are regression-based models, the number of dwellings in the database needs to be larger, and the models do not provide much detail and flexibility. As a result, they have limited capability to assess the impact of energy conservation scenarios. It is however possible to include socioeconomic parameters in the model if such data is available in the database. On the other hand, in most CDA models multicollinearity problem often makes it difficult to isolate the energy use of highly saturated appliances, such as the refrigerator [18]. Like the NN approach, the capability of the CDA models is limited to the variables used in the model equation, thus limited energy efficiency measures can be tested using the CDA approach.

In this paper, the CDA method is proposed to model the national residential end-use energy consumption. The paper presents the sources of data, methodologies used to develop the model, the accuracy of the predictions of the CDA model, and some sample results. The paper also presents the comparison of the CDA estimates with the NN and engineering model estimates.

2. Energy modeling for the building sector using conditional demand analysis

CDA was first introduced by Parti and Parti [13], in which the authors separated the total household energy consumption into 16 appliance categories based on the data from 5286 households. Their estimates of appliance energy use were reasonably close to the engineering estimates. Similarly, their estimates for price and income elasticities lied within the range of estimates presented in previous studies.

Aigner et al. [14] used the CDA method and 15-min demand data from 130 households to obtain hourly end-use load profiles. The authors used 24 regression equations, each representing an hour of the day, to estimate the consumption through the day. In order to generate more precise estimates, restrictions were imposed on the parameters of the hourly equations, assuming that some appliances were not used in the early hours of the morning and hence could be excluded from those particular equations. The estimated hourly loads for most of the appliances were reasonable, whereas some hourly load estimates of dishwasher, clothes washer, cooking range, and clothes dryer were negative (*i.e.* unreasonable).

In order to improve the accuracy of estimates, Fiebig et al. [15] reformulated the standard CDA model into a random coefficient framework [20]. During any particular hour, the intensity of use of a particular appliance varies from household to household, and the appliance dummy variables indicate only absence or presence of the appliance and do not allow for variation in size and capacity. Therefore, the authors treated the coefficients of appliance ownership dummy variables as random variables rather than fixed variables, which also enabled the integration of metered data. The authors used a sample of 348 households, with direct metering data for two appliances, namely main tariff water heater (125 out of 189 were directly metered) and off peak tariff water heater (21 out of 189 were directly metered) which was charged at a lower than normal rate. The model obtained positive hourly estimates for all appliances except the freezer. The results showed that estimates from the model integrated with metered data were closer to the direct metered data than the ones obtained from the model without the integration of the direct metered data.

Bauwens et al. [21] used a Bayesian approach [22] to integrate the direct metering data by viewing the data as prior information on the energy consumption of a specific appliance. A sample of 174 households from the dataset used by Fiebig et al. [15] was used to estimate end-use consumption on weekdays and weekends. Direct metering data for the main and off peak tariff water heaters were available for 21 and 87 households, respectively. The results from the standard CDA contained negative estimates for freezer and pool pump for both weekdays and weekends, whereas, the CDA model reformulated using the Bayesian approach to integrate the direct metered data obtained no negative estimates for any appliance.

The metered data information from a 1986 survey that included space heating direct metering data for 30 and water heating direct metering data for 23 households were used by Hsiao et al. [18] to form prior distributions of appliance consumption. Then, by applying Bayesian technique, these prior distributions were combined with the load, appliance ownership, and demographic data from a survey conducted in 1983 including 347 households. The model included the interactions of appliance dummy variables with socio-economic variables, and weather data as explanatory variables. The hourly water and space heating consumption estimates from standard conditional demand analysis approach were compared with the ones obtained from the proposed approach. The comparison done by evaluating the hourly consumption profiles showed that the proposed approach obtained more reasonable hourly estimates for water and space heating consumption.

One of the studies that integrated the estimates from the engineering models into the CDA model was performed by Caves et al. [23]. They treated the information from the engineering approach as prior evidence on usage patterns for specific appliances, and by using Bayesian analysis, engineering estimates were integrated into CDA model to estimate hourly appliance consumption. The sample data of the analysis contained electric consumption and appliance ownership information for 129 households for two summer months in 1977. The engineering estimates were generated from a simulation program running twelve scenarios. Average loads for each of these appliances for the sample data were constructed using a weighted average of the twelve scenarios, where the weights reflected the housing type and size characteristics of the sample, and the distribution of the sample households between the weather districts. For central air conditioning, both methods provided similar estimates, but for dishwasher the estimates differed considerably, since dishwasher consumption was more dependent on consumer behaviour than that of central air conditioning.

CDA was also used to model regional residential end-use energy consumption. Kellas [16] developed a CDA model with 38 independent variables to estimate the residential energy consumption in Manitoba Hydro's service area. The estimates of the model showed an error of 2.8% when compared with the billing data. Due to the multicollinearity problem, the author faced difficulties in predicting the energy consumption of highly saturated appliances, such as refrigerator and clothes washers. The estimates for space and domestic hot water (DHW) heating energy consumption were reasonable, but the estimates for space cooling were high, i.e. about 1360 kWh/yr/household.

Lafrange and Perron [17] used the data from three large-scale surveys to estimate residential end-uses using CDA approach in Quebec. The estimates for DHW heating energy consumption were reasonable, however space heating estimates were low and cooling estimates were high. The estimates showed similarities with engineering model estimates, but the space heating estimate was lower than the engineering model estimate. Like other researchers, the authors had difficulties to estimate the consumption of highly saturated appliances.

3. Overview of the conditional demand analysis model

The energy consumption of a household can be expressed as a summation of the energy consumed by each of the appliances present in the household. Thus, the energy consumption of a household is directly related to the appliance stock present in the dwelling, specific features of these appliances, dwelling characteristics, and utilization patterns such as thermostat settings on water/space heaters, and behavioural patterns relating to the use of appliances.

The basic CDA model can therefore be represented in algebraic form as [24]:

$$\mathrm{HEC}_{it} = \sum_{j=1}^{j} \mathrm{UEC}_{ijt} \times S_{ij} \tag{1}$$

where HEC_{it} is the energy consumption by household i in period t, UEC_{ijt} is end-use j unit energy consumption of household i in period t, and S_{ij} is a binary indicator of household i's ownership of appliance j.

To develop a CDA model, the data on household energy consumption (HEC_{it}) can be obtained from utility billing records and appliance stock (S_{ij}) information can be obtained through an appliance saturation survey.

The end-use energy consumption depends upon a variety of factors and this relationship can be formalized as

$$\mathrm{UEC}_{ijt} = f_j(\mathrm{AF}_{ij}, \mathrm{STRUC}_i, \mathrm{UP}_{ijt}, e_{ijt}) \tag{2}$$

where AF_{ij} is the features of household i's appliance j, STRUC_i is structural features of household i, and UP_{ijt} is the utilization patterns relating to appliance j, and e_{ijt} is a random error term for the end-use.

The effect of weather conditions (WC_{it}), market conditions (MC_{it}), and household's economic and demographic characteristics (EDC_i) on the end-use energy utilization pattern can be shown as

$$\mathrm{UP}_{ijt} = g_j(\mathrm{WC}_{it}, \mathrm{MC}_{it}, \mathrm{EDC}_i) \tag{3}$$

Substituting Eq. (3) into Eq. (2) yields

$$\mathrm{UEC}_{ijt} = F_j(\mathrm{AF}_{ij}, \mathrm{STRUC}_i, \mathrm{WC}_{it}, \mathrm{MC}_{it}, \mathrm{EDC}_i, e_{ijt}) \tag{4}$$

And finally substituting Eq. (4) into Eq. (1) gives the general equation:

$$\mathrm{HEC}_{it} = \sum_{j=1}^{j} F_j(\mathrm{AF}_{ij}, \mathrm{STRUC}_i, \mathrm{WC}_{it}, \mathrm{MC}_{it}, \mathrm{EDC}_i, e_{ijt}) \times S_{ij} \tag{5}$$

Since the individual error terms are additive within their respective unit energy consumption (UEC) functions, the household energy consumption equation can then be written as

$$\mathrm{HEC}_{it} = \sum_{j=1}^{j} F_j(\mathrm{AF}_{ij}, \mathrm{STRUC}_i, \mathrm{WC}_{it}, \mathrm{MC}_{it}, \mathrm{EDC}_i) \times S_{ij} + e_{it} \tag{6}$$

where $e_{it} = \sum_{j=1}^{1} e_{ij} \times S_{ij}$.

