Hourly Prediction of Building Energy Consumption: An Incremental ANN Approach

Aulon Shabani, Orion Zavalani

Abstract—Rapid growth of buildings energy consumption puts the focus to improve energy efficiency by building engineers and operators. Energy management through forecasting approaches using machine-learning algorithms is an increasing research domain. Most of algorithms focus on predicting energy consumption when a considerable amount of past-observed data exists. In this paper, we focus on the case when small amount of available data exists and the amount of data increases incrementally by time. Artificial Neural Networks used as the learning algorithm take as the training data mini batches of different sizes. Algorithm is evaluated on different batch sizes and compared to baseline learner.

Index Terms—Energy Prediction; Incremental Learning; Neural Networks; Energy Efficiency.

I. INTRODUCTION

Concerns of energy security and a rapid destruction of natural resources have caused energy solutions to gain worldwide attention. Statistics provided by International Energy Agency (IEA) show an increase in energy consumption especially in residential and commercial building sector [1]. Increase in energy consumption leads to a dis-balance between production and demand, and load increasing during peak hours.

Also in Albania, residential, public and commercial buildings occupy 76% of the total electricity consumption of the country as shown in Fig.1. Zavalani and Luga underline in [2] that public facilities in Albania result with great electrical energy losses and do not fulfill the National Energy Building Code.

Prediction of building energy consumption plays an important role in building management through advanced control strategies and building operation control [3]. Existence of prior information helps building management staff to take in advance measures reducing consumption and helping minimizing peak loads. One of the most used approaches intending to improve energy efficiency are machine-learning approaches. Fast development of smart metering devices is the key feature that pushes the application of artificial intelligence approaches to predict energy consumption based on the necessity of past historical measurements.

In literature, many statistical and machine learning models used are: time series models, regression models, artificial neural networks (ANN), support vector regression (SVR) and other relevant methods [4]-[14], [15]-[20], [21]-[34]. In our focus is ANN as one of the most used approaches in energy prediction. The methods mainly focus on batch learning (static), when the model is trained once with whole training dataset and then tested. The static approach is restricted to a certain amount of data, which usually are not available for every building and other drawback is that model does not use new observed measurements when they become available. Starting at that point we try to implement an online ANN approach to predict energy consumption in buildings.

Fig.1. USA electricity consumption by sector [1] left figure, Albania electricity consumption right.

Next section follows with description of existing approaches used to predict energy consumption, followed by description of neural network framework. In section IV experimental setup is described, concluding with final section.

II. EXISTING APPROACHES TO PREDICT BUILDING ENERGY CONSUMPTION

We devote this section to describe traditional machine learning (ML) and recent methods used to predict building energy consumption. Machine learning methods fit model or pattern, from given input examples to desired targets, and the learned rule predicts new outputs for given new (unseen) examples. There are various ML algorithms addressing problem of energy prediction in buildings.

Energy consumption has a complex nonlinear behavior, strongly related to outdoor climatic factors and since ANNs have a strong ability dealing with nonlinear problems, make them the most used approach to predict energy demand in buildings. ANNs applications consists in analysis of building energetic behavior, predicting building energy consumption, predicting HVAC systems and sub-level components energy.
consumption. Following, we shall briefly describe application framework of neural networks to building energy consumption.


Likewise, Yokoyama in [7] use a back propagation ANN to predict cooling energy demand. Preventing drawbacks of local optimization methods, they implement a global optimization method called “Modal Trimming” to optimize ANN parameters. Cooling demand of buildings is also in focus of Ben-Nakhi and Mahmoud in [8], where they use a general regression ANN. The network performance was evaluated with data generated from building simulation software ESP-r, where three-year of 24-hour measurements were used as training and one year was used as testing. Back propagation ANNs proposed by Cheng-wen and Ying [9] to predict heating and cooling demand use building envelope performance in combination with heating degree day HDD and cooling degree day CDD as training input parameters. They emphasize the importance of training such models to predict energy consumption in early design stages, and achieve accurate results.

Gonzales and Zamarrenro [10] predict hourly electricity consumption in buildings using a recurrent neural network. Simple model implemented using real observed data achieves significantly accurate results, where forecasted air temperature and current load are the input attributes. Li [11] compares traditional ANNs with hybrid genetic algorithm-adaptive network-based fuzzy inference system (GA-ANFIS), where GA is used to find best parameters and the ANFIS adjust the parameters to match training data. Predicted energy consumption estimated by GA-ANFIS shows higher accuracy in terms of coefficient of variation (CV). Olofsson [12] uses neural networks to predict space heating demand when limited number of performance parameters are available. They propose a quasi-physical approach to describe building performance based on outdoor and indoor temperatures and other simple parameters.

