Modelling thermal comfort in residential buildings in Portugal

Marek Utracki
marek.utracki@gmail.com

Instituto Superior Técnico Universidade de Lisboa, Portugal
November2021

Keywords

thermal comfort, building simulation, simplified models, extreme weather events, energy poverty

Abstract

For several years it has been possible to observe a strong correlation between the mortality rate and Extreme Meteorological Events (EWE) in Portugal. This relationship results from low levels of thermal comfort in buildings, a result not only of the building's constructive characteristics, but also of the meteorological conditions where it is located. The problem of energy poverty is further aggravated by the aging of society and low average income, which translates into insufficient levels of air conditioning. Some attempts to predict EWE have been made, but considering either heat waves or cold waves. The RELIABLE project aims to create a model for forecasting EWE for the whole year. Thus, within the scope of this thesis, building energy simulation models are compared in terms of reliability of results and calculation times. This research work compared Energy Plus, the 5R1C model and optimization tools, in particular Random Forest and Neural Networks. The models considered the thermal characteristics of nine existing building archetypes in Montalegre, Portugal. Simulation results were compared with real temperature measurements. This work shows that the simplified 5R1C model serves the purpose of forecasting EWE effect on building indoor temperature, due to the short computation time and its suitability for large buildings.

1 Introduction

1.1 Context / Overview

For several years now a correlation between the death rate and temperature in Portugal is observed. The peak death rates can be observed in winter and summer periods even though Portugal has one of the mildest winter climates [1] [2]. The ratio between Excess Winter Deaths and Heating Degree Days for Portugal is 5.7 where for much colder countries like Finland it equals 4.0 and for Denmark 4.8 [3] [4]. Such a high death rate is explained by the fact that the buildings do not meet minimum thermal requirements [5]. This problem is being swelled by the ageing society and low average income. Elderly people have less ability to handle inadequate indoor temperatures which is the effect of fuel poverty. There is an unofficial definition of fuel poverty: “anyone who meets in its housing particular difficulties to have necessary energy to meet its basic energy needs because of the inadequacy of its resources or of its housing conditions” [6] [7]. Such a situation falls within the scope of energy poverty. Energy poverty has no universal definition since energy poverty is displayed differently in different countries. As most European countries have no official definition for the term “energy poverty”, this state is often described as “inability to keep homes adequately warm” [8].

1.2 State of the art

From the reasons presented in the previous section stems the need for a quick and nationwide solution. A solution that is cheap but does not require physical interference with the building. Over the past few years, two projects were developed in Portugal: ÍCARO and FRIESA. Project ÍCARO, developed by INSARJ, is a nationwide project which detects and monitors heatwaves and determines if the forecasted number of deaths is going to exceed the expected number of deaths [9]. Initially it was created based on time series statistical model using dynamic regression and thresholds which were calibrated for Lisbon. The calibration was done concerning 1981-1991 heatwaves. Every day a report is made with the forecast for the next three days. The project is run only in the period between May and September and thus it is used only for heatwaves prediction. Its downside is the very low resolution because it divides Portugal into 5 regions. The FRIESA project is in the development phase and it is used in the period from November to March. Currently, FRIESA covers only two districts: Porto and Lisbon [10].

1.3 RELIABLE project

The projects briefly described in point 1.2 have some major flaws. Namely very low resolution and the need to use two different applications for different parts of the year. Also, these applications show only outside temperature or the predicted excess of deaths. Project RELIABLE has been started to develop a tool – real-time internet application – which will provide reliable data about indoor temperature for different types of buildings across Portugal [11]. It will determine zones of dangerously high indoor temperatures in the summer and dangerously low indoor temperatures in the winter. The biggest advantages of this project are public availability through a website all year long and transparency of the data and the ability to monitor the well-being of the citizens in danger zones. Such an ambitious enterprise requires a lot of computational power. simulation for the whole of Portugal requires weather data from hundreds of points. The administrative division of Portugal is 18 districts which are divided into municipalities. 308 municipalities are further divided into 3092 parishes. Simulation for 2 weeks ahead for every hour for 9 archetypes for the 308 municipalities, assuming 1second for one iteration, would require 931392 iterations. Such a number of iterations - roughly calculating - would take almost 11 days. Therefore there is a need for a faster but reliable model