In most CDA models, multicollinearity problem arises, which is caused by correlation amongst the variables included in the CDA model, limiting the capability of the regression to distinguish the impacts of these variables. Thus, the influence of some individual appliances on the total end-use energy consumption becomes difficult to separate. Mostly, appliances with high saturation cause multicollinearity problems. Moreover, it is not

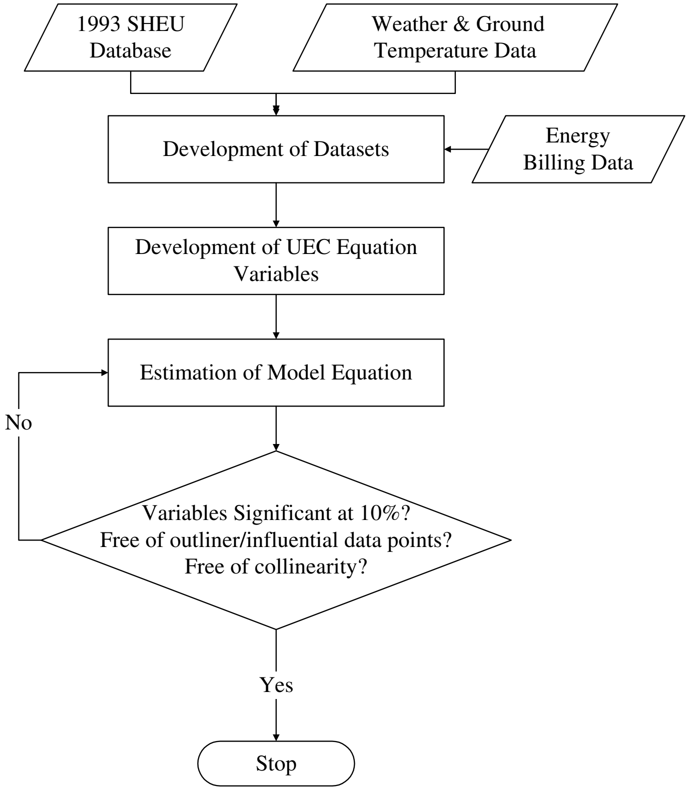

Fig. 1. Flowchart diagram depicting the methodology used for the development of the CDA model.

uncommon for this approach to yield unrealistic negative appliance consumption estimates, because of the high degree of multicollinearity.

The CDA model can be estimated statistically by standard multivariate regression analysis using data on household energy consumption, appliance saturation, and other variables given in Eq. (6).

The overall fit of the CDA model depends on the model specification and data quality. In general, the multiple coefficient of determination values of these models range from 0.55 [14] to 0.75 [16]. These values might seem low, but explaining the cross sectional behaviour of individual households is a difficult process since energy consumption is affected by many other factors that cannot be readily identified or quantified (tastes, habits, special circumstances), and consequently, can not be incorporated into the model. Similarly, it is not possible to incorporate all of the house characteristics (*e.g.* wall, roof, window, *etc.*, areas, insulation values, infiltration, solar heat gains, climatic factors, *etc.*) into the regression model.

Once the CDA model is estimated statistically, it can be used to predict the UEC of individual households, as well as a designated group of households.

4. Modeling of residential sector energy consumption using conditional demand analysis

As the review of the literature presented above indicates, CDA approach has not been used to model national residential energy consumption. In this paper, a new CDA model that was developed to model the end-use energy consumption in the Canadian residential sector is presented. This model is developed with three components. Each component of the model was used to disaggregate the energy consumption of the households with one type of energy billing data. These models are as follows:

- electricity model,
- natural gas model,
- oil model.

The approach used in the development of the each component of the model is similar and pursues the steps shown in Fig. 1. The SYSTAT 9.0 software [25] was used in the development of the CDA model.

5. Sources of data

Two sources of data were used for the development of the input units of the CDA model: the data from the 1993 survey of household energy use (SHEU) database [26] and the 1993 heating and cooling degree day data for the cities in which the households in the CDA data set are located [27].

The 1993 SHEU database contains detailed information on house construction, space heating/cooling and DHW heating equipment, household appliances, and some socioeconomic characteristics of the occupants for 8767 households in Canada. The electricity billing data of 2050 households, natural gas billing data of 1012 households, and natural gas billing data of 236 households in the 1993 SHEU database are used to develop the components of the CDA model.

6. Development of datasets

The households from the 1993 SHEU database with electricity, natural gas, or oil billing data were selected for the datasets of the CDA model. Hence, the dataset of the CDA electricity model includes the households with electricity billing data, while the datasets of the CDA natural gas and oil models include the households with natural gas and oil billing data, respectively.

7. Development of unit energy consumption equations

The total energy consumption of a household is the cumulative of all energy consumed for the various end-uses in the household. A UEC equation is developed for each end-use. The 1993 SHEU database [26], and the 1993 weather and ground temperature data [27] are the sources of information used in the development of the UEC equations. The input variables of each end-use UEC equation were chosen taking into consideration the major

Table 1
Definitions of commonly used variables in unit energy consumption (UEC) equations

Variable	Definition	Range
PROGT	Dummy variable: one if the household has programmable thermostat, zero if not	0–1
HRV	Dummy variable: one if the household has a heat recovery ventilation system, zero if not	0–1
EFF	Efficiency of the natural gas or oil furnace/boiler [%]	65–94
SHAGE	Age of the natural gas or oil furnace/boiler [yr]	1–25
AIT	Average indoor temperature [°C]	16–24
DTYPE	Dummy variable: one if the dwelling type is single detached, zero if it is single attached	0–1
AREA	Heated living area [m^2]	51–502
AGECAT	Dwelling construction year category	1–6
BSMNT	Dummy variable: one if the household has a heated basement, zero if not	0–1
GARAGE	Dummy variable: one if the household has a heated garage, zero if not	0–1
ATTIC	Dummy variable: one if the household has an attic, zero if not	0–1
TRIPLE	Number of triple glazed windows	0–30
DOUBLE	Number of double glazed windows	0–60
SINGLE	Number of single glazed windows	0–74
DOOR	Number of doors	1–11
HDD	Heating degree days [°C-day]	2930–6541
OWNER	Dummy variable: one if owner, zero if renter	0–1
INCOME	Household income [$10,000/yr]	10–85
CHILD	Number of children	0–7
ADULT	Number of adults	1–8
DAYTIME	Dummy variable: one if the dwelling is occupied daytime during weekdays, zero if not	0–1
POPUL	Size of area of residence: 1 if population is less than 15,000 2 if population is between 15,000 and 100,000 3 if population is 100,000 or over	1–3
TANK	Size of the DHW tank [L]	130–280
SYSAGE	Age of the DHW heating system [yr]	0.5–18
BLANKET	Dummy variable: one if there is an add-on insulation blanket around the outside of the DHW tank, zero if not	0–1
PIPEINS	Dummy variable: one if there is insulation around the DHW pipes, zero if not	0–1
AERATOR	Number of aerators	0–7
LOWFLOW	Number of low-flow shower heads	0–3
GT	Ground temperature [°C]	4–12
CWLOAD	Clothes washer [loads/week]	0–15
DWLOAD	Dish washer [loads/week]	0–15
CACUSE	Central A/C unit usage [h/yr]	0–1125
WACUSE	Window A/C unit usage [h/yr]	0–1125
VOLR1	Volume of the main refrigerator [L]	0–625
VOLR2	Volume of the second refrigerator [L]	0–625
FROSTR1	Dummy variable: one if the main refrigerator is frost-free, zero if not	0–1
FROSTR2	Dummy variable: one if the second refrigerator is frost-free, zero if not	0–1
HHSIZE	Number of occupants in the household	1–11
VOLF1	Volume of the main freezer [L]	0–710
VOLF2	Volume of the second freezer [L]	0–710
MICROW	Dummy variable: one if the household has a microwave, zero if not	0–1
CDLOAD	Clothes dryer [loads/week]	0–15

(*continued on next page*)

Table 1 (*continued*)

Variable	Definition	Range
LIGHTS	Total number of incandescent, fluorescent, and halogen lamps	7–132
WINDOW	Total number of single, double, and triple glazed windows	4–76
CDD	Cooling degree days [°C-day]	4–405

determinants of the end-use consumption. After identifying the major determinants of the end-uses, the UEC equation for each end-use was developed as given in Eq. (4).

7.1. Development of the electricity model UEC equations

The CDA electricity model was developed by combining the UEC equations of the end-uses as given in the following equation:

$$\mathrm{HEC}_i = \mathrm{Cons}\tan t + \sum_{j=1}^{N} \mathrm{UEC}_{ij} \tag{7}$$

where HEC_i is electricity consumption of household i [kWh/yr], UEC_{ij} is the end-use j unit energy consumption of household i [kWh/yr], N is number of electricity end-uses, i.e. main and supplementary space heating, DHW heating, space cooling, lighting, major and minor appliances.

The constant term in Eq. (7) was included in the CDA electricity model to represent the electricity consumption by unaccounted miscellaneous appliances.