A study how statistical methods can improve ANNs performance in hourly energy prediction is in focus of Karatasou [13], also they implement on real data combination of ANN with statistical approaches to gain highly desirable results. Neural networks are one of the most used approaches applied to predict energy consumption in buildings, as matter of fact Kalogirou [14] gives a brief review application of ANN methods with regard to building energy consumption.

Other machine learning approaches are used to predict energy consumption. We can mention support vector regression (SVR) approach as one of the most successful [15]-[18].

Likewise, recent methods as regression trees [19],[20]and case based reasoning (CBR) [21],[22] have started to grow attention. A brief review of most machine learning methods and other methods that predict building energy consumption reader can referred to [23] and [24].

The above-mentioned works focus on static energy prediction; they train once the model and then test it. In our knowledge there exist very few approaches that predict online energy prediction. For example, Yang [25] use adaptive and incremental learning approaches to predict building energy consumption. They achieve very good results using synthetic data, and less accurate when real observed data are used as model inputs. Other published work refers to usage of CBR model to predict energy consumption training the model incrementally. Platon [22] and Shabani [21] predict building energy consumption using CBR approach in case when new data are available. Results achieved using coefficient of variation of root mean squared error (CV-RMSE) show an improved model performance due to increasing amount of data.

III. ARTIFICIAL NEURAL NETWORK MODEL USED TO PREDICT ENERGY CONSUMPTION

In this section, we present ANN model proposed by us to predict building energy consumption.

A. Incremental learning

An incremental approach to predict energy consumption is used. The method arises from absence of enough data or when new data become available. We call this approach mini batch learning. It helps the model to be retrained and adapts seasonal weather conditions and daily trends. This training technique takes training data in portions of small batches and new data by throwing old data. Model learns incrementally at every new data portion by retraining.

Disadvantage of this model is that algorithm cannot follow seasonal trends and for this reason, additionally size of batch is very important. Selecting batch size is a crucial step to improve model performance. Small batch sizes remove time series trends or weather changes and large batch sizes need a considerable amount of data and the throws down the scope of method intending to predict online.

As shown in Fig. 2 sliding window approach learns by taking at every step a new batch and then after training the
network on this particular batch, predicts electricity consumption for the next hour.

In the next subsection we shall present the general ANN scheme on regression problems.

B. Artificial neural network general learning scheme

Attempts to use artificial neural networks for time series start at 1970’s, but the results were not so accurate and in that time some of the people interested in those fields abandoned the research [26]. Discovery of back propagation neural networks and successful application in time series forecasting was very impressive and nowadays come to play the deep learning which makes neural networks very competitive in time series forecasting. In this paper, we use as model for prediction a back propagation neural network with Levenberg-Marquardt algorithm which is a modification of Gauss-Newton method [20].

A single layer ANN takes a set of observed inputs $a = (a_1, a_2, ..., a_n)$ from input layer, when vector of $w = (w_1, w_2, ..., w_n)$ are the inputs associated weights. Weighted sum of multiplication (input, weights) form the pre-activation function $z$ as in (1). Term $b$ in (1) is called bias and most of the time has a value equal to one.

$$z = b + \sum_{i=1}^{N} a_i w_i$$

(1)

Network transforms pre-activation function using an activation function, which can be linear, or usually it is a nonlinear function (sigmoid or hyperbolic).

$$a_{out} = f(z) = \frac{1}{1 + e^{-z}}$$

(2)

In (2) is shown the output of hidden layer neurons using a sigmoidal activation function. ANN model shown in Fig.3 describes a simple network model with many inputs and single outputs. Network inputs are defined by the number of model attributes and the output is the number of time steps interested to predict in case of time series regression. In our case the output is next hour energy consumption. Fig.3. describes a general scheme of neural network. It is very important before using a neural network to optimize its architecture like number of hidden layers and number of neurons in hidden layer.

IV. COMPUTATIONAL EXPERIMENTS

This section analyses the ANN setup used for empirically evaluation online hourly energy prediction for publicly available datasets.

A. Data preprocessing

Energy consumption data (publicly available) are provided from a commercial building. The data are measured from 1st of January 2010 until 1st January 2011. Observed information within a frequency of 15 min contains power consumption and air temperature values. Same dataset is used to predict hourly building energy consumption using the ANN model which is presented by authors in [26].