1.4 Aim of this work

This work aims to compare methods for indoor temperature predictions. Build a 5R1C model, test Machine Learning (ML) models: Random Forest (RF) and Artificial Neural Networks (ANN) capabilities in this area. The models will be created based on energy certification data from buildings from different periods of construction - archetypes. The results from the 5R1C, RF and ANN models will be then compared with results from EnergyPlus software and with real data. In theory, the 5R1C model should provide a faster computation of the results. In theory, the 5R1C model should provide a faster computation of the results. If the accuracy of 5R1C, ML or ANN is reasonable and the computations times faster or comparable with Energy Plus, one of these models will be used for further proceedings in the RELIABLE project.

2 Models description

2.1 5R1C model

The model described in international standard EN ISO 52016-1:2017 consisting of five resistances and one capacity (5R1C) is proposed as a simplified way for simulation of thermal behaviour of the building zone given the external weather data. The envelope is described as an equivalent electric circuit consisting of five resistances and one capacity (Figure 1) which solution requires the employment of limited input parameters. Roberto Bruno et al. [12] investigated the performance of the 5R1C model for multi-house building and concluded that over bigger periods of time the fluctuations and time-shift between simplified model and real data are too high especially in houses with high thermal capacities. The case described in this paper is different since data from dozens of buildings are averaged which already theoretically attenuates the possible fluctuations but also introduces some degree of inaccuracy.

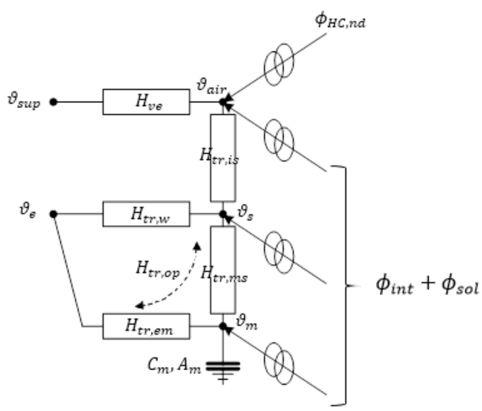


Figure 1 5R1C model represented as an electric scheme

Data for calculation – materials and thermal properties – has been taken from EnergyPLUS library. Dimensions of the envelope and surfaces have been taken from the 3D model created for EnergyPLUS based on Energy Performance Certificates. 5R1C model takes into account direct and diffuse radiation. The algorithm calculates the angle at which the sun shines on the windows of the dwellings based on the latitude and longitude of the location of interest and the altitude of the sun at every hour. Latitude and longitude have been provided manually as follows: 41.828 N, - 7.787 W. The simulation was conducted for a sample weather file for Braganca municipality for the year 2019.

2.2 Energy Plus

A free, open-source program developed by the United States Department of Energy (US DOE). Used mainly for the simulation of energy and water consumption. One of the main advantages of this program is user-defined, sub-hourly time steps. Inputs and outputs are given in text files. The programme simulates air movements and interactions between zones. Energy Plus is using the Transfer Function Method (TFM) for the calculation of thermal loads. This function has been adopted by ASHRAE and is regarded as one of the most accurate methods for the calculation of heating and cooling loads [13]. The TFM method is first computing heat gains from walls, windows, floors and roofs, and then by multiplying the heating loads by heat transfer coefficients the cooling loads are calculated [13].

2.3 Machine Learning

In this work machine learning has been implemented in two approaches. Random Forest Regression and Artificial Neural Network has been consider. This has been done for two reasons. First, to verify if Machine Learning can be applied for simulation of EWE using general information about archetypes and second, as an attempt to make the model more autonomous and simplify the analysis.

2.3.1 Random Forest Regression

Machine Learning is a part of Artificial Intelligence branch. Random Forest is a supervised machine learning algorithm that is constructed from Decision Trees algorithms. It is applied in various industry fields such as e-commerce, health care and banking to predict outcomes or behaviour [14]. Random Forest is used to solve regression and classification problems. It utilizes ensemble learning which is a technique that employs multiple times one algorithm in multiple combinations to provide a solution to a problem. Random Forest is constructed as blackbox which means that the internal workings are unknown which is different to linear regression where one can trace te subsequent steps [14].