7.1.1. Main and supplementary space heating UEC equations

The input variables of the space heating UEC equations were chosen considering the structural features of the dwellings, economic and demographic characteristics of the occupants, and the weather conditions. The main and supplementary space heating UEC equations used in the CDA electricity model are given in Eqs. (8) and (9), respectively, and the definitions of the commonly used variables in the equations are given in Table 1:

$$\begin{aligned}\mathrm{UEC}_{\mathrm{SH}S} = \mathrm{SH} * [&a_0 + a_1\mathrm{PROGT} + a_2\mathrm{HRV} + a_3\mathrm{AIT} + a_4\mathrm{DTYPE} + a_5\mathrm{AREA}\\ &+ a_6\mathrm{AGECAT} + a_7\mathrm{BSMNT} + a_8\mathrm{GARAGE} + a_9\mathrm{ATTIC}\\ &+ a_{10}\mathrm{TRIBLE} + a_{11}\mathrm{DOUBLE} + a_{12}\mathrm{SINGLE} + a_{13}\mathrm{DOOR}\\ &+ a_{14}\mathrm{HDD} + a_{15}\mathrm{OWNER} + a_{16}\mathrm{INCOME} + a_{17}\mathrm{CHILD}\\ &+ a_{18}\mathrm{ADULT} + a_{19}\mathrm{DAYTIME} + a_{20}\mathrm{POPUL}]\end{aligned} \tag{8}$$

where $\mathrm{UEC}_{\mathrm{SH}}$ is space heating unit energy consumption [kWh/household/yr], SH is dummy variable: one if the household has electricity space heating equipment, zero if not, and $a_0, \ldots, a_{20}$ is the regression coefficients of each variable.

$$\begin{aligned}\mathrm{UEC}_{\mathrm{SSH}} = \mathrm{SSH} * [&a_0 + a_1\mathrm{AIT} + a_2\mathrm{AREA} + a_3\mathrm{HDD} + a_4\mathrm{CHILD}\\ &+ a_5\mathrm{ADULT} + a_6\mathrm{DAYTIME}]\end{aligned} \tag{9}$$

where UEC_{SSH} is the supplementary space heating unit energy consumption [kWh/household/yr], SSH is dummy variable: one if the household has electricity supplementary space heating equipment, zero if not, and $a_0, \ldots, a_6$ is the regression coefficients of each variable.

7.1.2. DHW heating UEC equation

The input variables of the DHW heating UEC equation were chosen based on the available information on DHW heating system and equipment properties, DHW consumption patterns, economic and demographic characteristics of the occupants, and the weather conditions. The DHW heating UEC equation used in the CDA electricity model is given in Eq. (10), and the definitions of the variables used in the equation are given in Table 1:

$$\begin{aligned} UEC_{DHW} = {} & DHW * [a_0 + a_1 TANK + a_2 SYSAGE + a_3 BLANKET \\ & + a_4 PIPEINS + a_5 LOWFLOW + a_6 AERATOR + a_7 GT \\ & + a_8 CWLOAD + a_9 DWLOAD + a_{10} DTYPE + a_{11} OWNER \\ & + a_{12} INCOME + a_{13} CHILD + a_{14} ADULT] \end{aligned} \tag{10}$$

where UEC_{DHW} is the DHW heating unit energy consumption [kWh/household/yr], DHW is dummy variable: one if the household has electricity DHW heating equipment, zero if not, and $a_0, \ldots, a_{14}$ is the regression coefficients of each variable.

7.1.3. Space cooling UEC equations

The input variables of the central and window air conditioning (A/C) UEC equations were chosen based on the available information on space cooling system usage, structural features of the dwellings, economic and demographic characteristics of the occupants, and the weather conditions. The central and window A/C UEC equations used in the CDA electricity model are given in Eqs. (11) and (12), respectively, and the definitions of the variables used in the equations are given in Table 1:

$$\begin{aligned} UEC_{CAC} = {} & CAC * [a_0 + a_1 CACUSE + a_2 DTYPE + a_3 AREA + a_4 AGECAT \\ & + a_5 ATTIC + a_6 TRIBLE + a_7 DOUBLE + a_8 SINGLE \\ & + a_9 DOOR + a_{10} \mathit{CDD} + a_{11} OWNER + a_{12} INCOME \\ & + a_{13} CHILD + a_{14} ADULT + a_{15} DAYTIME] \end{aligned} \tag{11}$$

where UEC_{CAC} is central A/C unit energy consumption [kWh/household/yr], CAC is dummy variable: one if the household has central A/C unit, zero if not, and $a_0, \ldots, a_{15}$ is the regression coefficients of each variable.

$$\begin{aligned} UEC_{WAC} = {} & WAC * [a_0 + a_1 WACUS\mathit{E} + a_2 AREA + a_3 AGECAT + a_4 CDD \\ & + a_5 INCOME + a_6 CHILD + a_7 ADULT + a_8 DAYTIME] \end{aligned} \tag{12}$$

where UEC_{WAC} is the window A/C unit energy consumption [kWh/household/yr], WAC is dummy variable: one if the household has window A/C unit, zero if not, and $a_0, \ldots, a_8$ is the regression coefficients of each variable.

7.1.4. Major appliances UEC equations

The major appliances include the main and secondary refrigerators, main and secondary freezers, electric ranges, dishwashers, clothes washers, and electric clothes dryers. The input variables of the major appliances UEC equations were chosen based on the available information on appliance properties, and economic and demographic characteristics of the occupants. The major appliances UEC equations used in the CDA electricity model are given in Eqs. (13)–(20), and the definitions of the variables used in the equations are given in Table 1:

$$\mathrm{UEC_{REF1}} = \mathrm{REF1} * [a_0 + a_1\mathrm{VOLR1} + a_2\mathrm{FROSTR1} + a_3\mathrm{INCOME} + a_4\mathrm{HHSIZE}] \tag{13}$$

$$\mathrm{UEC_{REF2}} = \mathrm{REF2} * [a_0 + a_1\mathrm{VOLR2} + a_2\mathrm{FROSTR2} + a_3\mathrm{INCOME} + a_4\mathrm{HHSIZE}] \tag{14}$$

$$\mathrm{UEC_{FREZ1}} = \mathrm{FREZ1} * [a_0 + a_1\mathrm{VOLF1} + a_2\mathrm{INCOME} + a_3\mathrm{HHSIZE}] \tag{15}$$

$$\mathrm{UEC_{FREZ2}} = \mathrm{FREZ2} * [a_0 + a_1\mathrm{VOLF2} + a_2\mathrm{INCOME} + a_3\mathrm{HHSIZE}] \tag{16}$$

$$\mathrm{UEC_{COOK}} = \mathrm{COOK} * [a_0 + a_1\mathrm{HHSIZE} + a_2\mathrm{MICROW}] \tag{17}$$

$$\mathrm{UEC_{DISH}} = \mathrm{DISH} * [a_0 + a_1\mathrm{DWLOAD}] \tag{18}$$

$$\mathrm{UEC_{CLOTH}} = \mathrm{CLOTH} * [a_0 + a_1\mathrm{CWLOAD}] \tag{19}$$

$$\mathrm{UEC_{DRYER}} = \mathrm{DRYER} * [a_0 + a_1\mathrm{CDLOAD}] \tag{20}$$

where $\mathrm{UEC_{REF1}}$ is the main refrigerator unit energy consumption [kWh/household/yr], $\mathrm{UEC_{REF2}}$ is secondary refrigerator unit energy consumption [kWh/household/yr], $\mathrm{UEC_{FREZ1}}$ is main freezer unit energy consumption [kWh/household/yr], $\mathrm{UEC_{FREZ2}}$ is Secondary freezer unit energy consumption [kWh/household/yr], $\mathrm{UEC_{COOK}}$ is electric range unit energy consumption [kWh/household/yr], $\mathrm{UEC_{DISH}}$ is dishwasher unit energy consumption [kWh/household/yr], $\mathrm{UEC_{CLOTH}}$ is clothes washer unit energy consumption [kWh/household/yr], $\mathrm{UEC_{DRYER}}$ is electric clothes dryer unit energy consumption [kWh/household/yr], REF1 is dummy variable: one if the household has a refrigerator, zero if not, REF2: is dummy variable: one if the household has a secondary refrigerator, zero if not, FREZ1 is dummy variable: one if the household has a freezer, zero if not, FREZ2 is dummy variable: one if the household has a secondary freezer, zero if not, COOK is dummy variable: one if the household has an electric range, zero if not, DISH is dummy variable: one if the household has a dishwasher, zero if not, CLOTH is dummy variable: one if the household has a clothes washer, zero if not, DRYER is dummy variable: one if the household has an electric clothes dryer, zero if not, and $a_0, \ldots, a_4$ is the regression coefficients of each variable.

