Applied Energy 212 (2018) 372–385

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks

Aowabin Rahman[a], Vivek Srikumar[b], Amanda D. Smith[a,*]

[a] Site-Specific Energy Systems Laboratory, Department of Mechanical Engineering, University of Utah, Salt Lake City, UT 84112, USA
[b] School of Computing, University of Utah, Salt Lake City, UT 84112, USA

HIGHLIGHTS

- Two deep recurrent neural network (RNN) models are proposed for electricity forecasting.
- The models are applied to electricity forecasting over medium-to-long term time horizon.
- The models are also used to develop a missing data imputation scheme.
- The electricity consumption forecasts for commercial and residential buildings are analyzed in detail.
- The deep RNN models, in general, perform better than 3-layered perceptron neural network models.

ARTICLE INFO

Keywords:
Building energy modeling
Machine learning
Recurrent neural networks
Deep learning
Electric load prediction

ABSTRACT

This paper presents a recurrent neural network model to make medium-to-long term predictions, i.e. time horizon of $\geqslant 1$ week, of electricity consumption profiles in commercial and residential buildings at one-hour resolution. Residential and commercial buildings are responsible for a significant fraction of the overall energy consumption in the U.S. With advances in sensors and smart technologies, there is a need for medium to long-term prediction of electricity consumption in residential and commercial buildings at hourly intervals to support decision making pertaining to operations, demand response strategies, and installation of distributed generation systems. The modeler may have limited access to information about building's schedules and equipment, making data-driven machine learning models attractive. The energy consumption data that is available may also contain blocks of missing data, making time-series predictions difficult. Thus, the main objectives of this paper are: (a) Develop and optimize novel deep recurrent neural network (RNN) models aimed at medium to long term electric load prediction at one-hour resolution; (b) Analyze the relative performance of the model for different types of electricity consumption patterns; and (c) Use the deep NN to perform imputation on an electricity consumption dataset containing segments of missing values. The proposed models were used to predict hourly electricity consumption for the Public Safety Building in Salt Lake City, Utah, and for aggregated hourly electricity consumption in residential buildings in Austin, Texas. For predicting the commercial building's load profiles, the proposed RNN sequence-to-sequence models generally correspond to lower relative error when compared with the conventional multi-layered perceptron neural network. For predicting aggregate electricity consumption in residential buildings, the proposed model generally does not provide gains in accuracy compared to the multi-layered perceptron model.

1. Introduction

Residential and commercial buildings in 2015 were responsible for approximately 73% of the electricity consumption and 41% of primary energy consumption in the U.S, with the values projected to increase over the next 20 years [1]. There has been a growing emphasis on the development and implementation of smart grids and smart buildings in order to meet these electricity demands in an efficient and cost-effective manner while minimizing greenhouse emissions [2,3]. The case for smart grids are further strengthened by the increasing intermittent, renewable energy resources such as wind and solar, as well as a growing number of small-scale distributed generation systems [2,4].

As such, dynamic planning and management of smart buildings and smart grid systems, while integrating intermittent renewables and

* Corresponding author.
E-mail address: amanda.d.smith@utah.edu (A.D. Smith).

https://doi.org/10.1016/j.apenergy.2017.12.051
Received 10 October 2017; Received in revised form 6 December 2017; Accepted 8 December 2017

Nomenclature

DL	deep learning
DNN	deep neural network
EI	expected improvement
LSTM	long short term memory
ML	machine learning
MLP	multi-layered perceptron
NN	neural network
PSB	Public Safety Building
RNN	recurrent neural network
SMBO	sequential-model based optimization
TPE	tree of parzen estimator
γ	learning rate in gradient descent algorithm
σ	sigmoid function serving as a gating function
$\circ$	element-wise vector multiplier
τ_i	characteristic timescales in a periodic energy consumption profile
λ	parameter for weight regularization
μ	number of training epochs, i.e. number of runs for which the model is trained using the entire training set.
θ	set of hyper-parameters
θ^*	optimal set of hyper-parameters
$\mathbf{c_t}$	transient 'memory' value in LSTM function
e	mean squared error in predicting electricity consumption
$\mathbf{f}$	frequency-related variables used as inputs to the deep RNN model
$\mathbf{g}$	input activation function in LSTM
$\mathbf{h_t}$	output of LSTM function at given timestep t
h_j^m	value of hidden node in a neural network in node j, layer m
DNN	deep neural network
$\mathbf{i}$	input gate in LSTM
n	number if residential buildings considered when aggregate electricity consumption profile in residential buildings
$\mathbf{o}$	output gate in LSTM
p	fraction of consecutive missing data points relative to the size of the entire training set
q_1	fraction of data points prior to the missing block p, relative to the size of the entire training set
q_2	fraction of data points after the missing block p, relative to the size of the entire training set
s	parameter to describe discrepancy between electricity consumption in test data and that in the corresponding training data
w_{ji}	weight connecting j in layer m to node i in layer $m{-}1$
w_{ji}	weight connecting j in layer m to node i in layer $m{-}1$
$\mathbf{w}$	weather variables used as inputs to the deep RNN model
$\mathbf{x_t}$	input to LSTM activation corresponding to a previous layer and current timestep t
$\mathbf{X}$	feature vector used as inputs to the deep RNN model
y_p	predicted value of electricity consumption
y_a	actual value of electricity consumption
$\mathbf{s}$	date-related variables used as inputs to the deep RNN model
w_{ji}	weight connecting j in layer m to node i in layer $m{-}1$
$\mathbf{w}$	weather variables used as inputs to the deep RNN model
$\mathbf{x_t}$	input to LSTM activation corresponding to a previous layer and current timestep t
$\mathbf{X}$	feature vector used as inputs to the deep RNN model
y_p	predicted value of electricity consumption
y_a	actual value of electricity consumption