2.3.2 Artificial Neural Networks

A Neural Network is made of artificial neurons – perceptrons as suggested by M. Minsky and S. Papert [15] – that receive and process input data. Artificial Neural Network (ANN) consists of hidden layers of neurons. The most basic ANN has 3 layers: input layer, hidden layer and output layer. The process starts when input data is fed into the model. Data is then processed via its layers to provide desired output [16]. Every input signal has an assigned weight. Weights are the point in which ANN learn by calculating a function called "cost function" which compares the actual output value with predicted one and adjusting the weights. The goal is to minimize the cost function.

3 Methodology

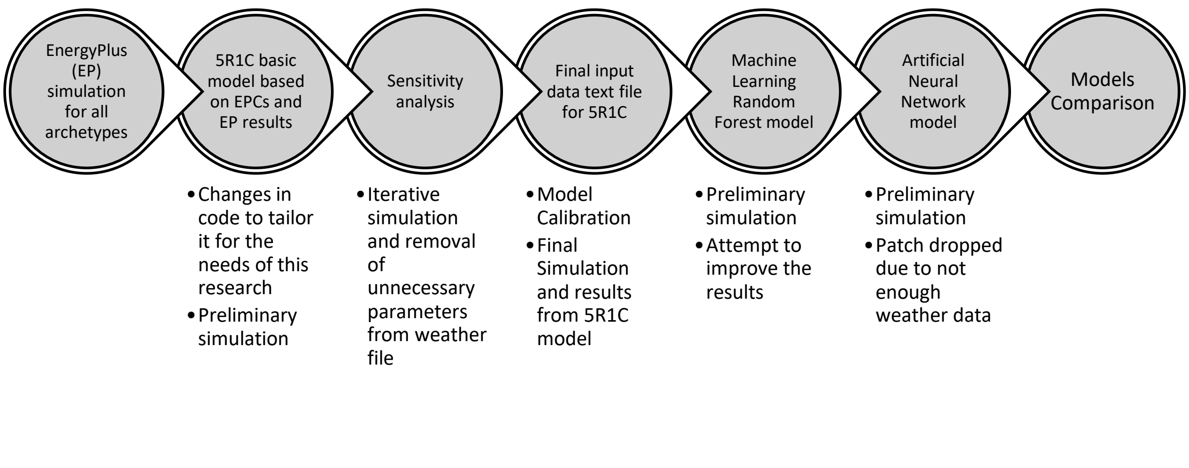
3.1 Montalegre case study. Energy Performance Certificates into models

The town of Montalegre has been chosen for the case study an testing the first prototype. It is a town located on the very North of Portugal in district Vila Real along the border with Spain. In Montalegre a monitoring campaign has been conducted to collect reference data about the indoor temperature. Energy Performance Certificates (EPCs) for buildings in this town has been collected by Portuguese Energy Agency (ADENE). Following the division per construction period followed by the ADENE based on construction booms, new milestones in building regulations, nine construction periods were considered: 1919-1945; 1946-1960; 1961-1970; 1971-1980; 1981-1990; 1991-1995; 1996-2000; 2001-2005; after 2006. As it is impossible to simulate each building of large urban area, there is a need to use this simplified approach in building energy simulation, usually called building archetypes [17] [18]. Such solution introduces an inaccuracy error. The RELIABLE project aims to forecast extreme weather events (EWE) therefore, there is no need for extreme accuracy

3.2 The basic model

When performing a building energy simulation, it is crucial to define several inputs regarding the building geometry, construction solutions, schedules of occupancy and internal gains of the building. To define the geometry of the building model (one per archetype) EPCs provide information about the area and height of the dwelling as well as the number of floors. From this information it is possible to calculate the width of square box and draw the building zone. Nine zones/cubes has been modelled in Google SketchUp with materials and thermal properties assigned to the surfaces according to the specification for given archetype. Then a methodology has been applied as shown on the Table 1.