7.1.5. Minor appliances UEC equations

The minor appliances include microwaves, color TVs, black and white TVs, VCRs, furnace fans and boiler pumps. The input variables of the minor appliances UEC equations were chosen based on the available information on the demographic characteristics of the occupants and weather conditions. The minor appliances UEC equations used in the CDA electricity model are given in Eqs. (21)–(26), and the definitions of the variables used in the equations are given in Table 1:

$$\mathrm{UEC_{FF}} = \mathrm{FF} * [a_0 + a_1\mathrm{AREA} + a_2\mathrm{HDD}] \tag{21}$$

$$\mathrm{UEC_{BP}} = \mathrm{BP} * [a_0 + a_1\mathrm{AREA} + a_2\mathrm{HDD}] \tag{22}$$

$$\mathrm{UEC_{MICROW}} = \mathrm{MICROW} * [a_0 + a_1\mathrm{HHSIZE}] \tag{23}$$

$$\mathrm{UEC_{CTV}} = \mathrm{CTV} * [a_0 + a_1\mathrm{HHSIZE}] \tag{24}$$

$$\mathrm{UEC_{BWTV}} = \mathrm{BWTV} * [a_0 + a_1\mathrm{HHSIZE}] \tag{25}$$

$$\mathrm{UEC_{VCR}} = \mathrm{VCR} * [a_0 + a_1\mathrm{HHSIZE}] \tag{26}$$

where $\mathrm{UEC_{FF}}$ is furnace fan unit energy consumption [kWh/household/yr], $\mathrm{UEC_{BP}}$ is boiler pump unit energy consumption [kWh/household/yr], $\mathrm{UEC_{MICROW}}$ is microwave unit energy consumption [kWh/household/yr], $\mathrm{UEC_{CTV}}$ is color TV unit energy consumption [kWh/household/yr], $\mathrm{UEC_{BWTV}}$ is black and white TV unit energy consumption [kWh/household/yr], $\mathrm{UEC_{VCR}}$ is VCR unit energy consumption [kWh/household/yr], FF is dummy variable: one if the household has a furnace, zero if not, BP is dummy variable: one if the household has a boiler, zero if not, MICROW is dummy variable: one if the household has a microwave, zero if not, CTV is number of color TVs owned by the household, BWTV is number of black and white TVs owned by the household, VCR is number of VCRs owned by the household, and $a_0, \ldots, a_2$ is the regression coefficients of each variable.

7.1.6. Lighting UEC equation

The input variables of the lighting UEC equation are the number of halogen, incandescent, and fluorescent lamps. The lighting UEC equation used in the CDA electricity model is given in Eq. (27).

$$\mathrm{UEC_{LIGHT}} = a_1\mathrm{HALO} + a_2\mathrm{INCA} + a_3\mathrm{FLOU} \tag{27}$$

where $\mathrm{UEC_{LIGHT}}$ is lighting unit energy consumption [kWh/household/yr], HALO is number of halogen lamps owned by the household, INCA is number of incandescent lamps owned by the household, FLOU is number of fluorescent lamps owned by the household, and $a_1, \ldots, a_3$ is the regression coefficients of each variable.

To simplify the UEC equations, the variables representing the number of halogen, incandescent, and fluorescent lamps were combined into one variable that represents the total number of lamps (LIGHTS). This new variable was used to replace the corresponding variables in Eq. (27).

7.1.7. Estimation of the electricity model equation

The end-use UEC equations given in Eqs. (8)–(27) were combined to develop the CDA electricity model. The detailed output of the regression analysis is given in Fig. A.1 of Appendix. The model equation was reduced by removing the non-significant variables at 10% level, outlier and influential data points, and variables that increase multicollinearity. The resulting model equation is as follows:

$$\begin{aligned}\mathrm{HEC} = {} & 2128.65 + \mathrm{SH} * [1.60 * \mathrm{HDD} - 2516.01 * \mathrm{HRV} + 1891.75 * \mathrm{DTYPE} \\ & + 12.67 * \mathrm{AREA} - 785.40 * \mathrm{AGECAT} + 78.14 * \mathrm{INCOME}] + \mathrm{SSH} \\ & * [8.13 * \mathrm{AREA} + 569.33 * \mathrm{CHILD}] + \mathrm{DHW} * [16.86 * \mathrm{TANK} - 691.34 \\ & * \mathrm{LOWFLOW} + 215.04 * \mathrm{DWLOAD} + 752.83 * \mathrm{ADULT}] + \mathrm{CAC} * [1.27 \\ & * \mathrm{CACUSE}] + \mathrm{REF1} * [1030.23 * \mathrm{FROSTR1}] + \mathrm{REF2} * [1636.48 \\ & * \mathrm{FROSTR2} + 28.24 * \mathrm{INCOME}] + \mathrm{FREZ1} * [1.50 * \mathrm{VOLF1}] + \mathrm{COOK} \\ & * [421.25 * \mathrm{HHSIZE}] + \mathrm{DRYER} * [303.61 * \mathrm{CDLOAD}] + 50.18 \\ & * \mathrm{LIGHTS}\end{aligned} \tag{28}$$

As shown in Fig. A.1 of Appendix, the multiple coefficient of determination of the CDA electricity model is 0.66, which indicates that 66% of variation in the estimated household electricity consumption can be accounted for by the prediction based on the variables of Eq. (28).

7.2. Development of the natural gas model UEC equations

The CDA natural gas model was developed by combining the UEC equations of the end-uses as given in Eq. (29). Since all natural gas end-uses were reported in the 1993 SHEU database, the constant term was not included in the model:

$$\mathrm{HEC}_i = \sum_{j=1}^{N} \mathrm{UEC}_{ij} \tag{29}$$

where HEC_i is natural gas consumption of household i [m^3/yr], UEC_{ij} is end-use j unit natural gas consumption of household i [m^3/yr], and N is the number of natural gas end-uses, *i.e.* main and supplementary space heating, DHW heating, clothes drying, cooking, pool heating, and fireplaces.

7.2.1. Main and supplementary space heating UEC equations

The input variables of the space heating UEC equations were chosen considering the end-use efficiency of the space heating equipment, structural features of the dwellings, economic and demographic characteristics of the occupants, and the weather conditions. The supplementary space heating UEC equation includes the households with natural gas supplementary space heating units and fireplaces. The main and supplementary space heating UEC equations used in the CDA natural gas model are given in Eqs. (30) and (31), respectively, and the definitions of the variables used in the equations are given in Table 1:

$$\begin{aligned}\mathrm{UEC_{SH}} = \mathrm{SH} * [&a_0 + a_1\mathrm{EFF} + a_2\mathrm{SHAGE} + a_3\mathrm{PROGT} + a_4\mathrm{AIT} + a_5\mathrm{DTYPE} \\ &+ a_6\mathrm{AREA} + a_7\mathrm{AGECAT} + a_8\mathrm{BSMNT} + a_9\mathrm{GARAGE} \\ &+ a_{10}\mathrm{ATTIC} + a_{11}\mathrm{TRIBLE} + a_{12}\mathrm{DOUBLE} + a_{13}\mathrm{SINGLE} \\ &+ a_{14}\mathrm{DOOR} + a_{15}\mathrm{HDD} + a_{16}\mathrm{OWNER} + a_{17}\mathrm{INCOME} \\ &+ a_{18}\mathrm{CHILD} + a_{19}\mathrm{ADULT} + a_{20}\mathrm{DAYTIME} + a_{21}\mathrm{POPUL}]\end{aligned} \tag{30}$$

where $\mathrm{UEC_{SH}}$ is space heating unit energy consumption [m^3/household/yr], SH is dummy variable: one if the household has natural gas space heating equipment, zero if not, and $a_0, \ldots, a_{21}$ is the regression coefficients of each variable.

$$\begin{aligned}\mathrm{UEC_{SSH}} = \mathrm{SSH} * [&a_0 + a_1\mathrm{AIT} + a_2\mathrm{AREA} + a_3\mathrm{HDD} + a_4\mathrm{CHILD} \\ &+ a_5\mathrm{ADULT} + a_6\mathrm{DAYTIME}]\end{aligned} \tag{31}$$

where $\mathrm{UEC_{SSH}}$ is supplementary space heating unit energy consumption [m^3/household/yr], SSH is dummy variable: one if the household has natural gas supplementary space heating equipment and/or fireplace, zero if not, and $a_0, \ldots, a_6$ is the regression coefficients of each variable.

7.2.2. DHW heating UEC equation

The input variables of the DHW heating UEC equation were chosen based on the available information on DHW heating system and equipment properties, DHW consumption patterns, economic and demographic characteristics of the occupants, and the weather conditions. The DHW heating UEC equation used in the CDA natural gas model is given in Eq. (32), and the definitions of the variables used in the equation are given in Table 1:

$$\begin{aligned} \mathrm{UEC_{DHW}} = \mathrm{DHW} * [&a_0 + a_1\mathrm{TANK} + a_2\mathrm{SYSAGE} + a_3\mathrm{BLANKET} \\ &+ a_4\mathrm{PIPEINS} + a_5\mathrm{LOWFLOW} + a_6\mathrm{AERATOR} + a_7\mathrm{GT} \\ &+ a_8\mathrm{CWLOAD} + a_9\mathrm{DWLOAD} + a_{10}\mathrm{DTYPE} + a_{11}\mathrm{OWNER} \\ &+ a_{12}\mathrm{INCOME} + a_{13}\mathrm{CHILD} + a_{14}\mathrm{ADULT}] \end{aligned} \tag{32}$$

where $\mathrm{UEC_{DHW}}$ is DHW heating unit energy consumption [m^3/household/yr], DHW is dummy variable: one if the household has natural gas DHW heating equipment, zero if not, and $a_0, \ldots, a_{14}$ is the regression coefficients of each variable.