Our interest consists in predicting hourly energy consumption. For this reason; we simply convert from 15 min time steps to hourly values by just computing the mean value of measurements within the hour.

Another problem when dealing with real measured data is that they may contain outliers like zero values or missing values. Solution to this issue is done using a simple approach by visual inspection and in case when few missing values exist, they are substituted by the mean value of two neighbor points. In contrary when more than four continuous missing points exists they are removed from dataset.

Building operation time is from 8:00 until 23:00 as shown in Fig.4. We have select 1440 working hours (four months period) to use as experimental dataset.

Attributes that represent the problem are outside air temperature and power consumption, an extraction of additional information taken from data is applied.

$$X(t) = \{t, T_1, ID_1, ..., ID_7, P_{-24}, P_{-168}, P_{-3}, P_{-2}, P_{-1}, \bar{P}_{24}\}$$

(3)

Where $X$ is the input sample represented as: $T_1$ temperature at time $t$, $ID_1$, ..., $ID_7$ are dummy variables indicating the day of the week coded as (Mon-0000001, to Sun-1000000), $P_{-24}$ power consumption at same hour one day before, $P_{-168}$ power consumption at same hour a week before and $\bar{P}_{t-24}$ average power consumption the day before prediction takes place. Other time lagged energy consumption measurements respectively measured for three($P_{-3}$), two ($P_{-2}$) and an hour ($P_{-1}$) before prediction starts.

Fig. 3. Principal schematics of neural network.

Fig. 4. Building energy consumption dataset.

1 http://en.openei.org/datasets
B. ANN experimental setup

As mentioned above, the point of interest in this paper is to predict hourly building energy consumption, which is a function of measured outside air temperature and other relevant information added by expert knowledge. Other information included in the dataset as described in previous section contains temporal variable like time of the day, time lagging energy consumption and a dummy variable explaining day of the week and automatically the holiday or working day.

Those variables are the inputs for the ANN model used to predict energy demand. In total number of outputs consists of eighteen input variables and one output variable. Therefore, the network consists of input layer with 18 neurons, one neuron output layer and a hidden layer selected. The architecture with one hidden layer was selected as the most used and efficient by literature. Number of hidden neurons was set to 10.

ANN was configured, trained, validated and trained using MATLAB Neural Network toolbox. This toolbox offers a variety of training algorithms and different network architectures. As one of the most used and efficient solving nonlinear problems on small datasets Levenberg-Marquart (LM) algorithm was selected as the appropriate one. Backpropagation neural network using LM algorithm was trained and validated using datasets of different bath sizes. Therefore, for every mini batch the network is trained and then predicts the next hour energy consumption. Batches move in direction of time as new data become available, as new data point is added to the database the network takes the batch with same size as previous batch containing new measurement and then predicts for the next hour. This procedure is repeated until whole dataset taken in consideration is covered.

V. EXPERIMENTAL RESULTS

Conducting experiments for sliding window training of neural network, the architecture used as described in above section. Backpropagation neural network using as training algorithm Levenberg-Marquart and nonlinear sigmoid activation function of ten hidden neurons in hidden layer is the network used in our experiments. Proposed network is compared to one of the most competitive baseline learners, known in literature as target mean, which predicts the mean value of energy consumption for current batch dataset.

The results are evaluated using cumulative value of coefficient of variation of root mean squared error (CV-RMSE).

\[
CV - RMSE = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{P_i - \bar{P}}{P_i} \right)^2} 
\]
In (4) $P_\mu$ is the predicted power value and $P_i$ the observed power value.

Conducted experiments take different batch sizes starting from batches with length of 6 hours corresponding to half working day until to 60 hours corresponding to 5 working days.

Results shown in Fig.6-Fig.10 describe comparison between online ANN approach to target mean, where outperformance of ANN is more obvious in case of small batch sizes and slightly better in case of large batch sizes.

VI. CONCLUSION

The surveyed literature mostly focuses predicting building energy consumption by using static machine learning models. We modeled and tested an ANN dynamic prediction model. Our modelling approach compared to a baseline model shows competitive results, but there exist many points to be improved.

One of the major issues this approach faces consists in defining the appropriate size of batch that ANN learns. Despite drawbacks, the method performs really well in terms of online prediction, which is fundamental practical issue in application domain. Selected small size batches perform really well compared to bigger ones like taking daily energy consumption (batch size 12) in Fig.11. Moreover the method shows to converge due to increasing amount of data available.

As well, the model needs to be compared with other machine learning approaches, which is a future goal.

REFERENCES