Table 1Flow chart of the methodology used in research



3.3 5R1C

The basic code (application) for the model has been accessed through *github.com*. It has been created by the Department of Architecture and Building Systems of the ETH Zürich [19]. The code has been than tailored for the needs of the research. A loop has been created to allow the program to go through all archetypes one by one. After preliminary simulation the results were not satisfying – there was no distinction between archetypes. Therefore, there was the need to conduct a sensitivity analysis to better understand the simulation parameters and outputs.

3.4 Sensitivity Analysis

Sensitivity analysis has been conducted to get a sense of how various input variables impact the output. It has been done by running the simulation nine times for the same weather period, with identical values for parameters besides the one under investigation, which has been changed slightly in every simulation. Sensitivity analysis showed that introduction of infiltration parameter improves the results.

3.5 Data input procedure in models

To provide data into EP it is needed to specify the features of each component. The thermal properties of walls and windows. The weather file is supplied by specifying the patch to the file. The output parameters from the software can be specified a priori simulation and can be collected in a CSV file. Providing data into the RC model is quite different. RC model is a code written in Python, therefore the data is read from an external CSV file and when once the variables are set in the code only the csv file can be changed to play with the model output. The unique parameters that have to be provided to the program are the geometry of the envelope namely the areas of floors, walls and windows and the thermodynamic parameters: heat transfer coefficients for the walls and windows and thermal capacitance per floor area of the dwelling. Both Machine Learning methods require a data set for training with specified input parameters and the output they produce. After training only a weather file is needed to conduct the simulation.

3.6 Model calibration

To compare the results two statistical indicators were used: Mean Bias Error (MBE) and Coefficient of variation of the Root Mean Square Error (Cv(RMSE)). MBE measures how closely the simulated data coincide with the measured data. It is an overall bias measure of the data. The final values are compensated (reduced) by negative values if present. Due to this compensation effect, MBE is often represented together with Cv(RMSE). Cv(RMSE) measures either the variability between measured and simulated data or the goodness-of-fit of the model. It specifies overall uncertainty in prediction reflecting the size of the error. It is always positive and it is insensitive to the compensation effect. The lower Cv(RMSE) the better the calibration of the model. Acceptable tolerances defined by ASHRAE [20].

3.7 Machine Learning model: Random Forest

The next step, after having the 5R1C model tuned up, was to build a Machine Learning (ML) model based on the output from 5R1C. Machine Learning models need to be trained before being able to produce reliable results. To do that a set of carefully selected features, together with the output they produce, needs to be provided.

3.8 Feature selection

Feature selection has been conducted using P-value scores. An assumption was made – the null hypothesis - that given combination of variables does not have any effect on dependent variable. Than P-value is calculated and variable with P-value higher than the arbitrary selected threshold is removed from the dataset. Abovementioned steps are repeated until a set of features with which the overall performance of the model is in the acceptable range.

3.9 Artificial Neural Network

Due to limited time resources and results far from acceptable level a conclusion was made that data sample of one year is not enough to properly train the ANN, the investigation of ANN usage for EWE prediction has been stopped and the results will not be presented here.

4 Results

4.1 Discussion about the results

The results from the three models were compared for the period from 12.05.2020 13:00 to 24.06.2020 23:00 for which the real measurement was conducted. The house in which the measurement using HOBO was done is of the archetype 1946-1960. The chart presented below (Figure 1) shows a comparison between models and real measurement for archetype 1946-1960. What is interesting is that Energy Plus reacts in a much more "real" way – influence of thermal capacity – to the changes of outside temperature shown with a thin purple line. Moreover, Energy Plus did not react as drastically to the sudden increase of temperature – at the right-hand side of the chart – as the 5R1C model did.