7.2.3. Cooking, clothes drying, and pool heating UEC equations

The input variables of the cooking, clothes drying, and pool heating UEC equations were chosen based on the available information on appliance usage, and economic and demographic characteristics of the occupants. The UEC equations used in the CDA natural gas model are given in Eqs. (33)–(35), and the definitions of the variables used in the equations are given in Table 1:

$$\mathrm{UEC_{COOK}} = \mathrm{COOK} * [a_0 + a_1\mathrm{HHSIZE} + a_2\mathrm{MICROW}] \tag{33}$$

$$\mathrm{UEC_{DRYER}} = \mathrm{DRYER} * [a_0 + a_1\mathrm{CDLOAD}] \tag{34}$$

$$\mathrm{UEC_{POOL}} = \mathrm{POOL} * [a_0 + a_1\mathrm{INCOME}] \tag{35}$$

where $\mathrm{UEC_{COOK}}$ is the natural gas range unit energy consumption [m^3/household/yr], $\mathrm{UEC_{DRYER}}$ is natural gas clothes dryer unit energy consumption [m^3/household/yr], $\mathrm{UEC_{POOL}}$ is natural gas pool heating unit energy consumption [m^3/household/yr], COOK is dummy variable: one if the household has a natural gas range, zero if not, DRYER is dummy variable: one if the household has a natural gas clothes dryer, zero if not, POOL is dummy variable: one if the household has a natural gas pool heater, zero if not, and $a_0, \ldots, a_2$ is the regression coefficients of each variable.

7.2.4. Estimation of the natural gas model equation

The end-use UEC equations given in Eqs. (30)–(35) were combined to develop the CDA natural gas model. The output of the regression analysis is given in Fig. A.2 of Appendix. The model equation was reduced by removing the non-significant variables at 10% level, outlier and influential data points, and variables that increase multicollinearity. The resulting model equation is as follows:

$$\begin{aligned}\text{HEC} = {} & \text{SH} * [-207.75 * \text{PROGT} + 129.84 * \text{DOOR} + 26.30 * \text{WINDOW} \\ & + 46.03 * \text{SHAGE} + 437.69 * \text{GARAGE} + 6.61 * \text{AREA} + 181.61 \\ & * \text{ADULT}] + \text{DHW} * [13.21 * \text{SYSAGE} + 70.77 * \text{HHSIZE} + 720.97 \\ & * \text{DTYPE}] + \text{COOK} * [94.04 ast \text{HHSIZE}] + \text{DRYER} * [35.91 \\ & * \text{CDLOAD}] + 969.56 * \text{POOL}\end{aligned} \tag{36}$$

As shown in Fig. A.2 of Appendix, the multiple coefficient of determination of the CDA natural gas model is 0.92, which indicates that 92% of variation in the estimated household natural gas consumption can be accounted for by the prediction from the variables of Eq. (36).

7.3. Development of the oil model UEC equations

There are only two households with oil cooking ranges or pool heaters in the database. Therefore, oil cooking and pool heating energy consumption are not addressed in the model. The CDA oil model was developed by combining the UEC equations of the end-uses as given in Eq. (37). Since all oil end-uses were reported in the 1993 SHEU database, the constant term was not included in the model.

$$\text{HEC}_i = \sum_{j=1}^{N} \text{UEC}_{ij} \tag{37}$$

where HEC_i is oil consumption of household i [L/yr], UEC_{ij} is end-use j unit oil consumption of household i [L/yr], N is number of oil end-uses, *i.e.* main and supplementary space heating, DHW heating, cooking, and pool heating.

7.3.1. Space heating UEC equation

There are only two households with oil supplementary heating units in the CDA oil model dataset. Thus, the oil supplementary space heating UEC equation is excluded from the analysis. The input variables of the space heating UEC equation were chosen considering the end-use efficiency of the space heating equipment, structural features of the dwellings, economic and demographic characteristics of the occupants, and the weather conditions. The space heating UEC equation used in the CDA oil model is given in Eq. (38), and the definitions of the variables used in the equations are given in Table 1:

$$\begin{aligned}\text{UEC}_{\text{SH}} = {} & \text{SH} * [a_0 + a_1\text{EFF} + a_2\text{SHAGE} + a_3\text{PROGT} + a_4\text{AIT} + a_5\text{DTYPE} \\ & + a_6\text{AREA} + a_7\text{AGECAT} + a_8\text{BSMNT} + a_9\text{GARAGE} \\ & + a_{10}\text{ATTIC} + a_{11}\text{TRIBLE} + a_{12}\text{DOUBLE} + a_{13}\text{SINGLE} \\ & + a_{14}\text{DOOR} + a_{15}\text{HDD} + a_{16}\text{OWNER} + a_{17}\text{INCOME} \\ & + a_{18}\text{CHILD} + a_{19}\text{ADULT} + a_{20}\text{DAYTIME} + a_{21}\text{POPUL}]\end{aligned} \tag{38}$$

where UEC_{SH} is space heating unit energy consumption [L/household/yr], SH is dummy variable: one if the household has oil space heating equipment, zero if not, and $a_0, \ldots, a_{21}$ is the regression coefficients of each variable.

7.3.2. DHW heating UEC equation

The input variables of the DHW heating UEC equation were chosen based on the available information on DHW heating system and equipment properties, DHW consumption patterns, economic and demographic characteristics of the occupants, and the weather conditions. The DHW heating UEC equation used in the CDA oil model is given in Eq. (39), and the definitions of the variables used in the equation are given in Table 1:

$$\begin{aligned}\mathrm{UEC}_{\mathrm{DHW}} = \mathrm{DHW} * [&a_0 + a_1\mathrm{TANK} + a_2\mathrm{SYSAGE} + a_3\mathrm{BLANKET} \\ &+ a_4\mathrm{PIPEINS} + a_5\mathrm{LOWFLOW} + a_6\mathrm{AERATOR} + a_7\mathrm{GT} \\ &+ a_8\mathrm{CWLOAD} + a_9\mathrm{DWLOAD} + a_{10}\mathrm{DTYPE} + a_{11}\mathrm{OWNER} \\ &+ a_{12}\mathrm{INCOME} + a_{13}\mathrm{CHILD} + a_{14}\mathrm{ADULT}]\end{aligned} \tag{39}$$

where $\mathrm{UEC}_{\mathrm{DHW}}$ is DHW heating unit energy consumption [L/household/yr], DHW is dummy variable: one if the household has oil DHW heating equipment, zero if not, and $a_0, \ldots, a_{14}$ is the regression coefficients of each variable.

7.3.3. Estimation of the oil model equation

The space and DHW heating UEC equations given in Eqs. (38) and (39) were combined to develop the CDA oil model. The output of the regression analysis is given in Fig. A.3 of Appendix. The model equation was reduced by removing the non-significant variables at 10% level, outlier and influential data points, and variables that increase multicollinearity. The resulting model equation is as follows:

$$\begin{aligned}\mathrm{HEC} = \mathrm{SH} * [&40.05 * \mathrm{SHAGE} + 4.95 * \mathrm{AREA} + 41.68 * \mathrm{WINDOW} + 471.21 \\ &* \mathrm{DTYPE}] + \mathrm{DHW} * [5.28 * \mathrm{TANK} + 50.91 * \mathrm{DWLOAD}]\end{aligned} \tag{40}$$

As shown in Fig. A.3 of Appendix, the multiple coefficient of determination of the CDA oil model is 0.87, which indicates that 87% of variation in the estimated household oil consumption can be accounted for by the prediction from the variables of Eq. (40).

8. Comparison of the CDA, NN, and engineering models

The prediction performance of the CDA, NN, and the engineering models was assessed by comparing the estimates of the models with actual energy consumption data from the 247 households with appliance, lighting, and space cooling (ALC) energy consumption data, 141 households with the DHW heating energy consumption data, and 307 households with space heating energy consumption data. The results are presented in Table 2. As it can be seen, all three models are capable of predicting the ALC, space and DHW heating energy consumption with reasonable accuracy (multiple correlation coefficient, R^2, better than 0.77). The engineering and CDA models have lower multiple correlation coefficient (R^2) and higher coefficient of variation (CV) values than the NN model, which shows that the NN model have a higher prediction performance than the engineering and CDA models. The formulae used to calculate R^2 and CV are given in Table A.1 of Appendix.

