ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

# Temperature Forecast in Buildings Using Machine Learning Techniques

Fernando Mateo[1], Juan J. Carrasco[1], Mónica Millán-Giraldo[1 2], Abderrahim Sellami[1], Pablo Escandell-Montero[1], José M. Martínez-Martínez[1] and Emilio Soria-Olivas[1 *]

1- University of Valencia - Intelligent Data Analysis Laboratory
Avda Universitat S/N, 46100 Burjassot - Valencia, Spain

2- University Jaume I - Institute of New Imaging Technologies
Av. Sos Baynat S/N, 12071 - Castelló de la Plana, Spain

**Abstract**. Energy efficiency in buildings requires having good prediction of the variables that define the power consumption in the building. Temperature is the most relevant of these variables because it affects the operation of the cooling systems in summer and the heating systems in winter, while being also the main variable that defines comfort. This paper presents the application of classical methods of time series forecasting, such as Autoregressive (AR), Multiple Linear Regression (MLR) and Robust MLR (RMLR) models, along with others derived from more complex machine learning techniques, including Multilayer Perceptron with Non-linear Autoregressive Exogenous (MLP-NARX) and Extreme Learning Machine (ELM), to forecast temperature in buildings. The results obtained in the temperature prediction of several rooms of a building show the goodness of machine learning methods as compared to traditional approaches.

## 1 Introduction

In the recent years, the interest in energy efficiency in large buildings has grown considerably. This interest is due to several reasons: the price growth of energy sources such as oil, the increasing need to preserve resources by shifting to renewable energy sources, the economic reasons like the increase in overall machine performance and industry productivity and the appeals and mandates imposed by governments to protect ecological systems and to ensure environmental sustainability. One of the most energy-intensive elements is the HVAC (*Heating, Ventilation and Air Conditioning*) system. HVAC control systems should be able to predict the temperature response to changes in the input data (meteorological and environmental conditions, changing seasons, holiday periods, etc.).

In the literature it is possible to find comparisons of linear models with ANNs applied to daily temperature profile forecast [1] and indoor temperature forecast in buildings [2], which conclude that neural networks are more advantageous than linear models. In the present work, some emerging methods like Extreme

*This work was financed by the Spanish *Ministerio de Economía y Competitividad* (MINECO) and FEDER funds in the framework of the project EFIS: *Un Sistema Inteligente Adaptativo para la Gestión Eficiente de la Energía en Grandes Edificios*, with reference IPT-2011-0962-920000. In addition, we would like to thank Grupotec Servicios Avanzados SA for the provided data.

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

Learning Machines (ELMs), or the combination of Multilayer Perceptron (MLP) and Non-linear Autoregressive techniques (NARX) will be introduced and compared with other more traditional models based on linear approximations like Multiple Linear Regression (MLR), Robust Multiple Linear Regression (RMLR) and Autoregressive models (AR).

This paper is organized as follows. Section 2 describes the data employed in the experiments. Section 3 explains the methods that will be assessed for forecasting purposes. In Section 4 we present the most relevant results by analysing and comparing the behaviour of the simulated models. In Section 5 some concluding remarks are sumarized.

## 2 Data Used

The temperature prediction is based on data obtained by simulating a building using the TeKton 3D software[1]. The simulated building, which is used for educational purposes, is located in Málaga (Spain) and has an altitude of 10 meters above sea level. The data that will be used in the experiments contain outdoor climate variables, indoor climate variables and power consumption of four different areas inside the building. Each of these variables is sampled hourly. The variables are *month, day of the month, official time, relative humidity (%), outside temperature, setpoint temperature of the room, total thermal power, current temperature and predicted temperature*. Low and high setpoint temperatures are 22°C and 24°C, respectively. These temperatures are kept constant for the entire year. The environmental temperature is measured inside each room. The variable *total power* represents the value consumed by the air-conditioning appliance during one hour. Power with negative values corresponds to heating mode, whereas positive values indicate the cooling mode. Low and high setpoint temperatures were merged as a single variable because they are set to 1 when they are active and to 0 otherwise. Input variables are standardized which gives as a result a new variable with mean 0 and standard deviation 1.