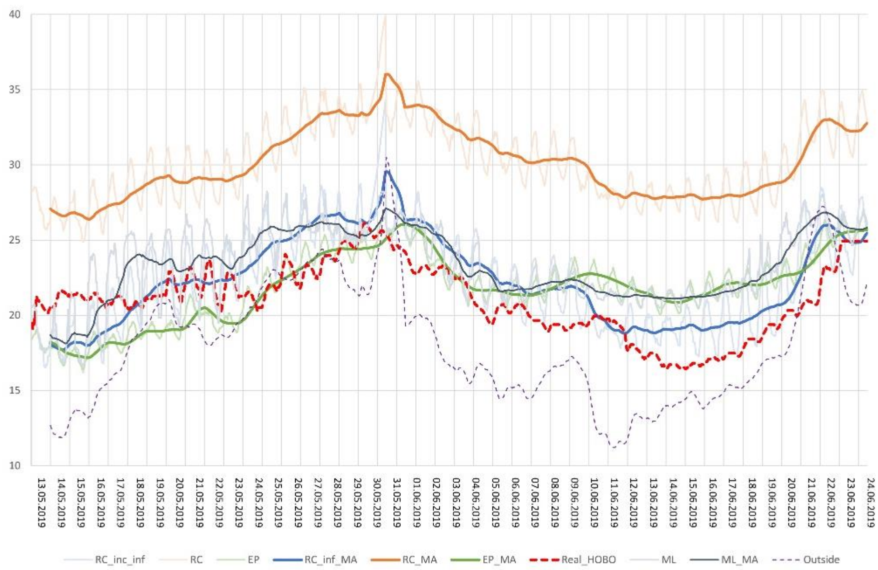


Figure 1 Comparison of models results with HOBO measurement and outdoor temperature

Comparing the results from different models with the real measurements a really good fit can be observed. Of course, there are some fluctuations between models but the overall outcome is very satisfactory. The ML Random Forest model also did a very good job although the beginning of the simulation was quite different from the two other models.

4.2 Cv(RMSE) scores comparison

Cv(RMSE) scores were calculated at different stages during the research. The comparison is presented on Table 2 for 2 archetypes when comparing models to the real data. Then a comparison between EnergyPlus and 5R1C is presented for the same two archetypes.

Table 2 MBE and Cv(RMSE) scores of comparison of the model with real measurement

`	EnergyPLUS		5R1C - 2 floors		5R1C - 1 floor	
	1919-1945	1946-1960	1919-1945	1946-1960	1919-1945	1946-1960
MBE	-11.06%	-4.02%	-12.29%	-12.35%	-18.08%	-18.07%
Cv(RMSE)	18.73%	12.90%	15.49%	15.60%	21.46%	21.43%

The scores from EnergyPLUS for the 1946-1960 archetype indicate that here the performance was the best and within the ASHRAE guidelines. Scores for the 5R1C model are different and not quite good. There is a distinction in the 5R1C model because it was first calculated as a single-floor building and then as a two-floor building as the real dwelling. In EnergyPLUS 2-floor construction is set by the 3D model configuration. In 5R1C this issue was resolved by doubling the wall and floor areas of the dwelling and increase in thermal capacity. Resulting score for 5R1C is satisfying and is on the verge of ASHRAE guidelines when it comes to Cv(RMSE) but not when MBE value is considered.

4.3 Computing time

All the simulations were done on a personal computer Asus TravelMate-P with Intel Core i7-5500U CPU at 2.40GHz clock speed, 8GB RAM and 64bit operating system. The computation times (Table 3) will depend on the hardware used. Times are given for whole year simulations for one archetype. 5R1C simulated all archetypes one after another.

Table 3 Computing times of the different simulation tools used

Model	Average Time
Energy Plus	~8.41 s
5R1C	~4.55 s
ML Random Forest	~5.58 s

5 Conclusions and recommendations

In this work different tool were used to simulate the indoor temperature of nine building archetypes. All of the models show potential, but some of them (like Machine Learning) need a bigger amount of training data and more time and effort to properly train and calibrate. Simplified model 5R1C has proven to be a good substitute to more sophisticated model like EnergyPlus. 5R1C showed outstanding performance in result-to-computing-time ratio. Energy Plus gave – as expected – the most accurate results. The natural choice when considering automatization is Machine Learning. Random Forest gave good results but not very reliable nor consistent due to shattered sample of data which was the effect of small sample data and the need for dividing the sample into train and test samples. ANN would be a great choice, but for that a much bigger sample of data is needed to train the model properly.

For the purpose of RELIABLE project 5R1C is good choice since it produced satisfactory results confirmed by calibration scores MBE and Cv(RMSE). In the future switch to ML or even ANN should be considered under the condition that enough data for training of the models is gathered. The almost twice shorter time of computation of 5R1C model shows that there is potential to conduct nationwide simulation

6 References

[1] J. Ballester, “European seasonal mortality and influenza incidence due to winter temperature variability,” *Nature Climate Change,* p. 927–930, 2016.