The CDA, NN, and engineering models were also used to predict the ALC, space and DHW heating energy consumption of the 1993 SHEU households that were not used in

Table 2
Prediction performances of the CDA, NN, and engineering (ENG) models

	R^2	CV
ALC energy consumption		
ENG	0.780	3.463
CDA	0.795	3.343
NN	0.909	2.094
DHW heating energy consumption		
ENG	0.828	3.898
CDA	0.814	4.052
NN	0.871	3.337
Space heating energy consumption		
ENG	0.778	2.877
CDA	0.892	2.008
NN	0.908	1.871

Table 3
Weighted average appliance, lighting and cooling (ALC), space and domestic hot water (DHW) heating energy consumptions by the NN, CDA, and engineering (ENG) models and average percent deviation

	Weighed average end-use energy consumption [GJ/yr/household]	Average deviation from NN estimate [%]
ALC energy consumption		
NN	32	–
CDA	30	−4.6
ENG	32	2.5
DHW heating energy consumption		
NN	26	–
CDA	25	−3.1
ENG	25	−4.5
Space heating energy consumption		
NN	80	–
CDA	75	−6.9
ENG	77	−3.4

the development of the models. The average ALC, space and DHW heating energy consumption estimates, and the average percent deviations are given in Table 3. As seen in this table, the averages of the CDA, NN, and engineering model estimates are close to each other.

Further comparisons of the estimates of the CDA, NN, and engineering models were conducted based on dwelling type, size and age, as well as the type of fuel used. The results of these comparisons, which are presented elsewhere [10], indicate that the CDA, NN, and engineering model estimates are generally in good agreement.

9. Household energy consumption in Canada

The total energy consumption of the households in the 1993 SHEU database were computed using the engineering model, by combining the ALC [11], DHW heating, and space heating [12] energy consumption estimates of the NN model, and by combining the elec-

Table 4
Weighted average household energy consumption of the households in the 1993 SHEU database and percent deviations from Natural Resources Canada's (NRCan) estimates

	Weighed average household energy consumption [GJ/yr/household]	Average deviation from NRCan estimate [%]
NRCan	134	
NN	139	3.7
CDA	132	−1.5
Engineering	135	0.7

tricity, natural, and oil energy consumption estimates of the CDA model. Since the 1993 SHEU database is representative of the Canadian housing stock, the weighted average household energy consumption estimates of the CDA, NN, and engineering models are the same as the weighted average household energy consumption of the housing stock. The household energy consumption estimates of the CDA, NN, and engineering models and their deviation from the 134 GJ/yr/household estimate by Natural Resources Canada (NRCan) [1] are shown in Table 4. As seen in Table 4, the estimates by all three models are close to the estimate by NRCan.

The average household energy consumption in each province was also calculated using the estimates of the CDA, NN, and engineering models. The results presented in Fig. 2 indicate that the estimates of the three models are in agreement. The average household energy consumption in Quebec is found to be the lowest, whereas Alberta and Saskatchewan have the highest household energy consumption. Since space heating energy consumption accounts for about 60% of the total household energy consumption, factors such as climate, end-use efficiency, and fuel type of the space heating equipment have significant effects on the total household energy consumption. Consequently, the trend seen in Fig. 2 is mainly due to the fact that, in the 1993 SHEU database, 79% of the households

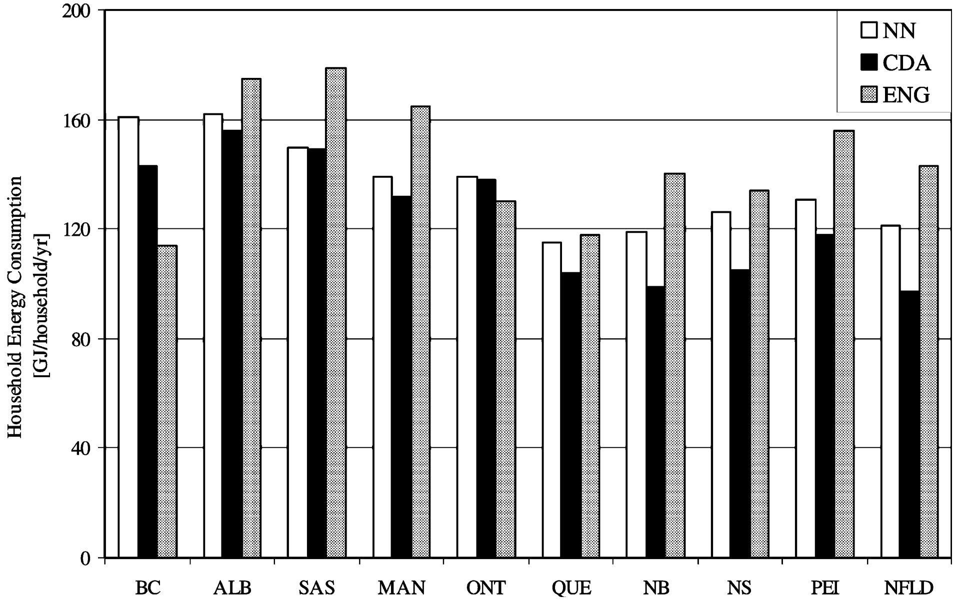

Fig. 2. NN, CDA, and engineering (ENG) models provincial household energy distribution.

in Quebec, and, respectively, 1% and 5% of the households in Alberta and Saskatchewan, have electrical space heating equipment that have 100% end-use efficiency. In addition, 66% and 69% of the households in Alberta and Saskatchewan, respectively, have standard (50–65%) efficiency natural gas, oil, or propane fueled space heating equipment. Along with the cold winters of Alberta and Saskatchewan, these factors explain the high household energy consumption trends in these two provinces.

10. Assessment of socio-economic factors

Since it is possible to incorporate socio-economic factors, such as income, dwelling ownership, size of area of residence, into CDA and NN models, it is possible to estimate the effects of such factors on the ALC, DHW heating, and space heating energy consumption. On the other hand, to include such factors into the engineering model, detailed occupancy and preference profiles are required, which are not available in the SHEU database, or elsewhere in the open literature.

As described in Section 5, data on number of socio-economic factors available in the 1993 SHEU database were incorporated into the CDA and NN model datasets to study the effect of socio-economic factors on ALC, space and DHW heating energy consumption.

The electric, natural gas, and oil CDA models initially included all the socio-economic factors, but most of them were eliminated as a result of statistical significance and multicollinearity problems. The CDA electricity model given in Eq. (28) includes the household income and the number of occupants as socio-economic factors on ALC electricity consumption, and household income on space heating electricity consumption. The CDA natural gas model given in Eq. (36) includes the number of adults as a socio-economic factor on space heating natural gas consumption, and dwelling type DHW heating natural gas consumption. The CDA oil model given in Eq. (40) does not include any socio-economic factors. This shows that the capability of the CDA model to evaluate the effects of a large number of socio-economic factors on the end-use energy consumption is significantly limited.

Some of the results obtained using the CDA and NN models are presented in Figs. 3–6, and more detailed results are presented elsewhere [10]. The ALC, space and DHW heating energy consumption estimates are affected by socio-economic factors as follows.

10.1. Household income

ALC, space and DHW heating energy consumption increases linearly as income increases. This is due to the fact that households with higher income levels have larger dwellings, more appliances and lighting, and use more DHW.

As seen in Fig. 3, CDA electricity model income estimates form a straight line indicating that ALC electricity consumption decreases with a constant rate (*i.e*. with a slope of 0.0064) as income decreases, and does not stay constant at low income levels. On the other hand, the slope of the NN model curve becomes less at low-income levels, indicating that the effect of income on the ALC electricity consumption is lower as shown in Fig. 4. In other words, as income decreases, the ALC energy consumption reaches a minimum and stays constant.

As shown in Figs. 5 and 6, as the income of the household increases, the electricity consumption for space heating increases in a linear fashion based on the estimates of the CDA electricity and NN models.

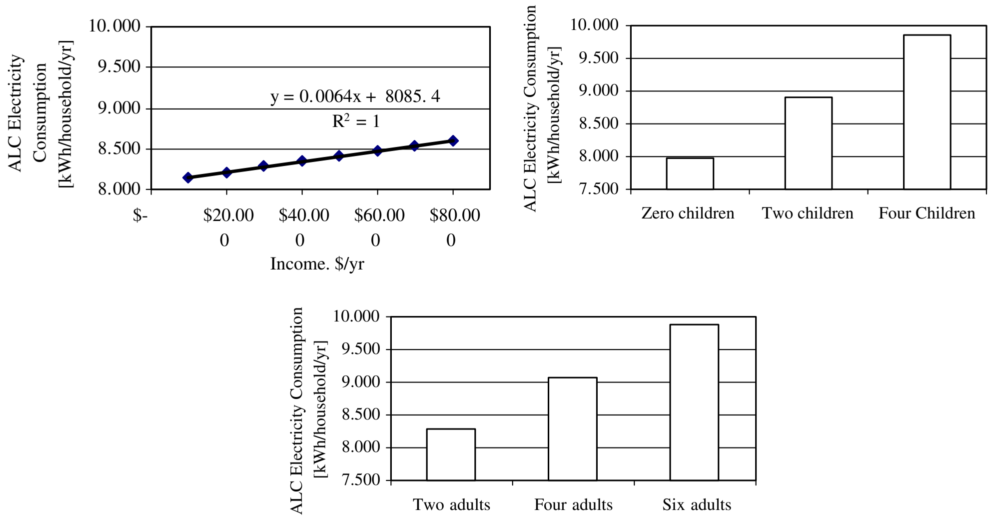

Fig. 3. Effects of socio-economic factors on the appliance, lighting, and cooling (ALC) energy consumption of the households estimated by the CDA electricity model.