distributed generation resources, requires accurate forecasting of electricity consumption over different time horizons [2]. Based on the time horizon of prediction, Mocanu et al. [5] grouped electricity demand forecasting into three categories: (i) Short-term forecasts ranging between one hour to one week, (ii) medium term forecasts between one week to one year and (iii) long-term forecasts spanning a time period of more than one year. Short-term forecasts are generally useful for generation capacity scheduling and short-term maintenance, evaluation of short-term energy storage usage, as well as real-time control of building energy systems and optimization of fuel purchase plans [5–8]. On the other hand, medium to long term forecasts are used to make decisions pertaining to the installation of new distributed generation and storage systems [9], as well as develop suitable demand response strategies [5]. At a regional level, forecasting of aggregated electricity consumption over medium-to-long term time horizons can be useful for planning and trading on electricity markets [10].

The approaches to estimating electricity demand in buildings can be physics-based or data-driven [11,12]. Physics-based or deterministic models, such as those employed by EnergyPlus and eQuest, usually formulate and solve heat and mass balance equations interconnecting the different zones, air handling and equipment systems inside a building [13]. However, these physics-based models often do not account for the complex energy consumption behavior in a building, and sometimes input parameters required by these models, are difficult to obtain in practice [11]. The resulting approximations often lead to a loss in accuracy, sometimes in excess of 100% [11,14], and as such, these models are often used as comparative tools rather than accurate predictors of building energy consumption.

Statistical and machine learning (ML) models provide an alternative to such physics-based models [11,12]. Previous work has employed simple linear regression [12,15,16], multi-variate linear regression [15], non-linear regression [15,12], support vector machines [11,12,17], Gaussian Process regression [18], multi-layered perceptron neural networks [19,12,11,16,17], and auto-regressive neural networks [11,20] in predicting building energy consumption. Hybrid models that couple physical models, i.e. thermal networks, with statistical and/or ML models have also been proposed [21]. These methods, in general, have been shown to achieve high accuracy for forecasting over a time horizon of one hour [20] to one week [22], the amount of work pertaining to medium to long term predictions at hourly or sub-hourly intervals has been relatively limited. The latter is a more difficult objective, with previous work showing that the relative errors corresponding to medium to long term predictions at one-hour resolution often in excess of 40–50% [5,15,23].

Deep neural networks [24] could potentially improve on the performances obtained using the aforementioned machine learning methods, as they allow for modeling of more complex functions by using multiple layers of abstraction [24], and are recently being employed in the energy forecasting context. Debinec et al. [7] used a deep belief network for electricity forecasting in Macedonia over a time horizon of 24 h, which consisted of stacks of restricted Boltzmann machines (RBMs) pre-trained layer-wise. Mocanu et al. [5] employed a conditional restricted Boltzmann machine (CRBM) and factored conditional Boltzmann machine (FCRBM) to predict electricity power consumption in a residential building. The two deep learning (DL) methods were used to obtained results for multiple cases, each case corresponding to a combination of time resolution and time horizon. The authors found that for a week-ahead prediction at one-hour resolution, the relative errors in predicting aggregate power corresponding to CRBM and FCRBM were 60.0% and 63.3%, whereas for a year-ahead prediction at one-day resolution, the corresponding errors were 18.2% and 17.0% [5].

The electricity consumption behavior is inherently transient in nature, and the consumption pattern (as detailed later in the paper) can