[2] T. F. R. J. Southgate, T. Waite, R. Harrell, S. Kovats, A. Bone, Y. Doyle and V. Murray, "Excess Winter Deaths in Europe: a multi-country descriptive analysis," *European Journal of Public Health,* p. 339–345, 2015.

[3] S. G. Simões, V. Gregório and J. Seixas, "Mapping fuel poverty in Portugal.," in *1st Energy Economics Iberian Conference, EEIC | CIEE*, Lisbon, Portugal, 2016.

[4] C. Liddell, C. Morris and H. Thomson, "Excess winter deaths in 30 European countries 1980–2013: a critical review of methods.," *Journal of Public Health, Volume 38, Issue 4. Oxford Academic,* 2016.

[5] R. Oliveira, C. C. Antonio, H. Santos and E. O. Fernandes, "THE PORTUGUESE HOUSING STOCK: AND INITIAL ANALYSIS OF THE NATIONAL ENERGY CERTIFICATION DATABASE," in *Proceedings of the 7th International Conference on Mechanics and Materials in Design*, Albufeira/Portugal, 11-15 June 2017.

[6] S. M. Magalhaes and V. M. Leal, “Characterization of thermal performance and nominal heating gap of the residential building stock using the EPBD-derived databases: The case of Portugal mainland.,” 2013.

[7] J. D. Healy, “Housing, Fuel Poverty and Health: A Pan-European Analysis.,” 2004.

[8] "An official website of the European Union," [Online]. Available: https://ec.europa.eu/energy/eu-buildings-factsheets-topics-tree/energy-poverty_en. [Accessed 27 June 2021].

[9] S. P. d. Silva, 28 October 2021. [Online]. Available: http://www2.insa.pt/sites/INSA/Portugues/AreasCientificas/Epidemiologia/Unidades/UnInstrObser/Paginas/ICARO.aspx.

[10] S. Antunes, "Modeling and forecasting the effect of extreme cold in the population health: the basis for the development of a real-time warning system," IPMA, 2014.

[11] IST-ID, "RELIABLE - REAL-TIME RISK WARNING SYSTEM FOR BUILDING OCCUPANTS DURING EXTREME WEATHER EVENTS," 2021. [Online]. Available: https://reliable.tecnico.ulisboa.pt/. [Accessed 27 07 2021].

[12] R. Bruno, G. Pizzuti and N. Arcuri, "The prediction of thermal loads in building by means of the EN ISO13790 dynamic model: a comparison with TRNSYS," *Energy Procedia 101,* pp. 192-199, 2016.

[13] O. M.A., Al-Rabghi and K. M. Al-Johani, "Utilizing Transfer Function Method for hourly load calculations," *Energy Convers. Vol. 38,* pp. 319-332, 1997.

[14] O. Mbaabu, "Section," 11 December 2020. [Online]. Available: https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/. [Accessed 15 07 2021].

[15] M. Minsky and S. Papert, Perceptron: An Introduction to Computational Geometry, 1969.

[16] J. Nduati, "Section," 25 October 2020. [Online]. Available: https://www.section.io/engineering-education/introduction-to-neural-networks/. [Accessed 28 July 2021].

[17] S. Claudia, C. Carlos, P. André, Y. Maribel and F. Paulo, "An urban building database (UBD) supporting a smart city information system.," *Energy and Buildings,* vol. 158, pp. 244-260, 2018.

[18] C. Sousa-Monteiro, F. Causone, S. Cunha, A. Pina and S. Erba, "Addressing the challenges of public housing retrofits. 9th International Conference on Sustainability in Energy and Buildings," in *9th International conference on sustainability in Energy and Buildings*, Chania, Crete, Greece, 2017.

[19] P. e. a. Jayathissa, "Optimising building net energy demand with dynamic BIPV shading.," *Applied Energy 202,* pp. 726-735, 2017.

[20] G. R. Ruiz and C. F. Bandera, "Validation of Calibrated Energy Models:," *Energies,* vol. 10, 2017.