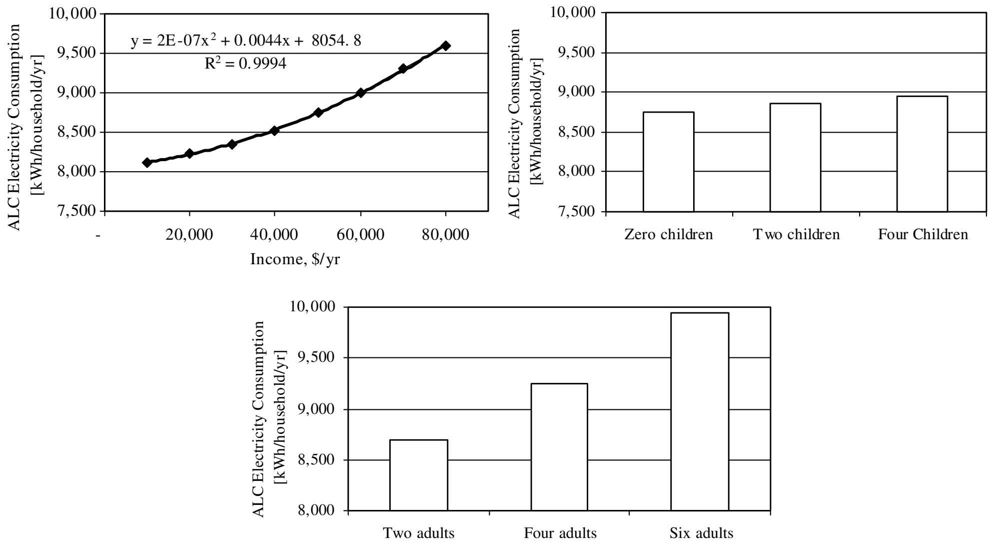

Fig. 4. Effects of socio-economic factors on the appliance, lighting, and cooling (ALC) energy consumption of the households estimated by the NN model.

10.2. Number of children and adults

ALC electricity consumption of a household increases as the number of children and adults increases. The increase in the number of adults has a more significant influence on the ALC energy consumption than the increase due to the number of children. As the number of adults increases, the number of TVs, VCRs, stereos, and computers increase; also, the number of bedrooms increases, increasing the living area. However,

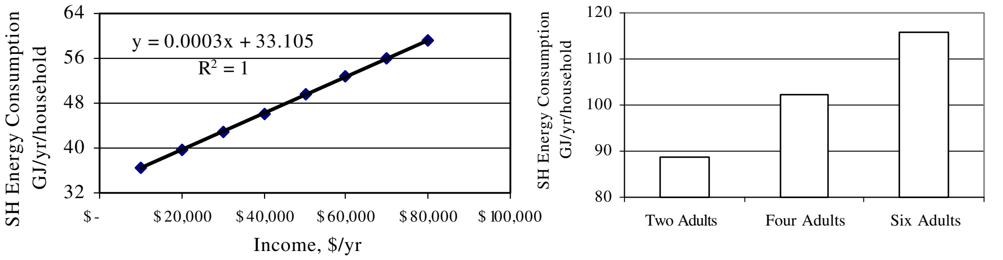

Fig. 5. Effect of income on electrical space heating energy consumption and effect of number of adults on natural gas space heating energy consumption estimated by the CDA model.

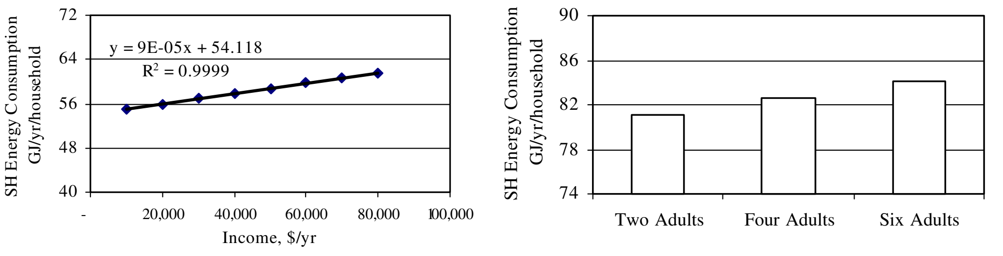

Fig. 6. Effect of income on electrical space heating energy consumption and effect of number of adults on natural gas space heating energy consumption estimated by the space heating NN model.

the increase in the number of children does not generally increase the number of appliances and bedrooms. Therefore, the effect of an increase in the number of adults on the ALC energy consumption is more significant than the increase due to the increase in the number of children. The CDA model estimates show a substantially higher increase in the ALC electricity consumption as the number of children increases as shown in Fig. 3, whereas the estimated effect of the increase in the number of adults is of similar magnitude for both models (as seen in Figs. 3 and 4).

As shown in Figs. 5 and 6, as the number of adults increases, the natural gas consumption for space heating increases for both CDA natural gas and NN models. Compared to NN model estimates, the CDA natural gas model estimates show a substantially higher increase in the natural gas consumption for space heating as the number of adults increases.

10.3. Dwelling type

The average DHW heating energy consumption of a single detached dwelling is higher than that of a single attached dwelling. This is due to the fact that in the 1993 SHEU database as the number of occupants increases, the living area of the dwellings increases. Single detached dwellings have larger living areas than single attached ones [10]. Consequently, the number of occupants is higher in single detached dwellings than in single attached ones, so is the DHW heating energy consumption.

Natural gas consumption of single attached and single detached dwellings for DHW heating were estimated using the CDA natural gas model to be 11.32 GJ/yr/household and 38.18 GJ/yr/household. This corresponds to a difference of 27 GJ/yr/household, which is unacceptably high since the average natural gas consumption for DHW heating was estimated to be 33 GJ/yr/household using the CDA natural gas model [10], and the dwelling type can not have such a large impact on DHW heating energy consumption. On the other hand, the difference in the DHW heating natural gas consumption of the single attached and single detached dwellings was estimated using the NN model to be 0.6 GJ/yr/household.

11. Conclusion

This paper investigates the use of CDA method for modeling residential end-use energy consumption at the national level. In this work, end-use energy consumption models were developed for the Canadian residential sector using the CDA method, and the extensive data available in the 1993 SHEU database of Statistics Canada [26]. The CDA method had not been used to model residential energy consumption at the national level, although there are several studies where CDA was used to model energy consumption at the regional level.

The CDA model developed in this work were compared with an engineering model developed earlier by Farahbakhsh [2] and Farahbakhsh et al. [3,4] and a NN model developed by Aydinalp [10] and Aydinalp et al. [11,12]. Both these models were also developed using the data from the 1993 SHEU database. The predictions of the CDA model were compared with those of the NN and engineering models to assess the comparative accuracy, as well as the versatility of the three models.

The comparison of the predictions of the models indicated that all three models are capable of accurately predicting the energy consumption in the residential sector. The household and end-use energy consumption estimates of the three models were found to be close to each other, and also to the estimates reported by Office of Energy Efficiency [1].

Although theoretically possible, the CDA model is unable to evaluate the effects of some of these socio-economic factors (dwelling ownership and size of area of residence), since the number of variables included in the CDA model is limited due to statistical considerations. On the other hand, the NN model is able to evaluate the effects of several socio-economic factors on end-use energy consumption. These include household income, dwelling type and ownership, number of children and adults, and size of area of residence. None of the socio-economic factors could be evaluated by the engineering model, since due to insufficient data on socio-economic factors in terms of occupancy and preference profiles in the 1993 SHEU database socio-economic variables were not included within the model structure of the existing engineering model. Thus from the perspective of assessing the impact of socio-economic factors, the NN model is superior to both the CDA and the engineering models.

The CDA model, while simpler to apply than the NN model and acceptably accurate in its overall predictions, is not flexible in evaluating end-uses, socio-economic factors and energy saving scenarios. Thus, the CDA model has limited utility for modeling the energy consumption in the residential sector. The fundamental difference between the NN and CDA models is that the NN model is less transparent to demonstrate the marginal effects of factors such as having or not having a particular appliance, DHW or space heating equipment. In comparison, the engineering model provides accurate estimates, has the highest level of flexibility in evaluating the impact of energy saving measures, but has difficulties with the inclusion of the socio-economic factors. The three models have their specific advantages and disadvantages. It may be possible to develop a hybrid model that uses the engineering model for physical and thermodynamic modeling and the CDA and NN models for modeling of socio-economic factors.

Acknowledgements

The authors gratefully acknowledge the financial support provided by Natural Resources Canada and Natural Sciences and Engineering Research Council of Canada (Research Grant 41739-95 RGPIN).

Appendix

See Table A.1 and Figs. A.1–A.3.