## 3 Methods

Table 1 shows the correlation matrix among the temperatures in the four rooms, which shows a strong correlation among them. For this reason, we will show only the results obtained for room 1.

The choice of a particular predictive method for a regression problem is not always obvious. In some cases, the simplest methods may not be enough to approximate the intricate relationships between input and output variables, while in others, the most complex methods would probably be an overkill, or will simply not be able to properly tune their parameters to approximate the data [3]. The classical methods that will be evaluated are MLR, RMLR and AR models, while the non-linear methods will be neural models based on ELM and a combination of NARX with MLP. Next, we describe these methods:

[1] `http://www.nemetschek.es/tekton3d/`

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

| Room | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1.0000 | 0.9411 | 0.9861 | 0.9771 |
| 2 | 0.9411 | 1.0000 | 0.8993 | 0.9027 |
| 3 | 0.9861 | 0.8993 | 1.0000 | 0.9914 |
| 4 | 0.9771 | 0.9027 | 0.9914 | 1.0000 |

Table 1: Correlation matrix of temperatures in the four rooms.

- Multiple Linear Regression (MLR) fits a model linear in the model coefficients [3]:

$$y(t) = b_0 + \sum_{i=1}^{p} b_k X_{(t,i)} + \epsilon(t), \quad t = 1, \ldots, n \tag{1}$$

  where $b_i$ are the regression coefficients, $X_{n \times p}$ are the input data, $y_i$ is the output variable and $\epsilon$ is the error vector. The robust version of the regression function (RMLR) uses an iteratively reweighted least-squares algorithm to give lower weight to points that do not fit well. The results are less sensitive to *outliers* [4].

- In the Autoregressive (AR) model, the output is based on a linear combination of its past outputs [5]. The formula that defines an AR process of order $p$ is:

$$y(t) = c + \sum_{i=1}^{p} \varphi_i y(t-i) + \epsilon(t) \tag{2}$$

  where $c$ is a constant, $\varphi_i$ are the parameters of the model and $\epsilon(t)$ is white noise. The AR parameters can be calculated using different methods [6].

- Extreme Learning Machine (ELM) is a Single-hidden Layer Feedforward Network (SLFN) with an optimized learning algorithm [7]. The ELM algorithm has the following stages:

  1. Coefficients of the hidden layer are randomly initialized.
  2. Calculate the optimal output weights using the pseudo-inverse.

  Many types of hidden nodes including additive/RBF hidden nodes, multiplicative nodes and non neural-like nodes, can be used as long as they are piecewise non-linear.

- Non-linear Autoregressive Exogenous models (NARX) are commonly used for the prediction of time series by approximating non-linear relationships among exogenous variables and the variable to predict, according to the equation [8, 9]:

$$\begin{aligned} y(t) = f(&x(t-1), x(t-2) \ldots x(t-d); \\ &y(t-1), y(t-2), \ldots y(t-d)) \end{aligned} \tag{3}$$

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

where $y(t)$ is the variable to predict at and $x(t)$ is an external or exogenous time series. Often, the non-linear function $f$ is simply a polynomial. It is also possible to create a dynamic MLP network to model the NARX function $f$ assuming that, at instant $t$, the $d$ past values of the variable to predict and the exogenous variable are available. This can be implemented in practice by applying a delay of $d$ samples to both variables.

## 4 Results

In the first place, a feature selection method, *RReliefF*, is applied to obtain the relevant inputs to the model [10]. The most relevant variables chosen by the *RReliefF* algorithm were *official time*, *setpoint temperature*, *current temperature*, *outside temperature*. The result of this analysis is used later. The past temperatures of the rooms appear listed as very significant. The official time is also selected in the first positions, except for room 4 (ninth place). The outside temperature appears in a higher position in the ranking for room 2, indicating that this room is probably more affected by outside climatic conditions than the rest.