Table A.1
Measures used to judge the performance of networks

Name	Symbol	Formula
Multiple correlation coefficient	R^2	$1-\frac{\sum_{i=1}^{N}(y_i-t_i)^2}{\sum_{i=1}^{N}t_i^2}$
Coefficient of variation	CV	$\frac{\sqrt{\frac{\sum_{i=1}^{N}(y_i-t_i)^2}{N}}}{\bar{t}}\times 100$

y_i is the predicted value of the ith pattern; t_i is target value of the ith pattern N is number of patterns; and $\bar{t}$ is mean of the target values.

Dep Var: HEC N: 2039 Multiple R: 0.813 Squared multiple R: 0.660

Adjusted squared multiple R: 0.657 Standard error of estimate: 5533.031

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	2128.647	508.485	0.000	.	4.186	0.000
HDD1	1.595	0.177	0.392	0.089	9.036	0.000
HRV1	-2516.011	1077.603	-0.032	0.922	-2.335	0.020
DTYPE1	1891.746	722.629	0.085	0.160	2.618	0.009
AREA1	12.666	5.341	0.077	0.160	2.372	0.018
AGECAT1	-785.396	173.621	-0.133	0.196	-4.524	0.000
INCOME1	78.142	12.795	0.168	0.223	6.107	0.000
AREA2	8.126	2.473	0.051	0.694	3.286	0.001
CHILD2	569.330	248.449	0.037	0.659	2.292	0.022
TANK3	16.860	2.842	0.179	0.184	5.932	0.000
LOWFLOW3	-691.340	305.421	-0.034	0.732	-2.264	0.024
DWLOAD3	215.042	69.005	0.050	0.656	3.116	0.002
ADULT3	752.825	199.481	0.105	0.218	3.774	0.000
CACUSE4	1.274	0.737	0.024	0.887	1.728	0.084
FROSTR1	1030.230	357.785	0.038	0.953	2.879	0.004
FROSTR2	1636.477	579.324	0.043	0.714	2.825	0.005
INCOME7	28.235	7.695	0.060	0.634	3.669	0.000
VOLF1	1.495	0.611	0.033	0.928	2.447	0.014
HHSIZE10	421.249	106.811	0.065	0.620	3.944	0.000
CDLOAD	303.606	40.751	0.113	0.738	7.450	0.000
LIGHTS	50.175	8.876	0.088	0.694	5.653	0.000

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	1.20051E+11	20	6.00253E+09	196.069	0.000
Residual	6.17799E+10	2018	3.06144E+07		

Fig. A.1. Regression analysis of the CDA electricity model.

Dep Var: HEC N: 1003 Multiple R: 0.961 Squared multiple R: 0.923

Adjusted squared multiple R: 0.923 Standard error of estimate: 984.001

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
PROGT1	-207.752	79.078	-0.026	0.778	-2.627	0.009
DOOR1	129.839	30.318	0.097	0.150	4.283	0.000
WINDOW1	26.302	5.505	0.102	0.168	4.778	0.000
SHAGE1	46.033	4.063	0.195	0.261	11.330	0.000
GARAGE1	437.688	111.559	0.037	0.884	3.923	0.000
AREA1	6.614	0.743	0.238	0.108	8.903	0.000
ADULT1	181.605	40.264	0.121	0.108	4.510	0.000
SYSAGE3	13.209	6.088	0.033	0.326	2.170	0.030
HHSIZE3	70.774	26.693	0.062	0.141	2.651	0.008
DTYPE3	720.967	85.233	0.184	0.164	8.459	0.000
HHSIZE4	94.039	53.298	0.016	0.919	1.764	0.078
CDLOAD5	35.910	21.770	0.015	0.903	1.650	0.099
POOL	969.555	377.900	0.023	0.969	2.566	0.010

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	1.15628E+10	13	8.89443E+08	918.600	0.000
Residual	9.58576E+08	990	968258.515		

Fig. A.2. Regression analysis of CDA natural gas model.

Dep Var: HEC N: 231 Multiple R: 0.935 Squared multiple R: 0.875

Adjusted squared multiple R: 0.872 Standard error of estimate: 1095.544

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
SHAGE1	40.046	8.384	0.206	0.300	4.777	0.000
AREA1	4.949	1.265	0.217	0.181	3.911	0.000
WINDOW1	41.681	10.282	0.205	0.219	4.054	0.000
DTYPE1	471.210	190.320	0.145	0.162	2.476	0.014
TANK3	5.277	0.814	0.246	0.388	6.480	0.000
DWLOAD3	50.907	29.073	0.050	0.682	1.751	0.081

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	1.88179E+09	6	3.13632E+08	261.313	0.000
Residual	2.70049E+08	225	1200217.528		

Fig. A.3. Regression analysis of the CDA oil model.

References

[1] Office of Energy Efficiency. Energy efficiency trends – summary tables. Available from: <http://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/analysis_ca.cfm?attr=0>; April 2006.

[2] Farahbakhsh H. Residential energy and carbon dioxide emissions model for Canada. Masters Thesis, Technical University of Nova Scotia, Canada; 1997.

[3] Farahbakhsh H, Fung AS, Ugursal VI. Space heating thermal requirements and unit energy consumption of Canadian homes in 1993. Canadian Residential Energy End-use Data and Analysis Centre, Halifax, NS; 1997. Available from: <http://www.dal.ca/~creedac/index_high.html>.

[4] Farahbakhsh H, Ugursal VI, Fung AS. A residential end-use energy consumption model for Canada. Int J Energy Res 1998;22(13):1133–43.

[5] Kreider JF, Wang XA. Improved artificial neural networks for commercial Building energy use prediction. Solar Eng ASM 1992;1:361–6.

[6] Anstett M, Kreider JF. Application of neural networking models to predict energy use. ASHRAE Trans 1993;99(1):505–17.

[7] Kawashima M. Artificial neural network backpropagation model with three-phase annealing developed for the building energy predictor shootout. ASHRAE Trans 1994;100(2):1095–103.

[8] Cohen DA, Krarti M. A neural network modeling approach applied to energy conservation retrofits. In: Proceedings of the building simulation fourth international conference; 1995. p. 423–30.

[9] Krarti M, Kreider JF, Cohen D, Curtiss P. Estimation of energy saving for building retrofits using neural networks. J Solar Energ Eng 1998;120:211–6.

[10] Aydinalp M. A new approach for modeling of residential energy consumption. Ph.D. Thesis. Dalhousie University, Halifax, NS, Canada; 2002.

[11] Aydinalp M, Ugursal VI, Fung AS. Modeling of the appliance, lighting, and space cooling energy consumption in the residential sector using neural networks. Appl Energ 2002;71(2):87–110.

[12] Aydinalp M, Ugursal VI, Fung AS. Modeling of the space and domestic hot water energy consumption in the residential sector using neural networks. Appl Energ 2004;79(2):159–78.

[13] Parti M, Parti C. The total and appliance-specific conditional demand for electricity in the household sector. Bell J Econ 1980;11:309–21.

[14] Aigner DJ, Sorooshian C, Kerwin P. Conditional demand analysis for estimating residential end-use load profiles. Energ J 1984;5(3):81–97.

[15] Fiebig DG, Bartels R, Aigner DJ. A random coefficient approach to the estimation of residential end-use load profiles. J Econometrics 1991;50:297–327.

[16] Kellas C. Presentation of work about conditional demand analysis in Manitoba 1993. Halifax: Canadian Electrical Association Meeting; 1993.

[17] Lafrance G, Perron D. Evolution of residential electricity demand by end-use in Quebec 1979–1989: a conditional demand analysis. Energ Stud Rev 1994;6(2):164–73.

[18] Hsiao C, Mountain DC, Illman KH. Bayesian integration of end-use metering and conditional demand analysis. J Bus Econ Stat 1995;13(3):315–26.

[19] Aydinalp M, Ugursal VI, Fung AS. Modeling of residential energy consumption at the national level. In: Proceedings of the 12th international symposium on transport phenomena, Istanbul, Turkey; July 2000.

[20] Swamy PA. Efficient inference in a random coefficient regression model. Econometrica 1970;38(2):311–23.

[21] Bauwens L, Fiebig D, Steel M. Estimating end-use demand: a bayesian approach. J Bus Econ Stat 1994;12(2):221–31.

[22] Zeller A. An introduction to bayesian inference in econometrics. New York (NY, USA): John Wiley & Sons, Inc.; 1971.

[23] Caves DW, Herriges JA, Train KE, Windle RJ. A bayesian approach to combining conditional demand and engineering models of electric usage. Rev Econ Stat 1987;69(3):438–48.

[24] EPRI. Residential end-use energy consumption: a survey of conditional demand analysis. Report No CU-6487. Palo Alto, CA; 1989.

[25] SYSTAT. SYSTAT- statistics. Chicago (IL, USA): SPSS Inc.; 1998.

[26] Statistics Canada, ‘‘Microdata user's guide’’, the survey of household energy use, Ottawa, Ontario, Canada; 1993.

[27] Environment Canada. The Canadian climate and water information and data site; July, 1999. Available from: http://www.msc-smc.ec.gc.ca/cmc/index_e.html.