All the methods were systematically evaluated on the dataset introduced in Section 2 and the Mean Absolute Error (MAE) on the test set was calculated as an average of 20 simulations, to make the results independent of the initial choice of parameters. The models were trained using annual data, with the first 20 days of each month for training and the rest for testing. The individual MAE for each month was obtained for every model, as well as the yearly error average. The number of hidden nodes for the ELMs was evaluated from 100 to 200 in steps of 20 nodes. The best architecture in terms of annual MAE (12 month average) for room 1 was 180 hidden nodes, which produced a MAE of 0.555°C.

Initially, the values of MAE obtained with ELM were disappointing because they were much higher than those obtained with linear models (see Table 2). However, the analysis realized with the *RReliefF* algorithm was used and the performance gain was substantial. The average monthly improvement in MAE with ELM when using variable selection is 67%. However, the fact of performing variable selection hardly influences the linear models' error. For the MLR models, the mean annual error increases in 0.8%, while for the RMLR models the average improvement is only 1.6%.

Regarding the AR, the model order was evaluated from 1 to 48. The order $p = 24$ was selected as the one which produced the lowest Akaike Information Criterion (AIC). Thus, one AR(24) model was built for each month and their coefficients were estimated by the Burg method [6].

For the MLP NARX implementation, separate open-loop NARX models were built for each month, using the most correlated room temperature with the desired one as exogenous variable and a lag of 24 hours. The MLPs for each monthly NARX model were given one hidden layer composed by 20 nodes with sigmoid activation functions. The monthly MAE attained for room 1 with the methods proposed in Section 3 are summarized in Table 2. Additionally, the

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

naïve forecast models $y(t) = y(t-1)$ and $y(t) = y(t-24)$ are added. The rest of the rooms' behaviour is similar so they have been omitted. From the inspection of the results, one can easily deduce that the MLP and NARX combination yields the best results in terms of MAE among all the methods. The MLP NARX predictions for the test set are graphically represented for a winter month (January) and for a summer month (July) in Fig. 1. The linear methods behave mostly alike, but the robust variant usually achieves a slightly better MAE.

ELM did not stand out clearly as an alternative to the much simpler linear methods for this particular problem. However, when features selection is performed using the *RReliefF* algorithm a significant improvement (more than 67%) is achieved with the ELM model. In general, the best errors are obtained for the months of April and October. This is due to the shorter range of temperature variations during these months of the year. As a consequence, these series are easier to predict by all methods.

| | *y(t)=y(t-1)* | *y(t)=y(t-24)* | MLR | RMLR | AR | ELM | ELM *RReliefF* | MLP-NARX |
|---|---|---|---|---|---|---|---|---|
| Jan. | 0.208 | 0.710 | 0.302 | 0.180 | 0.254 | 0.625 | 0.228 | **0.101** |
| Feb. | 0.149 | 0.298 | 0.149 | 0.123 | 0.158 | 0.413 | 0.130 | **0.088** |
| Mar. | 0.149 | 0.530 | 0.152 | 0.132 | 0.165 | 0.500 | 0.151 | **0.100** |
| Apr. | 0.125 | 0.383 | 0.146 | 0.115 | 0.139 | 0.463 | 0.101 | **0.064** |
| May | 0.225 | 0.285 | 0.290 | 0.237 | 0.296 | 0.605 | 0.198 | **0.095** |
| Jun. | 0.286 | 0.537 | 0.396 | 0.292 | 0.325 | 0.561 | 0.210 | **0.078** |
| Jul. | 0.377 | 1.063 | 0.554 | 0.376 | 0.447 | 0.749 | 0.292 | **0.164** |
| Aug. | 0.338 | 0.492 | 0.462 | 0.344 | 0.344 | 0.610 | 0.233 | **0.119** |
| Sept. | 0.228 | 0.562 | 0.304 | 0.241 | 0.316 | 0.536 | 0.212 | **0.143** |
| Oct. | 0.156 | 0.251 | 0.161 | 0.154 | 0.191 | 0.486 | 0.132 | **0.068** |
| Nov. | 0.164 | 0.392 | 0.166 | 0.138 | 0.151 | 0.552 | **0.137** | 0.150 |
| Dec. | 0.171 | 0.648 | 0.209 | **0.147** | 0.197 | 0.559 | 0.175 | 0.222 |
| Mean | 0.215 | 0.513 | 0.274 | 0.207 | 0.249 | 0.555 | 0.183 | **0.116** |

Table 2: MAE obtained in each month for room 1 with the different proposed methods.

## 5 Conclusions

Temperature prediction is crucial for the management of energy efficiency in large buildings. This prediction can be made using different linear and non-linear techniques, several of which have been compared in this paper and tested using a simulated one-year temperature record for a particular building. An exhaustive analysis of the data has been made as a first step before building these models. The results show that linear regression methods produce similar errors, with a slight advantage for robust models. AR models for short term prediction also offer similar performance, or slightly better, than multiple regression.

Non-linear machine learning methods offered mixed impressions. On one hand, the combination of MLP and NARX outperformed all the other methods in terms of mean absolute error. On the other hand, ELM did not perform very well initially. However, performing variable selection using the *RReliefF* ranking method did improve the errors of ELM by more than 67% in average, improving the errors of the linear methods. Furthermore, the results have indicated that the RMLR, ELM *RReliefF* and MLP-NARX models have outperformed the naïve

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

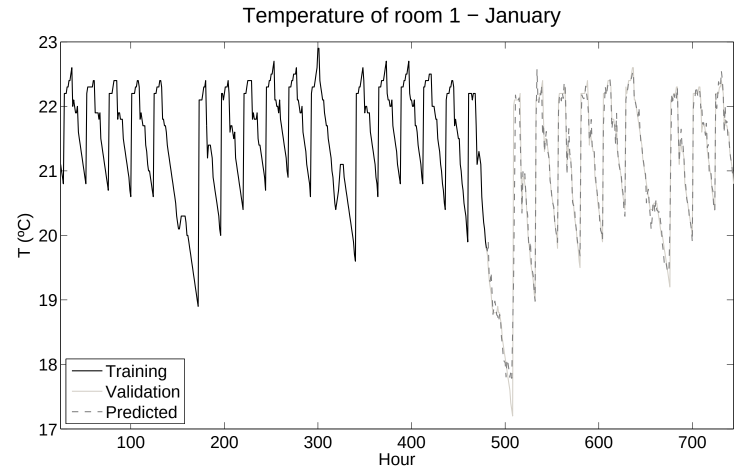


Fig. 1: Prediction of the temperature of room 1 with MLP NARX models for January.

models $y(t) = y(t-1)$ and $y(t) = y(t-24)$. As future work, systematic variable selection methodologies and ensembles of non-linear models will be researched.

## References

[1] H.S. Hippert and C.E. Pedreira. Estimating temperature profiles for short-term load forecasting: neural networks compared to linear models. *Generation, Transmission and Distribution, IEE Proceedings-*, 151(4):543–547, july 2004.

[2] B. Thomas and M. Soleimani-Mohseni. Artificial neural network models for indoor temperature prediction: investigations in two buildings. *Neural Computing & Applications*, 16:81–89, 2007.

[3] S. Weisberg. *Applied Linear Regression*. John Wiley, New York, second edition, 1985.

[4] P. W. Holland and R. E. Welsch. Robust Regression Using Iteratively Reweighted Least-Squares. *Communications in Statistics: Theory and Methods*, A6:813–827, 1977.

[5] Spyros G. Makridakis, Steven C. Wheelwright, and Rob J Hyndman. *Forecasting: Methods and Applications*. John Wiley & Sons, 1998.

[6] S.M. Kay. *Modern Spectral Estimation: Theory and Application*. Prentice Hall, 1988.

[7] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew. Extreme learning machine: A new learning scheme of feedforward networks. *Neurocomputing*, 70:489–501, 2006.

[8] I.J. Lentaritidis and S.A. Billings. Input-output parametric models for non-linear systems. Part I: deterministic non-linear systems. *Int. J. of Control*, 41:303–328, 1985.

[9] I.J. Lentaritidis and S.A. Billings. Input-output parametric models for non-linear systems. Part II: stochastic non-linear systems. *Int. J. of Control*, 41:329–344, 1985.

[10] K. Kira and L. A. Rendell. A practical approach to feature selection. In *Proceedings of the Ninth International Workshop on Machine Learning*, ML '92, pages 249–256, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.